A Novel Perceptive Robotic Cane with Haptic Navigation for Enabling Vision-Independent Participation in the Social Dynamics of Seat Choice

Shivendra Agrawal, Mary Etta West, and Bradley Hayes

I. Abstract

Goal-based navigation in public places is critical for independent mobility and for breaking barriers that exist for blind or visually impaired (BVI) people in a sight-centric society. Through this work we present a proof-of-concept system that autonomously leverages goal-based navigation assistance and perception to identify socially preferred seats and safely guide its user towards them in unknown indoor environments. The robotic system includes a camera, an IMU, vibrational motors, and a white cane, powered via a backpack-mounted laptop. The system combines techniques from computer vision, robotics, and motion planning with insights from psychology to perform 1) SLAM and object localization, 2) goal disambiguation and scoring, and 3) path planning and guidance. We introduce a novel 2-motor haptic feedback system on the cane's grip for navigation assistance. Through a pilot user study we show that the system is successful in classifying and providing haptic navigation guidance to socially preferred seats, while optimizing for users' convenience, privacy, and intimacy in addition to increasing their confidence in independent navigation. The implications are encouraging as this technology, with careful design guided by the BVI community, can be adopted and further developed to be used with medical devices enabling the BVI population to better independently engage in socially dynamic situations like seat choice.

II. Introduction

It is estimated that 295 million people have moderate or severe vision impairment, of whom 43.3 million are blind. Service dogs and white canes are the most commonly used orientation and mobility aids for blind or visually impaired individuals (BVI). While service dogs can cost upwards of \$50,000 initially, with \$1,200 annually in care costs; white canes are substantially more affordable costing \$20-\$60. Moreover, the skills from one dog cannot be transferred to another, making the training process labor extensive, expensive, and not scalable. For a BVI individual, learning to navigate safely is critical for independence.

While the white cane is the most popular assistive aid and does well in tracing along walls, curbs and entrances, it has very little utility in important social contexts including finding empty chairs in a crowded public area and avoiding contact with other pedestrians in a crowded environment.

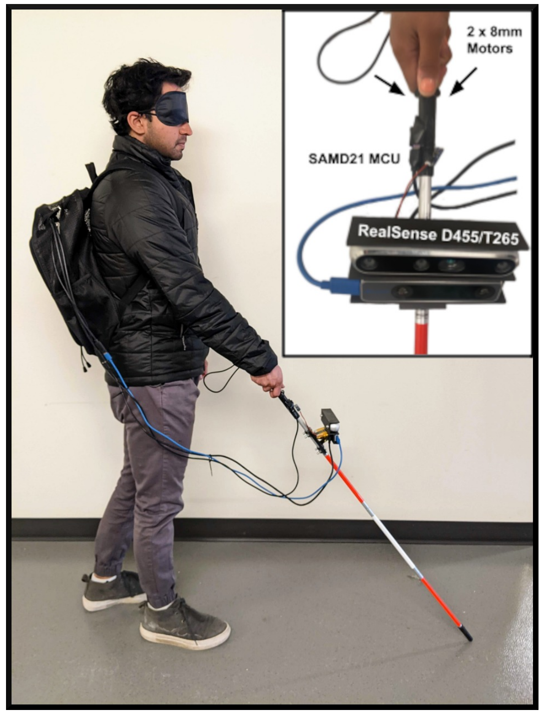

Fig. 1: Robotic cane system with a close-up view. The system includes a robotic cane equipped with RealSense D455 and T265 cameras to enable SLAM and object localization, as well as vibrotactile motors for haptic navigation guidance.

Wang et al. [1] report that **finding empty chairs in crowded public areas is the most important mobility task** with 100% of BVI people surveyed scoring it a 5 out of 5 for importance and that the standard cane's utility to perform that task is extremely low (utility score of 1.29 out of 5).

Staats and Groot show that people prefer to seat themselves in relation to others in a way that optimizes intimacy and privacy [2]. They found that people chose seats that were further away from others and more often chose anchored seats (e.g., a seat against a wall). With current technology, BVI people are not able to independently locate empty chairs without collision with the environment and possibly other people, let alone able to exercise the nuances of seat choices available to the sighted people. Moreover, the COVID-19 pandemic has increased the need for an independence-enabling technology, facilitating social distancing.

Through this work we present an end-to-end perceptive robotic cane system that enables purposeful navigation in unknown, indoor environments. We focus on the above mentioned challenge of finding *socially preferred* chairs in

The authors are all affiliated with the Department of Computer Science at the University of Colorado Boulder. {shivendra.agrawal, mary.west, bradley.hayes}@colorado.edu. This work was funded in part by NSF Award #1830686.

public areas, defined as those that optimize convenience, intimacy and privacy.

Specifically this paper contributes,

- A novel robotic cane system that leverages computer vision to enable socially preferred autonomous goal selection and navigation in indoor spaces.
- An algorithm for social norm-aware chair selection, optimizing for convenience, intimacy, and privacy.
- A novel, intuitive vibrotactile feedback system providing navigation guidance through the robotic cane's grip.
- A pilot validation demonstrating the system's success in finding socially preferred seats and providing effective navigation guidance to novice users.

III. Related Work

Computer-aided assistive technology (AT) has been researched and developed for over 70 years [3], yet still does not have significant adoption from the BVI community.

1) Form Factors: Robotic navigation assistance has been investigated in various form factors, including mobile platforms, quadrupeds, wearables, and handheld cane-like devices. In the CaBot project from IBM Research [4], the team created an autonomous mobile robot in a suitcase form factor. Their platform is capable of indoor navigation and localization, driving next to its user, and leading them to their destination through a haptic interface. While this platform was an excellent testbed for haptics research, its final weight was over 50 pounds which substantially limited its portability and suitability for real-world deployment. CaBot also required a map of the indoor space a priori, which may not always be feasible. In work by Xiao et al. [5], a quadruped is used for navigation assistance in narrow spaces with a leash that accommodates slack, but is a loud and potentially disruptive alternative that is a substantial hindrance when unpowered. Wearable interfaces such as the ALVU [1], [6] can provide a less disruptive alternative, but imposes significant personal instrumentation and has limited capability due to its positioning on the person. A recent work [7] developed a wheeled smart cane that can avoid obstacles and follow waypoints, but its powered wheel design significantly limits the feedback from the local environment that the cane can receive.

2) Smart Canes: Smart canes are a subset of assistive technologies that primarily rely on either ultrasonic sensors for obstacle detection and avoidance [1], [8]–[10] or computer vision to employ algorithms typically capable of path planning and goal-based wayfinding, as well as obstacle detection and avoidance [9]. The most popular smart cane systems have been developed by AssisTech[1] and WeWalk[2]. These products are equipped with a distance sensor and a single haptic motor and can only relay haptic feedback when there is an obstacle directly in front of the user and within a certain range. These devices do not support making or conveying navigational plans.

[1]https://assistech.iitd.ac.in/smartcane.php
[2]https://wewalk.io/en/

The majority of 'smart canes' surveyed utilized ultrasonic sensors for obstacle detection and communicated proximity and direction of the obstacle though vibrotactile motors [11], audio alerts [9], or a combination of the two [12]–[15].

3) SLAM-based Navigational Aids: Chen et al. [16] developed a smart cane using Google Tango that performs mapping and localization, enabling path planning. Many SLAM approaches are prone to failures in dynamic environments [17], and recent research has made progress towards improving performance in such environments. Alcantarilla et al. [18] used dense scene flow to improve visual SLAM and demonstrated greater accuracy in real world experiments with BVI individuals. We chose to utilize online SLAM in our system to build a persistent spatial model of the scene, learning relevant navigation targets that can be utilized for providing better context and for conveying denser information than ultrasonic methods.

4) Feedback Mechanisms: Vibrotactile motors are used to provide programmatic feedback to a BVI user as a form of sensory substitution to convey visual information through touch. Users feel vibration through the white cane and are able to interpret changes in texture and elevation as well as obstacles through direct collision, however programmable feedback can be used to alert users of obstacles before collision and provide navigational guidance in ways a white cane can not. We have observed several encoding schemes throughout the literature that use vibrotactile motors to convey visual information, such as obstacles and their distance and navigational guidance using relative position distributed across the body as directional cues [14]. Notably, Nasser et al., developed the ThermalCane [19], a smart cane which encodes directional cues (go, stop, u-turn, left, and right) into thermotactile feedback by controlling changes in temperature through Peltier modules affixed to the cane's grip. While prior studies investigating the efficacy of vibrotactile feedback have varied wildly in their outcomes, this is largely due to variations in motor mounting, encoding schemes, and number of motors used. We opted to use vibrotactile feedback rather than thermotactile due to potential environmental issues that may effect the perception of temperature changes (e.g., wearing gloves on a cold day).

Motivated by Wang et al.'s [1] identification of 'chair finding' as the most important and desirable missing capability in white canes (and derivatives) and Staats and Groot's [2] finding that humans prefer to seat themselves in a way that optimizes privacy and intimacy, we created a novel perceptive robotic cane with vibrotactile navigation feedback. To the best of our knowledge, this is the first self-contained robotic system that can find socially preferred seats autonomously without any external human guidance. The presented system is simultaneously portable, useful without power, and benefits the user with situated environmental and social interactions that are infeasible with modern wearable and mobile robotics approaches.

IV. Design Considerations

We have identified four considerations for the system design based on user feedback from two studies and observations we noted in research papers within our literature review. The first consideration, *modularity*, allows for iterative improvements without necessitating large integration costs.The computer vision, goal scoring algorithms, the sensors, the compute, and the vibration motors are swappable modules in the system. The second consideration, *usability*, involved a survey of work on modifications to white canes, many of which render the cane unusable for its original intended purpose. The addition of wheels [7], [20], requiring a fixed angle of the cane from the body [10], [21], and excessive use of electronics along the cane's length [22] are identified as potential challenges to the user when operating the white cane with tapping or hovering techniques. This surfaced a third consideration, that *a robotic cane should still be able to function as a white cane for local environment feedback*, especially in case of system failure or loss of power. Fourth, *vibration motors should be collocated and perceptile through one fingertip* because of their high spatial acuity, which mitigates the spatial and temporal effects of perception when motors are distributed over greater distances across the body (e.g., palm, torso, thigh) [23]. This also allows for vibrations like those naturally experienced through the use of an uninstrumented cane to be perceptible because of the amount of unobstructed contact between the hand and cane, helping ensure the third design consideration is met.

V. System Design

We developed a robotic cane capable of identifying socially preferred seating [2] and planning an obstacle-free path to reach it. The system is capable of communicating with and guiding the user along the navigational path through highly localized yet distinguishable vibrotactile motors that are in contact with the tip of the thumb when placed on the cane's grip. The system follows a serial architecture with three primary components — perception, planning, and conveyance. They are designed to be modular so that they are easily replaceable as the supporting hardware and foundational algorithms mature.

A. Hardware Overview

The hardware as shown in the system diagram in Fig 2 consists of a cane-mounted RealSense T265 and RealSense D455, connected to a Dell G15 laptop with an RTX 3060 GPU carried in a backpack worn by the user. The RealSense T265 was chosen for its odometry and the RealSense D455 was chosen for RGB-depth sensing. Additionally, an Adafruit Trinket M0 SAMD21 microcontroller is connected to the laptop via USB to receive serial commands for controlling the vibration motors. The microcontroller was chosen for its small form factor. This extra hardware and its mounting support increases the weight of the cane by just (approximately) 0.33lbs, primarily located near the grip to minimize added fatigue.

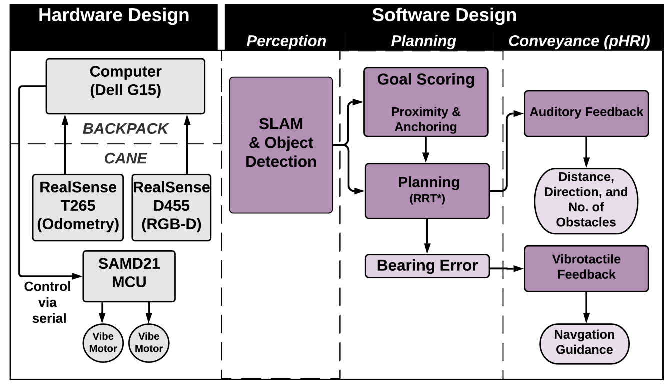

Fig. 2: Design diagram. Perception and navigation algorithms are executed on a backpack-worn laptop, while all sensing and haptics are mounted on the cane.

B. Software Overview: Perception

The perception of the scene begins with a SLAM algorithm that creates an initial 2D occupancy grid from RGB+D camera input as the user scans the room by turning in place with gradual movement. [3] The odometry provided by the Intel RealSense's IMU has low drift, allowing for very accurate user pose estimation. This enables accommodation of arbitrary handle tilt, as depth readings not within a specific height range can be discarded. The RGB+D data are also used to find objects and project them onto the 2D occupancy grid. The system uses Detectron2 for object detection and Mask-RCNN to obtain masks for classification. Classification masks are superimposed on the depth channel to estimate the depth of identified objects. Using the camera intrinsics, objects are mapped onto the 2D occupancy grid (Fig 3), completing the pipeline.

Challenge 1 (Data Association): Due to the system's high rate of movement, establishing proper data association between object observations and previously observed landmarks is a significant challenge. Particularly for objects that commonly exist in groups, such as chairs, it is important that the system be able to disambiguate between cases of nearby similar objects and a single object detected multiple times. For our reference implementation, we utilize an Extended Kalman Filter to model object positions as 2D Gaussians, providing a distribution against which to check new observations.

Challenge 2 (Exploration vs. Exploitation): Because the system is mapping its environment online and is intended to be usable in environments novel to the user, the decision of how to balance exploration of the space (scanning) with exploitation of the current map (choosing a target and navigating to it given the available free space) is critical for the user experience. As it is impossible to know whether there are more (possibly more desirable) objects of interest in the scene without performing the impractical step of fully exploring it, we parameterized the factors involved in this

[3] We use a publicly available SLAM implementation by RealSense at https://github.com/IntelRealSense/realsense-ros/tree/occupancy-mapping

decision. The system utilizes a time-based and goal score threshold to determine when to move from exploration to exploitation.

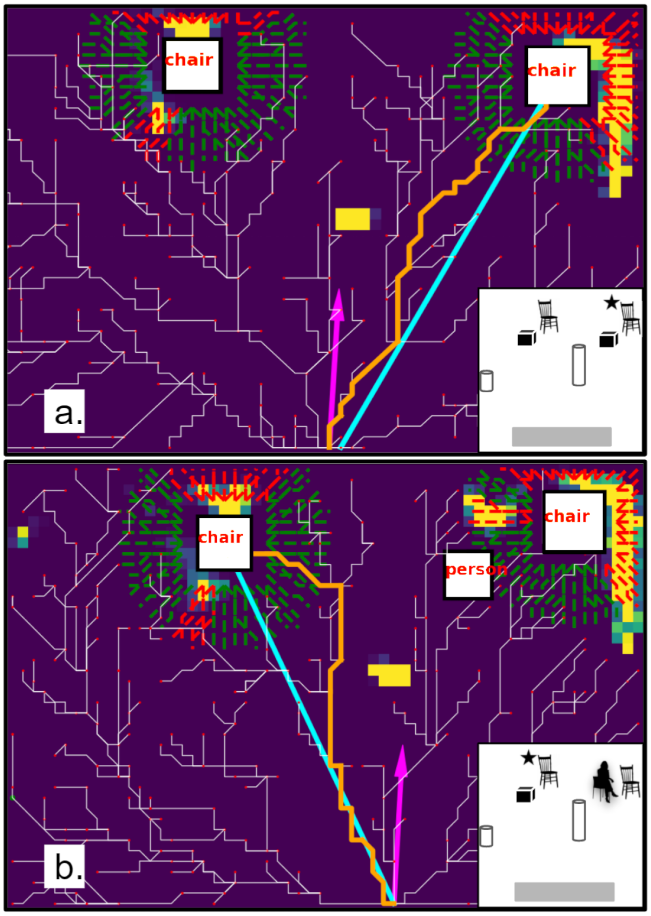

Fig. 3: Visualization of live data from the system. 1) SLAM creates a 2D occupancy grid. Darker (purple) spaces indicate higher probability of free space. 2) Recognizable objects are projected onto the 2D grid with labels in white. 3) Goal scoring is done via the *anchor score* and *proximity score* to determine a navigation target. *Anchor score* rays are shown in green and red, with red indicating the rays contributing to a higher anchor score. 4) A path is found towards the goal using the Rapidly-exploring Random Tree (RRT*) shown in white. The unmodified RRT* path is shown in orange, and a pruned, optimized version of that path is shown in cyan. 5) The user's orientation is shown in magenta, which is used to generate the bearing error to be conveyed using the vibrotactile motors. (a) The system selects the right-hand side chair that has the *higher anchor score* due to walls on two sides, thus *increasing privacy*. (b) The system selects the left chair because there is a person close to the other chair, thus *decreasing intimacy*.

C. Software Overview: Planning

We divide planning into two steps, goal selection and path planning:

1) Goal Selection: As objects are detected and mapped by the system, they are considered as potential navigation goals. Each object of the target class (e.g., chair) is scored by an objective function accounting for levels of *convenience*, *privacy*, and *intimacy*. Ideal goals for the seating selection task maximize convenience and privacy while minimizing intimacy. The system utilizes algorithms for evaluating *anchoring* to regulate privacy and *proximity* to balance convenience against intimacy. The objective function is defined as the sum of normalised anchor and proximity scores for each candidate goal object. A minimum threshold score (referenced in Challenge 2) can be used to inform the minimum score for a satisfactory goal candidate.

Anchor score: People prefer seats that are anchored in the environment to increase their privacy [2]. It is a common design choice for chairs to be anchored to large and contiguous objects, like a table that itself is anchored to a wall. Leveraging this, the system uses a novel anchor measurement ray-casting algorithm (Algorithm 1) to measure the relative anchoring of chairs using the 2D occupancy grid. The algorithm runs a sliding window of rays emanating from each detected chair in the scene (Green and red rays in Fig 3) to accumulate scores for each chair. Self-intersections with the chair being evaluated are ignored (defined by an `ignore_radius`), otherwise the rays emanating the center would not escape the chair itself (line 2). The sliding window approach assigns higher scores to contiguous obstacles (e.g., walls). A subset of rays is defined (line 7), counting how many of these rays intersect with occupied grid cells within `ray_cast_radius` of the object origin (line 9-13). If a window crosses the threshold (`wall_threshold`), we increase the `wall_counter` for the chair by 1 (line 14-15) (Fig 4). The normalized `wall_counter` value is returned as the object's anchor score (line 16-17). An example from real world data is visualized in Fig 3.

Proximity score: People also prefer seats that are at a greater distance to others in order to decrease intimacy [2]. Considerations related to proximity are incorporated into the goal selection optimization process through a normalized linear cost model (Equation 1) with terms for *intimacy* ($C_{intimacy}$), *travel convenience* ($C_{convenience}$), *objects correlated with signs of occupancy* ($C_{occupied}$), and *objects correlated with success* ($C_{success}$). The intimacy term penalizes goals that are closer to other humans, while the convenience term penalizes goals (g) that are further from the user. The object-occupancy term penalizes goals that are near objects likely to have owners, including backpacks or laptops, as these can be indicators that the goal is occupied or unavailable. Finally, the object-success term increases scores for goals that are close to alternatives or objects correlated with more selection (e.g., a table correlates with the existence of multiple chairs). $\mathcal{D}$ is used to represent the set of all detected objects or entities, with notation $\mathcal{D}_{human}$ indicating the subset of detections classified as humans, $\mathcal{D}_{occupied}$ indicating the subset of detections classified as object classes correlated with occupancy, and $\mathcal{D}_{success}$ indicating detections classified as object classes correlated

Algorithm 1: ANCHOR SCORE PROCEDURE

Input: Chair Objects $Chairs$, 2D occupancy grid
Output: Normalised anchor scores
Parameters: `ray_cast_radius`, `window_size`, `ignore_radius`, `wall_threshold`

foreach $chair \in Chairs$ **do**
  $rays \leftarrow [\,]$ of rays cast from $chair.position$ within `ignore_radius` to `ray_cast_radius`
  $n\_rays \leftarrow$ Total number of rays
  $wall\_counter \leftarrow 0$
  for $i \leftarrow 0$ **to** $n\_rays - window\_size$ **do**
    $intercepted\_rays \leftarrow 0$
    for $j \leftarrow i$ **to** $i + window\_size$ **do**
      /* *ray is made up of cells that it spans over*
      $ray \leftarrow rays[j]$
      Iterate over ray
        if *ray is intercepted by an occupied cell* **then**
          $intercepted\_rays \leftarrow intercepted\_rays + 1$
          break
    if $intercepted\_rays/window\_size >$ $wall\_threshold$ **then**
      $wall\_counter \leftarrow wall\_counter + 1$
  $chair.anchor\_score \leftarrow wall\_counter$
return *Normalised anchor scores*

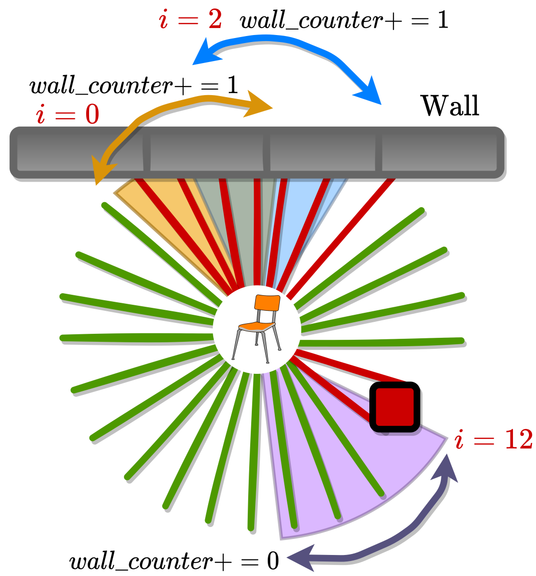

Fig. 4: Anchor scores are calculated using a sliding window to track object-intersection density with radially cast rays. This method gives more votes to contiguous anchoring obstacles like walls. In the above example, the sliding window at $i = 0$ and $i = 2$ only contain object-intersecting rays (colored red, non-intersecting in green) by the wall and thus contribute 1 each to the $wall\_counter$. Whereas the sliding window at $i = 12$ doesn't contain a sufficient density of object-intersecting rays and thus doesn't contribute to the $wall\_counter$.

with success. The model parameters are shown in Table I, utilized in Equation 1, and visualized in Fig 5. While it is possible that this approach can still select an occupied goal if either the final 'acceptable score' threshold or $C_{intimacy}$ is kept very low, in practice, we found that occupied goals (e.g., chairs) tend to receive very unfavorable scores due to $dist(\text{goal}, \text{human})$ being very small. Parameter values were manually tuned against reference scenes representative of normal use cases. Once computed for each candidate goal, the proximity scores are normalized. A sample result from real-world data is shown in Fig 3b.

TABLE I: Parameters used for proximity scoring.

Parameter	Role
$C_{intimacy} = -6$	Penalizes nearness to other humans
$C_{convenience} = 1$	Rewards nearness with the user
$C_{occupied} = -1$	Penalizes nearness to failure-correlated objects
$C_{success} = 1$	Rewards nearness to success-correlated objects

$$S_{prox}(g, \mathcal{D}) = \frac{C_{convenience}}{dist(g, user)} + \sum_{d \in \mathcal{D}_{human}} \frac{C_{intimacy}}{dist(g, d)^2} + \sum_{d \in \mathcal{D}_{occupied}} \frac{C_{occupied}}{dist(g, d)^2} + \sum_{d \in \mathcal{D}_{success}} \frac{C_{success}}{dist(g, d)^2} \quad (1)$$

2) Path Planning: Once a goal is selected from the set of detected objects (e.g., the chair with the best overall score), RRT^* [24] is used to find a collision-free path from the user's current position. The resultant path is successively pruned to avoid unnecessary motion through extraneous waypoints (Fig 3). We remove extraneous waypoints by recursively calculating the feasibility of the path between a waypoint's predecessor and successor.

D. Conveyance

Plans are conveyed to users through two modalities: a verbal goal overview and vibrotactile haptic guidance.

1) Verbal goal overview: Once the goal is found, the system generates a semantic description of the goal's relative location. The overview has the following template: *"{Goal Object} found about {} meters away in the {} o'clock direction"*, based on user patterns observed in popular human-guided assistive applications (e.g., *Be My Eyes*). The distance and direction estimates are calculated from the relative positioning of the last determined user pose and the goal. The verbal overview is presented to set user expectations about the navigation duration. The phrase *"with obstacles in the path"* is appended if there is not a collision-free straight-line path to the goal to set expectations regarding the complexity of the navigation plan. This overview is delivered via the laptop speaker, and precedes the vibrotactile guidance.

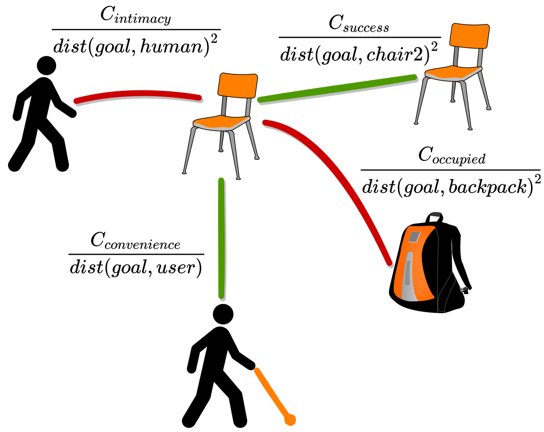

Fig. 5: A chair's proximity score is affected by it's relative position with respect to other socially meaningful objects, humans and the user. We add the scores shown by the green lines and subtract the ones shown by the red.

2) Vibrotactile guidance: The robotic cane utilizes a novel vibrotactile haptic communication system, providing navigational guidance through the grip. This is informed by prior work that has shown that two-point vibrotactile discrimination on the fingertips ranges from 2.1mm–8mm [25]. Accordingly, the robotic cane has two vibrotactile motors (8mm ERM) affixed on the grip such that during normal use the user will naturally position their thumb between them. This design removes the need for additional wearables (e.g., haptic belts) and minimizes the area of modification on the cane itself to maintain familiarity of feel.

Navigation information is relayed to the user by encoding waypoint bearing error into two distinct haptic animation patterns on each motor (left and right), creating five possible codes for conveyance: hard-right, soft-right, straight/no animation, soft-left, and hard-left. The design decision to have more than a single pattern per direction is rooted in literature indicating a preference for two levels of turn (e.g., *sharp left* and *soft left*) [26]. The soft-right and soft-left animations are programmed at one-half the vibration intensity of the hard-right and hard-left animation. The motors vibrate at 200 Hz. Intended use involves the user first correcting their bearing error by turning in-place in the direction of the corresponding vibration motor, then walking towards the next waypoint. A distinct animation pattern (both motors cycle between actuating at 100% duty cycle for 750ms and then at 0% duty cycle for 750ms for three seconds) signals the start and stop of the plan.

One significant challenge encountered during the design process was the observation that mounting vibrotactile motors directly to the cane causes the entire cane to vibrate, making it very challenging to tactilely discriminate between the vibration source as either left or right. Our system addresses this by mounting the wires of the motors directly to the cane grip, but allowing the motors to hover near the surface so that the cross-section of direct contact with the cane grip is negligible.

VI. Pilot Study

We conducted a pilot study (n=6; 2 male, 4 female) with novice users for preliminary testing of the device, which included navigating through six scenarios while blindfolded. Participants were all sighted graduate students and had no prior experience with using a probe (e.g., white cane) to navigate. For testing at this stage, we follow prior work in performing preliminary validation using blindfolded participants [7], [27]–[31] as a precursor to engaging with the BVI population.

The pilot tests took place in a configurable $12ft \times 17ft$ room as shown in Fig 6, representing more complex variants of room configurations used for similar studies of visually-impaired navigation [1]. The room ordering is randomized during the study to limit learning effects. Room 1 is designed to compare a user's proficiency in finding and navigating to a seat in a simple single chair scenario. Room 2 and Room 3 are designed to test the algorithmic goal-scoring capability, locating the more socially preferred seat based on proximity (which influences choices based on intimacy and convenience) and anchoring (which influences seat choice based on privacy). Cardboard boxes, bins, and suitcases are used as obstacles throughout each of the rooms to increase navigation difficulty.

During the experiment, users navigated through each room twice, using the robotic cane both with and without the verbal overview. The order of presentation of room layout and verbal condition was randomized, with users unaware of the number of unique layouts being presented. To begin a trial, the participant was blindfolded and led to the starting area (shown in gray in Fig 6) with random orientation.

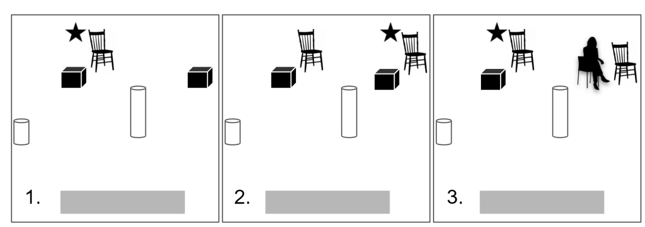

Fig. 6: The three room layouts: The Room 1 layout includes one chair and four obstacles. The Room 2 layout includes one highly anchored chair, one unanchored chair, and four obstacles. 3. The Room 3 layout consists of a highly anchored chair next to an occupied chair, one unanchored chair, and three obstacles. Stars denote chairs that should be chosen according to anchoring/intimacy-based western social norms and the gray rectangles represent the starting locations.

A. Procedure

Each user completed six scenarios while blindfolded, each scenario comprised of a combination of Room (1,2,3) and robotic cane with verbal overview enabled/disabled. For each scenario, the blindfolded user was guided to a starting location. When given the verbal "START" cue, the user was instructed to begin by performing a scan of the room by rotating in place. Once a plan is developed, the system's haptic start sequence is conveyed to the user, indicating that they can begin navigating to a chair following the navigation guidance provided by the vibrotactile interface. In between each scenario, the user is disoriented by being guided randomly around the testing area and while the room is rearranged into the configuration for the next scenario.

B. Results

Task Completion: Within each scenario, users were tasked with finding a suitable chair to sit in within 2.5 minutes. This task was considered successfully completed if the user navigated to a chair and said ”FOUND”.

We obtained promising results from our algorithm that optimized for privacy, intimacy, and convenience. Users were able to find the highly anchored chairs in **10/12** scenarios. We saw a similar success rate for Room 3, which had a person (quietly) sitting in a chair next to the highly anchored chair. The system was able to guide the users to the chair that minimized intimacy in **10/12** scenarios, thus effectively demonstrating the proposed proximity scoring formula's ability to guide users toward socially appropriate seat choices. Even in the four scenarios where the system was not able to find the more socially preferred chair, users were still successful in finding *a* chair, thus achieving a **100%** success rate in finding a seat and **83.3%** success rate for finding a socially preferred seat. In these less socially appropriate cases, we observed that users had obtained diminished scan coverage of the room, triggering the system to stop further exploration of the scene and suggest the only available choice to avoid idling for too long.

Navigation Time: Navigation time is measured starting at the moment the user is able to begin moving, after the plan has been calculated and the start sequence has been communicated to them. The navigation timer is stopped once the user comes in contact with the goal using the robotic cane, to avoid confounds due to delays in providing the verbal “found” cue. We used a cut-off time of 2.5 minutes for each trial, but participants were able to find a seat within 45 seconds on average with none exceeding the cut-off.

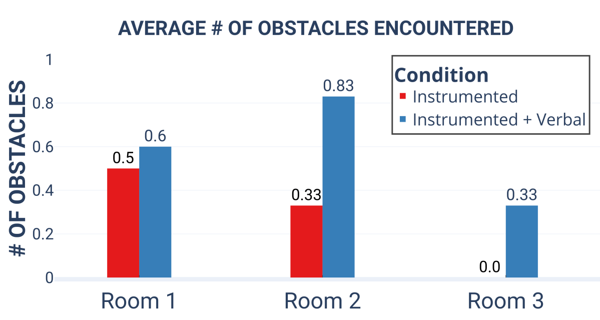

Fig. 7: Even though the rooms had obstacles and walls, the novice users in our study often were able to avoid collisions with the environment entirely while navigating to their goal.

Obstacle Avoidance: A summary of the average number of obstacles the users physically interacted within each scenario is presented in Fig 7. Despite the fact that the rooms were populated with obstacles and walls, we observe that due to the vibrotacile navigation's effectiveness users were often able to avoid causing even a single collision while using the robotic cane. This is an encouraging result as BVI individuals have expressed the desire to avoid cane contacts with other people [1]. Through this pilot study we have shown, with inexperienced users, that one benefit of our system is to help eliminate the socially undesirable situation of collisions with people seated in chairs (Room 3).

User Experience: We conducted post activity surveys to measure aspects of the user experience focused on usability and confidence in independent navigation. We found that the users were confident accomplishing the task with our system, more so with the verbal overview enabled. Using a 5-point scale, participants rated their *confidence in navigation* at 4.83 ± 0.41 (verbal enabled) and 4.00 ± 0.63 (verbal disabled), their *confidence in travelling fast* at 4.33 ± 0.82 (verbal enabled) and 4.00 ± 0.63 (verbal disabled), and their *confidence in finding the goal* at 4.5 ± 0.84 (verbal enabled) and 3.83 ± 1.17 (verbal disabled). Users also rated the *verbal overview's helpfulness* at 4.67 ± 0.82 and the *ease of use* of the system at 4.17 ± 0.98 (verbal enabled) and 4.67 ± 0.52 (verbal disabled). The users perceived their performance to be adequate with little effort. Using the NASA Task Load Index survey's battery of questions (20-point scale), users rated their *performance* at 16.67 ± 3.39 (verbal enabled) and 17.00 ± 2.61 (verbal disabled), and their *effort* at 7.83 ± 4.12 (verbal enabled) and 4.67 ± 3.43 (verbal disabled).

Overall, pilot study participants rated the system as being highly usable and highly performant. In a post-activity interview, they mentioned the following:

- Verbal overview - *“I used the verbal instructions as a global plan to reorient myself”,“Just having a mental picture of how far I need to go, that was good”, “[The] head start with the verbal announcement was a little comforting/assuring”* .
- Haptics - *“Hard and soft turns were easy to distinguish”,“The soft vs hard vibrations helped me a lot in determining how much to turn”,“[the] feedback was pretty fast”*.

VII. Discussion

We caution the reader to exercise care to avoid drawing conclusions about device readiness based solely on pilot tests with blindfolded, non-BVI participants, as these are not a reliable proxy for the (eventual) target population of this device [28], [32]. With this pilot study we have shown a proof of concept with regard to the hardware and software features of the system alongside a validation of the novel vibrotactile navigation interface, demonstrating that the cane-mounted system is able to construct a usable environment representation, plan within it, and effectively guide users to an autonomously determined desired goal location based on higher-order social features in real-time.

We observed that each user scanned the rooms uniquely, capturing various amounts of coverage which led to variations in the amount of information obtained for each room.

We found early on that a much slower scan sequence did not allow the algorithm to view the complete scene within a reasonable time frame. To avoid an assumed annoyance to the user, the algorithm calculates a path to the best available chair ('local optima') after a fixed timeout is reached, even if this means that the entire area hasn't been scanned and the 'globally optimal' chair hasn't been detected. This leads us to believe there are opportunities to improve user experience by exposing some of the system's inner workings to the user by allowing them to select their own scan time.

Interestingly, we also observed that all users performed best (as rated by task completion, reduced navigation time, and number of obstacles avoided) when using the robotic cane without verbal overview, yet 5/6 users stated they preferred the robotic cane *with* the verbal overview. It also became clear that the kinesthetic feedback from the cane resulting from the few object collisions that occurred was a valuable signal, reinforcing the importance of maintaining this core capability of the white cane form factor. These observations will need to be further explored within a user study with BVI individuals.

VIII. Conclusion

In this work we present a novel robotic cane system that combines computer vision and online SLAM with a novel vibrotacile interface to enable blindfolded participants to identify and navigate to socially preferred seats. We tested aspects of the system through a pilot study that provided an initial validation for our algorithm for social-norm aware chair selection, optimizing convenience, intimacy, and privacy. These results further demonstrated the effectiveness of the unobtrusive, vibrotactile haptic communication channel based on metrics of task completion, navigation time, obstacle avoidance, and participant free-form feedback.

References

[1] H.-C. Wang, R. K. Katzschmann, S. Teng, B. Araki, L. Giarré, and D. Rus, "Enabling independent navigation for visually impaired people through a wearable vision-based feedback system," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017.

[2] H. Staats and P. Groot, "Seat choice in a crowded café: Effects of eye contact, distance, and anchoring," *Frontiers in Psychology*, 2019.

[3] S. Real and A. Araujo, "Navigation systems for the blind and visually impaired: Past work, challenges, and open problems," *Sensors*, 2019.

[4] J. a. Guerreiro, D. Sato, S. Asakawa, H. Dong, K. M. Kitani, and C. Asakawa, "Cabot: Designing and evaluating an autonomous navigation robot for blind people," in *The 21st ACM SIGACCESS Conference on Computers and Accessibility*, NY, USA, 2019.

[5] A. Xiao, W. Tong, L. Yang, J. Zeng, Z. Li, and K. Sreenath, "Robotic guide dog: Leading a human with leash-guided hybrid physical interaction," 2021.

[6] R. K. Katzschmann, B. Araki, and D. Rus, "Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 3, pp. 583–593, 2018.

[7] P. Slade, A. Tambe, and M. J. Kochenderfer, "Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision," *Science Robotics*, vol. 6, no. 59, 2021.

[8] Y. Niitsu, T. Taniguchi, and K. Kawashima, "Detection and notification of dangerous obstacles and places for visually impaired persons using a smart cane," in *2014 Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, 2014, pp. 68–69.

[9] M. Saaid, A. Mohammad, and M. Megat Ali, "Smart cane with range notification for blind people," in *2016 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, 2016.

[10] H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, and S. Mizuno, "Kinect cane: An assistive system for the visually impaired based on three-dimensional object recognition," in *2012 IEEE/SICE International Symposium on System Integration (SII)*, 2012, pp. 740–745.

[11] J. Sakhardande, P. Pattanayak, and M. Bhowmick, "Smart cane assisted mobility for the visually impaired," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 10, pp. 1262 – 1265, 2012.

[12] R. K. Megalingam, A. Nambissan, A. Thambi, A. Gopinath, and M. Nandakumar, "Sound and touch based smart cane: Better walking experience for visually challenged," in *2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)*, 2014, pp. 1–4.

[13] M. H. A. Wahab, A. A. Talib, H. A. Kadir, A. Johari, A. Noraziah, R. M. Sidek, and A. A. Mutalib, "Smart cane: Assistive cane for visually-impaired people," 2011.

[14] B. Singh and M. Kapoor, "Assistive cane for visually impaired persons for uneven surface detection with orientation restraint sensing," *Sensor Review*, vol. 40, no. 6, p. 687–698, 2020.

[15] I. Y. Chung, S. Kim, and K. H. Rhee, "The smart cane utilizing a smart phone for the visually impaired person," in *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*, 2014, pp. 106–107.

[16] Q. Chen, M. Khan, C. Tsangouri, C. Yang, B. Li, J. Xiao, and Z. Zhu, "Ccny smart cane," in *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems*.

[17] S. Bateman, K. Harlow, and C. Heckman, "Better together: Online probabilistic clique change detection in 3d landmark-based maps," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 4878–4885.

[18] P. F. Alcantarilla, J. J. Yebes, J. Almazán, and L. M. Bergasa, "On combining visual SLAM and dense scene flow to increase the robustness of localization and mapping in dynamic environments," in *IEEE International Conference on Robotics and Automation*, 2012.

[19] A. Nasser, K.-N. Keng, and K. Zhu, "Thermalcane: Exploring thermotactile directional cues on cane-grip for non-visual navigation," in *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, New York, NY, USA, 2020.

[20] I. Ulrich and J. Borenstein, "The guidecane-applying mobile robot technologies to assist the visually impaired," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 2001.

[21] V. V. Meshram, K. Patil, V. A. Meshram, and F. C. Shu, "An astute assistive device for mobility and object recognition for visually impaired people," *IEEE Trans. on Human-Machine Systems*, 2019.

[22] S. Murali, R. Shrivatsan, V. Sreenivas, S. Vijjappu, S. J. Gladwin, and R. Rajavel, "Smart walking cane for the visually challenged," in *IEEE Region 10 Humanitarian Technology Conference*, 2016, pp. 1–4.

[23] R. W. Cholewiak, "The perception of tactile distance: Influences of body site, space, and time," *Perception*, 1999, pMID: 10664778.

[24] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The international journal of robotics research*, 2011.

[25] C. A. Perez, C. A. Holzmann, and H. E. Jaeschke, "Two-point vibrotactile discrimination related to parameters of pulse burst stimulus."

[26] M. A. Williams, C. Galbraith, S. K. Kane, and A. Hurst, "" just let the cane hit it" how the blind and sighted see navigation differently," in *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, 2014, pp. 217–224.

[27] A. Wachaja, P. Agarwal, M. Zink, M. R. Adame, K. Möller, and W. Burgard, "Navigating blind people with walking impairments using a smart walker," *Autonomous Robots*, vol. 41, no. 3, 2017.

[28] A. D. dos Santos, F. O. Medola, M. J. Cinelli, A. R. Garcia Ramirez, and F. E. Sandnes, "Are electronic white canes better than traditional canes? a comparative study with blind and blindfolded participants," *Universal Access in the Information Society*, vol. 20, no. 1, 2020.

[29] E. E. O'Brien, A. A. Mohtar, L. E. Diment, and K. J. Reynolds, "A detachable electronic device for use with a long white cane to assist with mobility," *Assistive Technology*, vol. 26, no. 4, pp. 219–226, 2014.

[30] A. J. Fukasawa and K. Magatani, "A navigation system for the visually impaired an intelligent white cane," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.

[31] R. Okazaki and H. Kajimoto, "Perceived distance from hitting with a stick is altered by overlapping vibration to holding hand," in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, 2014.

[32] A. M. Silverman, J. D. Gwinn, and L. V. Boven, "Stumbling in their shoes: Disability simulations reduce judged capabilities of disabled people," *Social Psychological and Personality Science*, 2015.