

MULTIMODAL TRAJECTORY PLAYBACK FOR TEACHING SHAPE INFORMATION AND TRAJECTORIES TO VISUALLY IMPAIRED COMPUTER USERS

ANDREW CROSSAN AND STEPHEN BREWSTER

Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow
Glasgow, UK.

[ac, stephen]@dcs.gla.ac.uk

<http://www.dcs.gla.ac.uk/~stephen>

There are difficulties in presenting non-textual or dynamic information to blind or visually impaired users through computers. This paper examines the potential of haptic and auditory trajectory playback as a method of teaching shapes and gestures to visually impaired people. Two studies are described which test the success of teaching simple shapes. The first study examines haptic trajectory playback alone, played through a force-feedback device and compares performance of visually impaired users with sighted users. It demonstrates that the task is significantly harder for visually impaired users. The second study builds on these results, combining force-feedback with audio to teach visually impaired users to recreate shapes. The results suggest that users performed significantly better when presented with multimodal haptic and audio playback of the shape rather than haptic only. Finally, an initial test of these ideas in an application context is described, with sighted participants describing drawings to visually impaired participants through touch and sound. This study demonstrates in what situations trajectory playback can provide a useful role in a collaborative setting.

Categories and Subject Descriptors: H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces ---Haptic I/O, Auditory (non-speech) feedback.

General Terms: Experimentation, Human Factors.

Additional Key Words and Phrases: Accessibility, Multimodal, Evaluation, Trajectory Playback.

1 INTRODUCTION

Without using the visual channel, it is difficult to present information where the spatial relationships between the data are important. Presenting non-textual information to blind and visually impaired people is therefore a challenging problem but is an important one to solve. For many everyday tasks, a person needs to be able to explore or create data where the spatial component of the data is key to its understanding. One example of such a task would be handwriting, where learning to generate the correct shape of a letter is difficult to do without visual feedback.

In particular for younger visually impaired children at school, mathematics can prove very challenging (e.g. Archambault *et al.* [2004], Edwards *et al.* [2006]). They are required to learn about shapes, geometry, graphs, charts, and formulae, which traditionally rely on the spatial layout of the data to provide important information. Sighted children, through a visual glance, can use the layout of a diagram, graph or chart

to very quickly gain information about it, whereas it is difficult for a visually impaired child to gain a quick overview of the same data. Similarly for geography, maps are important tools that are most easily presented visually. Presentation through text alone makes it difficult, for example, to convey the relative position, size, or shape, of countries.

Many of the information access problems experienced by visually impaired people are now being solved with computers. Although computer users generally rely heavily on visual feedback (with the graphical user interface playing an important part in interactions), users with little or no vision must rely on other modalities to access the same information. Screen readers, such as JAWS from Freedom Scientific (www.freedomscientific.com), have proved to be a successful solution for accessing the textual information required to interact with a computer. Dynamic Braille displays perform a similar function for situations where more discreet communication is required. However, these technologies generally only allow access in a linear manner (from the top left corner of the screen). Further to this, non-textual information such as pictures and diagrams are not easily displayed in this manner. The goal of the work described here is to examine techniques to enable users to explore information or shapes non-visually and to navigate and control computer interfaces in a non-linear manner.

As well as playing an important role in the perception of some types of information, shapes and trajectories can prove to be useful in control applications. Gesture has the potential to be a useful alternative control mechanism for blind and visually impaired users. A gesture can be performed with any device that allows continuous interaction - such as a mouse - and can be used to perform a wide range of control actions within an interface. The user would perform the gesture which the system would recognise, and perform the appropriate actions. Here the gestures can be trajectories relative to the user's current position with the shape of the gesture being the key factor. This means that users do not need to maintain an overview of their position within an environment to perform the gesture.

A previous study from Klok *et al.* [2005] examined gesture as an interaction mechanism for visually impaired users. The study looked at the performance of a visually impaired group as well as the group's acceptance of the technology. Results suggested certain gestures (such as a v-shaped gesture) were easy to remember and perform when the control action for the gesture was chosen carefully. This study demonstrates the potential for gesture as a useful interaction technique for blind or

visually impaired computer users. There are still questions about the best methods of teaching visually impaired people about shapes, and trajectories that might be important for gesturing.

1.1 Aims

The main aims of this work are to test the suitability of trajectory playback as a method of transferring spatial information to blind or visually impaired computer users. Users may be interested in the shape of an object on the screen or the path between two points on a map, or in situations where learning a particular motion is important such as learning to sign his or her name. We examine both haptic trajectory playback which drags a user through a trajectory and audio playback where varying the pan and pitch of the sound represents X-Y coordinates.

1.2 Outline

Section 2 examines related work in the area. This includes previous work on the presentation of spatial data and tactile diagrams with raised paper or computer technologies as well as describing previous work on communication through trajectory playback or gestures. Section 3 describes a study comparing the performance of sighted and visually impaired participants when recreating trajectories felt through haptic trajectory playback. Section 4 builds on this study, examining a multimodal solution to improve the performance of visually impaired participants in the task. A study examining the benefits of multimodal haptic-audio trajectory playback over haptic trajectory playback alone is described. Section 5 describes a short observational study on the use of these trajectory playback techniques presented in a more realistic setting, with the playback available as a communication channel between a sighted and blind user. Finally, Section 6 discusses and draws conclusions from the previous experiments.

2 RELATED WORK

2.1 Tactile Diagrams

Traditional methods of accessing diagrams use raised paper, which raises certain parts of the image to allow users to explore shapes or lines presented through tactile relief. Providing accessible tactile diagrams through this method is not a trivial task however. Many authors have noted that a direct translation of a visual diagram to a tactile diagram is in most cases not sufficient to provide accessible tactile diagrams (for examples see Challis and Edwards [2001], Kurze [1996], and Crossan and Brewster [2006]).

Eriksson [1999] examines this problem and describes several principles that are important when presenting a blind or visually impaired person with a tactile picture. One

key feature emphasised by Eriksson is ensuring that a textual description is available for both an overview of the image and a more detailed description of the individual components within the image. This work also found that it is important to ensure that users can build up a spatial representation of the image by allowing them to explore the relative positions of objects within the image.

While tactile diagrams provide an invaluable tool for allowing visually impaired people to browse non-textual information, they suffer a number of disadvantages. They are static representations that are difficult to modify without reprinting. They rely solely on tactile relief and cannot take advantage of any computer based technologies such as screen readers or dynamic tactile devices to aid comprehension. As such, a number of attempts to provide computer based or hybrid alternatives to raised paper have been investigated.

Bliss *et al.* [1970] describe an early commercially available device that was designed to make printed information accessible to visually impaired people by combining a camera and vibrotactile array. The user moved the camera over a document with his or her dominant hand. The printed information was then displayed to the user's non-dominant hand through a vibrotactile pin array with the dark areas of the document represented as vibrating pins and the light areas as stationary pins. Recent work by Ladner *et al.* [2005] and Horstmann *et al.* [2004] takes a different approach to making printed information accessible. Their work concentrates on the automatic generation of accessible diagrams from the visual printed diagrams. A diagram is scanned, and image processing techniques are used to identify features from the diagram that can be translated into an accessible form and rendered using touch and sound.

Kurze [1996] describes a drawing environment that combines swell paper – to create a physical line based representation of the drawing – with a stylus and digitiser to provide positional information to a computer. Verbal cues are used to label different lines on the image which can subsequently be read back by the computer as the user explores the drawing. Landua and Wells [2003] similarly describe the Talking Tactile Table, a hybrid system which combines a raised paper representation with a touch sensitive tablet for making diagrams accessible. Users can explore the raised paper representation as normal, but can also press down on an area of interest and hear information through speech about that particular area of the image. By using different pre-created raised paper diagrams that the Tablet system can distinguish, context sensitive speech information can be given to the user for multiple diagrams.

Wall and Brewster [2006] present a computer based system for accessing bar charts that shares many features with a raised paper diagram. The user navigates the image by moving a stylus over a graphics tablet representing the physical piece of raised paper. The user's non-dominant hand rests on a raised pin tactile display that provides a simple binary up/down signal to the user for the area around the user's cursor depending on whether they are above a dark area of the screen or a light area. One immediate advantage of this system over a traditional raised paper representation is that it is computer-based. Charts can easily and quickly be reloaded. The system can take advantage of the computer-based representation to track the user's movements and provide further feedback to aid the user to navigate the environment. Yu and Brewster [2003] describe a system to display bar graphs non-visually through a combination of force feedback, non-speech audio and speech. They were able to demonstrate significant advantages of their multimodal graph system over traditional tactile diagram approaches. McGookin and Brewster [2006] build on these results when describing a system that allows a user to explore a bar graph felt in negative relief through force-feedback. The novel feature here was the incorporation of multiple views of the same information that could be browsed in different manners for different purposes. A direct translation of the visual representation was presented (similar to Yu and Brewster [2003]) to allow a common frame of reference for exploring and discussing with a colleague with a visual representation of the same graph. An easy to browse 'sound bar' was also presented that allowed the user to get a quick overview of the data values through non-speech audio by running the cursor over the base of the graph.

The above examples demonstrate the benefits of using computer technologies to display spatial data such as diagrams, chart and maps. The ability to easily create, alter and save these diagrams is an important advantage over traditional technologies. Further to this, information can be presented in multiple modalities allowing users to make use of all available senses. The different forms of communication can be used to present different types of information such as non-speech sounds for quick browsing, or speech for detailed values.

2.2 Non-Visual Cueing for Teaching Shape Information

Audio has been used extensively to present spatial information to blind and visually impaired users. There are many examples of using audio to represent the shape of graphs and charts which are a traditionally heavily visual method of presenting information. For example, Brown and Brewster [2003] describe how pan and pitch can be used to

represent x, y values respectively when displaying an audio line graph. They were able to demonstrate that users could accurately draw the shapes of the graphs that they heard. Alty and Rigas [1998] describe their AUDIOGRAPH system that can be used to display simple diagrams to users. They vary the pitch of two tones of different timbres to represent x, y position in a two dimensional space. They demonstrate that it is possible to display simple shapes to a user through varying the pitch of audio tones. Zhao *et al.* [2005] use audio to display geographic information through active exploration using a tablet (or keyboard) based map system. They divide a map hierarchically into regions varying timbre, pitch, and pan to display a value for a region, along with the altitude and azimuth information. Percussive sounds are played to alert users when they move between regions. Kamel *et al.* [2001] combine audio feedback with a tablet for input to display simple graphical scenes. Users navigate the environment using the tablet, with audio cues alerting them when they enter different areas of the diagram. Shapes are represented by non-speech sound sources moving through space (using 3D audio techniques to move the source horizontally and vertically). Users can track the movement of the sound to recreate the shape. Changing the pitch of the sound is used to support the user for vertical movements, as it has shown that users have difficulty placing sources in the vertical dimension, when they are presented using standard 3D audio rendering techniques.

Training users to move through a path non-visually using tactile feedback has also been examined previously in the literature. Noble and Martin [2006] present a directional cueing study that examines a series of tactile guidance cues designed to guide a blind computer user around a shape. The user interacts with the VTPlayer tactile mouse with a series of eight tactile patterns indicating which direction the user should move in next. Similarly, Crossan and Brewster [2006] used tactile directional cues to guide a user through a complex path represented as a maze. Blind and visually impaired users interacted in a two-handed manner using a PHANTOM device (from SensAble Technologies) in the dominant hand to move around the environment while receiving tactile directional cues on the left hand to guide the movements. Both of these examples use tactile cueing to display spatial information, but there are also many examples where force-feedback is used to guide users around a trajectory.

Feygin *et al.* [2002] conducted a study into the possibility of providing training in moving along a trajectory using either visual only, haptic only, or visual-haptic guidance. Further to this, there were 2 conditions in which participants recalled the trajectories,

which were haptic-visual (where the participants saw their cursor as they attempted to perform the trajectory), and haptic only (where the participants attempted the trajectory with no feedback of cursor position). Results showed significant improvement in recreating the trajectory in all conditions between the first and the last attempt. The haptic only training mode performed significantly worse than the haptic-visual training mode, but not significantly worse than the visual only training mode.

Dang *et al.* [2001] discuss a constraint-based surgery training system that provides guidance to users by restricting their movements from deviating from a path. This method allows a user to follow the path taken for a procedure by an expert surgeon, but allows the user to apply the forces to perform the surgery. Yokokohji *et al.* [1996] similarly examine haptic force playback for the purposes of training in a simple task. The system they studied actively dragged a user through the motions required to perform a task to provide training in performing that task. However, no significant training effect was noted here as the task chosen for this study proved so simple as to not require training.

There have been several systems designed to develop skills in calligraphy. Teo *et al.* [2002] demonstrate a system where the position of a teacher can be recorded and played back to a student to aid in forming Chinese calligraphy characters. A separate Chinese calligraphy training system developed by Wang *et al.* [2006] examines the performance of users when drawing characters both with and without force-feedback guiding their actions, with results suggesting that users were more accurate when drawing with the force-feedback. Henmi and Yoshikawa [1998] discuss a haptic system to allow learning of Japanese calligraphy. The method used is to record the position and rotation of the teacher's pen along with the forces used and replay this to the student.

Gentry *et al.* [2003] demonstrate a system which allows the user and computer to collaborate on a dancing task. Here, users had to synchronise their moves to music and with the movements of the device. Force-feedback is used to guide the user and improve the synchrony between the user and the computer.

The above haptic and audio cueing examples show that these techniques have the potential to be used to transfer spatial information to blind and visually impaired users. Although there are many examples of haptic trajectory playback using force feedback to teach shapes of characters or physical skills, there is a lack of work examining trajectory playback as a method of transferring spatial information to visually impaired people. This paper attempts to address this issue.

2.3 Communication Through Gesturing

The key aim of this work is developing a system to allow the computer (or a second user) to convey spatial information to a user. Two notable closely related works are that by Graham and Argyle [1975] and Oakley [2003]. Both performed diagram perception studies (with sighted participants only) using a 'describe and draw' paradigm. Graham and Argyle studied empirically the effect that hand gestures had on the transmission of shape information. They examined participants describing complex abstract scenes using verbal description alone or verbal description and hand gesturing. Through independent rating of the closeness of the drawn pictures to the actual pictures, they were able to show significant improvements in communication when hand gesturing was allowed.

Similarly, Oakley [2003] studied the effect of gesturing on transmission of shape information but through a collaborative computer system. Computer mediated trajectory playback was used in this study. Oakley examined three conditions for communication of shape: through verbal discussion and haptic playback alone, through verbal discussion and visual feedback of the other user's cursor, and finally a combined verbal discussion, haptic feedback and visual feedback condition. The haptic only trajectory playback of the image was achieved by hiding both users' cursors during playback. The visual condition examined visual cursor playback combined with verbal description of the image. The participant drawing the image could see the describer's cursor as he/she was describing the image. Finally, a combined haptic, visual and verbal condition was also studied. The results suggested that visual only and visual/haptic conditions produced significantly better drawings than the haptic only condition although no significant differences were noticed between the visual only and the visual/haptic conditions. Oakley's results suggest that understanding of a shape felt through haptic trajectory playback can be improved by providing additional forms of feedback. As the visual channel is not available to our user group, we must attempt to compensate with additional feedback such as speech and non-speech audio.

2.4 The Trajectory Playback Library

Haptic trajectory playback is not a trivial problem. The two main issues to be addressed are stability of the algorithm and safety of the user, particularly when some haptic devices can apply enough force to injure a user. Loss of control of the end effector is a particular safety problem when the user may not be able to see the device. In the following studies, an implementation of a playback system based on the bead pathway

developed by Amirabdollahian *et al.* [2002], is combined with a PID Controller (Astrom and Hagglund [1995]).

A proportional-integral-derivative (PID) controller is a standard algorithm from the control engineering literature. The purpose of using the controller is to minimise the error between the current value of a system and a target value. In this case, we control forces sent to a force feedback device in order to minimise the distance between the user's current cursor position and the target position on the trajectory. As the user's cursor approaches the target position and gets within a threshold distance, the target position is moved along the trajectory by a preset amount. This is repeated until the target position is the end point of the trajectory. By carefully tuning the parameters of the PID controller and the playback system, the user will be dragged through a close approximation of the trajectory in a smooth and stable manner.

This trajectory playback system used for these studies is available as an open source library (more details can be found in Crossan *et al.* [2006]). The preset PHANTOM OMNI settings available in the library were used for the playback controller. Forces from the playback controller were capped at a maximum 3 Newtons for safety reasons.

3 EXPERIMENT 1

Previous work has suggested that there is potential to provide non-textual information through haptic trajectory playback. As this technique requires no sight, one as yet unexplored area where this is potentially useful is providing non-textual information to blind and visually impaired users. An experiment was therefore conducted to examine the performance of visually impaired users in recreating trajectories felt through haptic trajectory playback. The results were then compared against a sighted group on the same set of trajectories to provide a baseline for performance.

3.1 Methodology

3.1.1 Equipment. The device used for the trajectory playback in this study was the PHANTOM OMNI device developed by SensAble Technologies (www.sensable.com). The PHANTOM was chosen as the device for this study as it is a pen-based interface that provides a sufficiently larger workspace for the task, as well as high resolution force feedback strong enough to drag a user along a trajectory. When sighted users were taking part in the experiment, a curtained box was placed over the device to hide the device and the user's hand movements during the playback. The box provided a large enough gap not to restrict movement, while not allowing a participant to view the device or his or her

interactions with it. No visual feedback relating to the shape of the trajectory was given to participants at any time during the study.

3.1.2 Participants. Two groups of participants took part in the study. The first group (G_{VI}) consisted of 9 blind or visually impaired users from the Royal National College for the Blind in Hereford in the UK. Four participants were members of staff and five were students with four being female and five male. The age of these users ranged from 17 to under 50. There were four participants under the age of 30 (P_2, P_3, P_7, P_9), one under the age of 40 (P_4), and four under the age of 50 (P_1, P_3, P_6, P_8). Two participants (P_5 and P_8) were blind from birth and seven had some vision at some time in their life. Of these seven, two had no remaining sight (P_1, P_9) and five participants still had a small amount of residual vision (P_2, P_3, P_4, P_6, P_7). Two of these five participants (P_3, P_7) had enough vision to be able to use screen magnification software to interact with a computer and three did not (P_2, P_4, P_6). All participants had had a vision impairment for five years or longer.

The second group (G_S) contained 6 sighted participants. All participants were from the Department of Computing Science at the University of Glasgow. There were an equal number of male and female participants, and all were under the age of 35. All participants in both groups had prior limited exposure to the PHANTOM having previously taking part in short demonstrations or unrelated studies.

3.1.3 Task. The task set for both groups was to feel a 2D trajectory through haptic playback and then recreate that shape. Before playing each shape, the device was held at the centre of the workspace by the PHANTOM motors. This was to ensure that the user started in a position that had sufficient space in each direction to complete the trajectory. The shape was then played to participants by dragging their hand through it at an approximately constant rate (depending on the user's varying resistance to the movement). If the user resisted the movement of the device, the trajectory playback was paused until the participant's cursor got within 1cm of the current playback position. Pausing the playback above a 1cm error threshold ensured that the trajectory did not degrade and lose detail with time if the user resisted the movement. However, participants were strongly encouraged not to resist the movement.

At the end of the trajectory, there was a pause of 1.5 seconds where no forces were presented to the participant before the device was returned to its central workspace position. This pause was to provide the user with a chance to separate the playback trajectory from the centring movement.

Each trajectory was played to participants five times, which provided the participants with multiple attempts to experience it while still keeping the overall learning time short (~20s in total). Participants were then asked to recreate the movement three times to reduce the effects on the results of random variations in movement. When drawing the trajectory, participants held down a button on the PHANTOM stylus for the duration of the movement and released it when complete.

All participants had a small amount of previous exposure to the PHANTOM. However, each participant went through a training period before starting the experiment to familiarise the participants with the experimental task. Three sample trials were conducted before the study was started with simple, nameable shapes. The training shapes chosen were shapes 1, 2, and 3 shown in [Figure 1](#), (a circle, a triangle and a square). Participants were encouraged at this stage to hold the PHANTOM stylus as they would a pen.

All shapes played to the user during the study were two dimensional and set in the vertical plane. The vertical plane was chosen as it allows a larger workspace than the horizontal plane when using the PHANTOM OMNI. Nine trajectories were then played to the participants in a random order. These trajectories are shown in [Figure 1](#), as shapes 4 to 12. These shapes were designed by the experimenter to be more abstract than the simple, easily nameable training examples. They were also chosen to be sufficiently distinct from each other to aid analysis of the results (described in Section 3.1.4). A wide range of directions of movement and features (such as turning points) of varying size and directions were therefore incorporated into the shapes.

Deleted: Figure 1

Deleted: Figure 1

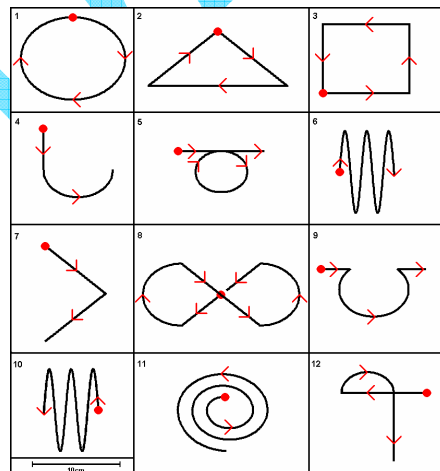


Figure 1: The trajectories used in the study. The circle on the trajectory indicates the start position with arrows indicating direction of movement.

3.1.4 *Analysing the Results.* The cursor trace data was analysed *post hoc* through a recognition algorithm adapted from a three layer MLP Neural Network (for more information see Bishop [1995]). The Neural Network was initially trained on one ideal example of the twelve trajectories where each trajectory was represented by the angle between 37 equally spaced points ordered chronologically over the entire trajectory. Recognition was considered successful if the appropriate trajectory was the largest output activation out of the twelve possible outputs of the Neural Network. The inputs to the Neural Network were the angle between 37 points that were an equal distance apart on the drawn trajectory. Shape and rotation of the trajectory were both factors affecting the Neural Network recognition. The recognition was not affected by the size of the drawing. Judging the performance of a user in generating a trajectory is not a trivial task for which a clearly defined, widely accepted, solution does not exist. The Neural Network is used here to provide an objective measure of closeness of fit to the ideal trajectory.

Standard sensitivity analysis techniques were used to test the robustness of the Neural Network for measuring relative closeness of fit of a trajectory to the ideal trajectory. For this analysis, five hundred examples of each trajectory with varying levels of added random noise were generated and classified by the Neural Network. Coloured noise (generated through a Gaussian Process) is used as a better approximation to user generated error than white noise, as smooth trajectories that take account of random variation and drift can more easily be generated. Figure 2 shows three examples of the noisy trajectories generated for a low, medium and high level of noise. Figure 3 shows the recognition rates for increasing levels of noise both overall and for individual trajectories. As shown, the recognition rates gradually drop as the trajectories become noisier suggesting that the Neural Network can be used in this instance as a relative measure of closeness of fit. It is important to note that the absolute recognition rate values are not important for this study as the analysis will concentrate on relative recognition rates between groups.

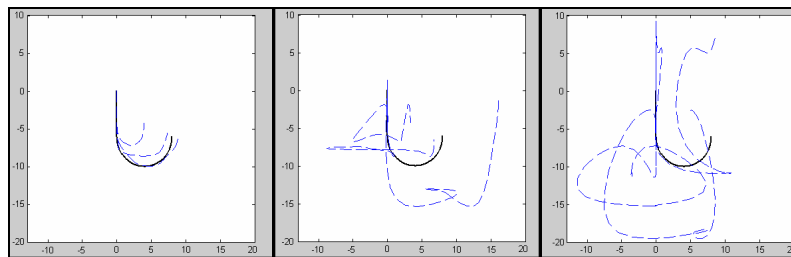


Figure 2. Nine examples of the noisy trajectories generated for trajectory 4 in Figure 1 are shown. The solid black line indicates the ideal shape. Three trajectories are shown with dashed lines for each of three different noise levels, incorporating drift and a standard deviation level of noise of 2cm (left), 6.4cm (middle) and 10cm (right).

Post hoc visual analysis of the cursor trace data was also performed by the experimenter to provide information about where variations in the trajectories were occurring and to identify situations where errors were consistently occurring.

3.1.5 Hypothesis. Hypothesis 1 (H1) was that the sighted group of participants will achieve a significantly higher proportion of correctly recognised trajectories than the visually impaired group due to the sighted group's greater experience at working with and visualising shapes and images.

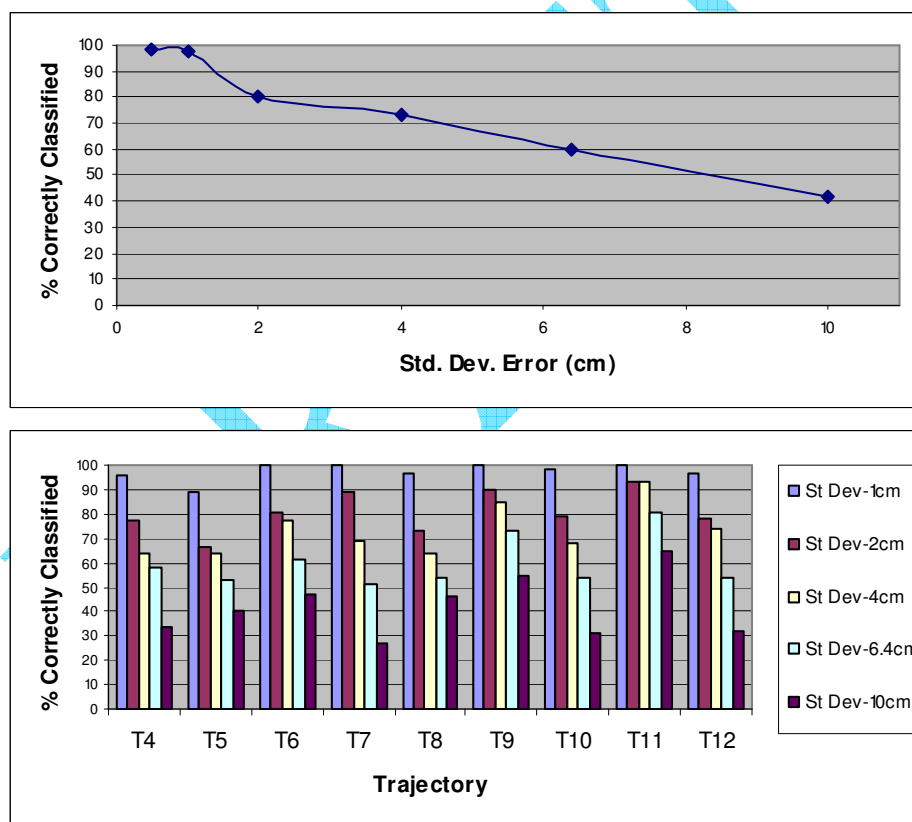


Figure 3. The overall recognition rates achieved by the Neural Network used during the experiment is shown for all trajectories (top) and for individual trajectories (bottom) when different levels of random coloured noise are added to the ideal trajectory.

3.2 Results

For each participant, the percentage of attempts correctly recognised by the Neural Network was measured. The mean percentage of correct responses for G_{VI} was 51.1%

(Std Dev = 26.3) compared with 74.7% (Std Dev = 5.9) for G_s . These data were analysed using a non-parametric Mann-Whitney test and a significant difference was found ($W = 53.0$, $p < 0.03$). Figure 4 shows the percentage of correctly recognised shapes for each participant in the visually impaired group. The percentage correct results for each participant in the sighted group are shown in Figure 5. Figure 6 shows the percentage of drawn shapes correctly recognised by the Neural Network for each individual trajectory.

3.3 Discussion

Results suggest that the sighted group was able to perform the task significantly better. This may be expected as the sighted group has everyday experience of gesture as a form of communication. Sighted people use gestures such as hand gestures everyday to aid communication. It is possible that the differences found here are due to the sighted group's greater experience visualising shape. The task also bears some resemblance to a drawing or handwriting task. Again the sighted group's greater level of experience at these tasks may lead to a higher level of performance. It is interesting to note that the highest level of performance recorded was by a participant in the visually impaired group. There is far more variability in performance in this group, compared to the sighted group whose results display low variability. The results suggest that some visually impaired participants found the task more difficult than their peers. The lowest recognition rates were noted for P_5 and P_8 who were both blind from birth. With only two participants in this category it is difficult to draw firm statistical conclusions here, but these results could indicate that the task is significantly harder for those who have never had sight.

From Figure 6, it can be seen that trajectories 5 and 11 were poorly recognised by both groups. This could either indicate the difficulty for the user in recognising and drawing the shape or could be the result of poor performance of the recognition algorithm for these two shapes. Figure 3 does not indicate a lower level of recognition for either trajectory, however, suggesting that it was the users that found it difficult to generate the shape.

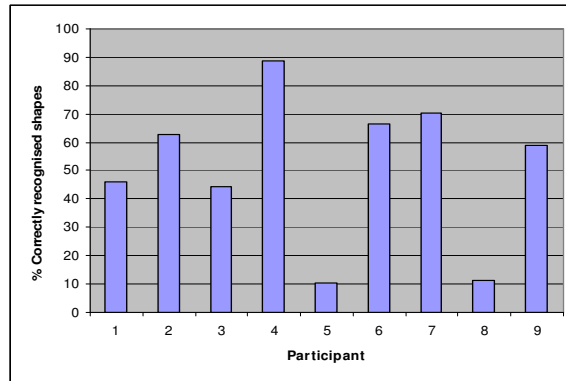


Figure 4: The percentage of correctly recognised trajectories for each participant in the visually impaired group.

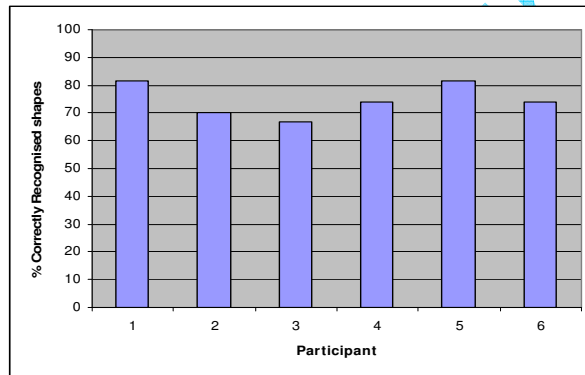


Figure 5: The percentage of correctly recognised trajectories for each participant in the sighted group.

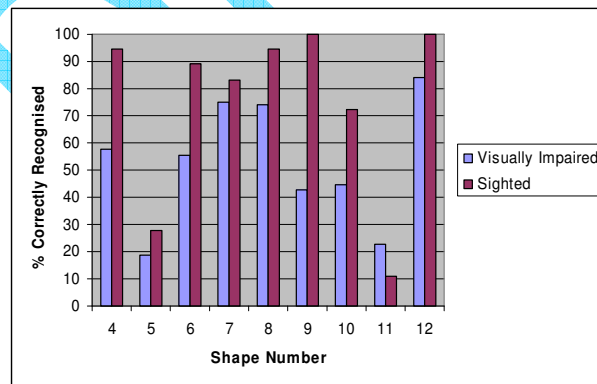


Figure 6: Percentage of correctly recognised shapes for both groups for each shape.

3.4 Observational results

After each of the training examples, participants were asked if they recognised the shape of the trajectory. For trajectories 1 and 2, participants reported the correct shapes 7 out of

9 times for the visually impaired group and every time in the sighted group. The square represented more of a problem. During playback, the control algorithm attempts to minimise the error between the user's cursor position and the playback position. However, the user's cursor position will lag slightly behind the playback position. When the playback direction turns through a sharp angle, this will have the effect of slightly rounding the corner of the path that the user has been dragged through. This effect is more noticeable when the turn is such that the new direction of movement is downwards. Here, the playback force combines with gravity to emphasise the rounding effect. The square was often represented more as a hemisphere when drawn by participants. Figure 7 shows one example of the actual position of the user's end effector when dragged through the square trajectory. It can be seen from the user's trajectory (the dashed line) that in particular, the sharp change of direction in the downward direction causes a rounding of the corner. One consequence of this is that designers must be aware that sharp corners will not be perceived by users when presented using this method.

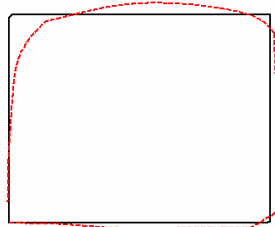


Figure 7: The dotted line shows the trajectory of a user being pulled along in a square shape (starting from the bottom left and initially moving right).

One factor mentioned by almost all participants after the experiment was the problem of segmenting the trajectories. To indicate the end of the trajectory, the device stopped applying a force through the device for 1.5 seconds before returning the user to the central position for the next playback. In some instances, the trajectory returning the user to the centre was included by participants in the playback (illustrated for shape 4 in Figure 8 with one participant's attempts to recreate the trajectory). Although this could only be seen in a small minority of occasions during *post hoc* analysis of the cursor trace, participants still reported confusion.

There were further confusions with the segmentation of the trajectories that were not reported by participants but observed in some cursor traces. The trajectories were all set in the vertical plane. At the end of every trajectory playback, the device motors were turned off allowing the user to move freely. Once the playback forces were turned off,

there was no force from the device supporting the user against gravity. In some occasions the participant's hand then moved downwards due to gravity adding the perception of a downward tail being added to the trajectory that was at times included in the user's drawings.

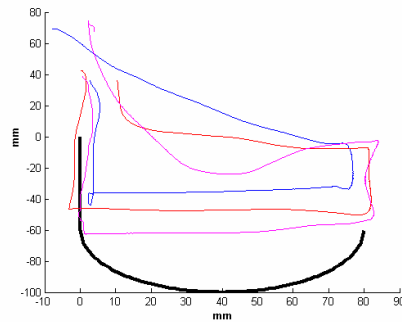


Figure 8: Three trajectories from one participant recreating Shape 4 where the centring force played at the end of the shape was mistaken for part of the shape.

3.5 Summary

This section presented a study examining perception of shape information through trajectory playback for blind and sighted users. From this study, it can be seen that while learning trajectories through trajectory playback is possible for both sighted and blind participants, there is a far greater variability in the level of performance for visually impaired users. It therefore becomes important to explore methods for improving the performance of blind and visually impaired people in this task.

Results also suggest that perceived shape of the trajectory can be altered when the user is released from the playback constraint once the trajectory is complete. The user's hand has a tendency to sink slightly before compensating for gravity giving the perception of vertical downwards tail to the end of the trajectory. A supporting force should be used to compensate in this event.

Segmentation of multiple shapes felt purely through haptic trajectory playback sometimes proved to be problematic. Additional feedback here could significantly reduce confusion.

4 EXPERIMENT 2

This second experiment builds on the findings from the previous section. It was clear from the initial study that the task was particularly difficult for some visually impaired participants. The goal of the second study was to test whether combining the haptic trajectory playback with complementary information from another modality could

improve the transmission of the shape information to the user. In this case, audio was combined with the haptic modality.

4.1 Combining Audio and Haptic Trajectory Playback

A similar approach to the haptic trajectory playback was used for the multimodal trajectory feedback system. However, to aid the user in determining the end point of a trajectory, the haptic feedback received at the end of the trajectory was changed. Once the trajectory playback was completed, users were held in place at the end point by the PHANTOM motors for two seconds before being returned to the central position. The goal here was to avoid users adding a tail to the trajectory as was noted in some traces from the previous playback experiment.

The goal of the auditory feedback was twofold: to help users segment the trajectory by providing more discernable start and end points, and to aid the user in discerning the shape of the trajectory. To help the user segment the trajectory, short distinct audio cues were played at the start and at the end of the trajectory. Further to this, during the drag back to the centre at the end of each playback, a wood scraping sound was played.

Audio cues were developed to help users to better determine the shape of the trajectory. Each trajectory in this study was limited to a two dimensional vertical plane. The audio cues were therefore developed to provide an indication of the current horizontal and vertical position in the environment. A sinusoidal tone was played to the user during the trajectory playback. The pitch of the tone was used to indicate vertical position. A higher pitch indicated a higher vertical position and a lower pitch a lower position. All shapes were either 12cm or 6cm in the vertical plane which corresponded to a one octave or a half octave change in pitch respectively. The frequency range over all the ideal trajectories was 200Hz to 400Hz with the midpoint (and trajectory start position) being mapped to a pitch of 300Hz. Audio pan was used to indicate horizontal position in the environment. As the user moved further left, the audio tone panned to the left and *vice versa*. The maximum pan was at 6cm horizontal distance from the centre (the maximum horizontal distance from the centre of the ideal shapes).

Two options were available when considering audio playback. The pitch and pan of the audio tone could either be related to the current controller position or the current user's position in the environment. The choice was made that the audio feedback corresponded to the user's position. Although the inevitable small deviations from the ideal trajectory ensured that a slightly different audio tone was heard during each playback, this ensured that the users could relate the sounds heard to their movements.

4.2 Methodology

4.2.1 Equipment. Similarly to the previous study, the device used for the trajectory playback in this study was the PHANTOM OMNI. In this study, the participants also had access to the keyboard for recording their drawing attempts. In the previous study the PHANTOM OMNI stylus buttons had proved problematic. In some instances, multiple button-down events occurred as the user held the PHANTOM button down causing multiple results files to be generated. While not affecting the participants or the results, it complicated the analysis of the trajectories as some results files needed to be concatenated to recreate the full trajectory. To avoid these issues, participants were asked to hold down the spacebar while drawing (instead of holding down a stylus button) with their non-dominant hand and release when they had finished. Two speakers placed to the front left and front right of the participants were used to play the audio. As in the previously described experiment, PID Playback LIB (described in Crossan *et al.* [2006]) was used for this study. Furthermore, the preset PHANTOM OMNI control parameters were again used for the playback controller.

4.2.2 Participants. Ten visually impaired participants from the Royal National College for the Blind in Hereford in the UK took part in the study. Four were female, with the age of the users ranging from 18 to under 45. Eight were aged under 30 (P1, P3, P5, P6, P7, P8, P9, P10), with one being under 40 (P4) and one under 45 (P2). Two participants were blind from birth (P1, P10) and eight had some vision at some time in their life. Of these eight, six participants still had a small amount of residual vision (P2, P3, P4, P5, P7, P9) with only two of those participants having enough residual vision to be able to interact with a computer through screen magnifier software (P5, P7). All participants had had their visual impairments for five years or longer. Five participants from the previous study took part again (due to limited access to visually impaired participants), although with a gap of three months between the studies. Experience in the task was not considered a major factor for this experiment due to the fact that there was a long gap between the studies. The within-subject design also ensured that all participants took part in both conditions to provide a fair comparison between the conditions.

4.2.3 Conditions. A within-subject design was used with all users performing two conditions in a counterbalanced order. The two conditions for this experiment were the haptic playback condition (similar to the previous experiment) and the combined haptic and audio playback condition, referred to as C_H and C_{HA} respectively.

4.2.4 *Training*. All but two participants had some previous exposure to the PHANTOM. These two participants were given time before the experiment to feel the device and try a standard demo distributed with the device. Each participant initially went through a training period before starting the experiment. Three sample trials were conducted before the study was started with simple, nameable shapes. The training shapes were similar to the shapes in the previous study. This training took place before each condition. Additionally in C_{HA} , participants took a short time (~1 minute) to freely explore the limits of the space to learn the mapping between their movements and the sound generated.

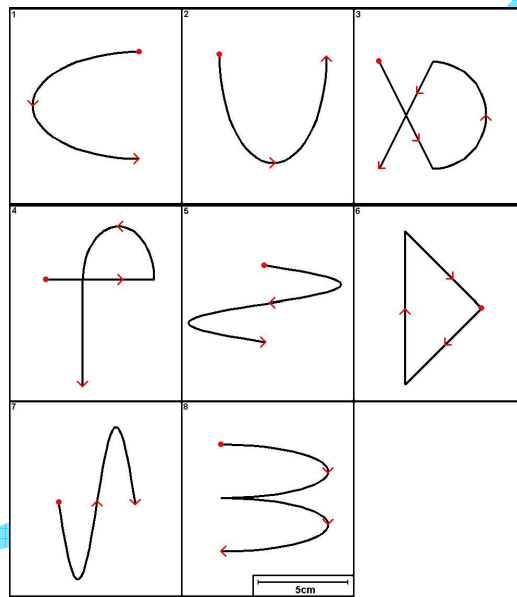


Figure 9: The trajectories used for the experimental task. The circle indicates the start point and the arrows the direction of movement.

4.2.5 *Task*. The task set for the participants was again to recreate the trajectories that were played to them. There were eight trajectories for each condition with the complexity of the trajectories in each condition being kept constant by using mirror image trajectories (flipped in the horizontal such that a leftwards movement becomes rightwards and vice versa) in either of the conditions. To aid analysis, the trajectories were again chosen to be distinct, with varying directions of movement and differing numbers of features (such as turning points). The order of presentation was randomised. Each participant would experience a trajectory five times and draw it three times. However, unlike the previous study, the drawing attempts took place after the 3rd, 4th and 5th trajectory playbacks. This design was chosen to examine user improvement as the

number of learning events increased. The eight shapes used in a condition are shown in Figure 9.

4.2.6 Analysis of the Results. Cursor trace data were recorded from this study at a rate of 100 samples per second. As in the previous study, the initial success of the participants in recreating the shapes was determined by a three layer MLP Neural Network. The Neural Network was trained on one ideal example of all sixteen shapes (the eight shapes shown in Figure 9 and their mirror images) used in the experiment. A shape was considered successfully recognised if the Neural Network returned the correct shape as the highest recognition probability for all sixteen shapes. To ensure a fair comparison, the cursor traces from the mirror image trajectories were flipped in the horizontal. Recognition for the corresponding images in each condition was therefore relying on the probability of the same output of the Neural Network. Similarly to Experiment 1, sensitivity analysis was carried out to test the performance of the Neural Network when different levels of coloured noise were added to the ideal trajectories. A Gaussian Process was again used to generate five hundred trajectories of varying noise level for each ideal trajectory. As shown in Figure 10, the recognition rates gradually drop for all trajectories as they become noisier, suggesting that the Neural Network can be used in this instance as a relative measure of performance.

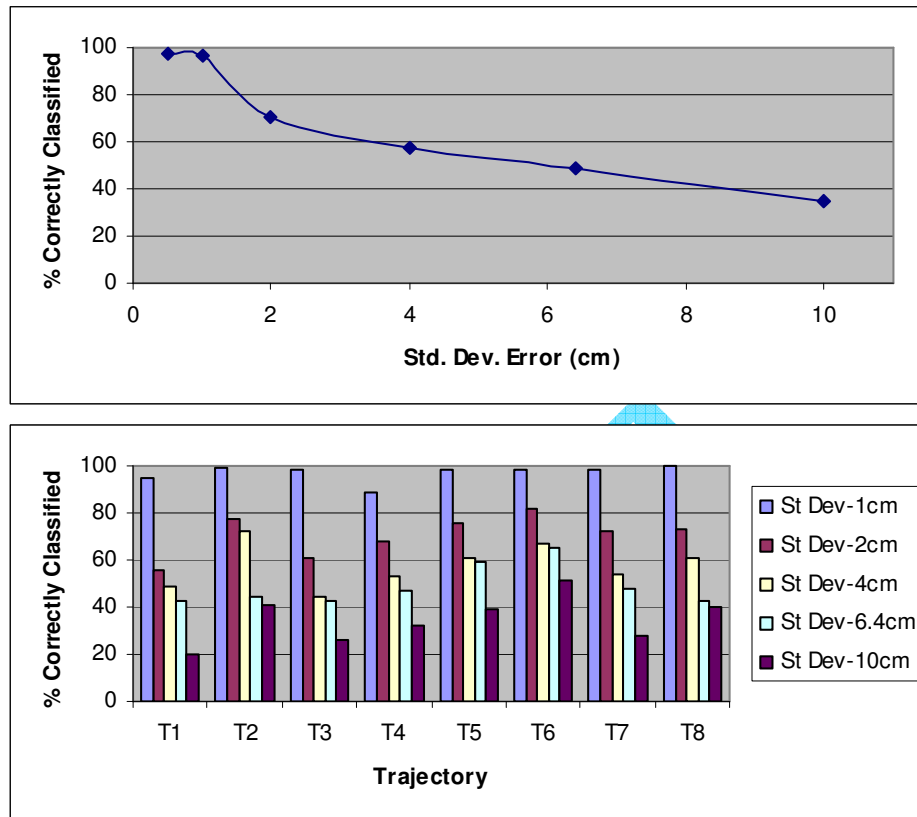


Figure 10. The overall recognition rates achieved by the Neural Network used during the experiment is shown for all trajectories (top) and for individual trajectories (bottom) when different levels of random coloured noise are added to the ideal trajectory.

Further to Neural Network recognition, three independent human raters were asked to judge the closeness of fit of each trajectory to the ideal trajectory. Human raters were used to give an estimation of closeness of fit rather than a simple binary 'recognised' or 'unrecognised' result. All drawing attempts were grouped by shape with the order of presentation randomised. The raters were able to browse on paper all attempts at drawing each trajectory and assign them a rating between zero and ten for closeness of fit to the ideal trajectory. The limits of the scale were 'No Match' (zero) and 'Perfect Match' (ten). No indication of participant identifier, experimental condition, or repetition number was given to the raters. The measures for closeness of fit were left for each rater to decide. Drawings were pre-scaled (while maintaining aspect ratio) to the maximum size of shape that would fit inside a fixed box. This was to remove size of drawing as a variable to allow a more valid comparison to the Neural Network results.

4.2.7 Hypotheses. The main hypotheses for the study were:

- H1 – The recognition rates for C_{HA} would be significantly higher than those for C_H
- H2 – There would be a significant improvement in recognition rates between the first and third drawing attempts in both conditions.
- H3 – The rating given to the drawings in C_{HA} would be significantly higher than in C_H

4.3 Results

4.3.1 Neural Network Recognition Results. The data for recognition correctness were tested for normality and found not to be normally distributed ($p < 0.02$). A non-parametric Paired Wilcoxon test was therefore used to test for significant differences between C_H and C_{HA} . Before being analysed by the neural network, the data for the mirror trajectories were flipped such that paired trajectories in each condition were compared to the corresponding eight trajectories in the Neural Network. For all Neural Network results, the flipped cursor data were used as it provides a fairer comparison between the conditions. Of 24 possible trajectories for each participant, there were a median of 21 (mean = 19.1, std dev = 5.40) correctly recognised trajectories for C_{HA} compared to a median 18.5 (mean = 17.0, std dev = 5.31) for C_H . A Paired Wilcoxon test was used to show that this difference was significant ($W_9 = 40.0$, $p < 0.05$).

To test order effects during the study, success of the participants due to order of presentation was also tested. As the conditions were counterbalanced, five participants performed C_{HA} first, and five performed C_H first. The median number of correctly recognised trajectories for the first condition presented was 19.5 (mean = 17.6, std dev = 4.74) compared to 20.5, (mean = 18.5, std dev = 6.08) for the second condition. These data were tested using a Paired Wilcoxon test, and no significant effect was found.

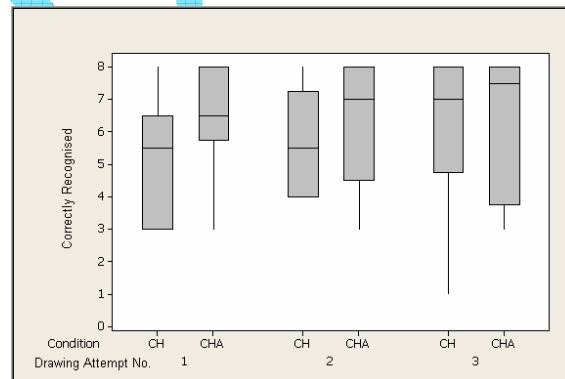


Figure 11: The median number of correctly recognised trajectories for all participants over all eight shapes for the three difference repetitions.

Learning due to multiple playbacks was also examined. At the first drawing attempt, a user had experienced a trajectory three times, at the second attempt four times and at the third five times. Figure 11 shows the median number of correctly recognised drawings for the eight trajectories performed in each condition over the three repetitions. A Paired Wilcoxon test was used to examine user performance for the different repetitions. There was no significant improvement in performance between the first repetition and the third repetitions in C_H ($W = 8.5$, $p = 0.11$) or C_{HA} ($W = 8.0$, $p > 0.9$).

The total number of trajectories correctly recognised for each participant in each condition is shown in Figure 12. Figure 13 shows the success of recognition by the Neural Network of each trajectory in either condition.

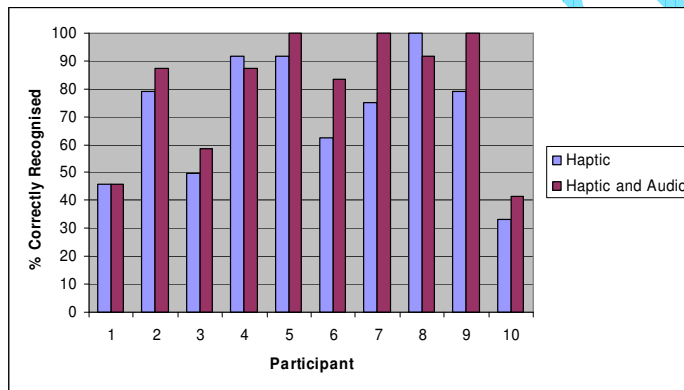


Figure 12: The percentage of trajectories correctly recognised for each participant in each condition.

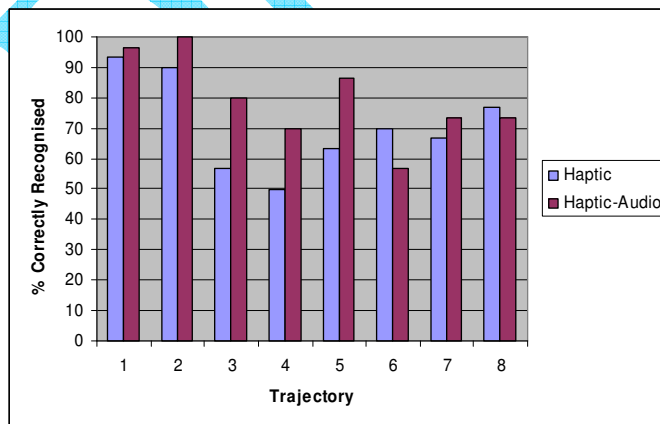


Figure 13: The percentage of correctly recognised trajectory's for all participants for each trajectory.

4.3.2 Rater Results. The results from three human raters were first tested for consistency between the raters to ensure the measures that the raters chose broadly

agreed. A Pearson Product Moment test for correlation showed a strong correlation between the ratings given by each rater when rating the participant trajectories with $p < 0.01$ in all cases (correlation coefficients: R_1 and $R_2 = 0.81$, R_1 and $R_3 = 0.83$, R_2 and $R_3 = 0.78$). As the results from the raters correlated, the rating given to each trajectory was then given as the mean of the three ratings given to the trajectory.

Initially the ratings given to the trajectories that were recognised by the Neural Network were compared to those that were not recognised (see Figure 14). The recognised trajectories were rated at a median of 6.67 (var = 2.89) compared with a median of 3.67 (var = 3.33) for the unrecognised trajectories. This difference was shown to be significant with a Mann Whitney test ($p < 0.001$).

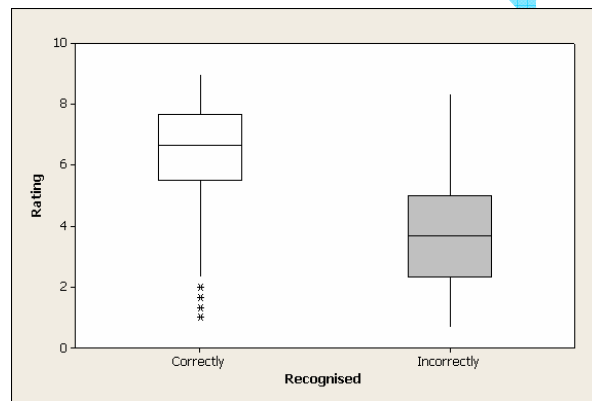


Figure 14: Graph of the ratings given to the trajectories separated by recognised and unrecognised trajectories.

When comparing ratings between conditions, a median rating of 6.33 was given to drawings in C_{HA} compared with 6.0 in C_H . A Friedman test was used to test for significance using paired comparisons of participants' ratings in either condition. There was no significant difference detected between C_H and C_{HA} ($S = 0.40$, $p = 0.53$).

Similar comparisons were made when looking at performance for the multiple drawing attempts. Figure 15 shows the median rating given to the first, second and third attempt at drawing a trajectory in both conditions. A Paired Wilcoxon test showed no significant difference in the rating between the first and third repetition for C_H ($W = 1408.5$, $p = 0.23$) or for C_{HA} ($W = 1311.5$, $p = 0.85$).

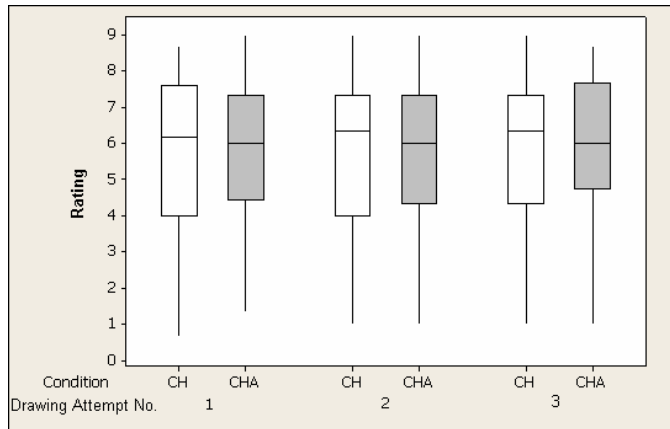


Figure 15: Median rating over the entire set of trajectories given to each attempt at drawing a trajectory in C_{HA}

4.3.3 Other Results. For both the Neural Network results and the human rater results, size was disregarded as a factor. The Neural Network recognised a small shape equally well as the same shape drawn larger and the drawings that the raters saw were pre-scaled (while maintaining the aspect ratio) to provide a valid comparison with the Neural Network results. However, sometimes the size of the trajectory is important for the task as well as the shape. The size of the trajectories drawn by each of the users was therefore compared to the ideal size. Cursor logs were analysed to determine the range of X (horizontal) and Y (vertical) values in each trajectory. The error in X and Y trajectory size was given as the percentage difference between the actual size and the ideal size with respect to the ideal size.

For C_{HA} , the median percentage X error with respect to the ideal size was 26% (conf. int. = 6.42) compared to 36% (conf. int. = 11.57) for C_H . This was tested using a Paired Wilcoxon test with no significant difference being detected ($W = 11.0$, $p = 0.36$). The median percentage errors in Y were 26.02% (conf. int. = 10.30) and 37.98% (conf. int. = 13.65) for C_{HA} and C_H respectively. These data were again tested using a Paired Wilcoxon test and the trajectories in C_{HA} were found to be significantly closer to the ideal in Y ($W = 0.0$, $p < 0.02$).

The additional audio cues proved successful for allowing users to segment the shape from the drag 'back phase'. From observation, only in one trajectory did one user confuse the 'drag back' section of a trial with the playback. The participant noticed the mistake at the second drawing attempt of the trajectory and reported it to the experimenter. Visual analysis of the cursor traces also indicated that the tail noticed in

the previous study (due to gravity affecting the user's movements at the end of a trajectory) was resolved by sticking the user temporarily to the end point of the shape.

4.4 Discussion

The results from the Neural Network recognition of trajectories indicate that combining haptics and sound can lead to significant advantages when learning the shape of trajectories non-visually. After three times experiencing the trajectory, visually impaired users were able to recreate the trajectory to such an extent that approximately 80% of trajectories drawn were recognised compared to 70% without the additional audio playback. These results support H1. Analysis of the multiple repetitions shows that user performance did not significantly change between the first and third drawing attempt in C_{HA} or C_H . This, combined with the high recognition results, suggests that users had a good idea of the trajectory even on their first attempt. The ceiling effect within the data that would then make it harder to detect significant differences between the groups. It was therefore not possible to support H2 with these data. One possibility for future studies would be to examine the performance of participants after a single trajectory playback instead of three as in this study. Alternatively, we could examine the performance of participants who performed less well in the task over a larger number of repetitions to determine whether performance can be improved with many more repetitions.

Similarly to the previous study, there was a high level of variability between participants. The two participants who had the least number of trajectories correctly recognised were the two participants who had never had vision. This suggests again that the task is most difficult for participants with the least experience of writing and of visualising shapes.

Despite the fact that mirror image trajectories were used in either condition, there was no significant improvement in performance detected between the first condition presented and the second. This suggests that little learning effect between the conditions took place due to the similarities of the shapes.

When comparing the Neural Network recognition results with the human rater results, there is agreement between more accurate and less accurate attempts at drawing a trajectory. This is indicated by the fact that the rating given to trajectories that were recognised wrongly by the Neural Network were given a significantly worse rating than those that were recognised correctly. However, no significant differences were detected between the ratings given to the drawings in C_H and C_{HA} . There are many possible

reasons for this difference. The Neural Network will be robust to some errors that may cause the raters to reduce their rating such as sharp corners being rounded or straight lines being slightly skewed or wavy. Conversely, the information given to the raters was different from that received by the Neural Network. The raters' task was a pattern matching task as opposed to a trajectory recognition task. The dynamic aspects of the trajectory were removed from the representation presented to the raters. With these results, it was not possible to support H3.

Analysis of the cursor trace data shows that users could more accurately represent the size of the trajectory using the audio and haptic feedback combined. Significantly greater accuracy was shown in the vertical dimension but not the horizontal. This can be explained by *post hoc* discussion with participants. The pitch was found to be far easier to determine than the pan for all users. Most users reported entirely disregarding the audio pan information as they found it difficult to concentrate on both pitch and pan at the same time.

These results suggest that participants were able to use the audio to achieve better performance. In the multimodal condition, significantly more trajectories were recognised, and trajectories were significantly closer to the ideal size in the vertical dimension. However, given user feedback about audio pan being difficult to determine in this instance, different methods of providing the second dimension of information should be considered. One option would be to adopt the method used by Alty and Rigas [1998] where the pitch of two distinct tones was varied to display two dimensional shapes. Alternatively, spatialised audio (similar to that used by Kamel *et al.* [2001]) has the potential to be useful in this instance. Blauert [1997] discusses a number of studies examining user performance when locating sound sources placed around the user. The results from these studies show that a high degree of accuracy can be achieved when locating a sound source in the horizontal plane (within approximately 4° when the source is directly in front of the user). User performance is poorer in placing sources in the vertical dimension. Therefore, by combining techniques - varying pitch to display vertical position and spatial audio moving the sound source horizontally around the front of the user - we may provide a more successful audio solution. However, a further study is required combining this technique with haptic trajectory playback before conclusions can be drawn here.

5 COLLABORATIVE TRAJECTORY FOLLOWING

5.1 Introduction

Sections 2 and 3 presented two studies examining the performance of users in the task of recreating a trajectory experienced through trajectory playback. This section uses the results of these studies to incorporate the techniques into an application such that they can be tested in a more realistic environment. The trajectory playback will now be tested as a tool to aid collaboration between a blind and sighted participant.

One area where trajectory playback can potentially contribute is in providing an extra channel of communication to help a sighted teacher work with and explain concepts to a blind student. Verbal communication will always be an important method of communication in this instance. However, the addition of haptic trajectory playback allows the teacher to interact with the student by directly moving them through the environment. This could be to guide them to areas of interest, or to explain a shape, concept or some tacit knowledge that is difficult to explain with words alone. Here, we use a teacher-student style of collaboration to examine such a situation. The task chosen was for the sighted teacher to describe to the visually impaired student a series of simple diagrams through touch, sound and speech such that the student can draw them.

5.2 Aims

Creation and perception of tactile diagrams by visually impaired users is a complex task. This work will attempt to study techniques for improving these processes by combining force, speech and non-speech audio feedback. Similarly to the collaborative gesture study described by Oakley [2003], the key research question will be, does the addition of haptic trajectory playback to the environment significantly improve creation of a diagram? However, a visually impaired user will be creating the image with additional forms of feedback not present during Oakley's study.

Some important key issues that are addressed in this work are collaborative learning, non-visual navigation of a 2D space and providing tools to help users to build up a spatial representation of and spatial relationships between objects in the environment. In this work, techniques from Kurze [1996], Oakley [2003] and Eriksson [1999] are combined to provide a prototype system that allows two users - one sighted and one visually impaired - to work together to interpret and build tactile images. A collaborative drawing environment is described that examines the learning of shape information through gesture (in the form of trajectory playback).

5.3 Collaborative Drawing

5.3.1 The Environment. The drawing environment (shown in Figure 16) allows users to draw simple line drawings within a rectangular workspace like a pen and paper. The user moves through the environment with the PHANTOM OMNI force-feedback device.

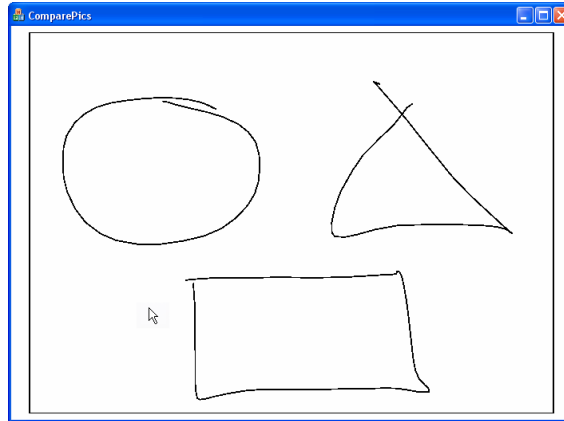


Figure 16: A screenshot from the drawing application showing the workspace, three drawn shapes and the user's cursor.

In the simplest case, with no interactions other than moving through the workspace, the basic force-feedback the user receives is simple. The walls of the workspace are represented as stiff springs such that the user is constrained by the motors to a rectangular area inside the workspace boundaries. The canvas is represented as a stiff wall in the vertical plane.

Two users can be present in the environment, and for the purposes of encouraging collaboration a teacher-student metaphor is adopted. We assign different roles to each user: one is the 'Describer' and one is the 'Drawer'. The Describer has some knowledge of what image is to be drawn (or is currently drawn) and the Drawer draws the image. Both can interact with the environment but in different ways.

The Describer acts as the teacher and may or may not be visually impaired (but with good residual vision). This user has some knowledge about a diagram that they must convey to the visually impaired Drawer. The Drawer acts as the student and has no previous exposure to the diagram such that the teacher must use words and gesture to describe the diagram. They each interact with the environment through a PHANTOM. The Describer can also see the Drawer's interactions on the screen.

5.3.2 The Drawer. The Drawer can draw a shape by holding down a button on the PHANTOM while touching the virtual paper surface. He or she can draw in the workspace like a pen on a whiteboard. When the user finishes drawing a shape and

releases the PHANTOM stylus button, a window is popped up allowing the user to enter a text label for the image. This is read out when the user enters the shape boundary while moving through the environment

The Drawer can also browse the objects in the scene by repeatedly pressing the left or right arrow keys to cycle through a list of the objects in the scene. The text label associated with the current object is read out as the user presses the key. For example, in Figure 16, as the user presses the right arrow key three times, the system may speak the labels “circle”, then “triangle”, and finally “rectangle”. Further, the user can hold down the down arrow key to be dragged by the PHANTOM to the current object’s location. Finally, the Drawer can play back the trajectory of the currently selected shape by pressing the second PHANTOM stylus button while inside the shape boundaries.

The Drawer also receives continuous non-speech audio cues to alert him or her to the cursor position within the environment. Similarly to C_{HA} in Experiment 2 above, the horizontal position is mapped to audio pan while the vertical position is mapped to the pitch of the audio. While better mappings may exist, this method had been shown in the previous experiment to improve the performance of users in recreating shapes over haptic playback alone.

5.3.3 The Describer. The Describer can traverse the image like the Drawer. However, he or she can also directly interact with the Drawer by grabbing the Drawer’s cursor and dragging it through a path. In this environment the Describer held down a button on his or her device and moved through the trajectory releasing the button at the end. Once the button was released, the Drawer was moved to the start of the trajectory and dragged through the Describer’s path at a constant speed. It was envisaged that this system would be used to aid the Describer when trying to describe a shape to the Drawer.

5.4 Evaluation

5.4.1 Method. An initial evaluation was carried out at the Royal National College for the Blind (RNCB) in Hereford in the UK. Although the techniques described here build on the previous experiments, this was the first attempt at integrating the techniques into a functional application. At this stage, it was therefore important to test all the features of the interface to determine whether they were usable and useful in an application context. This was achieved through an observational study.

Three pairs of individuals took part in the study. Two of the groups consisted of participants from the RNCB, and the third group consisted of one participant from the RNCB and one participant from the University of Glasgow. In each case, the Describer

was either sighted or had good residual vision, and the Drawer was visually impaired with very little or no residual vision. [Figure 17](#) shows the drawings used in the study. Drawing A contains simple and easy to describe shapes and was used to introduce users to the system. Drawings B and C were designed by the experimenter to contain more complex shapes with the relative positions of some of the shapes becoming more important.

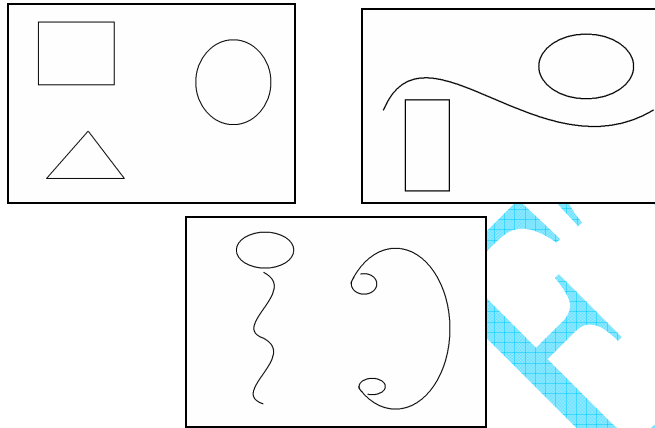


Figure 17: The three line drawings that were used in the study: A (left), B (right), C (bottom).

All interface features were described to participants, and a number of tasks were set. The three line drawings were given to the Descriptor on paper with the goal of the task being for the Drawer to use the software to create a representation of the paper drawings on the system. The Drawer and Descriptor were physically co-located so could talk easily. The methods used for this study were left to the participants to discuss. Participants were encouraged to discuss the drawings and their actions throughout the study. They were then observed solving the tasks using the interface features that they felt were appropriate. Video and audio recordings were then examined *post hoc* to determine the situations in which the various interface features were useful.

For the previously described experiments, the feature that was of interest was playback that allowed the Descriptor to drag the Drawer around a shape. Observations were made about when this feature was felt to be necessary to describe a shape and how it was used. Verbal communication between the users was also closely observed.

5.5 Observations

Through observation it became clear that the playback feature was useful in some instances but not in all. When the shapes to be drawn could be easily named, verbal

description was used in the vast majority of situations. This was the case when describing the shape, size and position of the object within the drawing.

For drawing A, the playback functionality was not used by any of the participants. Each shape was easily nameable, and easily positioned. The Describer simply stated, *'draw a circle'* or *'draw a triangle with the point at the top'*. Verbal guidance was used to position the objects within the scene. This took the form in all cases of dividing the scene into top, middle and bottom in the vertical and left, middle and right in the horizontal. The Drawers seemed comfortable positioning themselves within the environment using these terms. The virtual canvas boundary forces were used to build a frame of reference and the PHANTOM interactions allowed absolute position to be estimated from the proximity to the walls.

When shapes were not easily verbally describable like in Drawings B and C, the haptic playback became an important tool. Figure 18 and Figure 19 show the final drawings for all 3 groups of participants for drawings B and C respectively. In almost all instances all the shapes are included in the final drawings with appropriate relative positions. For all parts of these drawings, the playback mechanism was used to describe the shapes except for the oval in the top left of drawing C. In two instances, the Describer related this oval to the oval felt in drawing B and did not have to use the playback feature.

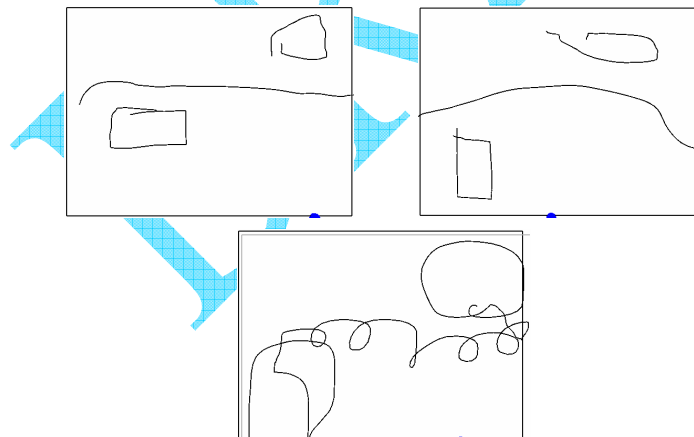


Figure 18: The results from the 3 groups when attempting to draw drawing B.

Trajectory playback was utilised in two different circumstances: sometimes instigated by the Describer and sometimes by the Drawer. When instigated by the Describer, it was due sometimes to being unable to describe the shape. In the other instances (and more commonly) it was due to a description not conveying the appropriate information. The

Describer describes the shape and the Drawer interprets it wrongly. The Describer then used the haptic playback as a different means of conveying information about the shape. One example for the right hand shape of drawing B where a Describer said *“It’s a bit like headphones on its side. I’ll draw it”*. After getting pulled around the shape and with no further verbal description, the Drawer replied *“I can feel that. It was like headphones on its side. Like a semi circle going to the right with two little circles on the end of it”*. Here, the Describer notices the difficulty in verbally describing the shape and draws it. The Drawer uses the playback to form a clearer mental picture and relates it to the original verbal description given.

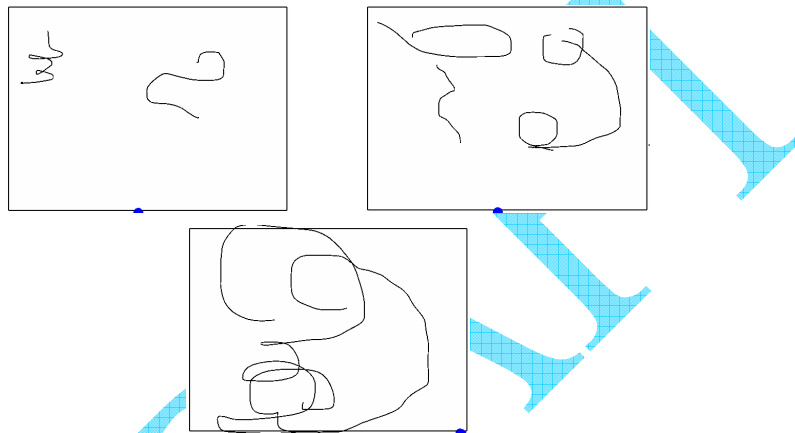


Figure 19: The results from the 3 groups when attempting to draw drawing C.

When instigated by the Drawer, it was used when he/she found the verbal description difficult to understand. In some instances, the Drawer would make an attempt to draw the image and not get close to the image. The Drawer would then request that the Describer draw the image. This was used to great success in some instances where the Drawer was then able to describe the shape to the Describer through what was felt. For the left hand shape of drawing C, one Describer described it as *“This one is a wee bit strange. There’s a circle on the end and it’s got a squiggle coming from it”*. After some unsuccessful attempts to describe the shape further, the Drawer asked how it felt. The Describer then dragged the Drawer through the shape. The Drawer then described the shape as *“Like a balloon on a bit of string”*.

These observations suggest that the collaborative feature of the drawing interface could be used to describe a diagram to a visually impaired user. Both the Describers and the Drawers found the playback a useful technique; the Describer when they found it

difficult to describe a shape in words (or the Drawer misunderstood the verbal description) and the Drawer when they did not understand the initial verbal description.

In all instances, verbal communication was used initially to describe the shape with the drawing functionality supporting the verbal communication. A common technique was to compare the shape to a similar object like a wave, a pair of headphones, a balloon on a string or a door. This provided a starting point for the discussion about the shape with the drawing tool being used to describe difficult sections of the image or clear up misunderstandings.

One novel use of the playback system was to direct the Drawer to the appropriate area of the canvas to start drawing. Instead of drawing a shape, the Descriptor tapped the draw button which moved the Drawer towards the Descriptor's position. This was used in particular where fine grain relative positioning of the objects was important.

The playback system was used in the task, although this study suggests that a number of improvements could be made. Here, the Descriptor would draw the shape and when completed, the Drawer would be dragged around it. This made multimodal communication difficult. The Descriptor could not talk the Drawer through the shape as they were drawing it since the Drawer would not feel the shape until the Descriptor had finished. By incorporating a real-time playback feature, the Descriptor could discuss the shape as well as any local context information while drawing, with the Drawer feeling the Descriptor's movements as they happened. It would also provide a more immediate method of interacting between users. The real-time interaction could be seen more as interacting with a 'shared pen' than a record and playback system, with the student moving when and where the teacher moves. Furthermore, in this study the trajectory was played back at the same speed, independent of the speed of movement of the Drawer. This led to problems in difficult areas of a shape where the Drawers might want to feel slower movements to gain a better understanding of the detail of the shape. However, it is yet to be seen if slowing down the speed of playback improves performance. Spatial integration of the trajectory could prove more problematic as the trajectory takes longer to play back and may place a greater burden on the user's memory.

All Drawers suggested that the audio positional cues were largely unused. Although this feature was popular in a drawing environment with one user and was shown to provide benefit, all three Drawers mentioned that they were concentrating on the haptics and were not using the audio. One possibility is that verbal communication interferes

with the audio positioning cues. The Drawer is listening to or talking to the Describer and may pay less attention to the audio cues as a result.

6 CONCLUSIONS & FUTURE WORK

This paper has described a series of studies examining non-visual trajectory playback as a method of transferring shape and trajectory information to a blind or visually impaired computer user. The first study compared the performance of blind and visually impaired users with a sighted group of users and demonstrated that shape recognition through haptic trajectory playback was significantly more difficult for blind and visually impaired participants. The second study looked to provide further support in the task through combined haptic and auditory playback, and was able to show significant improvements in performance for blind and visually impaired users over haptic only trajectory playback. Finally, the techniques were incorporated into a more realistic application. The third study (although preliminary) demonstrated how trajectory playback could be useful in a teacher-student situation. The teacher was able to use the playback feature of the interface to describe non-textual information to a blind user. It was shown to be particularly useful when verbal description was difficult.

In particular, environments that include these techniques should consider:

- Trajectory playback with a haptic device can be used to allow users to perceive shape.
- Audio cues can further be integrated to complement the shape information felt through the haptic device and to segment the playback from other haptic cues.
- Trajectory playback with a haptic device can be used as an additional communication channel to complement a verbal description when describing a trajectory or shape information for complex shapes.
- A collaborative environment containing trajectory playback should allow real time interaction between the participants such that users can discuss their movements as they move the other user through the environments.
- The rate of playback should be adjustable to allow the user to slow down for detailed areas of the trajectory if required.

We envisage that these techniques may be used in interfaces such as: a gesture-based system (e.g. to control an interface), a training system (e.g. to teach handwriting or geometry) or in a collaborative setting (to aid awareness and provide context information). The studies described suggests more work is required if a robust gesture control system is to be implemented. The long term training aspects of the system still need to be examined. Future work will examine the success of these systems in different contexts such as teaching blind children and adults to sign their names. Different methods of audio trajectory playback will be examined as the sonification chosen was far

more successful in one dimension than another. Further to this, constraint-based haptic trajectory playback will be examined as an alternative to the current system. Instead of dragging the user through the shape, the users will be constrained to the path by forces and provide the movement along the path themselves.

This paper has examined non-visual trajectory playback for displaying shapes to blind and visually impaired users. It was able to demonstrate how combining haptic and audio information could significantly improve performance of our user group in the task of recreating the shapes. Further, it has demonstrated how these techniques might prove useful in a real world application such as a teaching environment. Future areas of research will look to both improve the current playback system and integrate and evaluate the technique in more diverse environments. The experiments described in this paper suggest that there is great potential for multimodal trajectory playback as a useful tool for blind and visually impaired computer users in both single user and collaborative computer system.

ACKNOWLEDGMENTS

This work was funded by EU/IST eInclusion project MICOLE (IST-2003-511592 STP). The authors would also like to acknowledge John Williamson and Roderick Murray-Smith from the Department of Computing Science at the University of Glasgow, and Parisa Eslambolchilar from the Department of Computing Science at Swansea University for their technical input.

REFERENCES

- ALTY, J.L. and RIGAS, D. 1998. Communicating Graphical Information to Blind Users Using Music: The Role of Context. *In Proceedings of CHI*, Los Angeles, Ca, US, ACM Press, 574-581.
- AMIRABDOLLAHIAN, F., LOUREIRO, R. and HARWIN, W. 2002. Minimum Jerk Trajectory Control for Rehabilitation and Haptic Applications. *In Proceedings of the IEEE International Conference on Robotics & Automation*, Washington DC, 3380-3385.
- ARCHAMBAULT, D., FITZPATRICK, D., GUPTA, G., KARSHMER, A.I., MIESENBERGER, K. and PONTELLI, E. 2004. Towards a Universal Maths Conversion Library. *In Proceedings of 9th International Conference on Computers Helping People*, Paris, France, Springer Berlin / Heidelberg, 664-669.
- ASTROM, K.J. and HAGGLUND, T. 1995. *PID Controllers*. International Society for Measurement and Control, Research Triangle Park, N.C.
- BISHOP, C. 1995. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.
- BLAUERT, J. 1997. *Spatial Hearing*. MIT Press, Cambridge, MA, USA.
- BLISS, J.C., KATCHER, M.H., ROGERS, H.C. and SHEPARD, R.P. 1970. Optical-to-Tactile image conversion for the blind. *IEEE Transactions on Man-Machine Systems MMS-11*. 78-83.
- BROWN, L. and BREWSTER, S.A. 2003. Drawing by Ear: Interpreting Sonified Line Graphs. *In Proceedings of ICAD*, Boston, MA, 152-156.
- CHALLIS, B.P. and EDWARDS, A.D.N. 2001. Design Principle for Tactile Interaction. *Haptic Human-Computer Interaction. Springer LNCS*, 2058. 17-24.
- CROSSAN, A. and BREWSTER, S. 2006. Two-Handed Navigation in a Haptic Virtual Environment. *In Extended Proceedings of CHI*, Montreal, Canada, 676-681.
- CROSSAN, A., WILLIAMSON, J. and BREWSTER, S. 2006. A General Purpose Control-Based Playback for Force Feedback Systems. *In Proceedings of Eurohaptics* Paris, France.

- DANG, T., ANNASWAMY, T.M. and SRINIVASAN, M.A. 2001. Development and Evaluation of an Epidural Injection Simulator with Force Feedback for Medical Training. *Medicine Meets Virtual Reality*, USA, IOS Press, 97-102.
- EDWARDS, A.D.N., MCCARTNEY, H. and FOGAROLO, F. 2006. Lambda: a multimodal approach to making mathematics accessible to blind students. *Assets '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, Portland, Oregon, USA, ACM Press, 48-54.
- ERIKSSON, Y. 1999. How to make tactile pictures understandable to the blind reader. *In Proceedings of the 65th IFLA Council and General Conference*, Bangkok, Thailand.
- FEYGIN, D., KEEHNER, M. and TENDICK, F. 2002. Haptic Guidance: Experimental Evaluation of a Haptic Training Method for a Perceptual Motor Skill. *Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Florida, IEEE Computer Society, 40-47.
- GENTRY, S., WALL, S., OAKLEY, I. and MURRAY-SMITH, R. 2003. Got Rhythm? Haptic-only Lead and Follow Dancing. *In Proceedings of Eurohaptics*, Dublin, Ireland, 330-342.
- GRAHAM, J.A. and ARGYLE, M. 1975. A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology*, 10 (1), 57-67.
- HENMI, K. and YOSHIKAWA, T. 1998. Virtual lesson and its application to virtual calligraphy system. *In Proceedings of IEEE International Conference on Robotics and Automation*, 1275-1280.
- HORSTMANN, M., HAGEN, C., KING, A., DIJKSTRA, S., CROMBIE, D., EVANS, G., IOANNIDIS, G.T., BLENKHORN, P., HERZOG, O. and SCHLIEDER, C. 2004. TeDUB : Automatic interpretation and presentation of technical diagrams for blind people. *In Proceedings of Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment - CVHI'2004, EURO-ASSIST-VHI-2 : Accessibility, Mobility and Social Integration*, Granada, Spain.
- KAMEL, H.M., ROTH, P. and SINHA, R.R. 2001. Graphics and User's Exploration via Simple Sonics (GUESS): Providing Interrelational Representation of Objects in a Non-Visual Environment. *In Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 261-266.
- KLOK, M., UZAN, G., CHENE, D. and ZIJP, S. 2005. Blind Gestural Interaction: an Experimental Approach. *The 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*, Berder Island, France, May 18-20.
- KURZE, M. 1996. TDraw: A Computer-based Tactile Drawing Tool for Blind People. *In Proceedings of International ACM Conference on Assistive Technologies*, Vancouver, Canada, April 11-12, 131-138.
- LADNER, R.E., IVORY, M.Y., RAO, R., BURGSTÄHLER, S., COMDEN, D., HAHN, S., RENZELMANN, M., KRISNANDI, S., RAMASAMY, M., SLABOSKY, B., MARTIN, A., LACENSKI, A., OLSEN, S. and GROCE, D. 2005. Automating Tactile Graphics Translation. *ACM SIGACCESS Conference on Assistive Technologies*, Baltimore, Maryland, USA, ACM Press, 150-157.
- LANDUA, S. and WELLS, L. 2003. Merging Tactile Sensory Input and Audio Data by Means of the Talking Tactile Tablet. *Eurohaptics*, Dublin, Ireland, 414-418.
- MCGOOKIN, D.K. and BREWSTER, S.A. 2006. MultiVis: Improving Access to Visualisations for Visually Impaired People. *In Extended Proceedings of CHI*, Montreal, Canada, ACM Press.
- NOBLE, N. and MARTIN, B. 2006. Shape discovering using tactile guidance. *In Proceedings of EuroHaptics*, Paris, France, 561-564.
- OAKLEY, I. Haptic Augmentation of the Cursor: Transforming Virtual Actions into Physical Actions *Computing Science*, University of Glasgow, Glasgow, 2003, 283.
- TEO, C., BURDET, E. and LIM, H. 2002. A Robotic Teacher of Chinese Handwriting. *In Proceedings of the Symposium for Haptic Interfaces for Virtual Environment and Teleoperator Systems*, March, 335-341.
- WALL, S. and BREWSTER, S. 2006. Feeling what you hear: tactile feedback for navigation of audio graphs. *In Proceedings of CHI*, Montreal, Canada, ACM Press, 1123-1132.
- WANG, D., ZHANG, Y. and YAO, C. 2006. Stroke-based modeling and haptic skill display for Chinese calligraphy simulation system. *Virtual Reality*, 9, 118-132.
- YOKOKOHI, Y., HOLLIS, R.L., KANADE, T., HENMI, K. and YOSHIKAWA, T. 1996. Toward Machine Mediated Training of Motor Skills -Skill Transfer from Human to Human via Virtual Environment. *In Proceedings of RO-MAN*, Japan, 32-37.
- YU, W. and BREWSTER, S.A. 2003. Evaluation of Multimodal Graphs for Blind People. *Journal of Universal Access in the Information Society*, 2 (2), 105-124.
- ZHAO, H., PLAISANT, C. and SHNEIDERMAN, B. 2005. "I hear the pattern": interactive sonification of geographical data patterns. *In Extended Abstracts of CHI*, Portland, Or, US, ACM Press, 1905 - 1908.