Assessing the utility of dual finger haptic interaction with 3D virtual environments for blind people

K Gladstone[1], H Graupp[1] and C Avizzano[2]

[1]iSys R&D, Royal National Institute of the Blind,
105 Judd Street, London, WC1H 9NE, UK

[2]PERCRO, Scuola Superiore Sant'Anna, Pisa, Italy

{*KGladstone*, *HGraupp*}*@rnib.org.uk*, [2]*carlo@sssup.it*

[1]*www.rnib.org.uk*, [2]*www-percro.sssup.it*

ABSTRACT

Whereas the rapid growth in use of 3 dimensional computer images has proved to be of significant value to many computer users, they have remained totally inaccessible to blind people. The EU GRAB project is seeking to establish the degree to which a dual-finger haptic interface, augmented by audio input and output, can provide non-visual access to this important area of the information world. The haptic interface itself is an entirely new development controlled by a powerful haptic modelling tool. Validation of such a device is a complex procedure, not least because it falls entirely outside of the experience of most users.

1. INTRODUCTION

Of our five senses, vision is normally considered to be the most important in our interactions with our world. The eye has an amazing ability to rapidly adjust to different light levels and to zoom between viewing a great vista to examining a close object in detail. Of even greater significance is its capacity to work with the brain to collect, process and make sense of immense amounts of information. Any impairment to the visual path, whether caused, for instance, by a damage to the eye or by being in the dark, significantly reduces a person's ability to interact with their environment. The obvious implications of lost or diminished vision are difficulties in reading or moving around safely, with consequent obstacles to learning and being employed. Since increasing blindness afflicts many people in old age, the vast majority of blind people being over the age of 75, emotional effects such as social isolation and not seeing their grandchild grow up can be of far more consequence. One of the challenges for our use of technology is to develop solutions that address the issues that people really face in their lives and accept that these may not always coincide with the clear strengths of the technology itself.

The general approach to combating the effects of visual impairment has been to improve the visual environment or to shift modalities to the other senses. Any comparison will always show how limited the processing ability of hearing and touch is in relation to vision. That is neither to deny the specific and peculiar advantages of the other senses nor their ability to augment vision. This does mean, however, that people will generally use whatever residual vision they have in preference to resorting to audio or tactual material. The focus of this work is on those for whom vision is no longer or never has been an option for assimilating information by assessing to what extent a haptic and audio virtual environment can provide a valid and useful alternative.

This paper is based on contributions from members of the European funded IST project GRAB (IST-2000-27510). Further details of the project and the consortium can be found at www.grab-eu.com.

2. COGNITION

2.1 Haptic sensing

Perceptions of the world generally arise as a combination of correlated input across several of the senses. Haptic devices are frequently used in conjunction with sight, and information regarding object properties has been shown to be differentially salient under conditions of pure haptic exploration, and haptic and visual exploration (Wall and Harwin 2000). Information about an object's shape is easily extracted by visual means, whereas to gather this information haptically "requires the execution of the 'contour following' exploratory process" (Klatzky, Lederman and Reed 1987)), which puts a large demand on the user's memory for temporally varying signals, plus, the process is time consuming. In other words, exploring objects by sight results in an almost instantaneous understanding of its shape. But when objects are explored purely haptically, exploration is serial thus, more time-consuming and requiring more use of memory.

Exploration is an important part of the perception process of all senses. When exploring an object in a natural tactile environment all the fingers on both hands are involved. At each finger some extended skin area is in contact with the object providing the explorer with information. However, the exploration of virtual objects with a haptic device is different, as information at each moment is restricted to one point of the surface of the virtual object. In addition, the observer contacts it through either a stylus or a thimble meaning there is no skin contact. When the thimble is used only one finger is involved. The spatially distributed pressure on the finger tip within the thimble does not carry any information about the form of the virtual object as the latter is represented at each time by only one point and so the loss of cutaneous feedback will restrict the user (Jansson et al.1999).

One of the things determining the usefulness of a haptic device is whether it benefits, or adds to the visual experience of virtual environments. If haptic devices are to be assessed for their potential for blind people then usefulness "must be judged by partly different criteria, as haptics have to work without visual guidance" (Jansson et al.1999). But the usefulness is a combination of the haptic device and the software application, as stated by Kirkpatrick and Douglas (2000) "One of the most important attributes of a haptic device such as the PHANToM is its usability. However a haptic device by itself has neither good nor bad usability. Rather, usability is determined by the relationship between a task and an interface, which includes the haptic device as one hardware component of an integrated hardware and software system." Not all the applications may be suitable for a given haptic device. Special care has to be taken in order to define suitable applications for each haptic device.

When visual information about an object is readily available and the user is able to see it (i.e. they are not blind or severely visually impaired) then the global shape and structural cues become the defining properties of an object. On the other hand if an object is explored purely haptically then material cues such as texture and compliance have a greater significance (Wall and Harwin 2000). Similarly, Katz (1989) states that in human sensing and manipulation of everyday objects "the perception of surface texture is fundamental to accurate identification of an object."

Surface texture is clearly regarded as important by many researchers. The fact that devices such as that proposed for GRAB involve inserting the finger into a thimble reduces the amount of surface information directly available to the surface of the finger. In addition, the thimble of the PHANToM does not carry information about the form of objects because objects can only be presented through one point at a time. Jansson et al. (1999) point out also that "even if information is obtained from (changing) muscles, tendons and joints, the loss of cutaneous information may mean an important restriction". This is supported by other researchers, for example "the availability of an interface system capable of generating adequate cutaneous stimuli on the user's hand is a primary need when the problem of the recognition of features of an object is translated from a real to a virtual environment" (Bergamasco 1997).

2.2 Multimodal representation

With the increasing availability of auditory and haptic means of interaction it is not unusual for developers to try to incorporate many modalities in interfaces rather than the purely visual. Blind people must depend on non-visual senses to help in locating and identifying objects and people, and to help provide an overall conceptual organisation of spatial environments. Grabowski and Barner (1998) showed the feasibility of combining the sense of touch, using the PHANToM, and representative soundscapes to develop visualisation aids for the blind and visually impaired people.

Determining the role or roles of other senses in compensating for the loss of visual information requires careful examination, because it is not enough to merely replace visual information with audio and/or textural information. According to McGee, Gray and Brewster (2000), the user can be powerfully affected when information presented in different modalities is combined to become multimodal. They argue that “providing interface designers with the means to implement haptic-audio interfaces might result in adverse effects to interaction unless they are also equipped with structural knowledge on how to select effective combinations of such information”. They add that when two or more modalities are combined, the resulting perception may be weaker, stronger, or altogether different. Therefore effects of combining haptic and audio information “must be systematically explored to realise the potential of haptic-audio interfaces as well as to avoid creating interfaces that afford poor interaction”.

Multimodal representation of information is currently in an exploratory stage. The majority of research seems to conclude that more research is required in order to clearly identify how to improve such systems and representations of visual information. There are a number of areas that stand out as being important considerations for haptic exploration in a virtual environment with or without audio and textural enhancement. Firstly there is the issue of the degree of cutaneous feedback available through a limited number of cutaneous receptors. Exploration without sight is sequential and therefore issues about the time it takes to explore and the memory required to do so are pertinent issues. Additionally there is the issue of determining how multimodal feedback should be combined, in order to be helpful to the user, rather than weakening the feedback. It must be remembered that visual exploration that is enhanced by haptic and audio feedback is very different to haptic and audio exploration when visual feedback to the user is not possible. The effective representation of information that is normally presented visually is not a matter of directly transferring that information into a different mode. Contextual user issues and requirements need to be fully understood, and the problems that past researchers have experienced in presenting haptic, textural and audio information should be taken into consideration in the development of new devices and in cooping research.

3. USER NEEDS

Identifying users’ needs of a new technology is always difficult since few of us can articulate our needs of something we know little about. The approach within GRAB was to discuss areas of need and aspiration with groups of potential users.

All focus group participants were very interested in the GRAB device and its potential. Many had very high expectations of what the device would be capable. Some participants, for example, felt that the technology would enable them to explore the fine detail of sculptures, complex screen layouts or explore scanned photographs in 3D, even using the device to recreate parts hidden from the camera. Although some participants felt that the device would enable exploration of this kind, many participants questioned the ability of the device, expressing reservations for example about who would provide the content to such a system. The groups considered three particular applications:

- When discussing **3D Mobility Maps**, participants were unsure about the usefulness of the 3D element in what is, primarily, a symbolic representation. There was concern about how the content of such a device could be made complete and updated regularly enough and a feeling that such a device would need to be portable.
- Many participants were very enthusiastic when discussing **Games**. They were particularly excited about the prospect for blind people to play games that they cannot play in any other way. These games included target and searching games that could be combined with audio to indicate state of play and whether users were near target and so on. Adventure games, which, for example, comprise a series of rooms and tasks where objects need to be picked up before players can proceed were suggested, again with audio to indicate objects, state of play and so on.
- The more ‘technical’ participants showed great interest and high expectations of a **Haptic GUI**. Some participants liked the idea of being able to explore the entire screen in great detail, for example, being able to feel the word and emblems that are on icons. However, even with such interest in the concept, they seemed to feel that in day-to-day work they would not use the device for this kind of application, since quicker methods like the speech technology are already in use.

These focus groups added many potential applications to an already long list proposed by the developers. However, further analysis has called many of these into question either because they are not technically achievable, the cognitive load is prohibitive or because there are far simpler alternatives, many of which are already in use. Before these issues can be resolved, there is a need to gain a far better understanding about the capabilities of the device and its use.

4. INITIAL STUDIES

4.1 Simple object layouts

A series of studies were undertaken in order to obtain some indications of users aptitude in exploring objects and layouts through limited tactual sensing. Three layouts were constructed by gluing children's building blocks to a baseboard. A fourth abstract object consisted of a wooden ball threaded on a rope that was passed through a slot in a piece of wood and its ends terminated in wooden disks. Participants were asked to explore each layout in turn either with whole hand, wearing gloves, restricted to using two fingers or using thimbles. They then gave an overall description of the layout and stated how many items it contained. Further questions probed how easy they had felt the test to be, whether they were happy with the exploration time, if they thought that audio feedback would have helped and how useful they found information from touch alone. Sighted or partially sighted participants, who had been blindfolded, were asked to remove their blindfold and rate how good they felt their mental representation of the layout was in comparison with what they could see. Blind participants were given a verbal description and asked to re-explore it using two hands, and then asked to rate how accurate their original mental representation had been.

Although the average time to explore the layouts using thimbles was over twice as long as that with whole hands, it was more significant that participants did not feel that they had obtained an accurate representation. The thimbles were used since they partially simulated the proposed device by restricting the cutaneous feedback, which is the most sensitive part of the sense of touch. Participants found:

- Lack of sensation
- Difficulty determining size
- Curves being over-accentuated
- Difficulty determining shapes
- Difficulty determining angles
- Difficulty judging distance

Those who preferred thumb and index finger said that they were able to grip objects and judge distances better and they said it felt more 'natural' than with two index fingers. Those who preferred two index fingers seemed to be able to judge distance and 'arrangement' better by holding one finger still and exploring with the other.

Many participants felt that audio would help to some extent. The type of audio feedback described as being useful included a description of the layout and an indication of how many items the layout consisted of. Other participants did not feel that audio feedback would help, particularly in those conditions where thimbles were not used. Generally, they felt that the need for audio feedback depended on the situation and whether it was vital to get all the information (for example in an exam situation). Touch alone was considered to be useful as long as thimbles were not used. One blind participant said, "if there was an audio description, then what would I need to feel it for?"

No differences were evident in the findings with respect to visual status. It is important to remember in tests of this kind that it is not possible to understand completely what a blind person's mental representation is. On the other hand, sighted and partially sighted participants are able to 'check' their mental representation.

4.2 Three dimensional objects

In a second set of experiments, participants explored plaster busts of Bach and Mozart. They knew the type of object but had to glean information about the detail. Audio help was provided by identifying which part of the bust was being touched. Those who had some visual reference, i.e. had seen what a bust looked like, found it significantly easier to identify the difference between Bach's wig and Mozart's hair, though other distinguishing marks were missed entirely. The problem for a congenitally blind person was summed up by one of the participants:

> *"It's difficult to perceive. It's difficult enough to relate the real thing to an image, so this would be another few steps removed. Trying to get someone like me to understand art is like getting someone who only speaks a foreign language to appreciate my poetry. I do not possess colours and perspective, which is the language of sight. I don't have that."*

It is also clear that, while a bust may produce a good visual likeness, the feel of a real head is nothing like its plaster representation. A congenitally blind person has to learn visual shapes of real objects without ever sensing them directly. This may suggest that a haptic device of the type proposed in this project is unlikely to be suitable as an exploration tool. Effective applications may make greater use of the user's ability to interact with and modify the virtual environment.

5. THE GRAB SYSTEM

5.1 Description

The GRAB application is based on the integration of three tools:

- A two-finger 3D force-feedback **Haptic Interface** developed by PERCRO.
- **Audio Interaction** using speech recognition and voice synthesis (IBM ViaVoice**)**
- The **Haptic Modeller** developed by LABEIN to allow the interaction with any 3D virtual object through haptic stimuli, sounds aids and speech recognition.

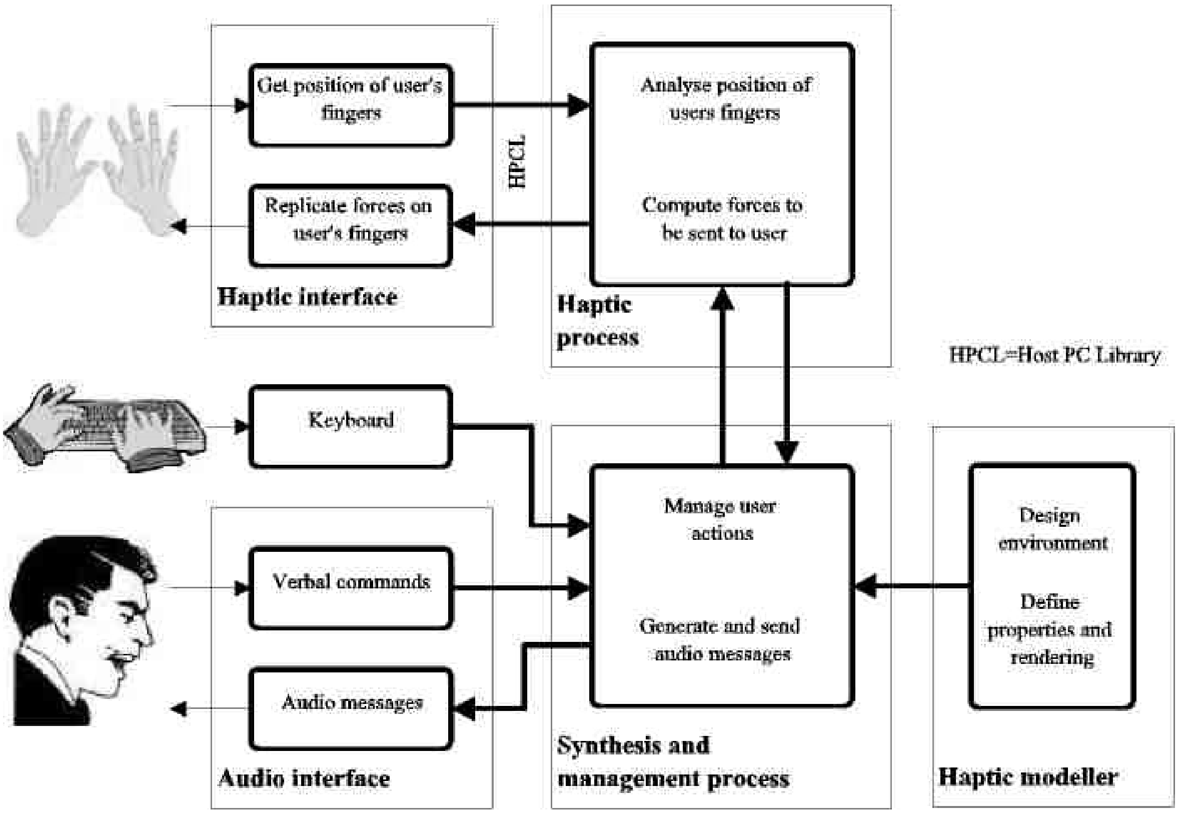

Figure 1 *System view of the GRAB Haptic and Audio Virtual Environment*

The user inserts two fingers (the thumb and index or both index fingers) into the thimbles of the haptic interface and moves them through the virtual workspace. The user can interact with the virtual objects, listen to audio messages and execute verbal or keyboard commands. The operator screen shows the virtual objects and the points that represent the positions of the user's fingers within the virtual space. The various elements of the system include:

- The properties and position of all objects within the virtual environment are defined by the **Haptic Modeller**.
- Using information about the position of the user's fingers, the **Synthesis and Management Process**:
  - Renders the virtual environment, including any dynamic modifications to, for instance, the placement of objects.
  - Manages the actions required by the user.

- Generates and sends both speech and non-speech audio messages either automatically, based on object properties, or at user request.

- User commands can be given either by voice through the **Audio Interface** or by keyboard.
- The control of the **Haptic Interface** is maintained by the **Haptic Process** executing the cycle of "get position – analysis position – replicate force" at a constant frequency of 1kHz.

5.2 Functionality

The first prototype of the GRAB application provides the following functionality:

- Simulation of different forces that: represent **contact** with an object, modified by properties such as stiffness, texture and stickineopjects whose size is very small or very large.
- **Panning** across the virtual environment if its size is greater than the physical workspace.

The initial implementation will include a number of test scenarios in which each function is, as far as possible, isolated to facilitate user validation. The functions will also be combined to produce an adventure style game requiring the user to navigate through several rooms on two floors. The user will encounter a variety of obstacles, including time delay triggers that will have to be reset within a given period.

6. HAPTIC INTERFACE

6.1 Mechanical Features

The haptic interface consists of two co-ordinated arms, each with six degrees of freedom, allowing relevant forces to be exerted on the user's fingers as shown in Figure 2. Each arm consists of a passive gimble attached to an actuated serial chain. The specific differential kinematics achieves a high degree of stiffness and isotropy while still exhibiting low inertia on the fingertips. The interface has been designed so that the joint workspace of the two arms can cover a large portion of the desktop (arranged as a parallelepiped of 600mm wide, 400mm height, and 400mm depth). The system is equipped with high performance DC motors that directly drive a cable transmission, thus avoiding problems of backlash. While operating in the centre of the workspace, a weight counterbalances the moving mass of the barrel, thereby reducing the amount of torque required from the motors. The motors have been designed such that the interface can exert a force of at least 4N throughout the workspace, though higher figures can be achieved when operating in the centre or the workspace. A peak force of 12N can be achieved for a short time. The system is equipped with a set of encoders for measuring the finger position with a spatial resolution of better than 100μm worst case. The device is designed to be transparent to the user by minimising extraneous forces generated by the interface.

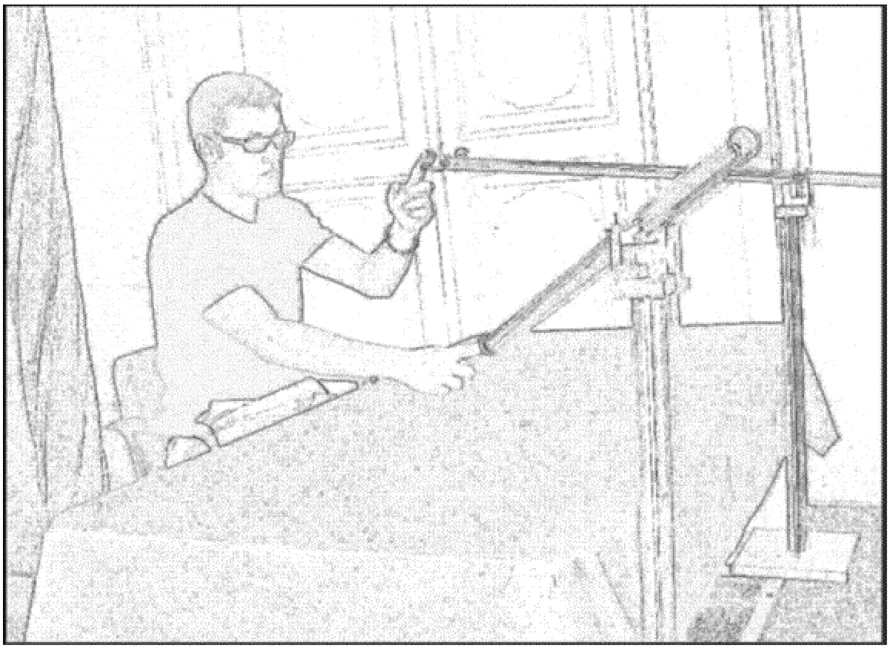

Figure 2 *The haptic interface in use*

6.2 Control Features

The moving arms are provided with semi-embedded drivers that take care of all the basic control aspects for managing the interface. Control and driver units generate the required forces on the fingertips while measuring their position in the workspace. Moreover, the control unit compensates for all the non-linear features related to the interface control and the force generation, such as the gravitation compensation, the active friction compensation and the interface kinematics. The whole control system implements a real time scheduler and runs at a nominal frequency of 5KHz, which was chosen to avoid stability problems at contact even when the arm is remotely controlled. The system can be driven by attaching its data port to the parallel port of a remote computer running any type of operating system. A set of libraries and drivers allow the control of the haptic interface at a frequency of 1kHz, though this could be increased to 14kHz if required.

The controller has been designed to incorporate several safety features, particularly bearing in mind the needs of blind people. When the interface is moving autonomously, the controller limits the maximum velocity and forces that can be exerted on the joints. This minimises risk of injury even if the user gets too close to the device or touches it improperly. The interface can be configured to warn the user of specific actions. The power to the motors is disabled if the emergency pushbutton is operated or the interface is been left unattended for a more than an hour.

7. VALIDATION PLAN

7.1 Aims

The developers of the GRAB device have designed several algorithms that simulate various properties of virtual objects through forces generated on two points. The aim of the validation is to confirm that the implementation of these algorithms is effective for users of the prototype. The findings of the validation testing will be used to decide whether the functional requirements have been met adequately and whether their usability meets the success criteria. The outcome will contribute to a wider feasibility report.

7.2 Participants for validation testing

In order to obtain representative results, participants will be drawn from groups of people with varying sight problems, including:

- Congenitally blind people
- Adventitiously blind people
- Partially sighted people
- Sighted people

Previous research suggests that people who have or have had some sight tend to use their visual memory even when carrying out a haptic task. The inclusion of congenitally blind people may indicate how important such visual memory is when using a system such as the GRAB prototype.

It has been suggested that a person's 'haptic aptitude', that is their ability to use touch, may significantly affect how well they can use the GRAB device. Factors that may indicate such an aptitude include whether they like tactual exploration and competency in reading raille or tactile graphs and diagrams. Although people who regularly use tactual material do not necessarily have a more sensitive sense of touch, that sense is trained to be effective. These factors will be checked as part of the participant characterisation.

7.3 General outline of procedure

The test sessions will consist of three stages.

- **Familiarisation:** During this stage the user will receive information about the main features of the system and broadly how it works. They will be given a short training session to become familiar with the device. This will mean that the system is tested after the initial 'novelty' period.
- **Test stage:** The user will be asked to work through a set of scenarios to test the main functionalities of the system, including the audio help features. Various aspects of their performance will be recorded including time, accuracy and error rates.
- **Subjective testing:** Participants will be asked how easy the system was to use and how satisfied they were with various parts of the experience. They will also be asked for their subjective

viewpoint of the potential usefulness of the system.

7.4 Professionals' validation testing

A second group of participants will include a sample of professionals composed of people from a variety of industry sectors, but all of who work closely with blind and partially sighted people. These participants will be drawn from groups such as rehabilitation workers, mobility trainers, teachers, IT professionals and psychologists. This group will help to assess the possible usefulness of the device in their work. Although some of the testing may include similar scenarios as used in the first set, there will also be opportunities to test the ability of the professionals to design virtual environments for use with the device.

The procedure for testing with professionals will be the same as for the other participants. That is, a familiarisation session, followed by an experimental testing session, followed by a gathering of subjective opinions from participants. This final session will focus on the perceived usefulness of the system, the perceived contribution of the system to their work and the perceived ability of the device to increase the quality of their work.

7.5 Reporting

The validation findings will be combined with a detailed review of possible implementation scenarios, the potential market size and estimated costs in order to produce the final feasibility report. One of the main outcomes of this phase of the project will be a list of proposed applications in the general areas of learning, employment, mobility, leisure and daily living. The first application being built is a realisation of an adventure game based on exploring several rooms that contain hidden dangers and methods to neutralise them.

Acknowledgements: Much of this work was carried out as part of the European GRAB project and we are grateful for funding under the Information Society Technologies Programme of the European Union Fifth RTD Framework Programme. The consortium members are LABEIN (Spain), ONCE (Spain), PERCRO (Italy), NCBI (Ireland), HAPTICA (Ireland) and RNIB (UK). Further details of the project and the participants can be found at www.grab-eu.com

8. REFERENCES

M Bergamasco (1997), Shape Perception of 3D Virtual Objects by Tactile Feedback Systems, *ERCIM News,* **31**

N A Grabowski and K E Barner (1998), Data Visualisation Methods for the Blind Using Force Feedback and Sonification, *Part of the SPIE Conference on Telemanipulator and Telepresence Technologies V*, Boston Massachusetts.

G Jansson, H Petrie, C Colwell, D Kornbrot, J FŠnger, H Kšnig, K Billberger, A Hardwick, and S Furner (1999), Haptic virtual environments for blind people: exploratory experiments with two devices, *International Journal of Virtual Reality*, **4**, 1, pp. 10-20

D Katz (1989) *The World of Touch*, (translated by Krueger, L.E.), Erlbaum, Hillsdale, NJ. (Original work published in 1925)

A E Kirkpatrick and S A Douglas (2000), A shape recognition benchmark for evaluating usability of a haptic environment, *HHICW 2000*

R L Klatzky, S J Lederman and C Reed (1987), There's more to touch than meets the eye: the salience of object attributes for haptics with and without vision, *Journal of Experimental Psychology,* **116,** 4, pp.356-369.

M R McGee, P Gray and S Brewster (2000), The Effective Combination of Haptic and Auditory Textural Information, *HHICW 2000*

S A Wall and W S Harwin (2000), Interaction of Visual and Haptic Information in Simulated Environments: Texture perception, *HHICW 2000*