Evaluation of multimodal graphs for blind people

Wai Yu[1], Stephen Brewster[2]

[1] Virtual Engineering Centre, Queen's University Belfast, Cloreen Park, Malone Road, BT9 5HN Belfast, Northern Ireland;
http://www.vec.qub.ac.uk; E-mail: w.yu@qub.ac.uk
[2] Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow, U.Room S131,
17 Lilybank Gardens, G12 8RZ Glasgow, UK; http://www.multivis.org; E-mail: stephen@dcs.gla.ac.uk

Abstract. This paper introduces the development of a multimodal data visualisation system and its evaluations. This system is designed to improve blind and visually impaired people's access to graphs and tables. Force feedback, synthesized speech and non-speech audio are utilised to present graphical data to blind people. Through the combination of haptic and audio representations, users can explore virtual graphs rendered by a computer. Various types of graphs and tables have been implemented, and a three-stage evaluation has been conducted. The experimental results have proven the usability of the system and the benefits of the multimodal approach. The paper presents the details of the development and experimental findings, as well as the changes of role of haptics in the evaluation.

Keywords: Haptics – Multimodal interaction – Assistive technology – Human computer interaction

1 Introduction

To compensate for the loss of sight, blind people use auditory and haptic modalities to perceive information from the world. Assistive devices have been developed to help blind people by using these two sensory modalities. Synthesized speech and Braille are two common examples of conveying information through the auditory and tactile senses, respectively. Our research is about utilising these two modalities to present various types of graphs to blind people.

Graphs such as line graphs, bar charts and pie charts are commonly used to present data in an easy-to-interpret way. Trends and distributions of data can be illustrated more effectively on a graph than in a table filled with raw data. Therefore, graphs are frequently used in economics, mathematics and other scientific subjects. Unfortunately, this kind of data visualization technique is not so useful to blind people. Being unable to access graphical information easily is a major obstacle to blind people in pursuing scientific studies and careers [1–3].

Traditionally, to make graphs accessible to blind people visual-to-tactile conversion is required on special paper. The content of a graph is raised to a different height than the background, so that it can be discriminated by the cutaneous sense on a person's fingers. Tactile diagrams exist in many forms and they mainly differ in the construction techniques and the materials used [4]. The most common type of tactile diagrams is a raised graph on swell paper. They are relatively easy to make and cost less when compared to other forms. Tactile diagrams provide blind people with the possibility to access graphical information. However, they are not very effective due to the characteristics of haptics and the limitations of the representation medium. Haptics is a much slower communication channel than vision in terms of the amount of information that can be perceived and processed in a given time. Perceiving information through the haptic channel is thus less efficient. It is not easy to make changes to tactile diagrams. Reheating the swell paper could overcook the already-raised parts, reducing the quality. Tactile diagrams are not very durable. After frequent use, the raised objects can be depressed or worn out. Moreover, there is a hygienic issue in that, after being used by many users, tactile diagrams often become very dirty, and they cannot be cleaned simply by using a wet cloth. In addition, it is very difficult for blind people to create tactile diagrams themselves as the production of tactile diagrams is quite tedious. Generally, training for blind people is required for successful use.

Attempts have been made to improve the representation of existing tactile diagrams by adding audio feedback. As a result, audio tablets have been developed to

use in conjunction with tactile diagrams [5–7]. A tactile diagram is placed on top of a touch-sensitive tablet that stores the content information about the diagram. Audio information can then be provided when users touch the object of interest on the diagram. Therefore, more information can be conveyed to users through the additional channel. Blind people can gain a better idea about the displayed graph than before. In spite of this, audio tablets still rely on the haptic representation of tactile diagram and thus suffer from the same limitations as tactile diagrams. In addition, the device has to be programmed before use, as the information about the graph has to be stored first. Any changes to the graph will require modifications to the program. Therefore, these devices are still not the best solution to solve blind people's accessibility problems.

In order to tackle these problems, we adopt a multimodal approach based on novel force-feedback devices, 3D spatialised sound and synthesized speech. Force-feedback devices, such as the SensAble PHANToM, Pantograph, and Logitech WingMan Force Feedback Mouse, have been proven to be useful in exploring 3D objects, graphic user interfaces and scientific simulations [8–13]. In our work, force-feedback devices are used to provide blind people with the sense of touch on the virtual graphs rendered by the computer. The use of these devices allows some shortcomings of tactile diagrams to be overcome. Combining haptic and audio modalities facilitates the presentation of information according to the strengths of each modality. Using computer-generated graphs has several advantages in comparison with traditional tactile diagrams, such as easier graph construction (they being dynamic and flexible to changes) and the use of durable and easy-to-store data in a digital form.

Currently, we have developed a multimodal data visualisation system that conveys graphical information to blind people through virtual touch, speech and non-speech sound. Various types of graphs have been implemented on this system and a series of experiments has been conducted to evaluate the system's usability. The evaluation can be divided into three stages, which are designed to investigate different aspects of the system. In the first stage, the effectiveness and usability of the haptic interface built on the novel force-feedback device was investigated. In the second stage, differences between two force-feedback devices and between multiple modalities have been identified. In the final stage, the usability of the multimodal system in comparison with the traditional tactile diagrams was investigated.

In this paper, the development and evaluation of the multimodal system are presented. The effectiveness of our approach and its advantages over the conventional method are demonstrated in the three stages of evaluation. Moreover, the paper provides a detailed discussion on the changes of haptics role in a user's exploration of the virtual graphs.

2 Multimodal system overview

The multimodal system consists of two main components: an IBM-compatible PC and a force-feedback device. The force-feedback device can either be a SensAble PHANToM (Fig. 1a) or a Logitech WingMan Force Feedback mouse (Fig. 1b). Each force-feedback device has its own advantages and limitations. The PHANToM provides high-fidelity force feedback, while the cost of the WingMan mouse is within the reach of most people. Therefore, the multimodal system has been developed based on these two devices with the aim of conducting haptic research as well as building user-affordable systems. The haptic modelling method for these two devices is different, due to the difference in their physical configuration and software API support. They both act as position input and force output devices. Users control the position of the pointer on the computer screen and feel the force generated by the electrical motors in the device. The amount of force depends on the object being contacted by the pointer. By varying the amount of force and the frequency of vibration, users can feel different physical properties of the objects simulated by the computer. Shapes, softness, and elasticity are a few properties that can be simulated by the devices.

Besides the haptic representation, information is conveyed to users through their sense of hearing. Speech and non-speech sounds are utilised to present detailed and abstract information. The audio representation is designed to give users assistance in understanding the graph contents. Detailed information is given by synthesized speech, which is implemented using the Microsoft Speech SDK 5.0, whereas abstract information is constructed by the MIDI notes synthesized on the PC's sound card. Both types of sounds are concerned with the value of data variables. Value of the data can be either spoken out by speech or played as a musical note. The pitch of the note is mapped with the data value: the higher the data value, the higher the pitch [14]. The types of sounds being played are determined by the position of the force-feedback pointer and the state of the switch mounted on the stylus. Users have the choice to decide which type of
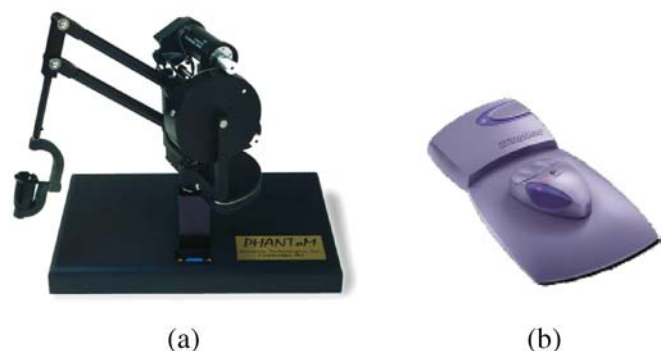

Fig. 1. a PHANToM from SensAble Technologies Inc. **b** Logitech WingMan Force Feedback mouse

information to listen to and thus can explore the graph in an interactive way.

3 First-stage evaluation

The first graph developed on the multimodal system was the line graph, which is usually used to present the trend of a continuous data series. The evaluation at this stage was mainly concerned with the effectiveness and usability of the novel haptic interface in presenting graphical information. Therefore, experiments were conducted on the haptic implementation of line graphs on the PHANToM. Both sighted and blind people took part in this evaluation.

3.1 Haptic line modelling

The virtual line graphs are constructed by using the GHOST SDK from SensAble. A virtual environment is created in the PHANToM workspace. Users feel the line graphs through the stylus of the PHANToM. The PHANToM-controlled pointer moves in a three-dimensional space in which all virtual objects are located on the background. The lines are assembled in polygons forming a V-shaped cross-section. An illustration of a line-graph model is depicted in Fig. 2. The V-shaped channels have inner and outer surfaces, and only the inner surface is defined as touchable by the PHANToM. As a result, the pointer can penetrate the channel from the outside and become retained in the inside. By moving the pointer along the channel, users can trace the path of the line.

The reason for making a V-shaped channel is to retain the pointer securely on the line. In our early studies, we found that users had problems with keeping the pointer on raised objects [15, 16]. A concave shape can effectively solve this problem. Instead of engraving the lines into the surface, we decided to construct them on the surface for several advantages. Firstly, this approach can take full advantage of GHOST SDK's support of polygons which can have a force model defined. Secondly, addition and removal of lines to and from the graph are relatively simple, without the need to recreate the whole haptic scene. Finally, this approach provides flexibility, since simple straight lines or smooth curves can be constructed by putting the desired number of polygons together.

In order to handle multiple lines on a graph, friction keys, which are the frictional property of the polygon surface, are used as a distinguishing feature. Therefore, by judging the degree of friction, users can distinguish different lines. On the haptic line graphs, three levels of friction are implemented. They are labelled as smooth, medium and very rough. The interval between different levels of friction is very wide, so that they can be easily distinguished by users.

3.2 Preliminary study

A pilot study was conducted to validate the design of the haptic interface. Fifteen sighted subjects took part. They were computing science students at the University of Glasgow. Two sets of six graphs were designed for the experiment. Each graph in the first set contained up to three separated lines, while those in the second set contained two intersected lines. Two sample graphs are depicted in Fig. 3. Participants had 4 min to explore each graph. Within this time limit, they needed to perform a set of tasks which included the following:

- Counting the number of lines
- Identifying the line friction
- Counting the number of bends on the line
- Counting the number of intersections, if any

Moreover, they needed to remember the relative position of the lines and their location on the graph, so that they could sketch the graphs after the exploration. This set of tasks was aimed at testing the accuracy of the in-
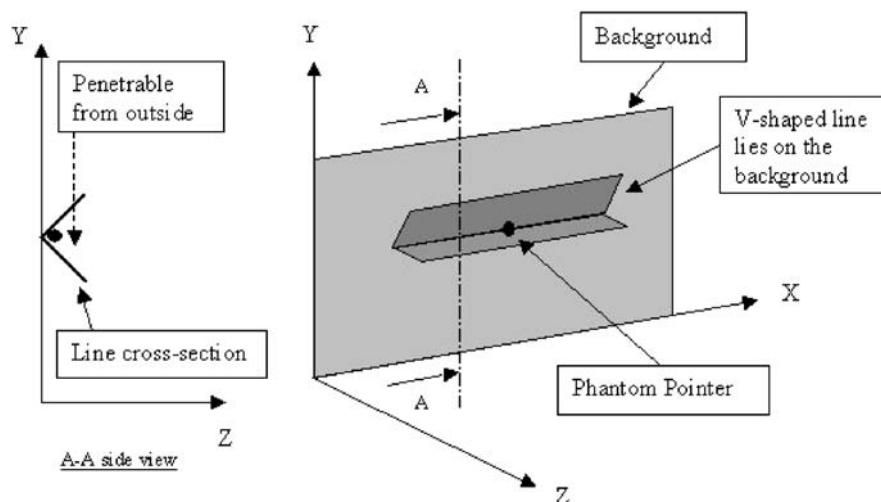
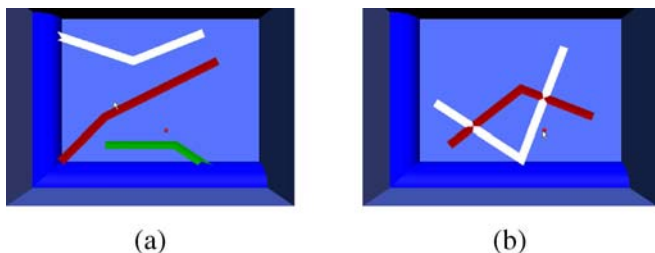

Fig. 2. Haptic line graph model

Fig. 3. Experiment graphs. **a** Non-intersected graph,
b Intersected graph

Table 1. Sighted participants' correct answers (all figures in %)

Questions	Non-Intersected		Intersected	
	Mean	Std Dev	Mean	Std Dev
No. of lines	86.67	15.69	96.67	9.34
Friction keys	94.44	6.03	98.89	4.3
No. of bends	76.11	16.33	88.89	15
No. of intersections	N.A.	N.A.	92.22	13.9

formation conveyed through the interface. Generally, line graphs are used to present data trends and provide difference comparisons. By asking about features and locations of the lines, the usefulness of the interface can be revealed. In addition, several measurements were performed during the experiment, including the following:

- Task-completion time
- A log file containing pointer positions (cursor log)
- A questionnaire regarding the interface design
- NASA task load index test

The NASA task load index (TLX) [17] indicates the workload placed on participants in the experiment, and contains six contributing factors: mental demand, physical demand, temporal demand, effort, performance and frustration. Three practice graphs were given to the participants before the experiment in order to familiarise them with the haptic interface and with the experiment procedure.

3.2.1 Results

Table 1 lists the summarized results of the experiment. The figures show the percentage of correct answers to the questions asked in the task. The overall correct answers in the first and second set of graphs were 85.56 and 94.07%, respectively. The difference between these two groups was significant ($T_{14} = 3.97$, $p = 0.0014$). A high percentage of correct answers was obtained from the identification of frictional properties Table (Table 1). Overall results obtained for the first and second set of graphs are compared in terms of correct answers, task-completion time and workload index (Fig. 4). From the graph, the mean task-completion time is similar in both sets (150.68 s in the non-intersected condition and 152.42 s in the intersected condition) while the workload is higher in the second set. The overall workload index is 10.6 and 12.21 in the first and second set, respectively, and the difference is significant ($T_{14} = 3.75$, $p = 0.0022$).

3.2.2 Discussion

The findings of the experiment show that the participants' performance is better with the second set of graphs, despite the higher complexity. Intersections appeared to be more difficult to deal with in our previous study [18], but they did not affect participants' performance much in this experiment. The amount of practice
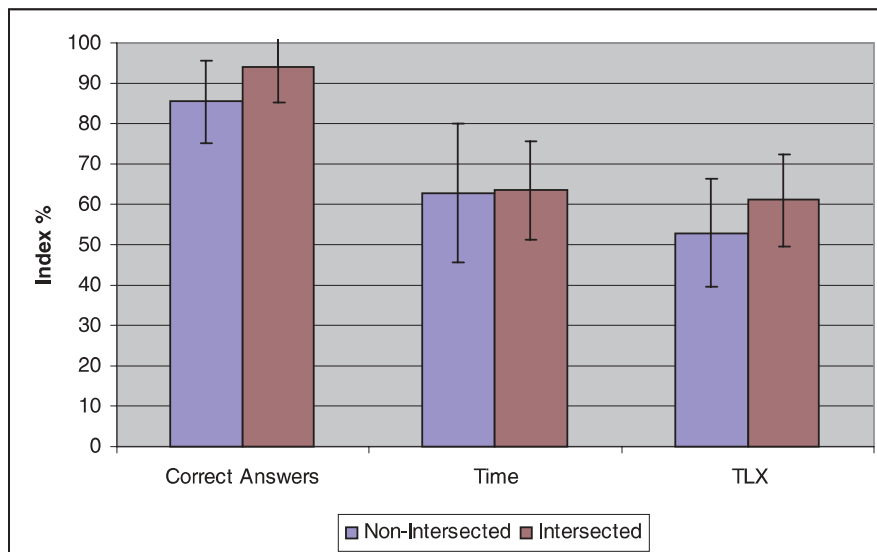

Fig. 4. Overall results of sighted participants (all figures are in percentage). Standard error bars are shown

participants had after completing the first set of graphs could be the main contribution. The significant difference in the workload index reveals that participants had to work harder in order to achieve better results. Among the six categories of workload index, mental demand and effort received highest ratings around 69.5 and 56.5%. This may indicate sighted people are not used to perceiving information by touch alone. In summary, results obtained in the pilot study are promising as participants were able to use the haptic interface to get a high percentage of correct answers.

Besides analysing the experiment data, several problems with the haptic line graph design were identified. Friction keys were regarded as an effective feature in an informal questionnaire. However, friction could hinder participants' movement on a line, especially when the line had sharp bends and strong friction. This combination could mislead the participants into misjudging the sharp bend as the end of the line. This gave participants an incomplete image of the graph. An example of this is shown in Fig. 5.

Figure 5a shows a graph on which two lines are intersecting at two points. The white line has no friction and the darker line has stronger friction. A user's exploration on this graph is captured in Fig. 5b. The user has explored most parts of the graph. However, he could not make a connection between the first segment on the darker line and the second segment on the same line, which he has encountered briefly. Therefore, on his sketch, the second segment of the darker line is missing. This demonstrates the problem with sharp bends on the lines with high friction. Participants often mistook the bend as the end point of the line, because it was difficult for them to find where the next part of the line was.

Perceiving information through touch is much slower than vision. The picture of the line graph is built up by pieces of information picked up through the finger tip over a period of time. Although the accuracy in counting the number of lines is high, 86.67 and 96.67% for the first and second set, respectively, it is very time consuming for the participants to determine whether a line has been explored before. In order to overcome this problem, additional cues are required to assist the user.

There were other problems encountered by participants during the experiment. Figure 6 shows a graph with three separate lines. The friction level increases from the top line to the bottom line. The cursor log only shows the traces of two lines. There are a lot of activities in the top half of the graph, while the bottom right region is almost free of any cursor marks, indicating that the user did not detect the third line at the bottom. This is confirmed by the user's sketch, on which only the top two lines were drawn. Moreover, the proportion of the first line and the shape of the second line are wrong. This shows that the users were attracted to the first two lines and did not pay any attention to the third line. A lack of overview of the graph is the major contributing factor to this problem. In the haptic domain, an overview is very difficult to achieve, due to the narrow bandwidth, especially on the PHANToM, which only provides one single point of contact; therefore, it is almost impossible to have a quick overview of the graph content. Other problems revealed in this example are the inadequacy of user haptic awareness of proportion and gradient change. As a result, our colleagues in the psychology department have conducted research to establish a clearer understanding of human haptic perception in these two areas [19].

Fig. 5. Cursor log analysis, **a** reference graph, **b** cursor log, **c** user's sketch

Fig. 6. Cursor log analysis, **a** reference graph, **b** cursor log, **c** user's sketch
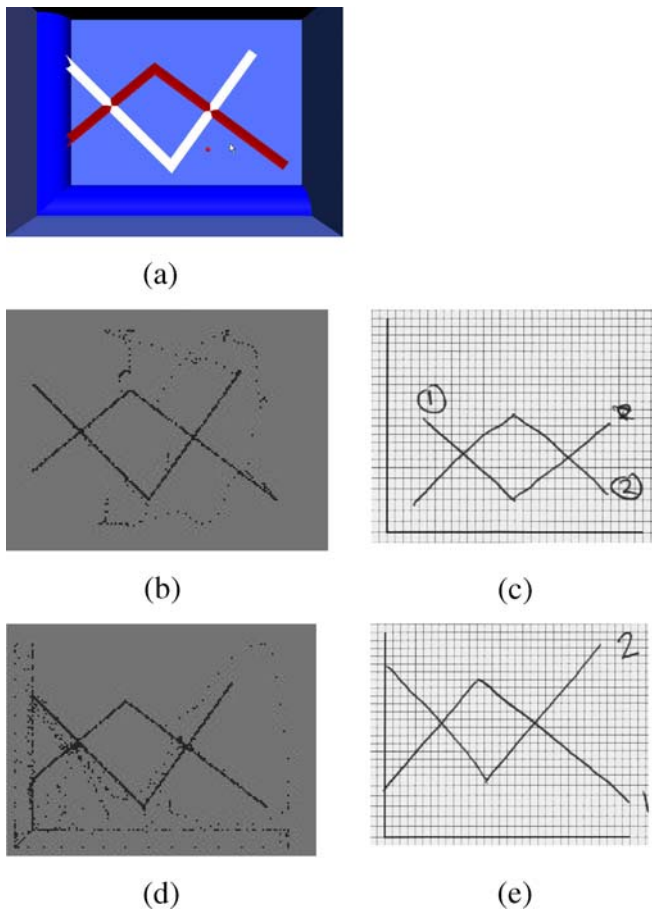
Fig. 7. Cursor log analysis, **a** reference graph, **b** Subject A's cursor log, **c** Subject A's sketch, **d** Subject B's cursor log, **e** Subject B's sketch

Figure 7 shows yet another problem that often occurs in a user's perception of the graph. The graph again shows two lines intersecting at two different places. Both users explored the lines fully and got the shape and number of intersections correctly. Their sketch also indicates that their mental images of the lines match the graph presentation in the reference graph. However, Subject A did not explore the surrounding area of the lines. Therefore, he did not realise that the lines were actually attached to the Y axis. Subject B on the other hand had spent time on the lines and the surroundings. This difference can be seen from the cursor log. As a result, Subject B produced a more accurate sketch than Subject A. On his sketch, the lines are clearly attached to the Y axis. Therefore, individual preference and behaviour have also affected their performance in the experiment.

3.3 Evaluation of improved haptic-line graphs

Based on the results and observations from the pilot study, two additional features were implemented to improve the haptic interface, namely, the instant line identity and the line end-point indication. The lines were numbered and whenever the pointer touched a line, the

PHANToM generated a sequence of clicks (a gentle movement in the Z axis). The number of clicks indicated the line number. The line end-point indication is a continuous vibration that only occurs when the pointer reaches the end of a line. It is designed to reduce the confusion between sharp bends and line end points. An evaluation was conducted to assess the effectiveness of these two techniques. This was done with blind participants to test the interface with a real user group.

The evaluation was conducted with 15 blind people, recruited from the Royal Blind College at Hereford and the Royal National Institute for the Blind at Peterborough. Some modifications were made to the experiment setup. The graphs used in the pilot study were used again, but the four simplest graphs were taken out to shorten the length of the experiment. They were arranged in two groups and each had two intersected and two non-intersected graphs. The difference between the two groups in this case is that one group had the enhanced features while the other did not. The order of performing tasks in the two groups was counterbalanced. Experimental procedure and measurements were similar to those in the pilot study except that participants were not asked to sketch the graph after explorations. Instead, a think-aloud method was used to test the graph image perceived by the participants.

3.3.1 Results

The number of correct answers achieved by the participants is listed in Table 2. The amount of correct answers obtained in groups one and two are very similar (87.45 and 86.2%, respectively). Again, participants did very well at identifying the frictional properties. Figure 8 shows the overall results in score, task-completion time and workload index. The scores obtained in group two are slightly lower than those in group one. It shows that enhanced graphs did not introduce significant advantages in getting accurate answers. Although task-completion time and workload index are slightly lower for group two, no significant difference was found in the t-test. The average task-completion time is 154.88 and 145.95 s in the standard and enhanced conditions, respectively, whereas the average TLX is 10.28 and 10.24 in these two conditions.

Table 2. Blind participants' correct answers (all figures in %)

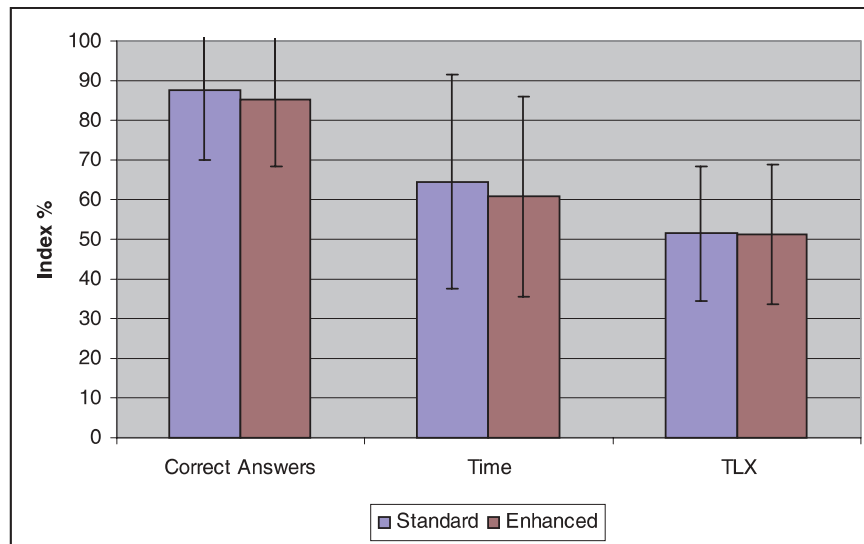| Questions	Standard		Enhanced	
	Mean	Std Dev	Mean	Std Dev
No. of lines	86.67	28.14	88.33	16
Friction keys	95.56	10.11	94.81	8.26
No. of bends	79.26	28.44	73.33	30.23
No. of intersections	88.33	18.58	88.33	16

Fig. 8. Overall results of blind participants in the evaluation (all figures are in percentage). Standard error bars are shown

3.3.2 Discussion

The participants achieved the same level of performance on the enhanced graphs, which means that newly implemented features had little effect in helping people to explore the graphs. However, the usefulness of the enhancement is reflected in other areas, such as variance of correct answers, mental demand of the workload index and users' feedback from questionnaire. A significance test on the variance of correct answers to number of lines produced $F_{14} = 3.09$, $p = 0.021$, which indicates that variations between participants' performance in the enhanced group is significantly less than the standard group. The enhancements helped people who did not perform that well in the standard group to improve. Similarly to the preliminary study, mental demand and effort receive highest overall ratings for subject workload. Mental demand has rated 61.67% for group one and 51% for group two and the difference is statistically significant ($T_{14} = 2.35$, $p = 0.034$). In the effort category, 57.33 and 53.67% were obtained for group one and group two respectively, but the difference is not significant. Therefore, although the effect of the enhancements on participants' performance is not so obvious, they have helped minimise the variance between individuals and reduce mental workload.

Blind participants did not achieve the same level of performance as sighted people. This may be due to the difference between sighted and blind people, as well as the wider range of blind people recruited for the experiment. Sighted participants were all computing-science students who are more familiar with graphs and capable of using the force-feedback device; whereas the majority of blind participants were following a diploma course at the Royal Blind College, and there was a big difference between their age, education background and familiarity with graphs. Although the blind participants' scores are slightly lower, their task-completion times are similar to those obtained by sighted people, and the workload is also lower.

Integrating auditory feedback into our haptic interface is necessary because problems that cannot be solved easily in haptics can be dealt with in another medium. From our experiments, haptic feedback is found to be more useful for guidance and assisting users to navigate the graphs. It is not good at presenting exact data values to the user. Moreover, if too much information is presented via haptics, the narrow bandwidth can be easily overloaded. Therefore, by using another sensory modality, workload can be shared and each sense can be used appropriately.

4 Second-stage evaluation

In the second stage of the evaluation, we investigated the effect of different modalities on user performance and compared the usability of two different force-feedback devices: the SensAble PHANToM and the Logitech Wing-Man Force Feedback Mouse. Each device has its own advantages and limitations, but their usability in real-world applications is seldom investigated and compared. Therefore, we conducted the experiments on these two devices under two experimental conditions: single mode and multimodal. The graphs used in the experiments were bar charts.

Bar charts are one of the most commonly used visualisation techniques and are encountered daily, such as in newspapers, journals and magazines. They usually show discrete and independent variables. Our multimodal interface provided audio and haptic representations. Due to their physical configuration differences, the PHANToM and WingMan FF mouse used different haptic rendering techniques. The audio representation remained the same for both force-feedback devices.
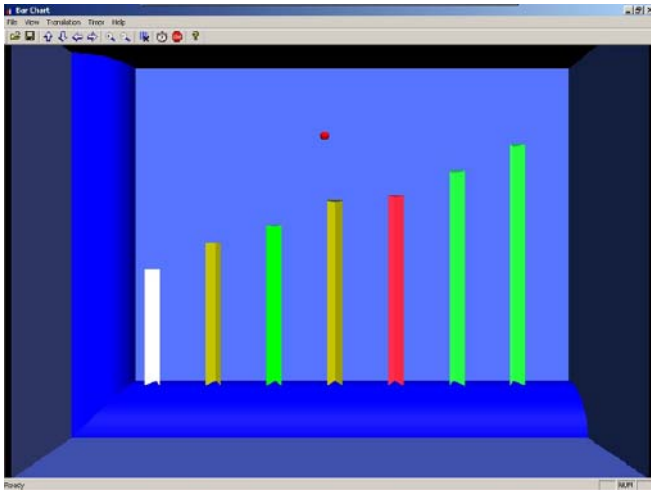
Fig. 9. A sample PHANToM bar chart (greyscale is assigned randomly)

Fig. 10. A sample WingMan bar chart

4.1 Haptic modelling on the PHANToM

The haptic modelling technique used on the PHANToM is based on the polygons supported in the GHOST SDK. A virtual V-shaped groove is constructed to represent a bar. The haptic property of the bar is defined as touchable on the inside but not on the outside. Therefore, the PHANToM pointer can enter the groove from the outside wall but become trapped inside the groove. This technique is the same used for the line graphs described in the previous section, and its effectiveness has been confirmed by the findings of the previous studies. Before rendering the haptic graph, data are scaled to fit into the frame and a gap is created between the bars. A sample bar chart is depicted in Fig. 9.

4.2 Haptic modelling on WingMan FF mouse

The haptic bars on the WingMan FF mouse are modelled by using the enclosure effects supported by the Immersion TouchSense SDK. A bar is simulated by an enclosed rectangular area. Once the mouse cursor enters the bar, it is forced to remain inside. Users will thus have the same type of force feedback on the bars as in the PHANToM case. In order to exit from one bar to move to another, users need to apply a bigger force to overcome the constraint force on the bar edges. A sample graph of the WingMan bar chart is shown in Fig. 10.

4.3 Audio implementation

The audio implementation includes speech and non-speech sound. Speech is generated by the text-to-speech engine provided by Microsoft's Speech SDK 5.0. Its purpose is to provide users with detailed information about the bar value. By pressing the right button of the mouse or the switch on the PHANToM stylus, the program will speak out the value of the bar on which the cursor or pointer is located. Speech information includes the bar number and the bar value.

Non-speech sound is constructed using MIDI notes. A "church organ" timbre was used as a musical instrument due to its continuous nature and wide pitch range. Longer audio feedback on the bars would hopefully give a better indication of the data value to the user. Moreover, a large number of MIDI notes available can improve the audio resolution. This means that more data values can be represented by the number of MIDI notes available. In order to present the data, the bar's height on the graph is mapped to the pitch of the MIDI note. A tall bar produces a high-pitch sound, while a short bar produces a low-pitch sound. The sound effect is triggered by detecting whether or not the cursor or pointer is on a bar. Whenever the pointer enters a bar, the assigned MIDI note will be played continuously unless the cursor or pointer moves away or the speech button is pressed.

4.4 Experiment setup

A series of experiments was set up to evaluate the interface developed on the two different force-feedback devices. Four experimental conditions were designed to investigate the effect of using different modalities in presenting bar charts. These conditions are as follows:

- WingMan audio
- WingMan multimodal
- PHANToM haptic
- PHANToM multimodal

In the WingMan audio condition, no force feedback was provided and the mouse was used as a pointing device. In the PHANToM haptic condition, only force feedback was available, and no audio representation was given. In multimodal conditions, both audio and haptic

representations were available. The first two experimental conditions on WingMan were considered as one group, while the latter two conditions on PHANToM constituted the other group. The reason for using WingMan audio instead of using WingMan haptic is because the results obtained in a pilot study have already shown a significant difference between the haptic only and multimodal conditions. User performance in the WingMan haptic condition was much worse than in the WingMan multimodal condition. Therefore, we decided to investigate the effect of audio feedback on a user's exploration to see whether it is the main contributing factor in a user's performance (detailed information about the pilot study can be found in the section where the results of this second evaluation stage are discussed).

We did not investigate the PHANToM audio condition either, because, without force feedback, both haptic devices become simple position input devices. The only difference is that the mouse works in a horizontal plane whereas the PHANToM works in the chosen vertical plane. The audio implementation on both devices is the same, therefore we only investigated the haptic and multimodal conditions.

Two groups of bar charts were developed based on data obtained from the UK Department of Health's web site [20]. The data describe the statistics of use of the hospitals in England from 1993/94 to 1999/2000. They include the number of beds, ward attendance and outpatient rates. Twenty graphs were produced and equally divided into two groups. There were seven bars on each graph.

The two groups of graphs were assigned to the experimental conditions on each force-feedback device in a random order. The order of conditions taken by each experimental participant was counterbalanced. Therefore, learning effects and any possible unequal difficulties between graphs were minimised. The experiment was conducted on two groups of 16 sighted people. They were recruited from the students at the University of Glasgow. One group of participants did the experiment on the WingMan FF mouse, while the other group did the PHANToM experiment. We did not use blind people in the experiment because of the difficulty of getting enough number of blind people for the experiment locally. Instead, sighted people were blind-folded in the experiment. Although there will be difference between sighted and blind people's results, the findings of the experiment still give some indication of the performance of two different devices. Moreover, the results of the first-stage experiment showed similarity between sighted and blind people's performances. Therefore, we used sighted people for our second-stage experiment.

A set of four questions was designed for each graph. Questions were related to the contents of the graphs and the general purpose of using graphs, such as trend detection and data comparison. The questions are listed below:

Q1. Describe the overall trend of the data.
Q2. Locate the highest bar on the graph.
Q3. Locate the lowest bar on the graph.
Q4. Find two bars which have the closest values.

For the last question, the two bars to be found may be either adjacent or separated by other bars. Answers given by each participant and the time taken to answer all four questions were recorded. At the end of the experiment, the participants filled in a questionnaire regarding the workload of each experimental condition. We used the NASA task load index (TLX) again to determine the workload placed on participants in the experiment. A cursor log was also taken to record the cursor movements during the experiment.

We first conducted the experiment on the WingMan FF mouse and then on the PHANToM. The experimental procedures were identical in these two experiments. Participants were given four practice graphs before the experiment to familiarise themselves with the experimental procedures.

4.5 WingMan FF mouse results

The overall results obtained in the WingMan FF mouse experiment are listed in Table 3, and the detailed analysis of the number of correct answers is given in Fig. 11. Average number of correct answers to each question and the total number are shown. The trend of participants' performance in both audio and multimodal conditions is similar. They managed to obtain accurate answers for the first three questions, but had difficulties in obtaining the right answer for the last question. Finding similar heights between bars seems to be the hardest part in both conditions. The overall number of correct answers in the audio and multimodal condition is 82.81 and 88.59%, respectively. There is a significant difference in the performance between the experimental conditions ($T_{15} = 3.278$, $p = 0.005$).

The significant difference indicates that participants could obtain more correct answers in the multimodal condition than in the audio-only condition. Question 4 has contributed to this difference. Audio seems to be effective at detecting the data trend, and maximum and minimum bars. The major difference between audio and

Table 3. Overall WingMan FF mouse results

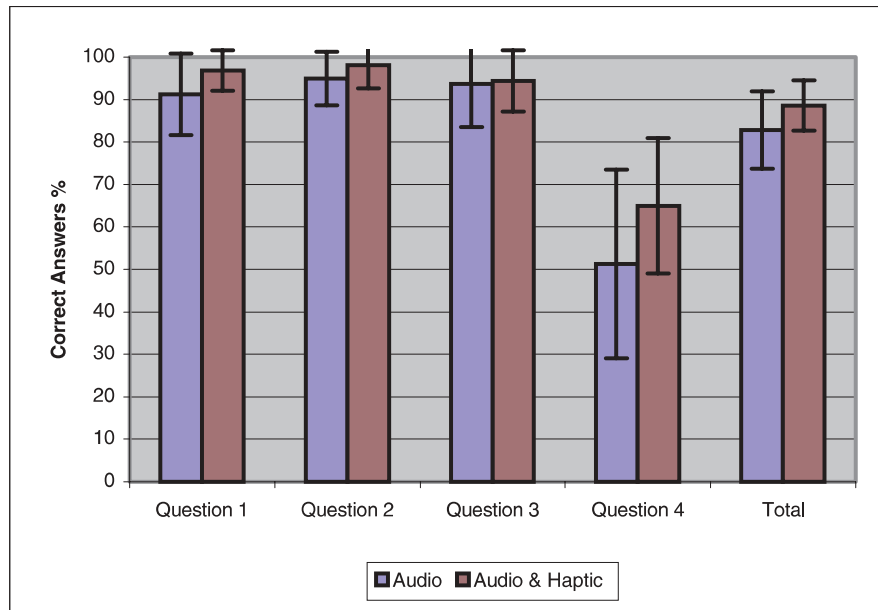|	Audio		Multimodal	
	Mean	Std Dev	Mean	Std Dev
Correct Answers (%)	82.81	9.12	88.59	5.91
Completion Time (seconds)	122	29.09	127	29.47
TLX (%)	50.92	11.28	44.08	11.12

Fig. 11. Correct answers in the WingMan experiment (standard error bars are shown)

multimodal appears in comparing and finding similarities between bars. A multimodal approach is better than audio only, as haptics can be used to compare different bar heights on the graph. A user's spatial perception and proprioception can be used to locate the correct answers. Therefore, using audio alone cannot solve all the problems in graph exploration. Combining haptics and audio has shown its benefits in this experiment.

The average task-completion time for each graph in the audio and multimodal condition is 122 and 127 s, respectively. Statistical tests do not indicate any significant difference. Using a multimodal approach had no major effect on the task-completion time as one more medium

was introduced to the experiment. Participants needed time to comprehend the information presented in both modalities. Some answers that could not be found on the audio interface, could be obtained through haptics. However, haptics is a slow medium in conveying information (due to the current state of technology). Therefore, the time spent on the haptic exploration counterbalanced the time saved from working out ambiguity of the audio feedback.

The data collected from the questionnaire filled in by the participants after the experiments is averaged and plotted in Fig. 12. Mental demand and effort again received highest ratings. The scale of performance is in-
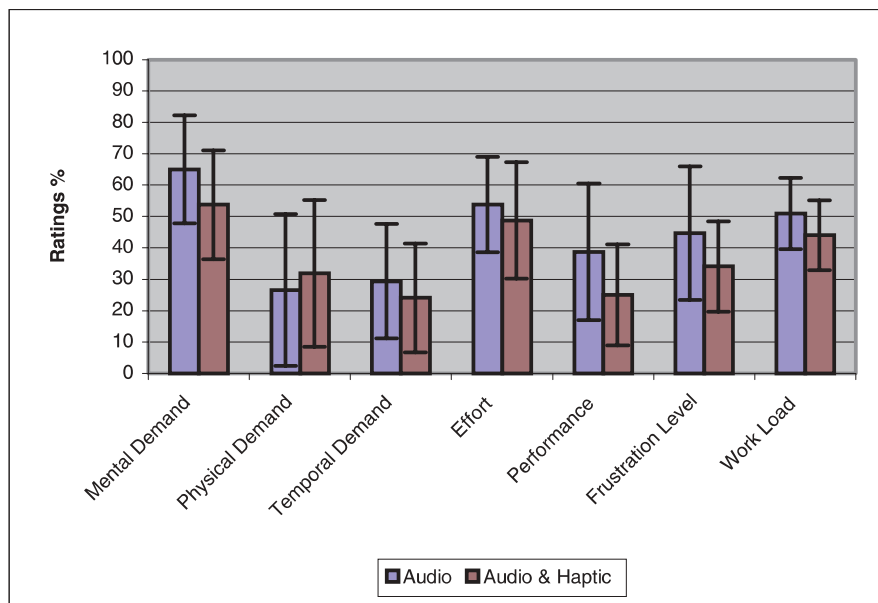
Fig. 12. Task load index in the WingMan experiment

verted so that the higher the bar the lower the performance. In general, participants rated less workload in the multimodal condition, except in the case of physical demand. This can be explained by the haptic interaction in the multimodal condition. Participants needed to apply more forces to compete against the feedback force.

The overall workload index again shows that the multimodal condition rating is significantly lower than the audio condition ($T_{15} = 2.542$, $p = 0.023$). The actual figure for the overall workload index is 50.92% in the audio condition and 44.08% in the multimodal condition. The lower workload index in the multimodal condition is crucial, as it indicates that participants did not need to work so hard when both audio and haptics were present. Participants' feedback confirmed the improvement of the number of correct answers in the multimodal condition.

4.6 PHANToM results

The overall results of the PHANToM experiment are listed in Table 4, and the average number of correct answers is shown in Fig. 13. A similar trend to the WingMan study can be observed. The first three questions again received higher scores, while the last question has a lower

Table 4. Overall PHANToM results

	Haptic		Multimodal	
	Mean	Std Dev	Mean	Std Dev
Correct Answers (%)	85.78	7.11	89.22	5.68
Completion Time (seconds)	139	33.24	115	29.76
TLX (%)	59.13	11.34	43.63	11.26

figure. Overall, the total number of correct answers in each condition is quite close, i.e., 85.78 and 89.22% in the haptic and multimodal condition, respectively. A *t*-test shows $T_{15} = 2.112$ and $p = 0.052$, which just misses the significance level.

From the results, the PHANToM showed its high performance in haptic representations. The force feedback alone is good enough to present information to participants. Moreover, its capability of providing 3 degrees of freedom force feedback is an advantage in user interaction with virtual objects. Users could get correct answers without visual feedback quite easily. Therefore, the number of correct answers in the haptic condition is not significantly different than in the multimodal condition.

The average task-completion time for a graph in the haptic and multimodal condition is 139 and 115 s, respectively. It shows that participants can finish the task much quicker in the multimodal condition ($T_{15} = 3.034$, $p = 0.008$). The task-completion time gives a convincing performance improvement in the multimodal condition. Participants could use audio to speed up the process of locating answers for the questions without affecting the accuracy of their responses.

Participants' ratings on the workload index are plotted in Fig. 14. The overall workload index of the haptic and multimodal conditions is 59.13 and 43.63% respectively. A significant reduction in workload is confirmed ($T_{15} = 7.538$, $p < 0.001$). The mental demand and effort again received higher ratings than the rest. The graph shows consistent reduction over all factors in the multimodal condition. The multimodal approach is again proved to be more effective and requires less effort from participants, leaving more cognitive resources for dealing with the graphs.
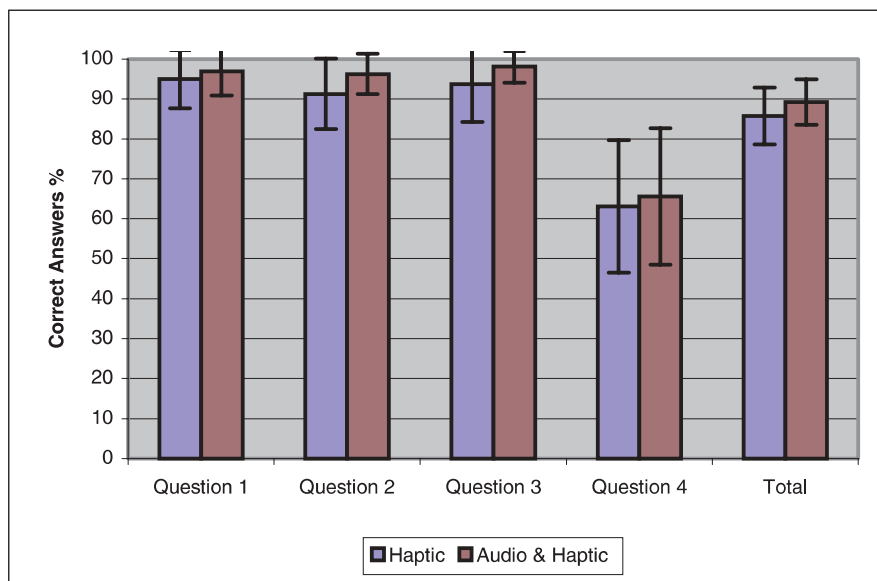

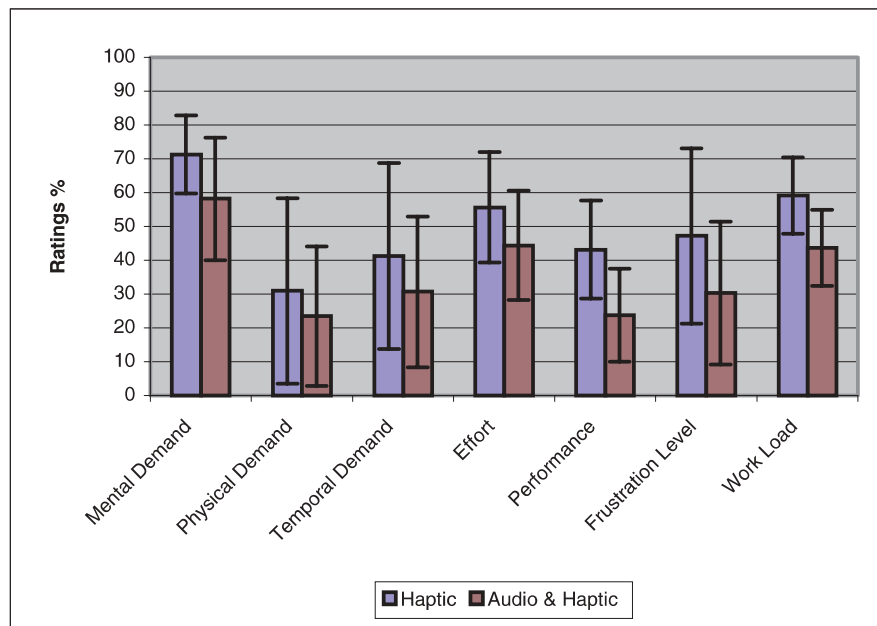**Fig. 13.** Correct answers in the PHANToM experiment

Fig. 14. Task load index in the PHANToM experiment

4.7 Cursor log analysis

Similar to the first-stage evaluation, participants' cursor activity was recorded into a log file. Based on this information, we can observe the users' behaviour and investigate the use of haptics. The main interests in the cursor log analysis include whether haptics has been used, how it was used, how often it was used and how it affects the users' performance. The analysis was only conducted on the log files taken in the multimodal condition of both PHANToM and WingMan, as the other conditions either had no haptic input or haptics was the only input. Haptics' role in the multiple-modalities condition is the main interest of this analysis.

Analyzing the cursor log is not an easy task, due to the difficulties in measuring and classifying the log files. There is a risk of being subjective and speculating on the user's intentions. In order to avoid this problem, we have established two general categories to classify the cursor log files. They are (1) haptics used as navigation tools alone and (2) haptics used to trace the graph. Within the second categories, two sub-classes have been defined: (i) no comparison on bar top and (ii) comparison on bar top (Fig. 15). The cursor logs that show traces of horizontal links between bars tops indicate that a participant used haptics to measure the height of the bars and make comparisons. These classifications are based on distinctive features of the cursor log in order to minimize any subjective view or bias. Moreover, three other measurements were taken based on the above classification:

- Number of people who used haptics as measuring tools
- Number of cursor logs that show the use of haptic input
- Number of people who improved performance
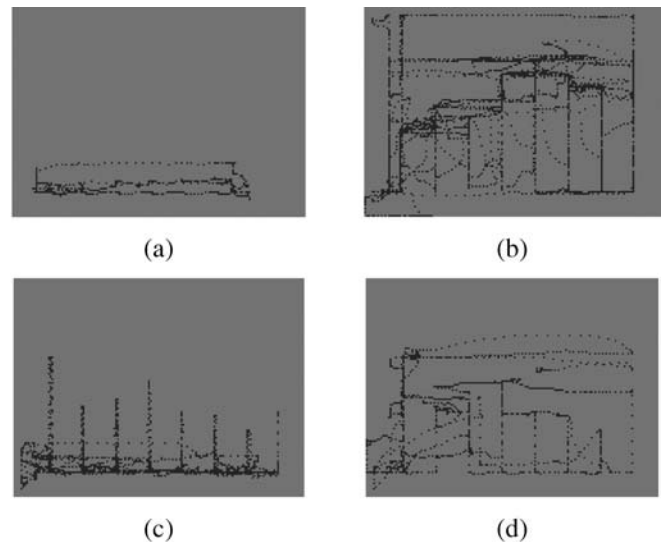
Fig. 15. Samples of cursor logs taken from the PHANToM experiment. **a** no haptic trace of bar chart, **b** haptic trace of bar chart, **c** no comparison on bar top, **d** comparison on bar top

Figure 16 shows the classification results. In the WingMan case, the number of cursor logs which show the use of haptics in tracing the bar charts and making comparisons is 88%. A similar figure of 85.33% is found in the PHANToM case, and there is no significant difference between them. Haptics being used as a measuring tool to compare the bar top in the WingMan and PHANToM case is 87.88 and 72.66% respectively. There are no significant differences between these two conditions. A high percentage of haptic involvement is observed even when other modalities are available. Device difference did not affect people's choice of using haptics, even though the PHAN-
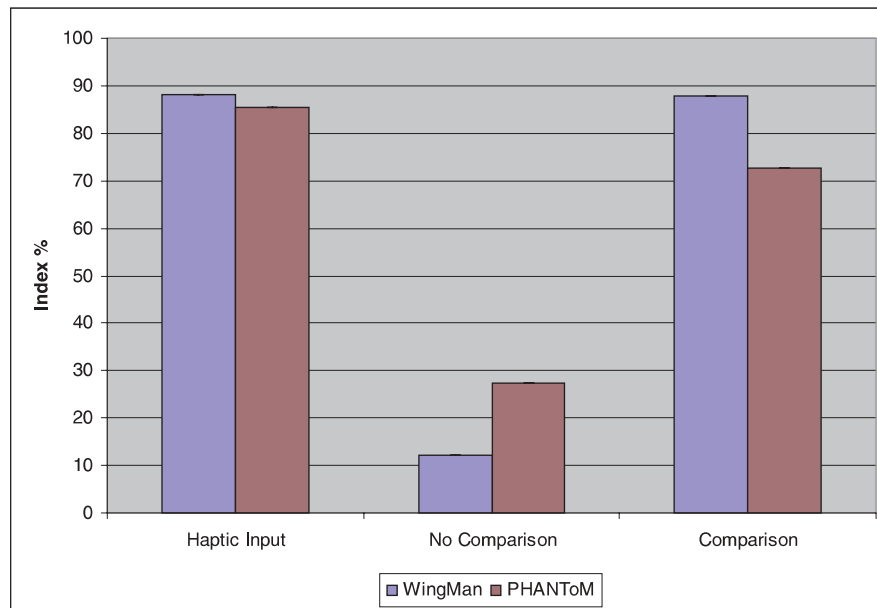
Fig. 16. Summary of cursor log classification

ToM seems to be a better force-feedback device. Differences only occur in the detailed classification of haptic input. Haptics is more pronouncedly used to compare bar tops in the WingMan case. This may mean that users needed to use haptics in an explicit manner so that accurate information could be obtained. In the PHANToM case, the higher percentage is on the no-bar-top comparison. More users are able to detect the bar height by just moving the pointer up and down on the bar, without explicitly comparing the bar tops. This could be a result of better force feedback and a larger workspace provided by the PHANToM.

The results of the three measurements are remarkably similar. In both PHANToM and WingMan cases, 15 out of 16 people used the haptic features on the graphs other than navigation cues. Moreover, 10 out of 15 people had improvements in their performance with respect to the single-mode condition. This shows that haptics had a major involvement in users' exploration and contributed to their performance improvement. Device differences do not affect user's preference in using haptics in the multimodal condition. The amount of haptic input and the role it played are very similar on both devices.

4.8 Discussion

Results obtained in the WingMan FF mouse experiment showed that the mouse is not so effective in conveying information without audio feedback. The amount of force feedback is not enough to inform users about the graph content. This can be seen from a pilot study in which four participants performed the same task in the haptic only and multimodal conditions. The results listed in Fig. 17 show that participants managed to extract less information in the haptic condition when compared with the multimodal condition, especially on Question 4. The substantial improvement in the multimodal condition raised a question about the amount of audio contribution, and specifically whether or not audio played a dominant role in the graph exploration and contributed to all the improvements in the multimodal condition. Therefore, we conducted the experiment on the WingMan in the audio and multimodal conditions, in order to obtain some indications of the audio influences.

Results from the experimental conditions on the WingMan revealed that audio was not the only contributing factor in the participants' performance. Haptics also had a role to play in graph exploration. Significant differences between participants' performance in the audio and multimodal condition, as well as the cursor logs, have proved this point.

The experiment results have revealed that both devices can be used by participants to extract data from bar charts. The multimodal approach is better than either single-modal approach. This can be seen from the objective measurements of correct answers and task-completion time, as well as from the subjective measurement of participants' workload index. The WingMan FF mouse, which is not a very strong force-feedback device, can give a better performance when audio feedback is introduced. This has a similar effect on the PHANToM; participants' performance can be enhanced by adding audio.

The most interesting thing is the similarity between participants' performance on the multimodal condition of the WingMan FF mouse and the PHANToM. Their performance is very close, and the summarised results show very little difference between these two conditions (Fig. 18). The average correct answers of the WingMan FF mouse and PHANToM are 88.59 and 89.22% respectively. The overall workload of these two conditions is
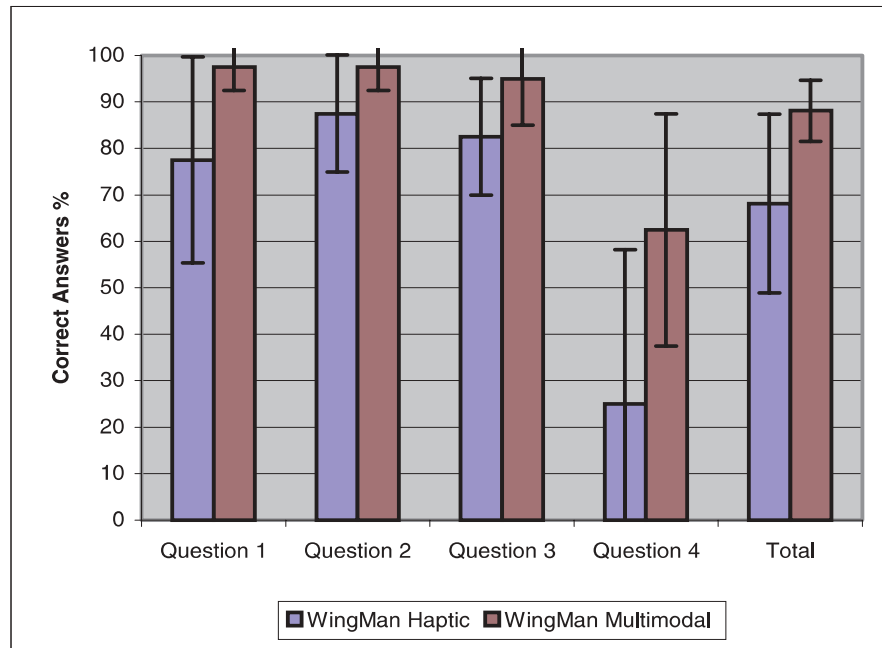
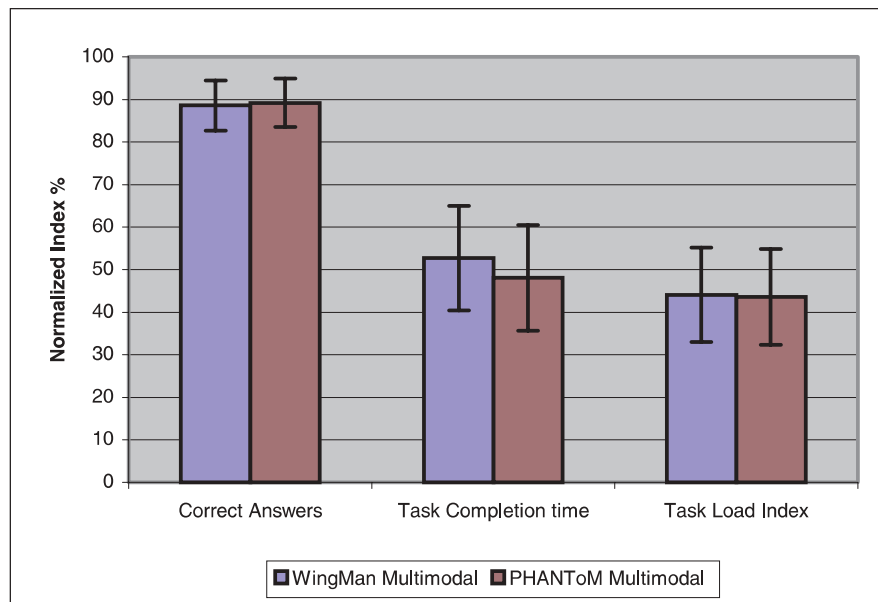**Fig. 17.** Correct answers in WingMan haptic only and multimodal conditions

Fig. 18. Summarised results of WingMan and PHANToM multimodal experiments. (All data are presented in percentage of obtained value against maximum possible value.)

also very similar, 44.08 and 43.63%. The only larger difference is on the task-completion time which is 127 s (52.71%) and 115 s (48.11%) for the WingMan FF mouse and PHANToM, respectively. Despite the huge cost and capability differences between these two devices, using a multimodal approach can actually minimise this difference and provide the same level of achievement in this situation.

The experiment results indicate that a cheaper device like the WingMan FF mouse can provide similar performance to the more expensive PHANToM in this particular application. This is important, as it means that the techniques we have developed can be used by ordinary users with low-cost equipment. As graphs are usually in 2D, a WingMan FF mouse will be capable of this kind of rendering. Moreover, when audio feedback is used in the representation, the haptic role is changed from extracting information to assisting in the detection and location of the interesting data. Therefore, even when the haptic cues from the mouse are not so strong,

users can still make use of the device and understand the graph.

As the amount of data on the bar charts used is not so large, audio can be used to represent the data very quickly. Users can obtain the answers for the trend information, and maximum and minimum values based on the audio feedback. This can be seen from the results in the WingMan audio condition. The place for haptics to be used to extract information is in comparing data to locate similar values. It is not so easy to get the answer based on the musical notes. Even when synthesized speech was available, participants could not use it successfully because of the very large values of the bars. These values could be thousands to millions. Again, the small number of correct answers in the WingMan audio condition illustrated this problem. Haptics become useful in this case; participants can compare the height of each bar by using the haptic cues. Therefore, a significant improvement can be found in the multimodal conditions.

In this set of experiments, although haptics took a major role in extracting information, audio could replace haptics to perceive information about the graphs. This is especially true when haptic feedback does not directly represent the data value to the user. In the bar-chart case, haptics is used to feel the boundary of the bars. The data value is determined by users' proprioception of the amount of movement on the bars. It is an indirect process of perceiving information compared with the instant indication of the MIDI notes. In other situations where the graphs represented are lines, curves or 3D surfaces, haptics represents the data value directly by different heights on the graph or attitudes in a 3D volume. The roles of audio and haptics may change according to the application. However, the cross-modal effect on these types of graphs needs to be investigated further.

5 Third-stage evaluation

In this final stage of evaluation, we investigated the usability of the multimodal system in comparison with the traditional tactile diagrams. The advantages and limitations of these two graph representation media have been discussed in the beginning of this paper. However, little is known on how they would compare in a real-world application, whether (and to what extent) their strengths could overcome their limitations, and how well can users use these two media in performing tasks and perceiving information. In order to answer these questions, we adopted the same experiment design used in the second stage and made some modifications to the multimodal bar-chart representation based on the feedback from the users. The PHANToM version of bar charts was chosen for the experiment due to its better force-feedback capability. The experiment was conducted with the blind people recruited from the Royal Blind College at Hereford.
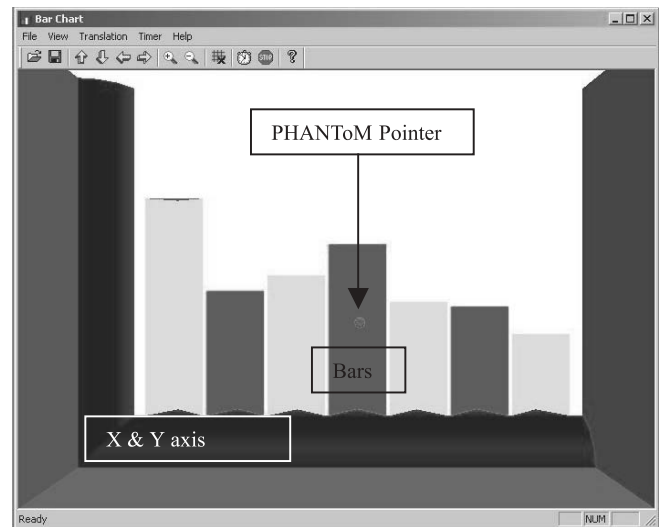

Fig. 19. A snapshot of the haptic bar chart

5.1 Multimodal-system modifications

The main items on the haptic bar chart are still the X and Y axes and bars. The axes are modelled as cylinders which are raised from the background, whereas bars are concave and engraved (Fig. 19). The distinctive feature aims to avoid confusions between these two items so that users would not mistake the Y axis for one of the bars. Bars are no longer individual objects. Rather, they form parts of the background in order to make the transition of the pointer between bars easier. Polygons are used to construct the bars that are placed closely together. This arrangement of bars is done according to the design guidelines for tactile diagrams [4, 21]. There is a small gap between the Y axis and the first bar on the left.

Modifications were also made to the speech and non-speech sounds. First of all, a piano has replaced the organ as the musical instrument to avoid the ambiguity for the high-pitch notes. The sound is triggered by the contact between the PHANToM pointer and a bar. Therefore, by moving the pointer across all the bars on the graph, a series of MIDI notes is played and the ups and downs of the pitch indicate the highs and lows of the bars. A quick overview of the data trend is perceived by users.

Two speech modes were implemented and their occurrence is determined by the number of clicks on the PHANToM's switch. On a single click, a bar's relative height on the graph is read out. The bars' height is the scaled value to fit into the display window. On a double click, the exact value of a bar is given. Thus, users have the flexibility to choose the type of data they are interested in. Single click provides easy-to-memorise information, whereas double click gives precise details.

5.2 Experiment set-up

The experimental conditions were the multimodal system and standard tactile diagrams. The tactile diagrams were
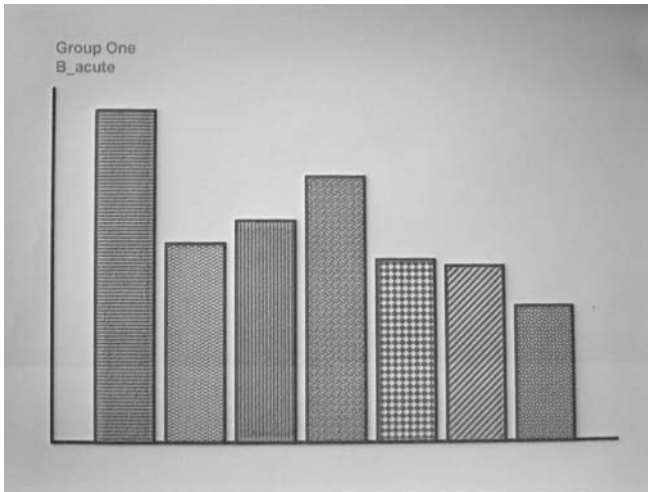
Fig. 20. A sample of tactile diagram used in the experiment

Table 5. Overall experimental results

	Tactile Diagram		Multimodal system	
	Mean	Std Dev	Mean	Std Dev
Correct Answers (%)	87.06	5.32	96.03	2.8
Completion Time (seconds)	59	19.58	92	23.23
TLX (%)	39.86	18.61	58.84	15.96

designed and raised by the National Tactile Diagrams Centre at the University of Hertfordshire [22]. A sample tactile diagram is shown in Fig. 20. There is a small gap of about 3 mm between each bar.

Seventeen participants took part in the experiment. They were students at the Royal Blind College at Hereford, and it was their first time using the multimodal system. Their visual impairment ranged from partially sighted to completely blind. Their ages ranged from 16 to 53. None of them had taken part in any of our previous studies.

The experimental procedure, tasks (four questions regarding the graph contents) and measurements were the same as in the second stage of the evaluation. In the train-

ing section, each feature was explained to the participants in detail. Therefore, they knew how to use all the features to extract information from the graph. In the experiment, it was up to them to decide which method of exploration to use. In the tactile-diagram condition, they were allowed to explore the graphs in the way that they normally use. There were no restrictions on whether to use one hand or both hands to touch the graphs. Complete freedom was given so that the difference between multimodal system and the tactile diagrams can be determined.

5.3 Results

Table 5 shows the overall results of the experiment. The number of correct answers reflects the accuracy of information extracted from the graphs by the participants. In the tactile-diagram condition, the average total number of correct answer was 87.06%, whereas in the multimodal system condition the figure was 96.03%. The difference between the results is significant (*t*-test gives $T_{16} = 5.914$, $p < 0.001$). The numbers of correct answers
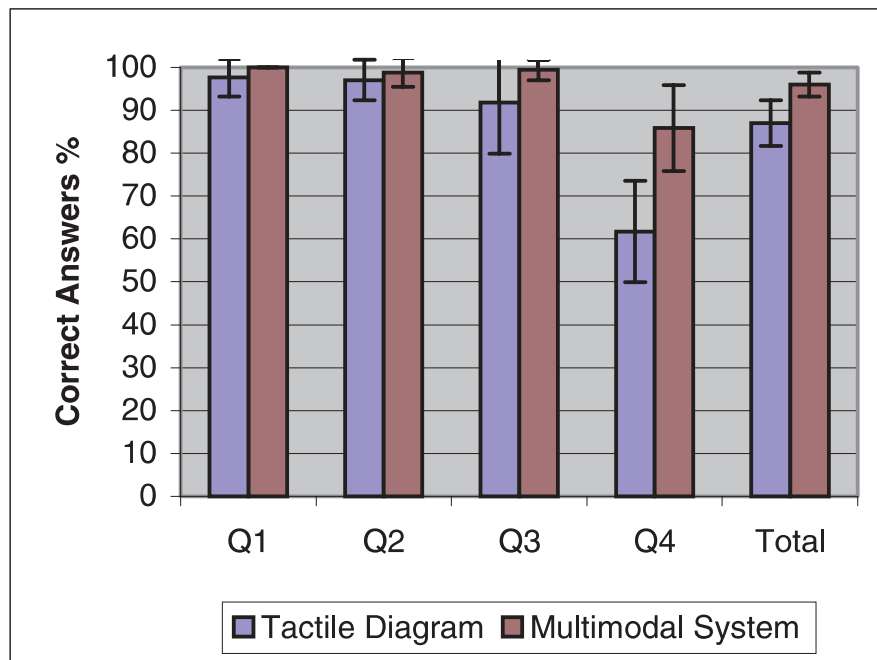

Fig. 21. Distributions of participants' number of correct answers
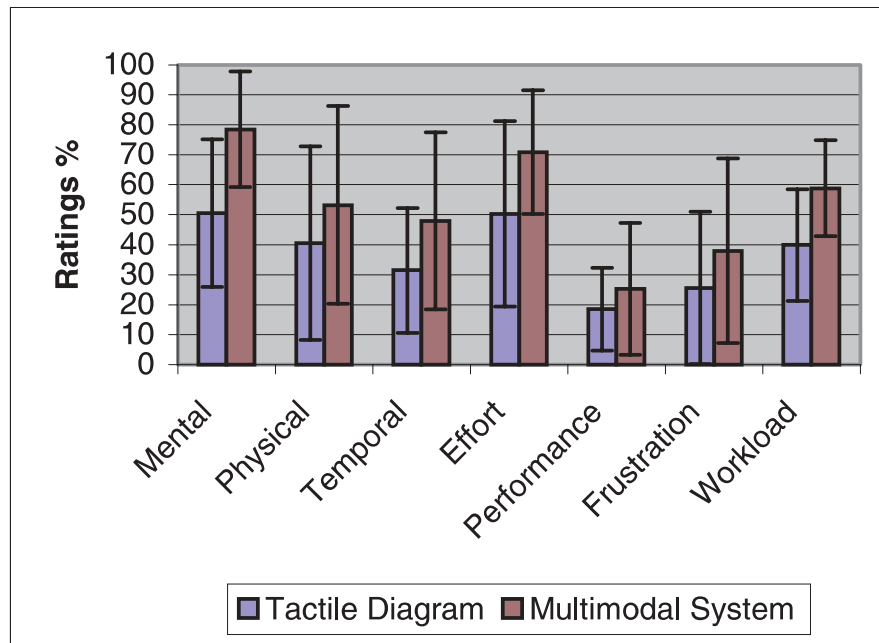
Fig. 22. Distribution of participants' ratings in workload assessment

to the first three questions are very high and similar. As in the experiments described in previous sections, this trend changes with answers to the last question in which data comparison is required. The accuracy drops significantly ($T_{16} = 5.734$, $p < 0.001$) in the tactile-diagram condition (61.76%). The accuracy in the multimodal system condition for the last question is 85.88%. A diagram showing the results of correct answers is given in Fig. 21.

In general, participants used much less time in the tactile-diagram condition than in the multimodal condition. The times taken in these two conditions are 59 and 92 s, respectively. The difference between these two conditions is significant ($T_{16} = 7.398$, $p < 0.001$).

The distributions of participants' ratings on the workload factors and the overall workload index are shown in Fig. 22. Participants gave higher ratings to the factors in the multimodal system condition. As a result the average overall workload index is 58.84%, which is significantly higher than the tactile diagram condition (39.86%). The t-test results shows $T_{16} = 3.742$ and $p = 0.0018$.

The cursor log shows very little use of haptic features during participants' exploration on the computer-generated graphs. Only 2 out of 17 people show signs of using haptics to detect the height of the bars. The majority of participants use haptics only for navigation purposes. MIDI notes and improved synthesized speech output had major contribution to participants' performance.

5.4 Discussion

The main focus of this final experiment was to evaluate whether the multimodal system can actually outperform traditional tactile diagrams in presenting graphical information to blind people. The answer to this question relies on the measurements taken in the experiment, and specifically the number of correct answers, task-completion time, workload index and the observations on participants' way of exploration. The results of the experiment have partially supported the hypothesis, as the number of correct answers produced by the participants is significantly higher in the multimodal system condition. This means that participants could obtain more accurate information through the system. On the other hand, participants used more time and spent greater effort to achieve this level of accuracy. In order to understand the causes and implications of these results, we need to look at each of them separately.

Two major types of tasks were involved in the experiment questions, namely, trend detection and data comparison. The first three questions required trend detection and simple data comparison, whereas the last question involved substantial data comparison. The results showed that participants did well in both conditions. Although the results in the multimodal-system condition are slightly higher, they were not proven to be significant as there is a ceiling effect. When substantial comparison is required, the multimodal system was better, as the accuracy of participants' answers is much higher. Therefore, the final score in the multimodal-system condition is much higher than the tactile-diagram condition. This is mainly due to the effect of speech output. Participants' frequent use of the speech output was observed during the experiment. Speech provides a precise form of information, which can solve the ambiguities in the haptic and non-speech audio representation. Therefore, most

participants used speech either to find their answers or confirm the answers that had been located. However, it takes time for the information to be read out and thus the time needed for answering the questions is longer.

The long task-completion time is also affected by the limitation of the force-feedback device used, which only provides one single point of contact. The information which can be conveyed through this one point of contact is very limited, whereas users could use their whole hands and all of their fingers in the tactile-diagram condition. With a single point of contact device, such as the one used here, a long time is necessary for participants to pick up information about a large area of interest. This inefficiency has thus prolonged the participants' exploration time.

Although non-speech audio is available to minimise the effort that has to be spent on the haptic interface, participants may not be as familiar with this type of audio representation. Even though they understand the relationship between the pitch and bar height, it is still not their usual way of accessing data. Blind people are more used to screen readers that read out in speech the information displayed on the screen. Speech is a more direct method to convey information, as no translation is required, unlike the haptic and non-speech audio cases. However, it takes time for participants to listen and understand the content of the speech. Moreover, they need to manipulate the received information in their short-term memory in order to tell the height difference between bars. This becomes even more demanding in the last question, in which different heights between bars had to be compared in order to find the two closest values. This requires good memory for storing data for later comparison. Therefore, a heavy burden is placed on the participants, and this has been reflected on their ratings on the mental demand and effort, which are much higher than the other contributing factors.

Participants indicated that there was more workload in the multimodal-system condition, but through their comments after the experiment, they generally thought that the interface was easy to use. They suggested that if they had more time to practise, they could perform better. As this was their first time in using the multimodal system, lack of familiarity definitely had a major effect on their workload assessment. Despite the relatively short training time compared with participants' previous experience on tactile diagrams, they managed to obtain substantially more correct answers in the multimodal system. This has partly proved the hypothesis of the evaluation. The multimodal system outperforms the paper-based medium and, provided that users have frequent use of the system, task-completion time and workload index can be expected to improve.

Some observations have been made on the participants' approach to the graphs. Most participants use haptics as a navigation tool to guide them entering and leaving the bars. Although they can use it to determine the height of the bars, they did not use it in this way. Some people were more musically skilled, so they could use the non-speech sounds to obtain most of the answers. They only needed to use synthesized speech to confirm or find answers to the last question. On the other hand, some participants used speech throughout the experiment. This complete reliance may show that they are more familiar with screen readers. On the tactile diagrams, most participants used two index fingers to feel the height difference between bars. These are usually the dominant fingers trained to read Braille. A gap between bars is useful for them to differentiate one bar from another, but the gap should be small enough to be covered by one finger. In the workload assessment, participants felt they had a better performance in the tactile-diagram condition. This again shows that they are more confident in the representation that is more familiar.

6 Summary

In the first stage of evaluation, we tested a haptic interface to present line graphs. The experimental results have proved that the haptic interface developed can effectively convey information to the users in terms of higher accuracy in the number of correct answers. The time taken to comprehend a graph could be quite long due to the limited bandwidth of the haptic channel. Putting more features in the haptic representation does not necessarily improve users' performance, as this channel can be easily overloaded. As a result, using additional modalities seems to be the solution to present more information to users.

As haptics was the only available modality in the interface, it was used fully. This can be seen from the cursor logs, which look similar to the layout of the graphs. Besides using haptics to locate lines, users also used it to detect the data trend and other graph contents. Due to the nature of the line graph, users have to move the cursor back and forth on the lines in order to determine the data trend. Haptics was used in navigation, line detection and trend spotting. Users' sense of proprioception and kinaesthesia have been fully utilised in order to perceive the graphs.

In the bar-chart experiments, haptics played the same role as in the line graphs, where it was the only modality on the interface. Users needed to explore every bar to determine the data trend. As the questions asked in the experiment were more concerned with comprehension of the information and not extracting features in the line graph, haptics was used to detect data trends and make comparisons between bars. Instead of going through each bar separately, users made connections between bars in order to know their height relationship. The cursor logs show a lot of cursor traces joining the tops of bars. This reveals that users are more concerned with the bar height instead than with the bar itself. Based on the users' behaviour on these two types of graphs, different graph representation methods affect the users' way of using haptics.

Using haptics to locate the answers for the questions is time consuming and less accurate, as shown in the results of the second stage of evaluation. Some information, such as exact value of the bars, is better represented in another modality. Therefore, in the multimodal bar-chart case, speech and non-speech audio were introduced to provide abstract and detailed information on the bars. The role of haptics has been affected by the introduction of the other modalities. Users could obtain information from audio alone and use force feedback as a simple tool to interact with the virtual graphs. Haptics was not only used to perceive information, it was also used for navigation purpose. However, based on the cursor logs, a majority of sighted users still relied on haptics to explore bar charts. Fifteen out of 16 participants used haptics to trace the bar heights and make comparisons on bar tops. From these people's cursor logs, over 85% showed evidence of using haptic information. Moreover, 10 out 15 people showed improved performance in the multimodal condition. Therefore, haptics still has a major contribution even when other modalities are available.

The results obtained from two different force-feedback devices are remarkably similar. Also, the usage of haptics is more or less the same. This demonstrated that when multiple modalities were used, limitations of force-feedback devices could be kept to a minimum and did not affect user performance in our application. A small and not significant difference was shown in the users' way of using haptics on these two devices. Fewer users compared the bar tops in the PHANToM case. This may due to the larger PHANToM workspace and the more definite force feedback produced. By just moving up and down on the bars, users seemed to be able to get a better idea about the bar height than in the WingMan case.

In the third stage of the evaluation, the multimodal bar charts were compared with standard paper-based counterparts. Advantages of the multimodal approach have clearly been shown in user performance in answering the experimental questions. Tactile diagrams are a more familiar medium to blind people, therefore they could perform the same task quicker and perceived less workload than on the multimodal graphs. Based on users' comments, the multimodal interface is easy to use and performance can be improved with more practice.

Although every feature of the multimodal interface was introduced to blind users during the training section, they did not really use the haptic features to extract information from the graphs. In contrast to the sighted users, they used haptics mainly to detect the bars and the cursor location on the graph. The cursor logs reveal the horizontal movement of the cursor made by the majority of blind users in the experiment. They seldom explored the whole length of the bars or used haptics to obtain the answers. By using the improvement of the audio features, and especially the simplified synthesized-speech representation, blind users were able to find the answers. Hap-

tics' contribution was limited to navigation. This could be caused by blind people's traditional way of interacting with computers. Screen readers are the most common tools used to access information displayed on the computer screen. In the main, the haptic interface is very new to blind people, who need some time to get used to this medium.

7 Conclusion

A multimodal system which uses haptics and audio modalities to represent information to blind people has been developed. A three-stage evaluation has been conducted to verify the effectiveness of the virtual graphs and the usability of two different force-feedback devices in this form of information presentation. The advantages of the multimodal system have been confirmed by the results obtained from the comparison study with the traditional tactile diagrams. Integration of haptic and audio modalities has clear advantages over the traditional information-representation methods. By using this multimodal representation, information can be conveyed to blind people in an easy and effective way.

Based on experimental findings, the role of haptics in a user's exploration of information and multimodal interaction has been discussed. The role of haptics varies according to four factors in our evaluation: application, device, modality, and user preference. Regarding the application factor, the change of role of haptics depends on the information representation. The nature of the application and the dimensions of the information being presented determine the amount of haptic input and the way it is used. The force-feedback device used also affects the use of haptics. The PHANToM is a 3D device, while the WingMan mouse only provides 2D force feedback. This greatly limits the type of data that can be presented to users. In our application, graphs are traditionally in a 2D form so that the difference between these two force-feedback devices does not appear as a major hindrance to their use.

The difference between devices is further reduced when an additional modality is introduced into the system. The purpose of the audio is to present abstract and detailed information about the data. It is designed to complement the haptic representation and reduce the limitations of haptics. In the experiment, audio was also used to extract information about the data. Audio was designed to provide a quick overview of the data and give exact values when requested by users. The role of haptics thus slightly shifted towards the purpose of assisting user navigation on the graph. There is a significant difference between sighted and blind users' way of using haptics in the multimdal condition. Haptics were largely involved in information extraction among sighted users, whereas it was replaced by audio among blind users. Personal preference affects the decision on how to use haptics. The limitations of the haptic modality and cur-

rent force-feedback devices also have some influence. In comparison, information can be conveyed to users more directly in audio than in haptics. For blind people, speech is the most common way of getting information and interacting with the surroundings. Therefore, they tended to use speech output to obtain answers even if it was not the most effective.

In order to present information effectively in haptic and audio modalities to blind people, the four factors described above have to be considered in the development stage. Using a multimodal approach has clear advantages over single-modality and traditional methods. Guidelines to coordinate multimodal inputs will be drawn up in our future research.

References

1. Dimigen G, Roy AWN, Taylor M (1996) The entitlement, expertise and availability of information technology to visually impared students and graduates in Great Britain. In: INSERM, Paris, pp 18–20
2. Dimigen G, Scott F, Thackeray F, Pimm M, Roy AWN (1992) Do visually impaired school-leavers have different expectations from those of their sighted age mates? New Beacon 76:293
3. Dimigen G, Scott F, Thackeray F, Pimm M, Roy AWN (1993) Career expectations of British visually impaired students who are of school-leaving age. J Vis Impairment Blindness 87:209–210
4. Edman, PK (1991) Tactile graphics. American Foundation for the Blind, New York
5. Lee S, Wiker SF, Vanderheiden GC (1993) Interactive haptic interface: two-dimensional form perception for blind access to computers. In: Proceedings of the 5th international conference on human computer interaction, Orlando, Fla., 8–13 August 1993
6. Blenkhorn P, Evans DG (1998) Using speech and touch to enable blind people to access schematic diagrams. J Network Comput Appl 21:17–29
7. Roth P, Richoz D, Petrucci L, Pun T (2001) An audio-haptic tool for non-visual image representation. In: Proceedings of the sixth international symposium on signal processing and its applications, vol 1, Kuala Lumpur, Malaysia, 13–16 August 2001
8. Ramstein C, Martial O, Dufresne A, Carignan M, Chasse P, Mabilleau P (1996) Touching and hearing GUI's: design issues for the PC-access system. In: Proceedings of the second annual ACM conference on assistive technologies. ACM Press, New York
9. Colwell C, Petrie H, Kornbrot D, Hardwick A, Furner S (1998) Haptic virtual reality for blind computer users. In: Proceedings of the third international ACM conference on assistive technologies. ACM Press, New York
10. Grabowski NA, Barner KE (1998) Data visualisation methods for the blind using force feedback and sonification. Proc SPIE 3524:131–139
11. Fritz JP, Barner KE (1999) Design of a haptic data visualisation system for people with visual impairments. IEEE Trans Rehabil Eng 7(3):372–384
12. Van Scoy F, Kawai T, Darrah M, Rash C (2000) Haptic display of mathematical functions for teaching mathematics to students with vision disabilities: design and proof of concept. In: Haptic human–computer interaction. Lecture notes in computer science, vol 2058. Springer, Berlin Heidelberg New York
13. Wies EF, Gardner JA, Sile O'Modhrain M, Hasser CJ, Bulatov VL (2001) Web-based touch display for accessible science education. In: Haptic human–computer interaction. Lecture notes in computer science, vol 2058. Springer, Berlin Heidelberg New York
14. Ramloll R, Yu W, Riedel B, Brewster SA (2001) Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users. In: Proceedings of BCS IHM-HCI 2001, Lille, France, 10–14 September 2001. Springer, Berlin Heidelberg New York
15. Brewster SA, Pengelly H (1998) Visual impairment, virtual reality and visualisation. In: First international workshop on usability evaluation for virtual environments, Leicester, UK, 17 December 1998
16. Yu W, Ramloll R, Brewster SA (2000) Haptic graphs for blind computer users. In: Haptic human–computer interaction. Lecture notes in computer science, vol 2058. Springer, Berlin Heidelberg New York
17. Hart SG, Wicken SC (1990) Workload assessment and predication. In: MANPRINT, an approach to systems integration. Van Nostrand Reinhold, New York
18. Pengelly HL (1998) Investigating the use of force-feedback devices in human–computer interaction. Masters dissertation, Department of Computing Science, University of Glasgow
19. Riedel B, Burton AM (2001) Perception of gradient in haptic graphs: a comparison of virtual and physical stimuli. In: Eurohaptics 2001 conference proceedings, Birmingham, UK, 1–4 July 2001
20. Hospital activity statistics. Department of Health. http://www.doh.gov.uk/hospitalactivity/index.htm. Cited 2 Febuary 2003
21. Challis BP, Edwards ADN (2001) Design principle for tactile interaction. In: Haptic human–computer interaction. Lecture notes in computer science, vol 2058. Springer, Berlin Heidelberg New York
22. National Centre for Tactile Diagrams, University of Hertfordshire, College Lane, Hatfield, Herts, AL10 9AB, UK