Interaction Sound Feedback in a Haptic Virtual Environment to Improve Motor Skill Acquisition

Christian Müller-Tomfelde

CSIRO - Information and Communication Technologies Centre
Locked Bag 17, North Ryde, NSW 1670, Australia
`Christian.Mueller-Tomfelde@csiro.au`

ABSTRACT

This paper describes the concept and the realisation of a research prototype of a haptic environment that is enhanced with sound feedback to impart implicit knowledge and to teach motor skills to trainees. The sound feedback is to be understood as an additional feedback that provides information to the trainee about his or her actions in relation to objects in a virtual haptic environment. Although the setup simulates the interaction with haptic objects, the additional auditory feedback goes beyond the imitation of real sound behaviour and enables new ways to convey information about user interaction, i.e., movements in space with or without forces applied to tangible objects. The goal of the interaction with additional auditory feedback is to enrich the virtual environment to a full multimodal interaction environment, to support precise movements and to optimise the motor skill learning curve of trainees. The prototype is focused on computer based learning in the context of surgical training and it uses a virtual document planner that provides the appropriate instructions and tailored feedback for the human computer interaction through reasoning on the basis of an explicit knowledge representation.

Keyword: auditory feedback, human computer interaction, implicit interaction, motor skill training, virtual document planner.

1. INTRODUCTION

In the area of medical applications and especially in the surgical procedures, the technology of haptic virtual environments, computer assisted surgery and medical robots will have a major impact on heath care in the next decade [1],[2]. As a side-effect of this development, the surgical training becomes no longer restricted to a) the apprenticeship model, where the trainees are supervised by an experienced mentor in the operation room or to b) courses in laboratories using models or animal tissue. Furthermore, new directions in skill acquisition can be identified by using haptic virtual environments [3], e.g., the teaching of special procedures with tools in a minimal invasive surgery, the simulation of a laparoscope in a cholecystectomy (gall bladder removal) [1], as well as skills, in case of suturing [4] and burring [5]. Although a skillfully performed operation is 75% decision making and only 25% dexterity [6], these 25% cannot be learned or improved by studying text books or visiting lectures. Instead, the dexterity has to be learned by performing the related movements several times. While reasoning refers to explicit knowledge, the motor skill represents the knowledge that is implicit or tacit [7], and that is established only during practice, e.g., during sporting activities.

In order to investigate new possibilities for skill acquisition for surgical training using computer technology, the current methods of training are imitated in the simulation, and additional features are added to exploit possible improvements beyond real environments. Therefore the contribution of sound feedback is considered to be useful during the user's interaction to support learning a certain sequence of movements and using tools. During the interaction with the environment, the user perceives visual and tactile cues delivered by the environment (exteroception) but also the position and movements of parts of his or her body (proprioception), like limbs and bones. This perception could be reinforced and the motor skill learning process improved through an alternative auditory feedback that provides supplementary indications to the user, for instance to adjust the position and orientation of arm, hand and fingers.

This proposed auditory feedback in the training situation can be compared to that of the play of a musical instrument, where, e.g., an unsteady tone of a violin indicates jerky movements, while a constant sounding tone refers to a smooth stroke of the bow by the player. Finally the sound feedback supplements and enriches the interaction to a full modal interaction that offers coherent perception of the virtual environment. In the following sections the context of the research prototype and the specific implicit forms of user interaction are explained. Different types of sound enhanced feedback are described and early experiences are discussed. Possible directions for future work are outlined at the end.

2. COMPUTER SUPPORTED SURGICAL TRAINING

Traditional surgical training is based on an apprenticeship model, where the surgical trainees learn by watching and participating in operations of their mentors. Surgical procedures are taught on the basis of the mentor's interpretation of current standards, and an objective assessment of motor skills is not possible [3]. The introduction of computer based medical simulations for surgical training is founded on a haptic virtual environment that provides new possibilities for training motor skills. This technology enables multiple new directions in surgical training. First, the way of learning and skill acquisition becomes non-threatening and risk-free for the surgical trainee, while the format or concept of the surgical training itself is not affected. Second, the learning and the curriculum of the trainees may alter, as the teaching method is changed from a teacher-centred to a more trainee-centred model. The education becomes more trainee-driven, depending on the individual initial capabilities and development.

The research prototype described in this paper focuses on two aspects of computer based motor skill training. A) The haptic environment is combined with a reasoning system based on a virtual document planner (VDP) [8], [9], which allows for the presentation of tailored instructions and appropriate feedback to the trainees to enable them to achieve the goal of the training. The reasoning is done on the foundation of the user response and several knowledge bases that model, amongst others, the domain, the user and the discourse history. B) Additional sound feedback provides the user with more information about the interaction in the haptic scene than an equivalent real environment would do. These unique characteristics of the resulting virtual environment can be employed to present useful information beyond those of real environments [10] as well as to obtain improvements in the quality of movements of a novice during the training phase.

3. AUDITORY FEEDBACK DURING TRAINING

The interaction of the user in the haptic virtual environment is organised by the relation between a force feedback stylus and anatomical landmarks. Landmarks represent prominent points in space that are usually associated with a special part or region of the body and that are invisible and not tangible. Landmarks provide help in orientation and navigation in an unknown environment. The absolute position of the landmark can vary from human to human but their meaning and function remain the same. The landmarks help novices during orientation and for experts, the landmarks help in communication and planning. Landmarks also reduce the complexity of a three-dimensional scene, so that reasoning with the VDP about the interaction becomes feasible.

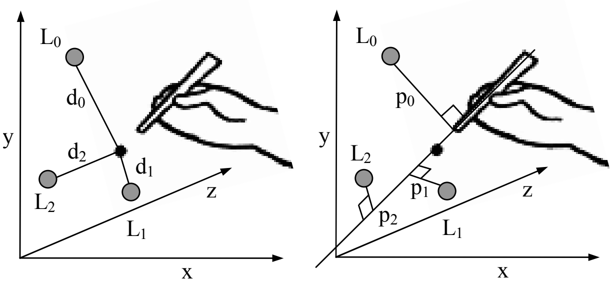
One of the basic tasks of an operation surgeon is to identify landmarks and to build up the mental model of the spatial relation of the anatomic structures in the present situation [11]. In the beginning of an exercise the trainee is instructed to successively identify the relevant anatomical landmarks. In contrast to an exercise supervised by an experienced surgeon who can observe and check the trainee's actions, the method in a computer based training changes: the trainee has to explicitly interact with landmarks to demonstrate knowledge about them.

The goal of the sound feedback at this stage is not to give a hint to the user about the kind of landmark in question, but to support the manual positioning of the tool tip at the interface object. Although the landmark is invisible and even non tangible, the user receives a clue about the existence of a landmark and can adjust the interaction with the object accordingly. One focus, so far, is on the support of different methods of this selection of landmarks, while another one deals with different qualities of the movement of the tool itself.

3.1. Approaching landmarks

To enable reasoning about a correct or false identification of a landmark during the orientation, the distance between the landmarks and the tool tip are calculated and updated continuously. This allows determining the closest of all landmarks to the current position of the tool tip (see figure 1a). In order not to process absolute distances, the measurements are set into relation to the furthest landmark and correspond to values of relative proximity. While the user is moving the tool the implicit interaction is transformed into an explicit interaction event, assuming the landmark remains the closest of all for a certain time span. To cause this explicit interaction event, the tip of the tool has to be held steady for approx. 1.5 seconds in the vicinity of the landmark under test (e.g., L_1 in figure 1a), and the resulting interaction event is passed to be examined by the reasoning system (VDP).

The distance between the landmark and the tool tip can be used to control a sound synthesis so that an audible feedback indicates the distance information. This sound feedback is a feature that usually does not appear in real environments. It therefore extends and enriches the possibilities of perception: the closeness of an invisible virtual object becomes audible; in other words, the landmarks gain a sonic aura, and, furthermore, the sound feedback allows the system to convey the distance parameter.



a) b)

Figure 1. a) The distances between the tool tip and landmarks in the 3d space of a haptic environment. 1.b) The distances of the landmarks according to the direction, in which the tool is pointing.

3.2. Pointing at landmarks

Pointing at a landmark is another way to refer to a location in space and thus selecting a landmark. In contrast to the approaching procedure, the tip of the tool could be now remote from the landmark in question, and the orientation of the stylus of the tool becomes important. An assumed line in the direction of the stylus has to be close to the landmark to select the landmark; brief, the user has to point at the landmark (laser pointer metaphor). The minimal and orthogonal distances between the extended stylus and the landmarks are calculated and updated (e.g., landmark L_2 and the minimal and orthogonal distance p_2 in figure 1b). The output values are set into relation to the furthest of all landmarks, and, as a result, a proximity factor in percentage is passed to the sound synthesis to feedback this information to the user.

3.3. Motion of the tool

Besides supporting interactions with landmarks the sound feedback can also be controlled by information that is gained exclusively from the analysis of the user's movement of the tool tip. In a first attempt the movement is processed in order to detect that the user holds the stylus steady at an arbitrary position in space. A minium radius defines a sphere in which the user must hold the stylus tip for a certain time span: after the spatial constraint for the hold detection is and stays fulfilled, a sound feedback is generated that is controlled by the relative displacement of the tip until the temporal constraint becomes fulfilled after approx. 2.0 seconds. Thus the feedback is sonifying the process of detection and could help to keep the tool steady, for instance in the situation in which the demand is to apply a constant force with the tool tip at a specific point in the scene for a certain time.

Under further scrutiny is a sound feedback for different forms of movements, like translation or rotation, more complex trajectories in space and also in a possible combination with applied forces during interaction, like pushing or pulling. We plan to give the users immediately feedback about certain qualities of their movements and therefore enable them to correct their action, if needed. At this point non-verbal sound feedback seems to be more suitable than other alternatives, like text advice or visual analogue value representation with a meter, because it is less distracting and demands the attention of the spare auditory modality. This *immediate* auditory feedback during the performance of a sequence of distinct movements has the potential to increase skill performance metrics, e.g., the 'economy' and 'purposefulness' or 'fluidity' of user's movements, like proposed in [12]. In contrast to a so called haptic guidance, that uses gentle force to keep the trainee stying on an ideal trajectory, the guidance by sound is less rigid, has a more indirect effect on the motor performance and draws attention of the trainee to the proprioception.

3.4. Sound Synthesis and Effect Strategies

Both forms of interaction with the landmarks are non-standard, e.g., there is no equivalent in natural environments of the application domain. The goal is not to simulate or imitate existing sound behaviour but to provide useful information for the user interaction [10]. The support of the discovery of the anatomical landmarks by auditory feedback can lead also to more engaging interaction and efficient training in computer based and trainee driven exercises.

Different abstract sound effects are employed to convey the sense of 'being tuned to', i.e., selecting a landmark by means of one of the two described methods (subsections 3.1 and 3.2). In a first approach the difference in frequency of two sinusoids, i.e., the beat frequency is controlled by the proximity factor in order to provide useful information to adjust the tool tip position. A tool position relatively far from the centre of the landmark sounds detuned with a high beating frequency, while the tool tip at the same position as the landmark sound like one single sinusoid.

Another sound effect addresses directly the spatial relation between the tool tip and the landmark: the distance is mapped into a reverberant auditory space where the tool tip functions as a microphone (probe) and the landmark as a loudspeaker. A synthetic reverb algorithm for late reverberation [13] is applied to the 'sound of the landmark' and an additional direct sound path is added to accomplish a simple impulse response and thus to imitate a plausible reverberant space [14]. The distance between the tool tip and the landmark is adjusting the mixing ratio between the amplitude of the direct sound path and the late reverb. Closeness is presented by a loud direct sound exceeding significantly the level of the late reverberation, while the direct sound path is dipped into the late decaying reverb when the distance increases. Here the landmark could emit an individual sound that is merely modified by the spatial effect, but does not effect the perception of the original character of the sound source. Finally the volume of the resulting sound is also controlled by the proximity factor to obtain a significant audio signal only in the vicinity of the landmark.

It should be mentioned that the suggested sound feedback could be further extended with perceptual relevant auditory information about the shape [15] or [16], and the material that are involved in the interaction (see therefore, e.g. [17], [18] or [19]). At the moment the anatomical landmarks are abstract objects and to prevent misinterpretations during interaction, a basic abstract sound was chosen for the 'sound of landmarks' (low frequency pulse train) to avoid any allusions and irritating connotations.

3.5. First Experiences

Early informal experiences were encouraging but detailed experiments have yet to supply evidence for the assumption that the additional feedback improves the interaction in terms of precision and accuracy. We plan to evaluate the proposed enhancements and to integrate appropriate extensions to the prototype for computer based surgical training.

At the moment, the sound feedback for the relative proximity between the tool tip and a landmark as described in section 3.4 (sound synthesis with beat frequency) gives the user a reasonable hint, in addition to a visual clue about the relation between the tip position and the landmark. The initial relative proximity threshold of 90% for the pointing detection without sound feedback was already sufficient to trigger an explicit interaction event. After adding the described sound feedback the threshold for proximity for explicit interaction had to be increased to complicate the positioning task and to collect experience with the proposed sound feedback. At that level of 98% rel. proximity it was reported that during the adjustment of the tool the user's attention starts to oscillate between the visual and the auditory modality.

Experiments with spatial sound effect processing lead to less improvements of adjustment but offer an interesting additional dimension of information representation by sound. Because the perception of the original character of the source sound is not affected by the spatial sound effect the source sound itself offers an extra 'channel' to convey information.

4. THE RESEARCH PROTOTYPE FOR TRAINING MOTOR SKILLS

The core system architecture of the skill training prototype is a haptic environment or Haptic Workbench (HWB,) that consists of a PHANTOM haptic force feedback device in a hand-immersive display setup that guarantees coherent spatial representation of haptic and visual objects (see for details [20], [21] and figure 2). The interactive scene is perceived fully three-dimensional due to the employment of shutter glasses for stereo vision. For this HWB an object oriented application programming interface (reachin API, [22]) is used that enables to develop the haptic and visual scene graph in a single document representation.

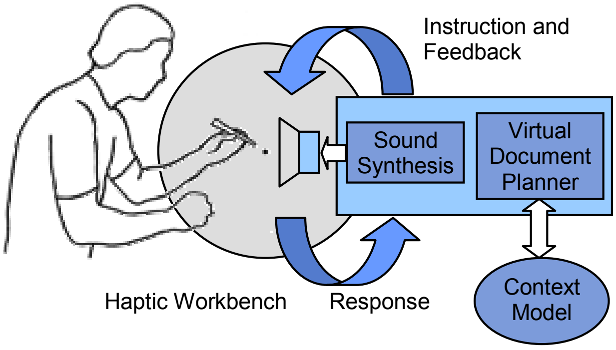


Figure 2. The user interacts in a haptic scene and the virtual document planner controls the exercise based on the context and the user response. Sound signals are generated and directly fed back to the user.

To experiment with different forms of sound feedback the setup was extended with the facility to present and control sound in real-time. Therefore the reachin API was used to build dedicated software objects to determine the user interaction. These objects pass the obtained information to other program environments using inter-process communication facilities. In a succeeding stage reasoning about the user interaction was achieved by employing and extending the VDP [8],[9]. Beyond the visual haptic feedback which is provided by the HWB, the VDP provides appropriate instructions and tailored feedback to the user in respect to the current state of the context. These instructions and feedback present and control information in different qualities and different modalities, like text output or graphical output on the computer screen, as well as text-to-speech output.

The dedicated interface objects calculate the distances and perform the transformation of a continuous implicit interaction into a discrete explicit event, in order to obtain information about the user interaction. Concurrently, the distances are continuously mapped to a relative percentage scale and sent to a sound environment. This control message stream with an update rate of about 60 Hz is further processed in a sound synthesis programming environment (pure data, [23]).

In the sound synthesis environment the messages are parsed and control parameters of the addressed sound synthesis units. Different synthesis and effects can be combined to experiment with the proposed sound feedback for the future evaluation purposes.

5. FUTURE WORK

The next steps are the extension of the catalogue of sound synthesis and effects to experiment and evaluate the support of the user interaction in a formal way. Another part of future work will be the development of a collection of real time detectors for different characteristics of movements during the interaction and associate them to sound effects. This will allow identifying possible reinforcing trends in skill acquisition due to the enhancement with sound feedback. It is assumed that this extra auditory perception enriches and supports the learning process of precise and accurate motor skills.

6. ACKNOWLEDGEMENTS

The author would like to thank Clara Teoh for her valuable contribution on extending the VDP to cope with the requirements of implicit interaction. Thanks also to Chris Gunn and Matthew Hutchins for their valuable advice and for their help to tame the Haptic Workbench. Many thanks also to Cecile Paris and Duncan Stevenson for the encouraging discussions and comments.

7. REFERENCES

[1] McCloy R. and Stone, R., Virtual reality in surgery, *British Medical Journal*, Volume 323, pp. 912-915, Oct. 2001.

[2] Tendick, F., Downes, M., Gokekin, T., Cavusoglu M., Feygin, D., Wu, X., Eyal, R., Hegarty M. and Way, L.W., A Virtual Environment Testbed for Training Laparoscopic Surgical Skills, *Presence* 9(3), pp 236-255, 2002.

[3] Cosman, P. H., Cregan, P. C., Martin C. J. and Cartmill, J. A., Virtual Reality Simulators: Current Status in Acquisition and Assessment of Surgical Skills, *ANZ Journal of Surgery* 72, pp 30-34, 2002.

[4] Moody, L., Baber, C. Arantis, T. N. and Elliott, M., Objective Metrics for the Evaluation of Simple Surgical Skills in Real and Virtual Domains, *Presence*, 12(2), pp 207-221, 2002.

[5] Agus, M., Giachetti, A., Gobbetti, E., Zanetti G. and Zorocolo A., Real-Time Haptic and Visual Simulation of Bone Dissection, *Presence* 12(1), pp 110-122, 2003.

[6] Darzi A., Smith, S. and Taffinder, N., Assessing operative skill, *British Medical Journal* 318, pp 888-889, 1999.

[7] Magill, R. A., *Motor learning: concepts and applications*, 4th ed. Madison, Wis.: Brown & Benchmark, 1993.

[8] Moore, J. and Paris, C., Planning text for advisory dialogues: Capturing intentional and rhetorical information, In: *Computational Linguistics*, 19(4), pp 651-694, 1993.

[9] Paris, C., Information Delivery for Tourism, In: *IEEE Intelligent Systems* Intelligent Systems for Tourism: Trends and Controversies, Stephen Staab and Hannes Werthner (Eds), pp 53 - 66, November/December, 2002.

[10] Müller-Tomfelde, C., Streitz, N. A. and Steinmetz, R., Sounds@Work - Auditory Displays for Interaction in Cooperative and Hybrid Environments, In: *Human-Computer Interaction: Theory and Practice (Part II)*, C. Stephanidis, J. Jacko (Eds), Lawrence Erlbaum Publishers. Mahwah, N.J., pp 751-755, 2003

[11] Wilhelm, D. M., Ogan, K., Roehrborn, C. G., Cadeddu J. A. and Pearle, M. S., Assessment of Basic Endoscopic Performance Using a Virtual Reality Simulator, *American College of Surgeons*, 195(5): 675-681, November 2002,

[12] Satava, R. M., Cuschieri, A. and Hamdorf, J. Metrics for objective assessment, *Surgical Endoscopy* 17, pp 220–226, 2003.

[13] Stautner, J. and Puckette, M., Designing multichannel reverberators, *Computer Music Journal*, 3(2):52-65, 1982.

[14] Jot, J.-M., Etude et Realisation d'un Spatialisateur de sons par Modèles Physique et Perceptifs, PhD thesis, Telecom Paris, September 1992.

[15] Dutilleux, P., Müller-Tomfelde, C., AML: Architecture and Music Laboratory, In: *Proceedings of the AES 16th International Conference: Spatial Sound Reproduction*, Rovaniemi, Finland, pp 191-206, April 1999.

[16] Rocchesso, D. and Dutilleux, P., Generalization of a 3-D resonator model for the simulation of spherical enclosures, *Applied Signal Processing*, 2001(1), pp 15–26, 2001.

[17] Gaver, W., Synthesizing auditory icons, In: *INTERCHI* '93,24-29, pp 228-235, April 1993.

[18] Klatzky, R., Pai, D., and Krotkov, E., Perception of material from contact sounds, *Presence*, 9, pp 399-410, 2002.

[19] Barrass S. and Adcock M., Interactive Granular Synthesis of Haptic Contact Sounds, *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, Helsinki University of Technology, Espoo, Finland, 15th-17th June 2002.

[20] Stevenson, D., Smith, K., Veldkamp, P., McLaughlin J., Gunn, C. and Dixon, M., Haptic Workbench: A Multisensory Virtual Environment, *The Engineering Reality of Virtual Reality, Electronic Imaging '99*, San Jose, January 1999.

[21] Phantom Sensable, retrieved September 01 2003, from http://www.sensable.com/

[22] Reachin API Programmers Guide, retrieved August 08 2003, from http://www.reachin.se/

[23] Pure data, retrieved July 23 2003, from http://www.pure-data.org/