

CIS II: Final Report
Multisensory Navigational Aid for Visual Prosthesis Users
An Chi Chen
Mentors: Dr Seth Billings, Chigozie Ewulum

Technical Summary

1. Background and Significance

The Argus II retinal prosthesis system is a system for patients with end stage Retinitis Pigmentosa [1] – a condition in which there is a loss or breakdown of retina cells. Many patients in the later stages of this condition are classified as legally blind [2]. The Argus II is composed of an electrode array that is implanted into the back of the users’ eye. This is accompanied by a camera built into a pair of glasses along with the required processing unit.

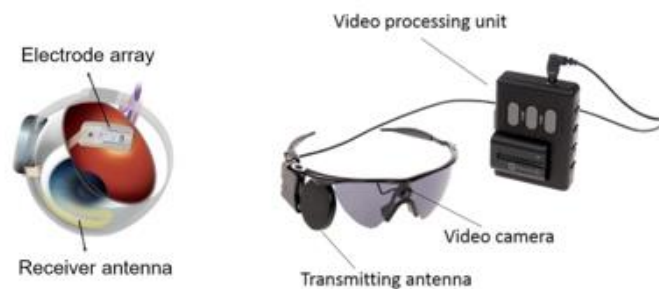


Figure 1: Argus II Retinal Prosthesis System Components

Information from the camera is used to create a stimulation image which is used to simulate the users’ viable retinal cells, using the electrode array. This simulation is perceived by the user as light patterns. This process is illustrated in the figure below.

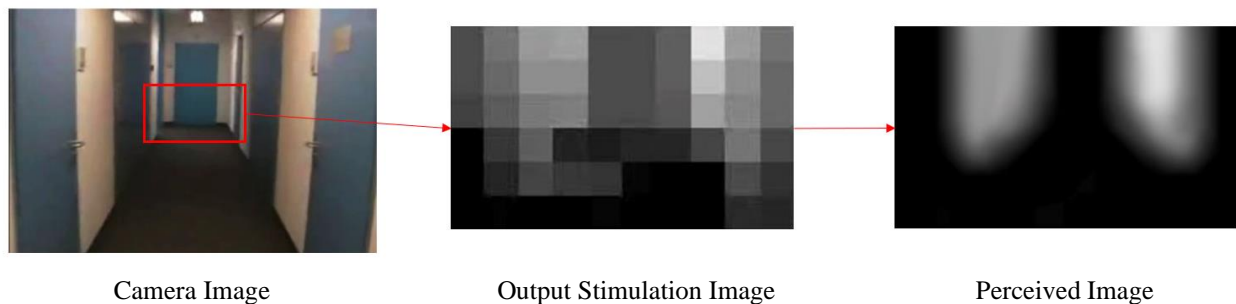


Figure 2: Process from Captured Image to Perceived Image for Argus II

There are limitations to this system that make it difficult for users to independently navigate a space. The artificial vision provided by the Argus II is not the same as actual sight as can be seen in Figure 2 above. In addition, the perception of this artificial vision varies between users. Some users are able to distinguish varying, though limited, light intensities whilst others can only discern the on/off states. Furthermore, as what the users see is determined by the camera, the field of vision is severely limited – users can only see the scene at which the camera is directed. There is also no peripheral vision which can be dangerous if the users turn their bodies before viewing the scene through turning their heads. This limitation also makes it challenging for users to find their targets as this requires physically moving their heads to scan their surroundings for their desired target.

Thus, the goal for this research project is to develop haptic and auditory feedback systems that work cooperatively with the Argus II retinal prosthesis system to aid users in independent navigation. The

supplementary systems are designed to assist with target navigation and object localisation. These additional systems should be intuitive, such as having a minimal learning curve, as well as maintain a low cognitive load.

2. Technical Approach

Prior Work

The Applied Physics Lab (APL) at Johns Hopkins University has already developed a SLAM-based navigational system for the Argus II system. This system makes use of an RGB and depth cameras to map a room and locate any obstacles. It is also able to determine the optimal route from the users' location to their target whilst avoiding the obstacles. For this project, the object detection capability will be of interest as it will be used for both the haptic and auditory feedback systems.

2.1 Haptic Feedback

The haptic feedback system's hardware consists of a headband fitted with 8 repositionable haptic actuators which connects to a custom 8-channel haptic driver. This is shown in Figure 3 below. The function and objective for this haptic system is to guide the users' gaze towards the direction of their target (e.g. a door). By doing so, this should assist the users in finding their targets and navigating to them more effectively.

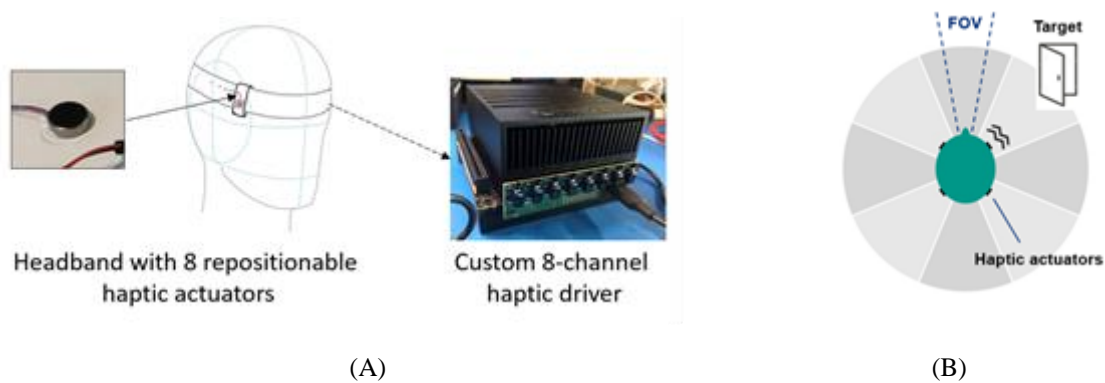


Figure 3: (A) Hardware and (B) Function and Objective of Haptic Feedback System

Headband Design

The haptic actuators that were used for this haptic headband are LRAs (linear resonant actuators). With respect to the placement of the actuators, in total, seven were placed around a headband as shown in the figure below. Four actuators located around the temporal and back of head – two on each side, and three actuators located along the top of the head.

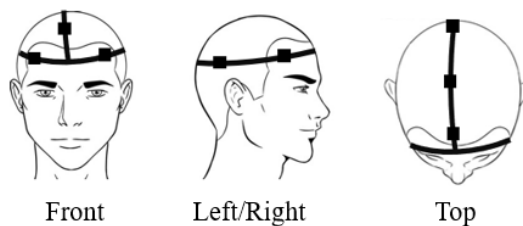


Figure 4: Haptic Headband Design and Actuator Placement

Haptic Pattern

Previous studies [3] have shown that there is some difficulty in determining the location of actuators in certain positions around the head. In order to evaluate the reality of this with the haptic headband described above a simple test was performed. In this test, the investigator attempts to determine which actuator was actuated with the headband on while the haptic actuators are activated randomly.

With respect to the directional haptic commands, there are various patterns which can be employed to relay a directional instruction. A static pattern involves only actuating motors on the side the users should direct their gaze. Whilst a dynamic pattern relays a direction by using perceived motion – the haptic motors are actuated in a specific order, such as right to left to relay a left direction. In order to determine which pattern is most intuitive, an iterative approach was adopted – testing various haptic pattern designs and making adjustments where necessary to produce an intuitive pattern.



Figure 5: Illustration of (A) Static and (B) Dynamic Haptic Pattern

Feedback Algorithm

The gaze guidance algorithm implemented is simply playing the appropriate haptic pattern, repeatedly, to direct the user towards their desired target, until their gaze reaches the target. A specific path guidance method was not implemented as a means to relay distance through the haptic headband was not investigated. However, pseudo path guidance can be achieved using the same algorithm as gaze guidance.

2.2 Auditory Feedback

The hardware for the auditory feedback system is a pair of open-ear bone conduction headphones (Aftershokz Aeropex). These headphones were chosen with the aim to not overwhelm the users' senses, due to the open-ear nature of these headphones they can still hear noises from the world. The function and objective for this system is to provide users with the location of potential obstacles using sound. The name of an obstacle will be played to the users through the headphones and sound like it is coming from the direction of the objects. This mitigates the need for users to manually scan their surroundings for obstacles before moving.

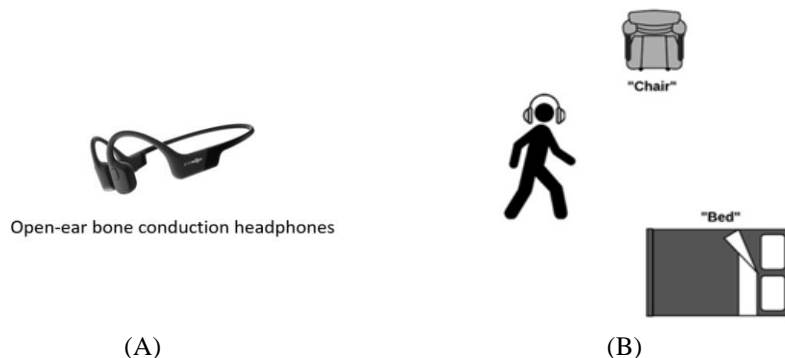


Figure 6: (A) Hardware and (B) Function and Objective of Auditory Feedback System

Individualising HRTFs

Head-related transfer functions (HRTFs) aid individuals in determining the locations of a sound as they describe how an ear perceives a sound originating from some point in space. HRTFs are unique to each individual as they are dependent on personal factors such as the shape of the ear and head. However, individualising HRTFs is currently not a well-developed process with various approaches suggested for creating personalisation however none producing the same results as using direct measurements. It is particularly challenging to produce an HRTF that allows users to localise sounds from the front and back accurately. This discrepancy – where sounds from the front sound like they are coming from the back and vice versa – is a common issue. For this aspect of the project, the focus was placed upon the feedback algorithm and thus the exact correctness of individualising HRTFs was not heavily examined. Though, two achievable individualisation methods were investigated. The first was a tournament style method. In this approach, a sound was played through headphones using various HRTFs (out of a set of generalised HRTFs) with the sound placed at various locations. After all the generalised HRTFs are tested the best fitting HRTF was chosen. The second method utilised an open-source application [4] that makes use of neural networks and measurements of the ear and head to produce an individualised HRTF profile. The inputs required for this application are a few measurements of the ear and head which are determined through direct measuring.

Spatialising Sound

Given an angle from a body (elevation and azimuth, see Figure 7(A)) an HRTF with the nearest matching angles can be found. In order to spatialise a computer-generated sound source the head-related impulse response (HRIR) was used. HRTFs are the Fourier transform of the HRIRs. The HRIR for the left and right ears were determined utilising the given HRTF to the system. After which, convolution of each HRIR with the sound source was performed to produce the correct signals for the left and right ear. The user would thus perceive the sound as if it were located at the simulated source location in real life – successfully producing a spatialised sound. Figure 7 illustrates the action of HRIRs as well as the algorithm followed to produce spatialised sound from a given HRTF.

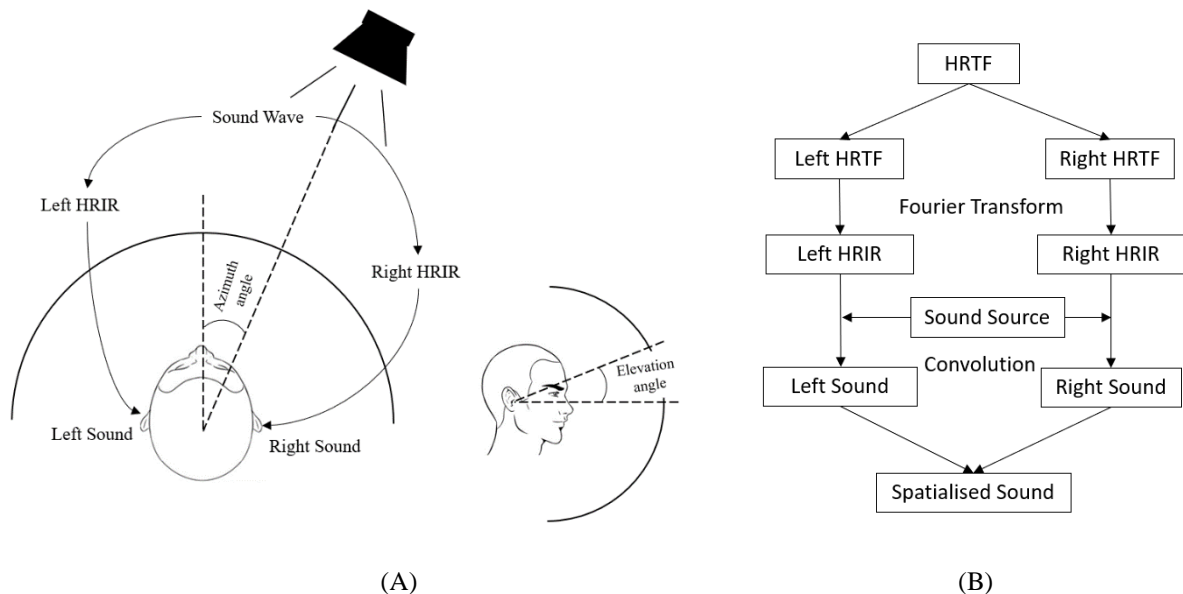


Figure 7: (A) Illustration of Sound Perception in Real World (B) Algorithm to Produce Spatialised Sound

Feedback Algorithm

To determine where an obstacle name should sound like they are originating from, the position of the obstacle with respect to the user was calculated (detailed in section 2.4). This was established for all obstacles identified by the SLAM system. To assist Argus II users with obstacle localisation, the name of each identified obstacle is played sequentially along with their distance from the user. The obstacles are played continuously at a set time frequency.

A fine localisation function was also implemented. The purpose of this function is to assist users in locating obstacles that are extremely nearby by playing the name of the obstacle followed by consecutive beeps. This function was implemented in this way as beeps allow for a faster feedback rate than obstacle names. When this function is being executed, no other obstacles are announced until the user steps further away from the nearby obstacle.

2.3 Visual Feedback

This feedback system will not form part of the supplementary system for Argus II users however it is a useful system to use when sighted users test the Argus II feedback system. This feedback system attempts to emulate what Argus II users see through the use of the virtual reality (VR) head-mounted-display, Oculus. A grid of 6x10 pixels was used as the viewing display – simulating what is seen by Argus II users. This system was used to display the landmarks that come into the field of view (FOV) of the user. In order to achieve this the horizontal and vertical angles between where the user is looking and the landmarks were determined (detailed in section 2.4). If the angles fell within the visible FOV they were displayed at the correct location determined by the horizontal and vertical angles. This is illustrated in the figure below.

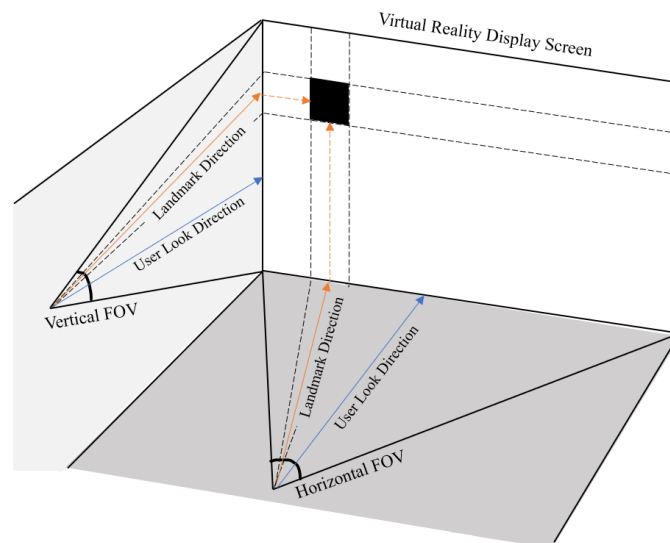


Figure 8: Positioning of Landmark on VR Headset Display

2.4 SLAM Integration and Mapping Calculations

For the purposes of this project ArUco markers were used to emulate real life landmarks – targets and obstacles. For testing the haptic feedback system, a landmark was manually selected using the computer interface. This would not be the implementation for Argus II users and was just used for the purposes of testing. For the audio feedback system, no selection of landmarks is necessary as all identified “obstacles” are played.

For the integration of the above systems with the developed SLAM system, the position of the landmark(s) with respect to the body (user) was determined. The calculations to determine this position as well as the required angles are detailed in this section.

The SLAM system returns the body and landmark pose with respect to a global world coordinate system. In addition, the look direction of the body is simply the x-axis of the body coordinate frame. These coordinate frames are illustrated in Figure 9.

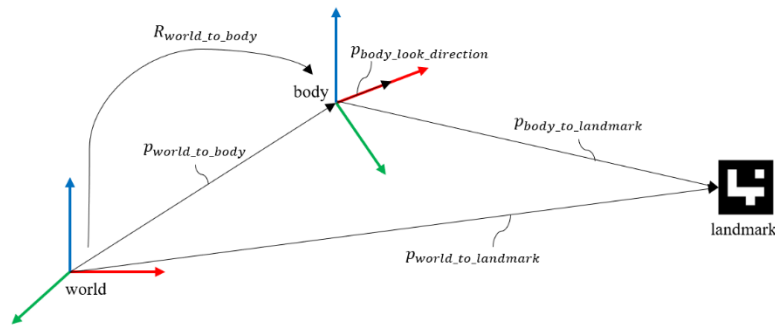


Figure 9: Relationship Between Coordinate Frames of Interest

Thus, the desired positional vector between the body and the landmark was determined using,

$$\vec{p}_{body_to_landmark} = \vec{p}_{world_to_landmark} - \vec{p}_{world_to_body}$$

Furthermore, the body look direction in world coordinates was found as follows,

$$\vec{p}_{body_look_direction} = R_{world_to_body} \times [0, 0, 1]^T$$

Finally, the horizontal and vertical angles between the body look direction and landmark position – with respect to the body – was determined using,

$$\theta = \tan^{-1}\left(\frac{\vec{u} \times \vec{v} \cdot \vec{n}}{\vec{u} \cdot \vec{v}}\right)$$

Where \vec{u} and \vec{v} are the vectors of interest (body look direction and landmark position) and \vec{n} is the normal of plane in which the desired angle lies; for horizontal angle, $\vec{n} = [0,1,0]^T$ and for vertical angle, $\vec{n} = [0,0,1]^T$.

3. Results

All the testing required for the methods described above was performed by and on the investigator for this project. This includes determining intuitiveness and accuracy of developed feedback systems.

3.1 Haptic Feedback

Haptic Patterns

The investigator found no trouble in determining the position of the haptic actuators during the random actuator activation test described in section 2.1. This provided useful information in creating the haptic patterns for the desired directional instructions as detailed in this section.

Through the iterative testing it was found that static patterns were optimal for relaying the left and right instructions. This was sufficient and intuitive enough to indicate to users in which direction they should look. All 4 of the actuators located around the temporal and back of the head were utilised for this group of instructions thus making it possible to indicate a direction located towards the back of the head (as seen in Figure 10).

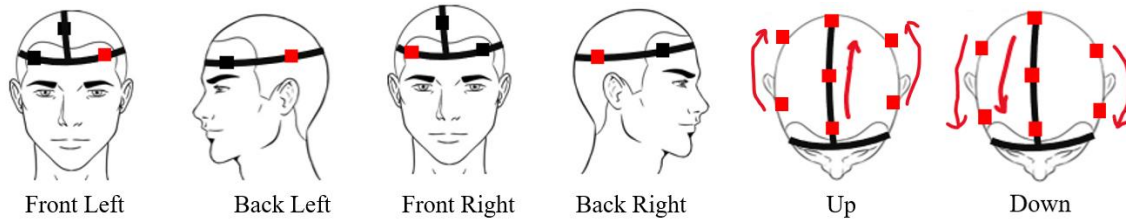


Figure 10: Haptic Patterns to Relay Directional Instructions

Dynamic patterns were employed to relay the up and down instructions. Interestingly, when only the 3 actuators located along the top of the head were used in the dynamic patterns the interpretation of the haptic pattern became unclear – potentially due to the lower density of actuators along the top of the head. Thus, these patterns make use of all 7 actuators actuated in a specific order. By using all 7 actuators, the feedback from the haptic pattern felt more distinct and clearer in relaying a directional instruction. The up pattern starts with the actuators closest to the front of the head and moves towards the back, whilst the down pattern starts at the back of the head and ends in the front – as illustrated in the figure above.

During the testing, it was realised that the duration of a haptic pattern should be as short as possible to allow for a fast update time. If this was not satisfied, there was a noticeable delay between a turning of the head and the next correct haptic instruction which results in an overshoot of head turning movement and creates confusion. Thus, for the static patterns the durations were set to 0.015 seconds. This may seem like a short duration but the haptic patterns are repeated continuously until the target gaze direction is reached. For the dynamic patterns, it is important that the durations are still set at a suitable time which makes them discernible as the directional instruction which they are relaying. A total duration of 0.18 seconds was found to be a sufficient duration that satisfies these criteria.

Furthermore, it was found that it is important that the haptic actuators are flush and are placed with a bit of pressure against the head in order to localise them properly. Additionally, hair density may result in a reduced ability to determine which haptic actuator is on. This potential issue could be solved by ensuring the haptic actuators are very close to the surface of the head.

Gaze and Path Guidance

For integration with the SLAM system, the horizontal and vertical angles described in section 2.4 were used to determine which haptic pattern to play. This algorithm is depicted in Figure 11.

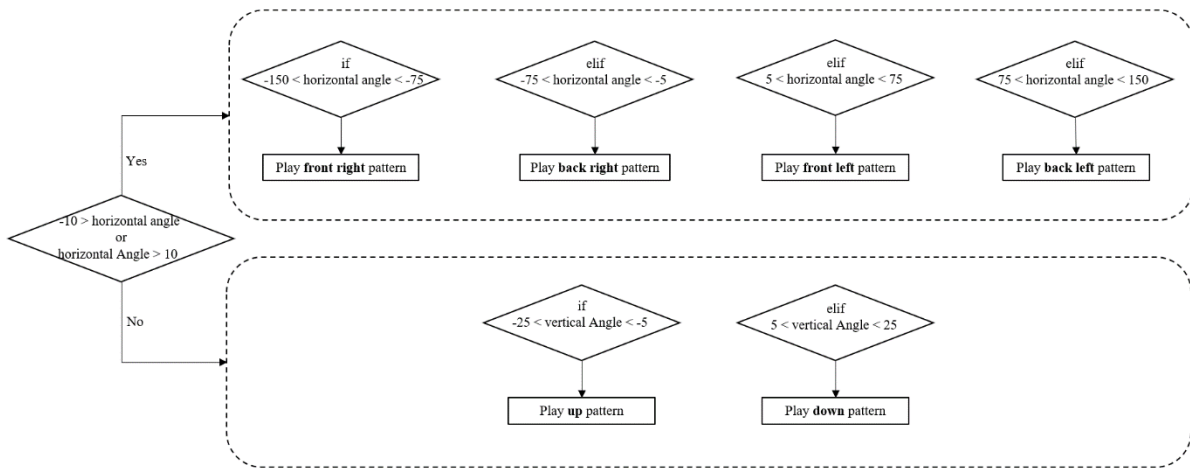


Figure 11: Algorithm to Determine Which Haptic Pattern to Play on Haptic Headband

In order to avoid confusion and maintain low cognitive load, two patterns were not permitted to be played concurrently. Thus, only after the horizontal angle between the user look direction and the desired landmark is within a certain threshold, the vertical position is guided.

The gaze guidance was tested using landmarks located at various position with respect to the user and was able to successfully direct the user's look direction towards and eventually find the desired target.

3.2 Auditory Feedback

Individualising HRTFs

A tournament to find a best fitting HRTF was set up using Unity with the SOFAlizer plugin [5] which facilitated switching between different HRTF profiles with ease. The HRTF contenders were obtained from the open-source databases by MIT, CIPIC and RIEC [6, 7, 8]. A total of 30 HRTF profiles were used during this evaluation. A completely accurate best fitting HRTF profile match was not found. In essence, of the HRTF profiles examined, none overcame the front/back discrepancy described earlier. Furthermore, this process was rather tedious as each HRTF had to be tested with the sound source located at varying position – the effort required did not equate to an extremely well fitted HRTF.

The second approach of using the application that generates an individualised HRTF based off of measurements of the ear was a much more concise process. After providing the input of a few measurements of the ear and head the application produced a comparable HRTF to that of the first individualisation method; although it was still not perfect. Nevertheless, as the focus for this project was on the feedback algorithm, and not the accuracy of individualisation of an HRTF profile, the HRTF obtained in the second method was used in the subsequent feedback algorithm testing.

Spatialising Sound and Feedback Algorithm

Similar to the haptic feedback system, the angles found using the equations outlined in section 2.4 were used when integrating the auditory feedback system with the SLAM system. With the addition of the method described in section 2.2, spatialised sounds were able to be accurately created and played through the bone conduction headphones. The position of the landmarks (ArUco markers) located along the side of the user could be successfully located solely based-off of the sound. Visual feedback provides assistance in locating sound sources from the front and back due to the front/back discrepancy, however, the use of a well-fitted HRTF profile would mitigate this issue.

The distances, in feet, between the user and the landmark(s) were also announced to the users. This was computed by simply calculating the norm of the vector between the user body and the landmark. It was found that if more than three identified obstacles (often will be the case), including their distances, were announced created a larger load on the senses and could potentially disengage the users. Therefore, in order to reduce the cognitive load, the algorithm was updated such that only obstacles within 10 feet of the body were announced. It was also found that the listing of landmark names along with their distances is slow and users can move rather quickly. Thus, the distances of the obstacles from the user were only announced every third iteration of the loop and not every single time the obstacles were listed in order to ensure adequate and accurate auditory feedback is provided.

3.3 Visual Feedback

The horizontal and vertical FOVs had to be fine-tuned to accurately reflect the landmark positions in real life. The final FOVs are:

Horizontal FOV = 80°

Vertical FOV = 60°

The figure below shows that the landmarks within the user's FOV were successfully placed and shown on the virtual reality headset display.

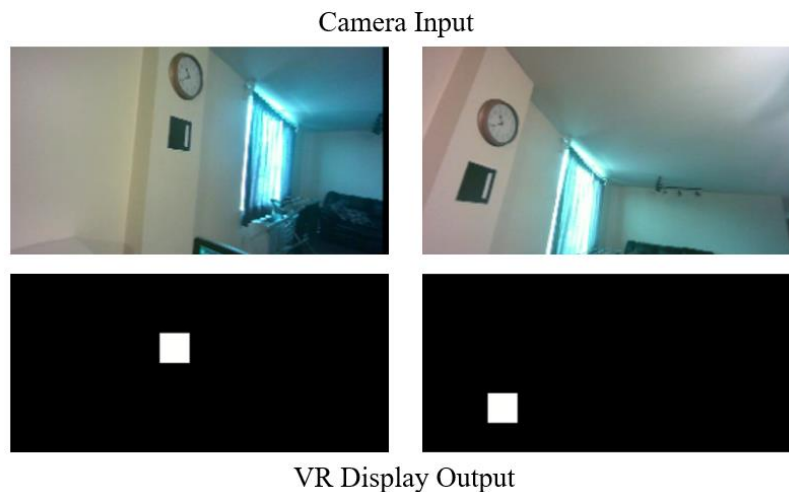


Figure 12: Resulting Display Shown on the VR from Camera Input

Management Summary

1. Work Distribution and Acknowledgements

The SLAM system including marker recognition and integration with the Oculus were provided to me by APL. The work outlined in the sections above were completed by myself as I am the only member in this group. I would like to thank my mentors, Dr Billings and Chi for their guidance and assistance throughout this project.

2. Deliverables

The project's planned minimum, expected and maximum deliverables are detailed below.

Minimum:

Operational haptic feedback system that can intuitively guide user's gaze direction

Expected:

Haptic feedback system with the addition of an auditory feedback system that is able to assist users with obstacle localization

Maximum:

Integrated retinal prosthesis system with haptic and auditory feedback that has been user tested. Evaluation on the system's performance in terms of target navigation and obstacle avoidance as a result of the participant testing

I have successfully completed the minimum and expected deliverables within this project timeframe as detailed in the earlier sections of this report.

The following table details the dependencies anticipated for this project.

Dependency	Responsible Person	Plan of Action	Fallback Plan	Date Achieved	Status
Mapping, SLAM navigation (incl. required hardware)	Mentors	Required parts have been made /ordered and delivered	Develop haptic and auditory system without SLAM navigation	28 Feb 2022	Achieved original plan
IRB	An Chi	Do the required training as soon as possible		-	Not Attained
Participant availability	An Chi / Mentors	Start the participant recruitment 4 weeks prior to planned testing	Test system with seeing participants using VR headset	-	Not Attained

As can be seen the IRB and participant availability dependency were not achieved. This was due to the fact that the existing IRB protocol only includes recruiting participants from APL or Argus II users. The developed supplementary feedback systems, though integrated, are not in a complete enough state to be tested with Argus II users at this stage. In addition, there was not enough time during this project timeline to execute the fallback plan for participant testing which was to test the developed integrated system with APL employees.

As a result, due to the time constraints and the IRB dependency not being met the proposed maximum deliverable was not achieved. The maximum deliverable was however adjusted to the following:

Maximum (updated):

Integrated haptic and auditory feedback system with simulated Argus II feedback on virtual reality headset

The above altered maximum deliverable was achieved as detailed in this report.

3. Future Work

The first next step for this project is to integrate the haptic and auditory feedback systems with the Argus II retinal prosthesis system. After which participant testing can be performed with Argus II users. An evaluation of the integrated system's performance can be completed – during which the system's effectiveness in target navigation and obstacle avoidance would be determined. Finally, a more complete investigation into the methods to produce individualised HRTFs can be performed as this could potentially improve the performance of the auditory feedback system.

4. Lessons Learnt

During the completion of the work outlined in this report I have learnt about the challenges in producing accurate individualised HRTFs without taking direct measurements. Furthermore, I also learnt how the placement of haptic actuators can greatly change how the feedback is perceived with respect to its intensity and accuracy in localising its position. With regards to project management, I learnt the benefits of having a schedule with intermittent milestones in addition to the main project deliverables. It greatly aids in gauging project progress and driving weekly agendas.

Technical Appendices

A more in-depth documentation of the hardware and software aspects developed in this project can be found in the documentation on the project wiki page [here](#).

The actual code developed during this project is located in the Bitbucket repository for the associated APL project. Access to the repository is only granted to those working on the project currently.

References

- [1] Arevalo, J. F., Al Rashaed, S., Alhamad, T. A., Al Kahtani, E., Al-Dhibi, H. A., & KKESH Collaborative Retina Study Group (2021). Argus II retinal prosthesis for retinitis pigmentosa in the Middle East: The 2015 Pan-American Association of Ophthalmology Gradle Lecture. *International journal of retina and vitreous*, 7(1), 65. <https://doi.org/10.1186/s40942-021-00324-6>
- [2] U.S. Department of Health and Human Services. (n.d.). *Retinitis pigmentosa*. National Eye Institute. Retrieved February 28, 2022, from [https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/retinitis-pigmentosa#:~:text=Retinitis%20pigmentosa%20\(RP\)%20is%20a,of%20side%20\(peripheral\)%20vision](https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/retinitis-pigmentosa#:~:text=Retinitis%20pigmentosa%20(RP)%20is%20a,of%20side%20(peripheral)%20vision).
- [3] Oliver Beren Kaul, Michael Rohs, Marc Mogalle, and Benjamin Simon. 2021. Around-the-Head Tactile System for Supporting Micro Navigation of People with Visual Impairments. *ACM Trans. Comput.-Hum. Interact.* 28, 4, Article 27 (August 2021), 35 pages. DOI: <https://doi.org/10.1145/3458021>
- [4] GitHub. 2022. *GitHub - davircarvalho/Individualized_HRTF_Synthesis: Synthesis of individualized HRTFs based on Neural Networks, Principal Component Analysis and anthropometry*. [online] Available at: https://github.com/davircarvalho/Individualized_HRTF_Synthesis
- [5] C. Jenny, P. Majdak, and C. Reuter, eds. SOFA Native Spatializer Plugin for Unity—Exchangeable HRTFs in Virtual Reality, AES 144th Convention, 2018, AES, Milan.
- [6] *HRTF measurements of a Kemar dummy-head microphone*. [Online]. Available: <https://sound.media.mit.edu/resources/KEMAR.html>.
- [7] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, “The CIPIC HRTF Database,” Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics, pp. 99-102, Mohonk Mountain House, New Paltz, NY, Oct. 21-24, 2001.
- [8] K. Watanabe, Y. Iwaya, Y. Suzuki, S. Takane, and S. Sato, “Dataset of head-related transfer functions measured with a circular loudspeaker array,” *Acoustical Science and Technology*, vol. 35, no. 3, pp. 159–165, 2014.