VISION (Visually Impaired Spatially Interactive Orientation Network)

Aaron Crawford, Alexander Parady, Arsene Landry Tatke, and Noah Harney

Dept. of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida 32816-2450

***Abstract* — This paper presents *VISION*, an assistive billiards aid for the visually impaired. *VISION* uses computer vision, artificial intelligence, Bluetooth beacons, and audio guidance to enable a visually impaired player to participate in a game of billiards. The *VISION* system is responsible for capturing the state of the billiards table, determining the best shot to take, localizing the user, and navigating the user to the location required to take the shot. *VISION* works with the SCRATCH team responsible for executing the shot once the user navigates to the proper location.**

***Index Terms* — Artificial intelligence, Bluetooth, computer vision, python, simultaneous localization and mapping, wireless communication.**

I. INTRODUCTION

VISION is an ambitious project that extends one of America's favorite pastimes to the visually impaired community. VISION's goal is to serve as the visual component of a player so that a player does not need to see the billiards table to play the game. The VISION system can capture the state of the game by taking a picture of the billiards table. The image is the input to a computer vision algorithm that identifies and localizes the billiard balls in the image. The billiard balls and their relative positioning on the table are input to an artificial intelligence algorithm that determines the most advantageous shot. The best shot for a player to take is described by the position the player must stand in, the amount of force applied to the ball, and the angle at which to aim the cue stick. The user guidance and user localization systems use the ideal location of the player. The user guidance system uses a speaker array positioned around the billiards table to provide the player with noises and commands to guide the player to the desired location. The system will continue to guide the user to the desired position until the user localization system has determined that the player is in the correct position. The user localization system tracks the player's position with a cellphone and three ultrawide-band Bluetooth beacons around the table. Once the player has arrived at the proper location, the user guidance system will use more sounds and commands to orient the player in the proper direction. Once the player is oriented, *VISION* will transfer control to the *SCRATCH* team by sending the angle and amount of force required for the shot. *SCRATCH* will assist the player with completing the shot and return control to *VISION* once the shot is complete. The computer vision system will capture another image of the billiards table and determine the shot outcome. The result of the shot will be communicated to the player through audio feedback.

VISION is a proof of concept to show that billiards for the visually impaired is possible. Currently, *VISION* only supports a single visually impaired player due to budget constraints. However, *VISION* can be extended to include another visually impaired player or an individual not suffering from visual impairments.

II. HARDWARE DETAIL

A. Camera

The computer vision system requires a camera to take pictures of the current state of the billiards table for input to a computer vision algorithm. The camera of choice was an Anker PowerConf C200 webcam because of its price, resolution, and field of view. The camera cost $50 which helps the project to remain within budget, has a 2k resolution to take clear pictures needed by the computer vision algorithm, and has an adjustable field of view to accommodate different camera positions. Fig. 1. is an image of the camera used by the computer vision system.

Fig. 1. Anker PowerConf C200 webcam used by the computer vision system.

Mounting the camera above the billiards table is an important design decision. The two main mounting methods are a ceiling-mounted camera and a movable fixture that extends above the table. The ceiling-mounted design does not interfere with the playable region but does not allow the system to be moved. The fixture-mounted design obstructs a small amount of the playable region but allows the system to be moved. VISION implements a fixture-mounted design so that the system can be set up in different locations around the campus. A ceiling-mounted design is recommended if a permanent version of VISION is implemented. Fig. 2. below shows the fixture-mounted camera stand used by the computer vision system.

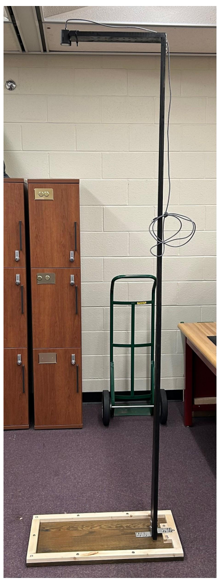

Fig. 2. Fixture-mounted camera stand used by the computer vision system.

B. Speakers

Crucial for the guidance of the visually impaired is a method that can assist across the wide spectrum of ailments to the blind. In the case of VISION, 40 mm internal magnet speakers are used to output guidance commands in the form of digitally generated beeps that will be distinct in origin to most impaired users. Pulse width modulated (PWM) waves are raised and lowered to create a sound, which can be localized by its frequent beeps and non-constant sound. Additionally, PWM signals were found to be more rigid and distinct in early testing and preferred in comparison to the harder-to-localize analog signals that were additionally examined. The speakers used by VISION are shown in Fig. 3.

Fig. 3. Speaker used for the user guidance system mounted to the billiards table.

C.UWB Beacons

The user localization system makes use of ultrawide-band (UWB) technology. A radio wave-based technology that uses low energy for short-range transmissions. UWB-capable beacons from Estimote are used due to their relatively cheap price compared to other alternatives, compatibility with U1 chips (making it compatible with all the latest iPhones), internal computing capabilities with an inherent ability to report a distance reading through a Swift-compatible SDK, a two-year long battery life, and inertial sensors to account for movable objects. Estimote offers the beacons in three packs which are shown in Fig. 4. The three beacons can be differentiated based on their colors: coconut, lemon, and caramel. The beacons are used by the user localization system to track the user around the table.

Fig. 4. UWB beacons from Estimote used in the user localization system.

The hardware setup will consist of three beacons mounted at strategic locations on the pool table. The beacons are mounted at corners or pockets, making them less intrusive for the player. Three beacons are needed so that trilateration can be done based on the user's distance from the beacons. The location of the beacons is then converted into our chosen coordinate system. As earlier mentioned, the compatibility of these beacons with the U1 chip allows VISION to introduce a second hardware element, the user's iPhone. The iPhone will run a mobile application discussed later in this paper.

D. Central Processing Unit (Jetson Nano)

VISION requires a significant amount of computational power, specifically for computer vision and artificial intelligence tasks. The central processor for the project is a third-party Jetson Nano development board from Seeed Studio. The Jetson Nano is a powerful embedded computer boasting a quad-core ARM processor, a 128-core Cuda GPU, and 4GB of RAM. The Jetson Nano does not support wireless connectivity by default. An aftermarket adaptor was added to provide the Jetson Nano with Wi-FI and Bluetooth 4.2 connectivity. Fig. 5. shows all major connections between the Jetson Nano and other system components. Fig. 6. depicts the exact Jetson Nano used by VISION.

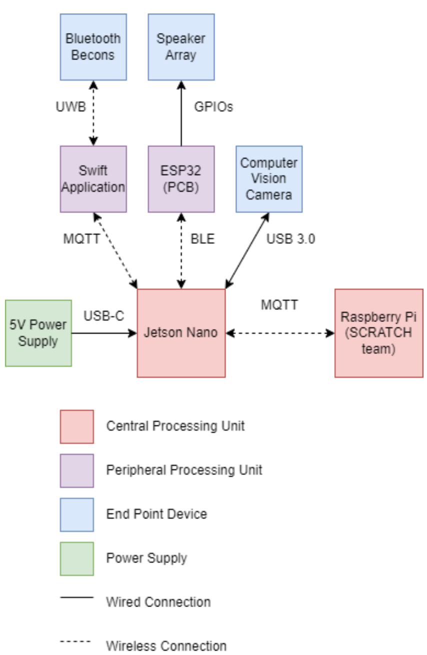

Fig. 5. Connection diagram between the Jetson Nano and peripheral devices.

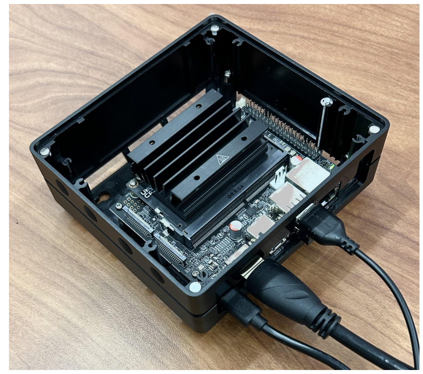

Fig. 6. Seeed Studio Jetson Nano used by VISION. The aftermarket Wi-Fi and Bluetooth module is connected to the bottom right USB port.

E. Printed Circuit Board (PCB)

To minimize wired connections to the processing unit of the user guidance system, a printed circuit board (PCB) was designed to facilitate several core subsystems of the design in a condensed format and location. Placed on the underside of the billiards table, the PCB includes a variety of components to ease the setup of VISION. These components include the core guidance processor, which is an ESP32, five, large push buttons for the user, a UART header for programming the ESP32, a linear power regulator, two 12-pin male headers, a 4:16 multiplexer/demultiplexer, and other small components.

The ESP32 is the heart of the guidance subsystem of VISION and is what allows the PCB to function. The ESP32 communicates via digital data pins to the demultiplexer, utilizes an internal pull-up interrupt for each push button, and contains a built-in Bluetooth antenna (shown hanging off the left side of Fig. 7. And Fig. 8.) to communicate with the Jetson Nano. The ESP32 receives commands from the Jetson Nano and uses a state machine to control the peripheral components. To properly power the ESP32 and other onboard systems, an ADP124 linear voltage regulator is wired to an inputted five-volt header that will be fed from the Jetson Nano 5V output line. This design allows for a consistent stepped-down voltage of 3.3V for system functions and keeps electronics on the PCB safe from any potential damages that may occur from the Nano or vice versa. Both the ESP32 and linear regulator also include bypass capacitors that will smooth out signal disturbances that may occur during the board's use.

The key to the speaker selection within VISION stems from the use of the demultiplexer and two 12-pin headers on the PCB. The multiplexer takes in four digital pins and can use them to determine which one of the 12 desired speaker outputs to send the discussed PWM signal through. Aiding in the process of guidance are five push buttons used by a user assistant to perform several commands. These commands include start, pause, end, redo, and a designated variable button for a custom feature if desired. Additionally, there is a sixth button which is needed in conjunction with a UART header to properly program the ESP32 in a similar way to what is done on more traditional development boards. This UART connection allowing for the data transfer from the computer to the ESP32 is protected by a pair of diodes on the data pins, preventing overcurrent draw in these critical avenues of data transfer.

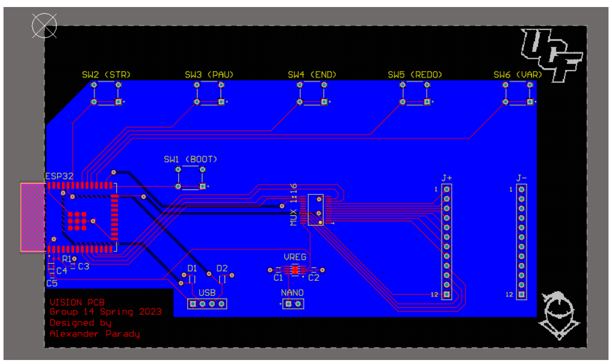

Fig. 7. 2-D Schematic View of the VISION PCB. This image was designed in Altium.

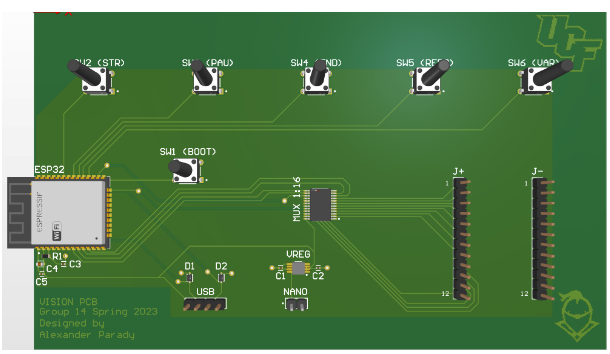

Fig. 8. 3-D Schematic View of the VISION PCB. This image was designed in Altium.

III. SOFTWARE DETAIL

A. Computer Vision System

The computer vision system is responsible for taking a picture of the current state of the billiards table, identifying the color of all billiard balls on the table, and localizing all billiard balls to a common location on the table. The output of the computer vision system is a list of the billiard balls on the table where each item in the list is the type and location of the billiard ball. This list is then used by the artificial intelligence system to determine the best shot to take. Once the shot has been completed, the computer vision system is used once again to determine the outcome of the shot and provide feedback to the user. Fig. 9. depicts the control flow of the computer vision system while taking the initial picture and while determining the outcome of the user shot.

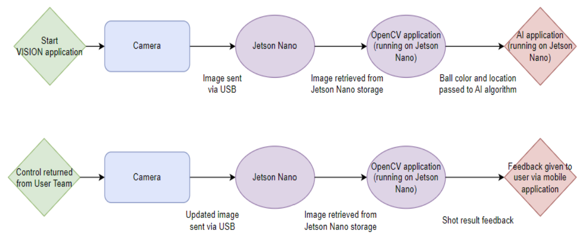

Fig. 9. Control flow of the computer vision system when taking a shot and determining the shot outcome.

The computer vision system utilizes OpenCV, a popular open-source computer vision library, to accomplish its goals. The Jetson Nano uses OpenCV methods to control the camera and take pictures of the table. The OpenCV implementation of the Hough Circle Transform detects the billiard balls in the captured image. Several measures are in place to validate the algorithm to ensure no billiard balls are missed and no false billiard balls are detected. Each ball in the image is sampled 121 times to extract the ball color from the RGB values of the image. The center location of each billiard ball is normalized to the upper left pocket on the table. This location is the common origin for all components of VISION. The ball color and location are added to a list subsequently used by the artificial intelligence system. Fig. 10. depicts a sample output image of the computer vision system after the billiard balls are localized.

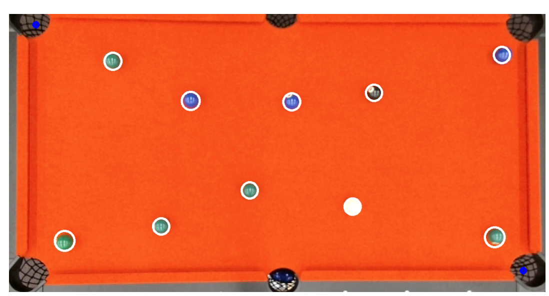

Fig. 10. Sample output of the computer vision system where detected balls are outlined with a white circle.

Once a shot is complete, the process is repeated to create a new list of billiard ball colors and locations. The new list of billiard balls is compared with the old list of billiard balls to determine what changes have occurred. The shot outcome is communicated to the user through audio feedback from the Swift application.

B. Artificial Intelligence System

The purpose of the artificial intelligence (AI) system is to assist players in making accurate and realistic shots in billiards. The system is heuristic-based and takes as input the (x,y) location and classification of each ball on the table. The output of the system is the relative angle of the shot as well as the desired force in newtons. A diagram of the artificial intelligence system is shown in Fig. 11.

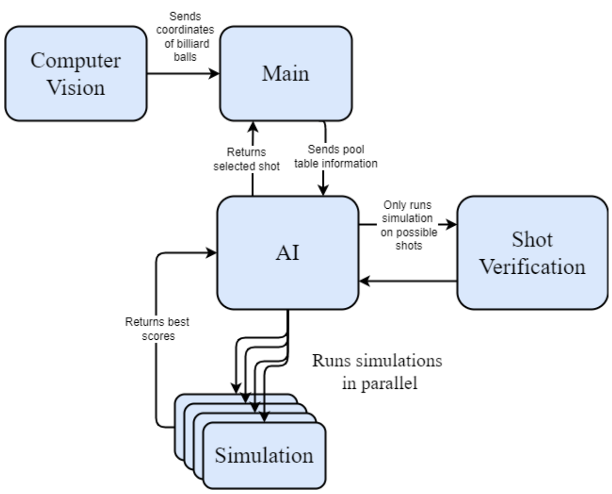

Fig. 11. Flow chart of the Artificial Intelligence system.

To begin the process, the AI system takes in the raw data from the computer vision algorithm and processes it to convert pixel values for x and y coordinates to actual units (feet and inches). The system then applies a correction algorithm to account for camera tilt, resulting in a simulation that closely reflects the actual state of the billiards table.

The system employs several key steps before outputting a realistic shot. First, the system verifies the possibility of the shot by using a ray-casting algorithm against a rectangle. The system then ensures that there are no obstacles within a 15-degree sweep of the cue stick by repeating a ray-casting algorithm against a circle. The 15-degree sweep of the cue stick is required by the SCRATCH team to ensure that the visually impaired user has sufficient room to adjust for their shot. If a show is not able to be taken due to insufficient space, the algorithm will choose a different location for the user to stand.

The next step for the system is to simulate thousands of shots using parallel processing for speed. Combinations of angles and forces for each shot are simulated and assigned a score. The highest-scoring shot is then selected and sent back to the main VISION program. The system includes short-circuiting logic to automatically stop all other processes and return the chosen shot if a certain score threshold is reached. The short-circuit logic is useful when a clear shot is available and an extensive search is not required.

The most critical step in this process is the heuristic-based scoring system used. The scoring system assigns points based on the number of pocketed balls, the number of opponent balls pocketed, whether the shot results in a scratch and other similar metrics. The scoring system also incorporates measures to limit the complexity of shots that are infeasible for a visually impaired user. As shot complexity increases the overall score decreases. Factors such as the distance the cue ball traveled before first contact, collisions with the table walls, pocketed ball collisions, total collisions, and total distance traveled are taken into account to decrease the score.

C. User Guidance System

The user guidance system uses a speaker array of 12 speakers positioned around the table to guide a user. 12 of the speakers shown in Fig. 3. are spaced equally around the table and represent landmarks for a user to use for navigating around the table. The user guidance system uses an algorithm running on the Jetson Nano to determine how many speakers and what speaker(s) to play next. The user is then guided around the table by the speakers which are ultimately controlled by a state machine running on the PCB. The Jetson Nano and PCB communicate through a Bluetooth low energy (BLE) channel.

The algorithm running on the central processing unit takes into account the target position and the current position, which are both expressed as speaker numbers. To improve accuracy, the area directly between two speakers is identified by a combination of the two neighboring speakers. For example, if a user is directly between speaker one and speaker two, the user's position would be represented as speaker one-two. The addition of such fine-grained positioning helps improve the accuracy of the user guidance system. If a user should be standing directly at one speaker, only that speaker will play. If a user needs to stand between two neighboring speakers, both neighboring speakers will play.

When guiding a user around the table, the user guidance system uses the four corners of the table as intermediate landmarks for the user. If a user needs to get to the other side of the table, they will first be navigated to a corner speaker, then another corner speaker, and finally to the necessary speaker on the opposite side of the table. This guidance scheme allows VISION to provide feedback to the player as needed.

To ensure the safety of the visually impaired user, the user guidance algorithm takes the location of the camera stand into account when guiding a user around the table. VISION will never instruct the player to walk around or over the camera stand to ensure that they do not trip on the camera stand. The user guidance algorithm will always guide a user away from the camera stand to guarantee the safety of a player.

Once the necessary speakers have been determined, the Jetson Nano will send the speaker numbers to the PCB over the BLE channel. VISION implements its own proprietary communication protocol to meet the needs of communication between its various systems running on a variety of devices. The guidance algorithm on the Jetson Nano will update as feedback from the user localization system provides feedback to the current location of the user.

Once it has been determined that the user is in the desired position (based on feedback from the user localization system), the user will be instructed to stop moving (also by the user localization system). The user will then be instructed to turn toward a speaker, without moving toward the speaker. A speaker will play for ten seconds to allow the user to orient themselves towards the sound to ensure that they are facing the proper direction. At this point, VISION will send information and hand control over to the SCRATCH team (via Bluetooth) to execute the shot.

D. User Localization System

The two key software components of the user localization system are the IOS application designed using Swift and a Python-based code that interfaces with the Swift application. The Swift application runs on the user's iPhone and communicates directly with the UWB beacons. The Python code running on the Jetson Nano uses MQTT to communicate with the Swift application. The Swift application sends the distance values of the three beacons to the Jetson Nano for more sophisticated processing to occur. The MQTT code that receives the data from the Swift application runs in a separate thread from the other VISION components.

The Swift app was designed to be user-friendly and tailored toward visually impaired individuals. In this sense, there are two tactile regions available to the user in the form of on-screen buttons. One that starts or resumes localization and the other that pauses localization. The buttons are wide enough to be easily differentiated, and there is audio feedback confirming which button has been touched. There is also tactile feedback through vibrations when either button is touched. The application reads the distance from the iPhone itself to the three beacons in meters. The current design of the application can be seen in Fig. 12.

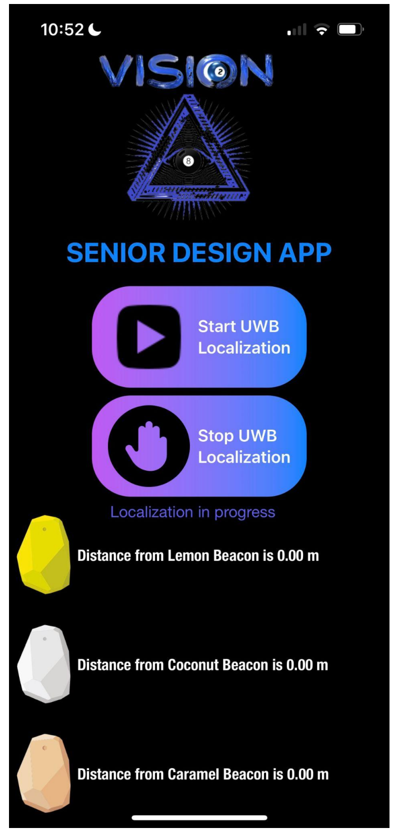

Fig. 12. Final design of the VISION mobile application used by the user localization system.

Once the distances have been acquired by the Swift application, the distances are transmitted to the Jetson Nano with MQTT. MQTT is a popular networking protocol for IoT devices that makes use of an external server to communicate between devices. MQTT is used over a direct connection between the two devices due to the restrictions placed on the UCF network. The Jetson Nano then computes the (x,y) position of the user through trilateration, smoothes the values to account for any inaccuracies in the reported beacon distances (due to multipath fading), and filters the values so the (x,y) position returned is a valid location around the billiards table.

The (x,y) location is then converted to a speaker location which can be a single speaker or a location between two neighboring speakers. A global lock is used to safely update a shared buffer so that the main VISION thread can have the updated user location. A simplified diagram of the user localization system is shown in Fig. 13.

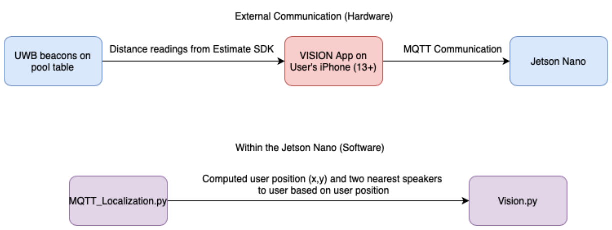

Fig. 13. Simplified flow diagram of the user localization system.

IV. SYSTEM DESIGN

A. Bluetooth Beacon Design

As mentioned earlier, the main system design for the Bluetooth beacons are linked to trilateration. For this reason, the beacons are positioned on the pool table at predefined locations consistent with the coordinate system used by the CV and AI systems. Beacon 1, the lemon beacon, is positioned on the top left corner of the table at (0,0) feet. Beacon 2, the coconut beacon, is positioned on the top right corner of the table at (6.3,0) feet. Finally, beacon 3, the caramel beacon, is positioned in the middle of the bottom side of the table at (3.2,3.37) feet. Trilateration uses all three positions and the reported distances to the user's iPhone to locate their user as shown in Fig. 14.

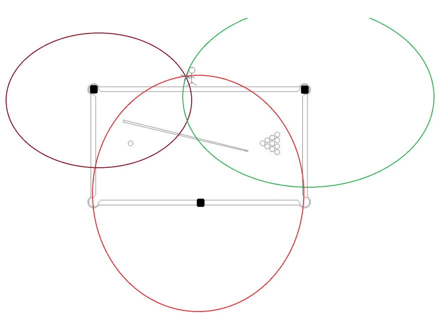

Fig. 14. UWB beacon trilateration schematic (trilateration circles not to scale).

B. Speaker Array Design

For the best realization of VISION's goals and given the dimensions of the VISION table, a 12-speaker array is utilized for the guidance of impaired users. The array system is used for well-defined navigation around the table and a proof of concept angular guidance mechanism as described earlier. The layout of the array, shown in Fig 15., lays speakers out every 19" across the 43"x76" table. There is one speaker at every pocket of the table and one speaker between each set of neighboring pockets. The use of 12 speakers and the capability of playing two speakers concurrently allows for a margin of error of approximately 8" which meets one of the toughest goals VISION had to accomplish.

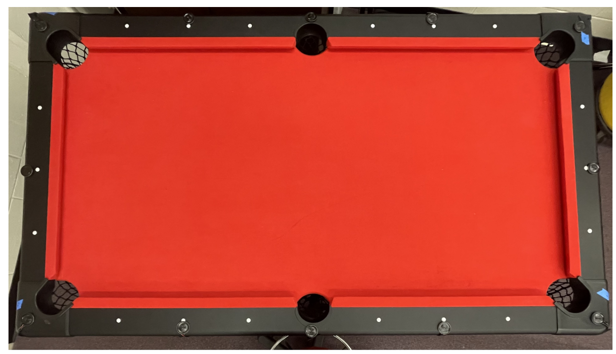

Fig. 15. VISION pool table with 12 guidance speakers

In addition to being used for navigating a user around the billiards table, the speaker array is also used for orienting a user in the proper direction to make a shot. VISION needs to orient a user within 15 degrees of the desired shooting direction so that SCRATCH can complete the actual shot. This is one of the most important goals of VISION because this is one of the major assumptions that the SCRATCH team designed their project around. Once a user has reached the desired location, the speaker array will be used to play a single speaker to have the player turn towards. VISION's speaker array produces a maximum margin of error of 8" in user location and 7.1 degrees in user orientation as shown in Fig. 16.

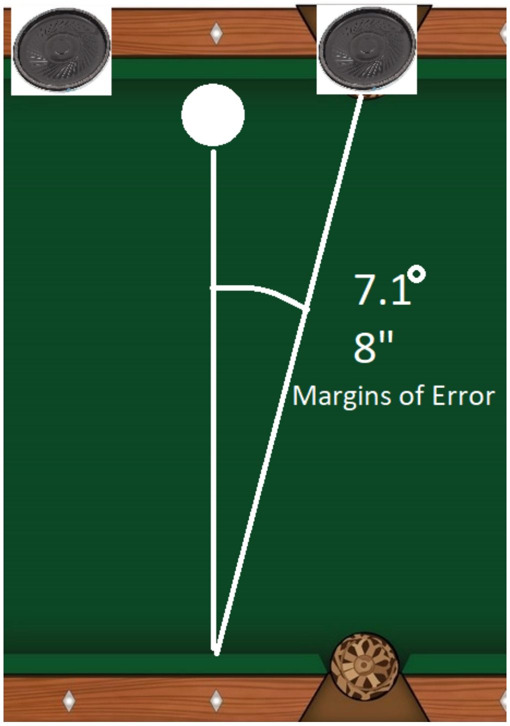

Fig. 16. Maximum margin of error of the user guidance system.

V. CONCLUSION

THE ENGINEERS

Aaron Crawford is a computer engineering student who interned at Leidos and Amazon as a software engineer. Aaron currently conducts cyber security research under Dr. Zou. Aaron hopes to work at a startup after graduation.

Alexander Parady is an electrical engineering student with a computer science minor who interned at Qualcomm, Tesla, and NEI. Alexander will be working on the technical sales and marketing team at Intel after graduation.

Arsene Landry Tatke is an electrical engineering student who interned at Siemens, I-CON Systems, and General Electric Digital. Arsene hopes to continue working with General Electric after graduation.

Noah Harney is a computer engineering student who interned at ICR inc. Noah currently works part-time at ICR inc. where he will continue working full-time after graduating as a software reverse engineer.

ACKNOWLEDGMENT

The VISION team would like to thank Dr. Chan, Dr. Richie, Dr. Wei, and the SCRATCH team for their hard work, dedication, and support throughout the project. The VISION team would also like to thank all of the other UCF faculty who have helped the VISION team grow to be a successful team of young adults who are prepared to enter the next phase of their professional careers. A very special thank you goes out to Dr. Abichar, Dr. Borowczak, and Dr. Suboh for volunteering their time to be the committee members for VISION's final presentation!