

Real-Time LiDAR-Based Stair and Elevation Change Detection on Consumer Mobile Hardware: A Single-User Study

Nikhil Herdt

herdtnikhil@gmail.com

University of Iowa

Terry Braun

University of Iowa

Tyler Bell

University of Iowa

Stephen Russell

University of Iowa

Article

Keywords: stair detection, LiDAR, assistive navigation, mobile computing, visual impairment, depth sensing

Posted Date: August 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10154859/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

For individuals who are blind or visually impaired (BVI), undetected stairways, curbs, and sudden elevation changes represent a leading cause of fall-related injury. Herein we present and evaluate a computationally efficient geometric algorithm for detecting stair and elevation changes in real-time implemented entirely on an LiDAR-equipped consumer smartphone (iPhone 17 Pro Max, iOS 26.2). The algorithm uses a set of eight downward- and forward-facing raycasts projected into an ARKit (Artificial Reality developer Kit)-reconstructed mesh and performs line-segment angular analysis at 60 Hz within a 16.7 ms per-frame budget, without requiring cloud processing, GPU inference, or custom hardware. A single-user evaluation across 120 trials spanning eight real-world hazard locations, five step heights (4.5 to 7.5 in), four stair grades (22 to 33 degrees), varied lighting conditions (indoor artificial, outdoor daytime, and outdoor nighttime), and both ascending and descending approaches yielded an overall detection rate of 92.5% with zero false positives and a mean alert distance of 2.31 ft (SD = 0.91 ft). Detection reliability generally increased with step height, with more consistent detection performance above six inches of step height. Descending approaches produced lower detection rates and shorter warning distances than ascending approaches, a limitation intrinsic to the geometry of depth-based sensing. These results demonstrate that accurate, low-latency elevation hazard detection is achievable on commodity mobile hardware, with implications for the design of accessible, low-cost assistive navigation tools.

1 Introduction

Falls from undetected stairways, curbs, and sudden elevation changes represent a disproportionate source of serious injury for individuals who are visually impaired. Unlike collisions with walls or furniture, a missed downward elevation change provides virtually no recovery opportunity once a step is taken into open space [1, 2]. Traditional mobility aids such as the white cane are effective at detecting ground-level obstacles but provide limited advance warning of floor discontinuities, particularly when descending. Wearable electronic travel aids have been developed to supplement these limitations, yet many rely on custom hardware, cloud connectivity, or specialized computing platforms that raise cost, weight, off-device computational access and social acceptability barriers that constrain real-world adoption [3, 4].

The widespread availability of time-of-flight LiDAR sensors on consumer smartphones since 2020 has opened a practical path toward on-device, real-time depth sensing for assistive applications. Prior work has demonstrated that smartphone LiDAR enables continuous three-dimensional environment reconstruction at 60 Hz within a range of 0.1 to 5 m [5, 6]. However, reliable detection of floor-level hazards such as stairs and curbs from a hand-held mobile device remains an open problem. Existing stair detection algorithms based on point cloud plane-fitting, RANSAC-based extraction, and deep learning classification each impose per-frame computational costs that exceed the 16.7 ms budget available at 60 Hz, and several additionally require the sensing platform to remain stationary during analysis [6, 7, 10, 12]. These constraints are acceptable for the robot platforms and controlled laboratory settings such methods were designed for, but they preclude continuous operation on a general-purpose

mobile CPU that must concurrently sustain spatial audio, haptic feedback, and voice recognition workloads. Section 2 reviews these methods and their specific limitations in detail.

This paper presents and evaluates a raycast-based geometric algorithm for stair and elevation change detection embedded within SensoryScape, a native iOS assistive navigation system that we designed. The algorithm operates entirely on device within the existing 60 Hz spatial mapping cycle, using downward- and forward-facing raycasts projected into an ARKit mesh. Rather than performing the dense mesh segmentation or plane fitting that underlies prior approaches, it constructs line segments between adjacent raycast intersection points and classifies hazard type from the angular relationships among those segments and intercepts. This design is motivated directly by the computational landscape reviewed above: the algorithm must fit within a 16.7 ms per-frame budget on a general-purpose mobile CPU while sharing that budget with spatial audio, haptic feedback, and voice recognition, a constraint that excludes all known point cloud and learning-based methods. In this report, we describe and evaluate a structured single-user empirical application of the algorithm across 120 real-world trials, characterizing detection performance as a function of step height, stair grade, lighting condition, and approach direction. The evaluation identifies practical reliability thresholds, characterizes the geometric constraints that drive false-negative patterns, and compares the results to prior depth-based stair detection methods.

2 Background and Related Work

2.1 Stair Detection Using Depth Sensors

Automated detection of staircases using depth sensors has been attempted using several algorithmic families. Range-sensor systems, including ultrasonic and infrared rangefinders deployed on early wearable navigation aids, provided binary proximity warnings but lacked the spatial resolution to distinguish step-like structures from flat floors [8]. With the introduction of structured-light and time-of-flight depth cameras, plane-fitting algorithms became the dominant approach: a set of candidate planes is extracted from the point cloud and step-like geometric relationships among pairs of planes are tested [6, 10]. RANSAC-based variants are widely used for this plane extraction step but are non-deterministic and do not guarantee a fixed time bound – the algorithm randomly samples candidate plane points and may produce different results across runs, with computation time varying unpredictably depending on scene complexity. This makes RANSAC-based methods incompatible with safety-critical real-time systems that must respond within each 16.7 ms frame [10]. The point cloud method reported by Westfechtel et al. achieves accurate geometry estimation among published work but requires the sensing platform (the physical device housing the depth sensor, including its position and orientation in three-dimensional space) to remain static during analysis and takes between 623 ms and 4282 ms per detection (or to put in context it takes 0.6 to 4.3 seconds for detection with a 1.8 second (avg) time to step off the edge) depending on stair type. This design reflects the intended service robot deployment context rather than a limitation of the approach itself, but it places the method outside the operational envelope required for continuous ambulatory mobile use [10]. Sriganesh et al. address this speed

limitation with an edge-based point cloud pipeline that reduces detection time to a mean of approximately 20 ms across stair types on a Boston Dynamics Spot and an unmanned ground vehicle, each equipped with a dedicated Jetson AGX Xavier embedded GPU [12]. While this represents an improvement of approximately two orders of magnitude over Westfechtel et al. and approaches real-time feasibility on robot hardware, it still consumes the entirety of the 16.7 ms per-frame budget available at 60 Hz, and does so on dedicated, purpose-built GPU co-processor hardware rather than on a general-purpose mobile CPU. On a smartphone, which must simultaneously sustain spatial audio rendering, haptic feedback generation, voice command processing, and ARKit mesh reconstruction, no computational reserve remains for even this optimized point cloud methodology.

Learning-based approaches train convolutional classifiers or depth-image segmentation networks on labelled stair datasets, achieving detection rates above 92% in controlled settings [7]. However, inference latency and GPU memory requirements make real-time deployment on resource-constrained mobile hardware challenging if not impossible, particularly because the neural model must share computational resources with parallel sensing and feedback calculation requirements [3].

A smaller body of work addresses stair detection specifically on smartphone LiDAR, most directly the EchoSee application by Schwartz and Bell [5]. EchoSee is discussed alongside its predecessor system LEOBelt in Section 2.2, in the context of the design lineage that motivated the present work. The limitations identified in that line of systems directly informed the multi-ray geometric approach developed here.

2.2 Predecessor Systems and Motivation

SensoryScape was developed as a successor to two prior assistive navigation systems within our research group. The Low-Vision Enhancement Optoelectronic Belt (LEOBelt) is a wearable haptic belt using a waist-mounted Intel RealSense D435i depth camera and eight torso vibration motors. User studies demonstrated a 44% reduction in navigation errors and a 77% increase in self-reported confidence for participants simulating advanced retinitis pigmentosa [4]. However, the waist-height camera mounting location resulted in limitations of detecting floor-level hazards and stairs proved particularly difficult due to the uni-planar field of view of the sensor at the critical target distance. EchoSee is a mobile augmented reality design that is reliant on Unity (Unity Technologies, San Francisco, CA, USA) that uses iPhone LiDAR for spatial mapping and delivers directional audio cues through spatialized sound [5]. User studies showed a 39.8% reduction in obstacle contacts compared with unassisted navigation. The Unity-based architecture, while enabling rapid development of the spatial mapping and audio feedback pipeline, introduced computational constraints that limited the density of raycasts available per frame. The single downward ray configuration was found to have limited reliability for stair detection across varied geometries, an observation that directly motivated the multi-ray approach developed in the present work [5].

Our current approach, SensoryScape, addresses these limitations by redesigning the architecture in native Swift (Apple Inc., Cupertino, CA, USA), which removed the constraints on per-frame raycast density

and enabled the expanded ray configuration central to this work. The stair detection algorithm described herein represents one of the primary technical and usability advances over both predecessor systems.

2.3 Consumer LiDAR Characteristics

The Apple LiDAR scanner integrated into iPhone Pro models uses direct time-of-flight (dToF) infrared sensing, projecting a sparse grid of approximately 576 sample points at up to 60 Hz with an effective range of 0.1 to 5 m under optimal conditions [9]. Depth accuracy of detection is strongest for matte, diffusely reflective surfaces; performance degrades on specular or reflective materials (glass, polished metal) and for very dark surfaces that absorb infrared energy. Independent evaluations of consumer iPhone LiDAR confirm these surface-dependent performance characteristics. Luetzenburg et al. [14] measured absolute accuracy of ± 1 cm for objects with features larger than 10 cm under controlled conditions but documented measurable degradation for smaller geometric features, consistent with the shallow step geometry encountered in this study. The Apple LiDAR receiver is built around a SPAD (Single-Photon Avalanche Diode) array – a grid of detectors each capable of registering individual photons by triggering an electron avalanche upon photon arrival. This extreme sensitivity enables the precise photon-arrival timing that dToF ranging requires but also makes the sensor susceptible to interference from any ambient photon source, including solar infrared radiation. Specular surfaces degrade depth measurements through two mechanisms: first, glass and polished metal deflect the infrared pulse away from the receiver rather than returning a diffuse echo, producing data voids or erroneous returns; second, on SPAD-based dToF sensors, high-reflectivity targets can saturate individual detector pixels, causing pile-up in the photon-arrival histogram and corrupting range estimates [15]. Dark surfaces attenuate the returning infrared signal below the detection threshold of the SPAD array, producing sparse or absent depth returns [14, 15]. Tondo et al. [16], characterizing iPhone 13 Pro LiDAR across lighting conditions, found that depth map standard deviation increased substantially at object boundaries in low-light conditions relative to normal ambient illumination, underscoring that the system's fusion of LiDAR with RGB imagery means dark environments penalize depth confidence even when the infrared pulse itself is unaffected. In outdoor environments, background solar infrared introduces noise that degrades the signal-to-background noise ratio (SBR) of the SPAD detector array. Solar irradiance at ground level contains substantial near-infrared power at wavelengths overlapping the sensor's 940 nm operating band, generating a continuous stream of background photons that compete with the reflected laser pulse in the dToF histogram. Morimoto et al. [17] demonstrated that high ambient background light degrades SPAD-based dToF ranging performance and reduces effective measurement range, a finding corroborated by Tondo et al. [16], who showed that the iPhone LiDAR accuracy degrades under outdoor sunlight conditions relative to indoor baselines. This effect can reduce effective range and increase depth variance, particularly for small geometric features such as shallow steps, where the vertical depth gradient across the step riser is small and the signal margin is already thin. ARKit fuses LiDAR depth sensing with camera imagery and inertial data to generate incrementally updated triangle meshes with semantic surface classification such as “floor”, “wall”, “ceiling”, etc., providing the environmental representation on which the algorithm operates [9].

3 Algorithm Design and Implementation

3.1 Raycast Architecture

The detection algorithm operates within the SensoryScape spatial mapping cycle, which maintains a continuous ARKit scene reconstruction session at 60 Hz. Eight raycasts are projected each frame from the current device position into the reconstructed mesh. Three horizontal raycasts (forward at 0 degrees, left at -30 degrees, right at + 30 degrees) and one upward raycast (20 degrees elevation) support general obstacle and audio feedback functions. The remaining four raycasts are dedicated to floor-level awareness and hazard detection: one at 30 degrees below horizontal (Down1), one at 60 degrees below horizontal (Down2), the Lower ray at approximately 75 degrees below horizontal, and a silent Floor Detector ray directed steeply downward (Fig. 1). ARKit's spatial acceleration sensors achieve raycast query times of 0.1 to 0.5 ms per ray, well within the 16.7 ms per-frame budget at 60 Hz.

The system maintains a calibrated floor height estimate derived from a rolling median of recent Floor Detector and downward ray intercepts (or intersections) that carry an ARKit “floor” semantic classification. The median value for each ray over a rolling window (30 frames, equivalent to 0.5 seconds at 60 Hz) is preferred over the mean to reduce sensitivity to transient mesh reconstruction noise. A confidence value is derived from the variance of the sample buffer. Floor height can also be set manually from the initial camera height in environments where semantic classification performs poorly.

3.2 Descending Elevation Change Detection

Descending hazards (stair edges, curbs, platform drops) are detected by evaluating the geometric relationships among three line segments constructed from pairs of raycast intercept points for each frame. The segment between the Down2 and Floor Detector intercepts is evaluated for downward slope away from the user: if the dot product of the segment direction vector with the forward camera vector is negative and the segment subtends a minimum vertical span of 4 cm, a descending candidate is flagged. The segment between the Forward and Lower intercepts captures the lip geometry at the edge of a descending step, where the Forward ray strikes the upper surface and the Lower ray intersects the drop face. The segment between Down1 and Down2 provides an additional descending indicator when it slopes away from the user (Fig. 2).

For a segment to qualify as an elevation change candidate, its tilt from vertical must fall within 5 to 50 degrees, a range characteristic of step and curb geometry as opposed to flat floors (near 90 degrees) or vertical walls (near 0 degrees). A general case additionally requires that at least two independent ray pairs satisfy the geometric criteria simultaneously, reducing false positives from single-ray noise.

To filter out transient mesh reconstruction artefacts, a candidate detection must persist across four consecutive frames before an alert is issued. Alerts are subject to a cooldown period (a fixed inhibition window of 3 seconds following each issued alert, implemented as a frame counter reset) to prevent repeated warnings as the user continues to approach the same hazard.

3.3 Ascending Hazard and Staircase Detection

Ascending steps and staircases are detected using the same line-segment approach applied to a different ray pairing. The algorithm constructs two segments: a lower segment between the Lower and Forward intercepts, and an upper segment between the Forward and Up intercepts. The angle of each segment from vertical is computed and their relationship determines the classification.

If the upper segment is within 25 degrees of vertical, the configuration is interpreted as a wall, and no elevation change alert is issued; walls are communicated to the user separately by the general obstacle audio feedback system. If the upper segment is 65 degrees or more from vertical (approximately horizontal) while the lower segment falls within 10 to 50 degrees from vertical, the configuration is classified as a tripping hazard: the lower segment captures the step riser and the upper segment captures the tread surface above it. If both segments fall within the 10 to 50-degree range and their angles agree within a tolerance of 15 degrees, the geometry is classified as a staircase (Fig. 3). Angular similarity between both segments indicates a repeating step pattern rather than a single isolated step (Fig. 4). A confidence value is derived from the deviation of the measured angles from a nominal stair angle of 34 degrees.

3.4 Alert Generation

When either a tripping hazard or staircase is detected, the system generates a spoken alert via AVSpeechSynthesizer subroutine indicating hazard type and angular offset from the user's current heading. A three-pulse haptic pattern is delivered through Core Haptics subroutine on the iPhone. The three-pulse pattern is perceptually distinct from the continuous proximity vibrations used for general obstacle distance feedback, enabling users to distinguish elevation-change warnings from routine distance cues without requiring visual confirmation (capability which an impaired user may not have). Ascending and descending alert channels are rate-limited independently. Users can select from four alert modes (descending only, ascending only, both, or none) through a high-level unified configuration setting, enabling the system to be tuned to navigation environment and reducing alert fatigue, especially during periods of minimal or non-ambulation. All geometric thresholds (angle bands, persistence frame count, minimum segment span, and elevation change depth) were determined empirically to balance detection sensitivity against false positives arising from normal gait dynamics and scanning motion.

4 Evaluation Methodology

4.1 Apparatus and Test Conditions

Testing was conducted using an iPhone 17 Pro Max running iOS 26.2. This device represents most capable currently- and commercially available consumer mobile hardware equipped with LiDAR sensing capabilities. Testing spanned eight real-world hazard locations: four indoor staircases under controlled artificial lighting, three outdoor staircases, and one outdoor curb. Outdoor locations were tested under

both daytime and nighttime conditions, yielding a total of twelve distinct location-by-condition pairings. Step heights ranged from 4.5 to 7.5 in, step lengths from 10.5 to 12.75 in, and stair grades from 22 to 33 degrees across the seven staircase locations.

4.2 Protocol

Each trial followed a standardized protocol designed to reduce variability from inconsistent device handling. Device scanning behavior was controlled using lateral sweeps combined with slow, steady forward motion. A single user conducted all trials; the participant had normal vision and was trained on the protocol prior to data collection. Each unique combination of location, direction (ascending or descending), and lighting condition was tested with five repeated trials, yielding 120 trials in total.

A trial was classified as successful (detection) when the algorithm generated a hazard warning before the user's feet reached the hazard boundary. Trials in which the user reached the hazard without receiving a warning were recorded as misses. False positives were defined as alert activations occurring with no hazard present within the sensing field. For all successful detections, the distance from the user to the nearest hazard boundary at alert onset was recorded using a tape measure to characterize available reaction time.

4.3 Limitations of Single-User Design

This evaluation was conducted by a single user with normal vision rather than participants from the target population. This design reflects the exploratory, technical focus of the study: the goal is to characterize the geometric and sensing-based constraints of the algorithm across environmental conditions, not to assess usability or safety for Blind or Visually Impaired (BVI) individuals. Single-user trials eliminate inter-participant variability in scanning behavior, isolating algorithm performance from human factors. The use of a sighted participant conducting controlled sweeps is consistent with prior technical evaluations of assistive navigation algorithms [5, 6]. Future work should include longitudinal user studies with BVI participants to assess real-world utility. All procedures described in this section were performed in accordance with the relevant institutional guidelines and regulations of the University of Iowa.

5 Results

5.1 Overall Detection Performance

The system achieved an overall detection rate of 92.5% (111 of 120 trials). The mean alert distance for successful detections was 2.31 ft (SD = 0.91 ft), with a median of 2.0 ft and a range of 0.5 to 4.5 ft. No false positive activations were recorded across any trial. The nine missed detections were distributed across five of the eight test locations (Table 1).

Table 1
Overall detection performance across all
120 trials.

Metric	Value
Total Trials	120
Successful Detections	111 (92.5%)
Missed Detections	9 (7.5%)
False Positives	0 (0.0%)
Mean Alert Distance	2.31 ft
Median Alert Distance	2.0 ft
Standard Deviation	0.91 ft
Range	0.5–4.5 ft

5.2 Effect of Lighting Conditions

The 80 outdoor trials were evenly divided between daytime and nighttime conditions. Daytime trials yielded a 95.0% detection rate (38 of 40) with a mean alert distance of 2.22 ft (SD = 0.71 ft). Nighttime trials produced a 90.0% detection rate (36 of 40) with a mean distance of 1.88 ft (SD = 0.63 ft). Indoor trials under controlled artificial lighting achieved 92.5% (37 of 40) with a mean distance of 2.82 ft (SD = 1.07 ft) (Table 2).

Ascending detection showed a larger day-to-night difference than descending. Daytime ascending trials achieved 100% detection (20 of 20, mean 2.58 ft), while nighttime ascending dropped to 90.0% (18 of 20, mean 2.19 ft). Descending detection was 90.0% under both outdoor conditions (18 of 20 each), with mean distances of 1.83 ft (daytime) and 1.56 ft (nighttime).

An exception to the general daytime advantage was observed at the shallowest outdoor location (4.5 in step height, 22-degree grade), where descending detection was lower in direct sunlight (3 of 5, 60%) than at night under artificial lighting (5 of 5, 100%). This reversal is consistent with the operating characteristics of the direct time-of-flight LiDAR sensor, which relies on infrared emission and timing: solar infrared loading degrades signal-to-noise ratio on small geometric features.

Table 2
Outdoor detection performance by approach direction and lighting condition.

Condition	Detection Rate	Mean Distance	SD
Day - Overall	38/40 (95.0%)	2.22 ft	0.71 ft
Day - Descending	18/20 (90.0%)	1.83 ft	0.64 ft
Day - Ascending	20/20 (100.0%)	2.58 ft	0.59 ft
Night - Overall	36/40 (90.0%)	1.88 ft	0.63 ft
Night - Descending	18/20 (90.0%)	1.56 ft	0.51 ft
Night - Ascending	18/20 (90.0%)	2.19 ft	0.57 ft

5.3 Effect of Step Height

Detection rate and mean alert distance both generally increased with step height. At 4.5 in, the system detected hazards in 85.0% of trials (17 of 20, mean 1.68 ft, SD = 0.88 ft). At 5.5 in, detection rose to 92.5% (37 of 40, mean 2.00 ft, SD = 0.55 ft). Steps of 6.0 in and 7.5 in achieved 100% detection with mean distances of 2.52 ft and 2.98 ft respectively. The 7.0 in steps achieved 85.0% (17 of 20, mean 2.85 ft, SD = 1.32 ft); both failures occurred at a single indoor location, suggesting a site-specific mesh reconstruction issue rather than a height-dependent effect (Table 3).

Table 3
Detection performance by step height across all lighting conditions and directions.

Step Height	Trials	Detection Rate	Mean Distance	SD
4.5 in	20	85.0% (17/20)	1.68 ft	0.88 ft
5.5 in	40	92.5% (37/40)	2.00 ft	0.55 ft
6.0 in	30	100.0% (30/30)	2.52 ft	0.66 ft
7.0 in	20	85.0% (17/20)	2.85 ft	1.32 ft
7.5 in	10	100.0% (10/10)	2.98 ft	0.79 ft

In outdoor trials, the interaction between step height and lighting varied. At 4.5 in, daytime detection was 80% (8 of 10) and nighttime was 90% (9 of 10). At 5.5 in, daytime achieved 100% (20 of 20) while nighttime dropped to 85% (17 of 20). At 6.0 in, both conditions achieved 100% (Table 4).

Table 4
Outdoor detection by step height and lighting condition.

Step Height	Day Rate	Day Mean Dist.	Night Rate	Night Mean Dist.
4.5 in	8/10 (80%)	1.75 ft	9/10 (90%)	1.61 ft
5.5 in	20/20 (100%)	2.23 ft	17/20 (85%)	1.74 ft
6.0 in	10/10 (100%)	2.60 ft	10/10 (100%)	2.35 ft

5.4 Effect of Step Length

Step lengths ranged from 10.5 to 12.75 in across the seven staircase locations. No consistent monotonic trend was observed between step length and detection rate, with rates ranging from 80.0% (11.0 in tread, 8 of 10) to 100.0% (10.5 in and 12.75 in treads). This is consistent with the algorithm's reliance on vertical depth gradient rather than tread depth analysis (Table 5).

Table 5
Detection performance by step tread length.

Step Length	Trials	Detection Rate	Mean Distance	SD
10.5 in	10	100.0% (10/10)	2.60 ft	0.91 ft
11.0 in	10	80.0% (8/10)	2.75 ft	1.13 ft
11.5 in	20	85.0% (17/20)	1.68 ft	0.88 ft
12.0 in	40	92.5% (37/40)	2.51 ft	1.05 ft
12.75 in	20	100.0% (20/20)	2.48 ft	0.53 ft

5.5 Effect of Stair Grade

Stair grades ranged from 22 to 33 degrees. At 22 degrees, detection was 85.0% (17 of 20, mean 1.68 ft). At 26 degrees, detection was 95.0% (38 of 40, mean 2.26 ft). At 31 degrees, detection was 85.0% (17 of 20, mean 2.85 ft), with all failures at two specific indoor locations. At 33 degrees, detection reached 100% (20 of 20, mean 2.79 ft). No monotonic relationship between grade and detection rate was observed, suggesting that site-specific mesh reconstruction quality played a larger role than grade alone at the tested range.

5.6 Effect of Approach Direction

Ascending approaches achieved a 95.0% detection rate (57 of 60) with a mean alert distance of 2.74 ft (SD = 0.82 ft). Descending approaches achieved 90.0% (54 of 60) with a mean of 1.86 ft (SD = 0.77 ft). Six of the nine total failures (67%) occurred during descending approaches (Table 6).

Table 6
Detection performance by approach direction and setting.

Condition	Desc. Rate	Desc. Mean	Asc. Rate	Asc. Mean
All Trials	54/60 (90.0%)	1.86 ft	57/60 (95.0%)	2.74 ft
Indoor	18/20 (90.0%)	2.18 ft	19/20 (95.0%)	3.42 ft
Outdoor Day	18/20 (90.0%)	1.83 ft	20/20 (100.0%)	2.58 ft
Outdoor Night	18/20 (90.0%)	1.56 ft	18/20 (90.0%)	2.19 ft

5.7 Missed Detection Analysis

Of the nine missed detections, three occurred indoors under controlled lighting, two during outdoor daytime conditions, and four during outdoor nighttime conditions. By step height, three failures occurred at 4.5 in, four at 5.5 in, and two at 7.0 in; no failures occurred at 6.0 or 7.5 in. By direction, six occurred while descending and three while ascending. The three 4.5 in failures all occurred at the same outdoor 22-degree location, with two during daytime descent under direct sunlight and one during nighttime ascent. The four 5.5 in failures were distributed across two outdoor locations under low-light or nighttime conditions. The two 7.0 in failures occurred at a single indoor location with 31-degree grade, suggesting a site-specific mesh reconstruction issue rather than a height effect.

6 Discussion

6.1 Detection Reliability and Safety Implications

A 92.5% detection rate with zero false positives across 120 diverse real-world trials represents a promising baseline for on-device, real-time elevation hazard detection on consumer mobile hardware. The mean alert distance of 2.31 ft provides sufficient margin for a user walking at a normal pace to initiate a stopping response before reaching the hazard boundary. The absence of false positives is particularly important for assistive applications: warning systems that generate frequent incorrect alerts produce alarm fatigue, causing users to discount or disable warnings and thereby defeating the safety benefit [3, 11].

The 7.5% missed detection rate, while modest in aggregate, carries greater significance in safety-critical deployment. Missed downward hazards, unlike missed lateral obstacles, offer little recovery opportunity and carry elevated fall risk [1, 2]. The concentration of misses during descending approaches (six of nine failures, 67%) underscores this concern: descending falls typically produce more severe outcomes than ascending stumbles. These results reinforce the intended role of this algorithm and of SensoryScape overall as a complementary safety layer rather than a replacement for established mobility aids such as the white cane.

A critical operational requirement for a BVI navigation aid is that hazard detection must function dynamically during normal ambulation, not only when the user is stationary. A blind or visually impaired user walking at a natural pace cannot be expected to pause at each potential hazard to allow a static sensing window to complete. The algorithm described here operates continuously at 60 Hz as the user moves, maintaining detection capability throughout the gait cycle. This dynamic operation distinguishes it from prior point cloud approaches that require the sensing platform to remain stationary during analysis and are therefore not viable for real-world ambulatory use.

6.2 Geometric Constraints on Detection

The increase in detection rate and alert distance with step height confirms that the algorithm responds primarily to the magnitude of vertical depth gradient in the reconstructed mesh. A practical reliability threshold appears near 6 in: both 6.0 in and 7.5 in steps achieved 100% detection across all tested lighting conditions, while 4.5 in and 5.5 in steps exhibited more variable performance. This finding is encouraging for real-world deployment, as most building codes in the United States specify stair riser heights of 7 to 7.75 in, and residential codes typically require a minimum of 4 in [13]. Non-standard elevation changes such as shallow outdoor steps, degraded curbs, and low platform edges remain more vulnerable to missed detection.

The absence of a trend between step length and detection rate is consistent with the algorithm's design: it evaluates the angular relationship between adjacent ray hits rather than performing tread depth analysis. This property means the algorithm is insensitive to tread geometry and sensitive primarily to riser height, which is well matched to the primary safety objective of detecting falls rather than characterizing stair geometry.

The directional asymmetry between ascending and descending detection is the most consequential geometric limitation identified by this study. Ascending stairs present positive depth discontinuities in the reconstructed mesh that the forward- and downward-facing rays encounter reliably as the user approaches. Descending hazards initially appear as empty space below the floor level: the lower floor surface only becomes observable as the user moves sufficiently close that the downward rays can intersect it. Warning generation during descent is therefore inherently delayed relative to ascent. Addressing this limitation may require forward angled raycasts with lower elevation angles, temporal reasoning about missing or truncated floor geometry ahead of the user, or boundary analysis of mesh edges rather than relying on positive depth returns.

A path toward achieving higher and eventually 100% detection rates for all vertical discontinuities is to systematically optimize raycast angles as a function of target stair geometry. The current fixed ray configuration (Down1 at 30 degrees, Down2 at 60 degrees, Lower at approximately 75 degrees) was determined empirically and may not be optimal for all combinations of riser height and tread depth. For shallow stair geometries (4.5 to 5.5 in riser, 22-degree grade), the angular separation between adjacent ray intercepts on the riser face is reduced, decreasing the signal available to the line-segment classifier.

Varying the Down2 and Lower ray angles dynamically based on estimated floor distance and step height could preserve sufficient angular separation across a broader range of stair geometries.

Achieving the goal of 100% detection across all building-code-compliant stair dimensions and real-world surface conditions will likely require a combination of technical and algorithmic improvements beyond ray angle optimization alone. On the hardware side, higher-power LiDAR emitters and more sensitive SPAD receiver arrays would improve signal-to-background noise ratio at shallow step risers, particularly under direct sunlight. On the algorithmic side, increasing ray density, dynamically adapting ray configurations in real time based on detected floor geometry, and incorporating mesh-grid analysis of the ARKit scene reconstruction could collectively reduce the geometric blind spots that drive the current false-negative pattern. A multimodal approach that fuses raycast geometry data with vision language model inference on the camera feed represents a longer-term avenue: such a system could identify stair candidates from visual context and use that prior to guide more targeted geometric verification, rather than relying on geometric classification alone. Future work should evaluate which combination of these improvements can close the remaining detection gap at shallow step heights and deliver the reliability required for deployment as a primary safety aid for BVI users.

6.3 Effects of Illumination

Lighting conditions influenced alert distance even when overall detection rates remained similar. Reduced ambient illumination decreased the mean detection distance by approximately 15% across outdoor trials (from 2.22 ft in daytime to 1.88 ft at night), compressing the available warning window. At typical walking speeds of approximately 3 ft/s, this corresponds to roughly 110 ms less reaction time, which may be meaningful for users with slower motor responses or those navigating with additional cognitive load.

The anomalous inversion of performance at the shallowest outdoor step under direct sunlight aligns with the known characteristics of the direct time-of-flight Apple LiDAR sensor: solar infrared energy increases shot noise (statistical measurement error arising from the discrete, random arrival of photons at the detector; in time-of-flight LiDAR, higher photon flux from solar infrared increases shot noise variance, degrading the precision of per-pixel depth measurements) and reduces effective signal-to-noise ratio, disproportionately affecting small geometric features where the depth gradient across the step riser is weakest [9]. Future hardware generations with higher LiDAR dynamic range and noise rejection may reduce this vulnerability. Adaptive threshold adjustment using onboard ambient light measurements is a potential software-level mitigation.

6.4 Comparison with Prior Work

The 92.5% detection rate achieved here is competitive with published depth-sensor stair detection results, particularly given the substantially tighter computational constraints of real-time on-device mobile operation. The point cloud pipeline of Westfechtel et al. requires 623 to 4282 ms per detection and a static platform, making frame-rate operation impossible [10]. The optimized edge-based approach of Sriganesh et al. reduces this to approximately 20 ms on a dedicated embedded GPU, which they

describe as real-time for robot autonomy tasks [12]. However, 20 ms already exceeds the 16.7 ms per-frame budget at 60 Hz, and their platform uses a Jetson AGX Xavier co-processor with no competing workloads. On a smartphone, the detection algorithm must share the CPU with ARKit mesh reconstruction, spatial audio rendering, haptic throttling, WatchConnectivity messaging, and voice recognition; under these conditions, even the Sriganesh et al. pipeline would cause sustained frame-rate degradation. The present raycast-based algorithm consumes under 0.5 ms per frame for geometric queries, leaving the remaining 16 ms for all other subsystems. Deep learning classification methods achieve 92 to 96% accuracy under controlled laboratory conditions but require GPU inference workloads incompatible with the real-time budget constraint described above [7]. The zero false positive result achieved here, while not directly comparable to prior work that does not consistently report false positive rates, is a critical metric for assistive applications where alarm fatigue drives users to disable warnings, defeating the safety benefit of the system.

7 Conclusions

We report the design and evaluation of a computationally efficient geometric algorithm for real-time stair and elevation change detection utilizing consumer smartphone LiDAR. The algorithm uses angular analysis of line segments constructed from adjacent raycast intercepts, operating within the existing 60 Hz ARKit spatial mapping cycle without cloud processing, GPU inference, or custom hardware. Across 120 single-user trials at eight real-world locations and under four lighting conditions, the algorithm achieved a 92.5% overall detection rate, no false positives, and a mean alert distance of 2.31 ft.

Additional key findings include a practical reliability threshold near 6 in riser height, above which detection was consistent across all tested conditions; an asymmetry between ascending (95.0%) and descending (90.0%) detection driven by fundamental geometry of depth-based sensing; and a 15% compression of alert distance under nighttime conditions relative to daytime, with an anomalous reversal at very shallow steps under direct sunlight consistent with solar infrared interference. Step length had no systematic effect on detection, confirming that the algorithm responds to vertical elevation change rather than tread depth.

These results demonstrate that accurate, low-latency stair detection is achievable on widely available consumer hardware within the computational budget of a complete multimodal assistive navigation application. The primary open challenges are improving descending hazard detection through boundary analysis or temporal reasoning about absent floor geometry, and validating performance across the full BVI population in realistic navigation contexts. The algorithm is implemented as part of SensoryScape, an open research platform for multimodal navigation assistance on iOS.

Declarations

Competing interests: The author has no competing interests to declare that are relevant to the content of this article.

Funding: No funding was received for conducting this study.

Author contributions: N. Herdt conceived and designed the algorithm, implemented the SensoryScape system, conducted all experiments, analyzed the data, and wrote the manuscript.

Data availability: The trial-level detection data supporting the findings of this study are openly available in Figshare at <https://doi.org/10.6084/m9.figshare.32180439> [18]. Data include per-trial detection outcomes, alert distances, step heights, stair grades, lighting conditions, and approach directions across all 120 trials.

Ethics approval and consent to participate: This study involved a single sighted participant, the author, conducting controlled technical trials of a mobile sensing algorithm; no other individuals were recruited or served as human subjects. Under University of Iowa Human Subjects Office policy, IRB review is required for projects meeting the federal regulatory definition of human subjects research (45 CFR 46.102), which requires a living individual other than the investigator from whom data is obtained through intervention, interaction, or the collection of identifiable private information. Because the sole participant in this study was also the investigator, the author determined, applying this published institutional criterion, that formal IRB review was not required. This determination was not submitted for independent IRB confirmation. In accordance with the guidelines of the University of Iowa Human Subjects Office policy, the study accordingly did not require IRB approval or informed consent. The study was nonetheless conducted in accordance with the relevant institutional guidelines of the University of Iowa and applicable ethical standards, including the principles of the Declaration of Helsinki.

Consent for publication: Not applicable. This manuscript does not contain any individual person's identifiable data, images, or videos.

References

1. Legood, R., Scuffham, P. & Cryer, C. Are we blind to injuries in the visually impaired? A review of the literature. *Inj Prev.* **8** (2), 155–160. <https://doi.org/10.1136/ip.8.2.155> (2002).
2. Ivers, R. Q., Cumming, R. G., Mitchell, P. & Attebo, K. Visual impairment and falls in older adults: the Blue Mountains Eye Study. *J. Am. Geriatr. Soc.* **46** (1), 58–64. <https://doi.org/10.1111/j.1532-5415.1998.tb01014.x> (1998).
3. Maidenbaum, S. et al. A wearable platform for closed-loop GNSS navigation for the blind. *Nat. Commun.* **16**, 758. <https://doi.org/10.1038/s41467-024-55223-x> (2025).
4. Brown, F. E. et al. A novel, wearable, electronic visual aid to assist those with reduced peripheral vision. *PLoS ONE.* **14** (10), e0223755. <https://doi.org/10.1371/journal.pone.0223755> (2019).
5. Schwartz, B. S., King, S. & Bell, T. EchoSee: an assistive mobile application for real-time 3D environment reconstruction and sonification supporting enhanced navigation for people with vision impairments. *Bioengineering* **11** (8), 831. <https://doi.org/10.3390/bioengineering11080831> (2024).

6. Strumillo, P. et al. Different approaches to aiding blind persons in mobility and navigation in the 'Naviton' and 'Sound of Vision' projects. In: *Mobility of visually impaired people*. Springer, Berlin, 435–468 (2018).
7. Smith, M. J. et al. Smart assistive technology for people with disabilities: a systematic review. *JMIR Rehabil Assist. Technol.* **11**, e55776. <https://doi.org/10.2196/55776> (2024).
8. Shoal, S., Borenstein, J. & Koren, Y. The NavBelt - a computerized travel aid for the blind based on mobile robotics technology. *IEEE Trans. Biomed. Eng.* **45** (11), 1376–1386. <https://doi.org/10.1109/10.730436> (1998).
9. Apple Inc. ARKit. (2026). <https://developer.apple.com/documentation/arkit>
10. Westfechtel, T., Ohya, A., Drouilly, R., Takemura, K. & Amano, H. Robust detection of stairways in point clouds for service robots. *IEEE Access.* **10**, 54088–54107. <https://doi.org/10.1109/ACCESS.2022.3174998> (2022).
11. Kristjansson, A. et al. Designing sensory-substitution devices: principles, pitfalls and potential. *Restor. Neurol. Neurosci.* **34** (5), 769–787. <https://doi.org/10.3233/RNN-160647> (2016).
12. Sriganesh, P., Bagree, N., Vundurthy, B. & Travers, M. Fast staircase detection and estimation using 3D point clouds with multi-detection merging for heterogeneous robots. arXiv:2211.00610. (2023). <https://doi.org/10.48550/arXiv.2211.00610>
13. International Code Council. *International Building Code* (ICC, 2021).
14. Luetzenburg, G., Kroon, A. & Bjørk, A. A. Evaluation of the Apple iPhone 12 Pro LiDAR for an application in geosciences. *Sci. Rep.* **11**, 22221. <https://doi.org/10.1038/s41598-021-01763-9> (2021).
15. Tontini, A., Gasparini, L. & Perenzoni, M. Numerical model of SPAD-based direct time-of-flight flash LiDAR CMOS image sensors. *Sensors* **20** (18), 5203. <https://doi.org/10.3390/s20185203> (2020).
16. Tondo, G. R., Riley, C. & Morgenthal, G. Characterization of the iPhone LiDAR-based sensing system for vibration measurement and modal analysis. *Sensors* **23** (18), 7832. <https://doi.org/10.3390/s23187832> (2023).
17. Morimoto, K. et al. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica* **7** (4), 346–354. <https://doi.org/10.1364/OPTICA.386574> (2020).
18. Herdt, N. Real-Time LiDAR-Based Stair and Elevation Change Detection: Trial-Level Dataset. Figshare. <https://doi.org/10.6084/m9.figshare.32180439> (2026).

Figures

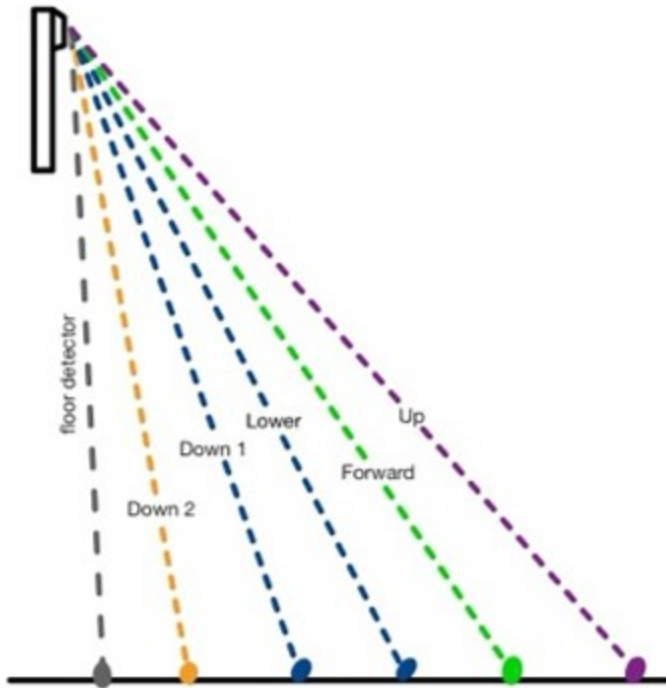


Figure 1

Ray configuration showing all eight raycasts projected from the device position onto a flat floor. Silent rays (Floor Detector and Lower) are shown dashed.

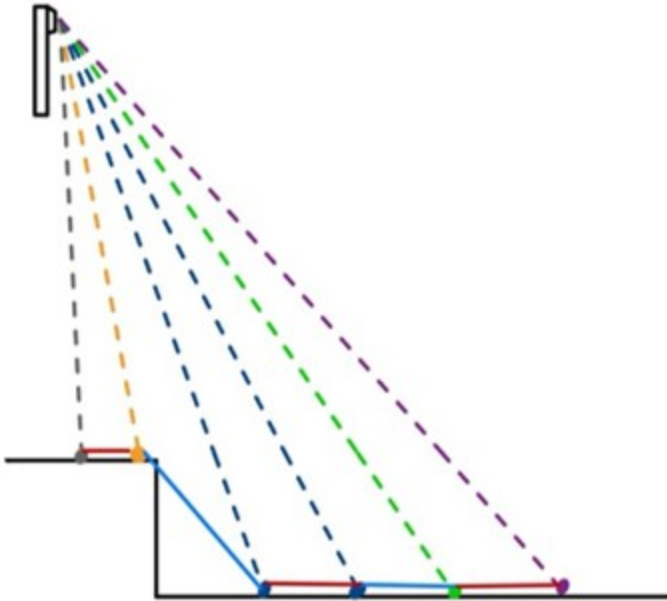


Figure 2

Ray positions during approach to a descending step edge. The Forward ray strikes the upper surface while the Lower ray intersects the drop face, creating a segment characteristic of descending hazard

geometry.

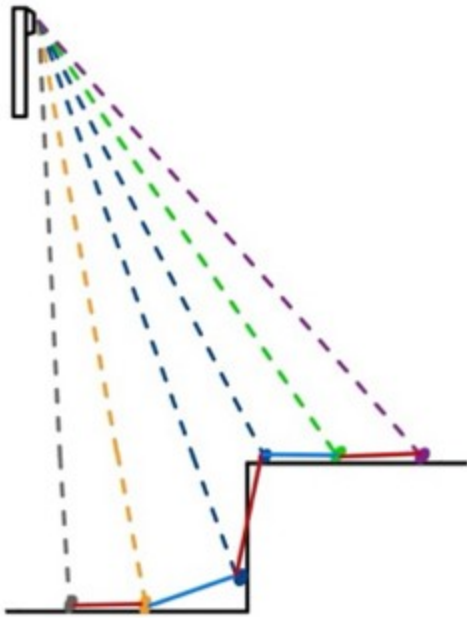


Figure 3

Ray positions during approach to an ascending step. The lower segment captures the riser, and the upper segment captures the landing, distinguishing a tripping hazard from a flat wall.

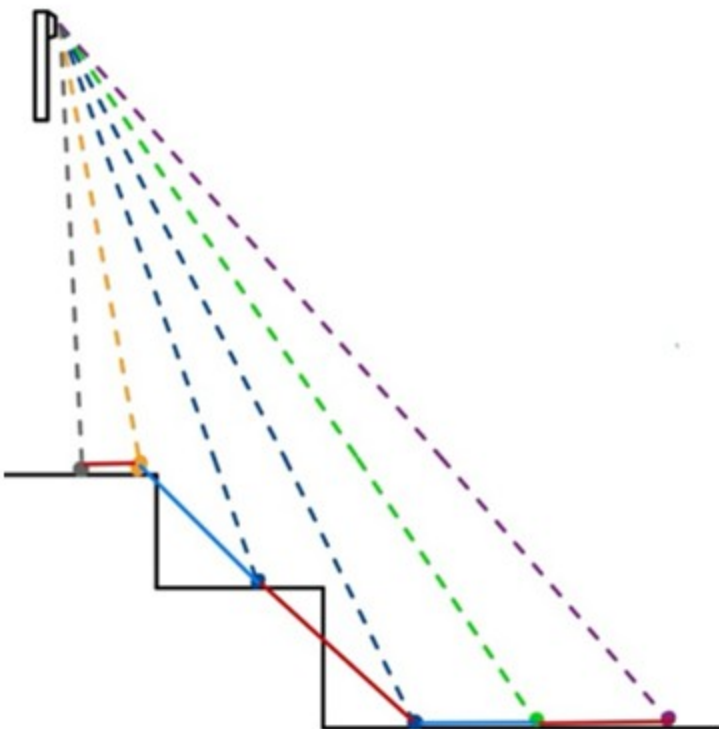


Figure 4

Ray positions for a descending staircase. Both the lower and upper segments exhibit the characteristic stair slope range, triggering the staircase classification.