

Interactive Landmarks: Linking Virtual Environments with Knowledge-Based Systems

Christian Müller-Tomfelde, Cécile Paris and Duncan Stevenson
CSIRO - Information and Communication Technologies Centre
Locked Bag 17, North Ryde, NSW 1670, Australia
Email: {Christian.Mueller-Tomfelde, Cecile.Paris, Duncan.Stevenson}@csiro.au

Abstract

This paper describes the concept of interactive landmarks for human computer interaction in a three-dimensional virtual environment enhanced with knowledge-based reasoning. The proposed concept of landmarks extends the traditional and more passive role of landmarks for orientation and navigation tasks. It enables the system to perform additional reasoning about the user's interactions in space. This approach bridges the gap between a simulation of physical system in a virtual environment and its potential representation in a knowledge-based reasoning system and couples the systems to form a knowledge-enhanced virtual environment. As a proof of concept, we present an initial prototype of an application for surgical training using the concept of interactive landmarks. The goal of the prototype is to demonstrate how training can be authored and interaction can be designed in a coherent and effective way.

Keywords

Human Computer Interaction, interactive landmarks, virtual environment, knowledge-based reasoning, surgical training.

INTRODUCTION

The term virtual environment refers usually to a user interface concept where the computer provides a real-time, interactive, computer-simulated three-dimensional environment, in contrast to traditional desktop environments with Windows, Icon, Menus and Pointing (WIMP). The fundamental user interaction in this virtual environment is based on the spatial metaphor and the simulation of objects within the space. It provides an intuitive access to the environment. While simulations have achieved considerable realism due to the available processing power of modern computers, more effort is needed to increase the quality of interaction, by, for instance, incorporating an explicit knowledge representation (West and Aylett, 1998; Jung *et al.*, 1998). A closer examination reveals differences in the structure of the representation of an object in a three dimensional interactive virtual environment and that in a knowledge base reasoning system. These differences have to be overcome when the goal is to create an interaction system that can provide explicit knowledge, e.g., to guide and direct the user (Luck and West, 2000).

The proposed concept of *interactive landmarks* avoids dealing with the trade-off of having either a virtual environment with some knowledge representation or a knowledge base reasoning system with some spatial components. So, instead of reinventing the wheel, existing approaches from both domains should be connected to exploit and bring together the achievements of the heterogeneous approaches. Therefore, a linking object is favoured rather than transferring one approach or system over the described structural barriers towards the other. This link must be reasonable with respect to the resulting user interaction and feasible in terms of realisation effort and reuse. We identified landmarks as possible candidates to link the scene-graph with the knowledge representation and to build a knowledge-enhanced virtual environment. The resulting system enables the creation of more engaging user interactions due to the incorporation of explicit knowledge (Bowman, 1999).

In the following, we describe landmarks in the context of interaction, starting with a definition and discussion of aspects of landmarks that are important regarding user interaction. Then we propose the concept of interactive landmarks and illustrate implicit methods of interaction with landmarks in space. The differences between the representation of objects in a virtual environment and a knowledge base are also discussed. Finally, the prototype and the basic components are described as well as the current realisation and interaction with interactive landmarks, and the paper finishes with suggestions for future work.

LANDMARKS FOR INTERACTION

In the traditional or literal meaning, landmarks refer to points in our visible surrounding that help us decide where we are and which direction we have to go to reach a destination or goal. They are marks or signs in the environment that are persistent and long lasting, and hence enable us to build a simple mental model of an otherwise complex environment, that is perhaps even unknown to us. This mental model allows us to simplify

the task of getting from A to B. It also forms the foundation of the communication with others about orientation and navigation problems. Besides the more metaphorical usage of the term landmark in literature, the term landmark is also borrowed from other disciplines for the orientation and navigation in an unknown environment. In the medical area, so-called anatomical landmarks are defined for large number of anatomically meaningful locations in the human body, e.g., limbs, bones, etc. Anatomical landmarks are also defined for the human brain, on the one hand, to facilitate orientation in the surgical situation and, on the other hand, to support discussions amongst experts or in an educational setting. While the absolute position of an anatomical landmark is depending of the individual anatomy, the role and meaning of the landmark remains common knowledge.

All considerations until now have in common that the landmarks are thought to be distinct points or districts of a complex environment, and that the landmarks are considered to be passive. The user can see them and can watch them from different directions. But, beyond the role of a constituent of a model of the environment, a landmark does not offer supplementary features that enable the watcher to interact with landmarks explicitly. Before furthering these arguments we first provide a definition of the term landmarks.

The definition of landmarks

The definition of three dictionaries (Oxford English Dictionary, 2004; Merriam-Webster Online, 2004; Hyperdictionary, 2004) have been analysed and summarised to formulate a general understanding of what constitutes a landmark. The most important characteristics are:

- A landmark is a conspicuous object;
- A landmark forms a district or boundary but may also be a temporal event;
- A landmark represents a point in space;
- A landmark has the ability to guide or support navigation;
- A landmark provides a way to measure distances and
- A landmark is used as a point of origin to locate other landmarks.

With the exception of an event, a landmark is basically an object in space. However, it is also meaningful to include temporal objects, because events can also segment different time zones or periods¹. The role of landmarks is mainly for orientation and navigation purposes in unknown environments and, additionally, to estimate distances in space or time. A relatively important aspect is that a set of landmarks can be grouped together to form one new landmark. This recursive property of landmarks can lead to a hierarchical description of all landmarks of an environment. Finally, a linear combination of landmarks can represent a path from location A to location B. A tentative definition of a landmark could be the following:

A landmark is a conspicuous object in space or time that marks a district or boundary. Multiple landmarks together can form another landmark and enable users to orient themselves and navigate in unknown terrain.

Other definitions can be found in the literature, like that of Sorrows *et al.* (1999) which says: “Landmarks are prominent, identifying features in an environment, which provide an observer or user of a space with a means for locating oneself and establishing goals.” An important and not mentioned property is that landmarks are usually static in time and space in order to represent reliable, useful and recallable information for different users. The indented definition above provides the foundation to discuss further aspects and possible extensions of the use of landmarks for interaction.

Aspects of landmarks: spatial knowledge, dimensions and affordance

In the context of the spatial knowledge of humans, Stephen M. Kosslyn (1987) proposed, as described by Ware (2000), two forms of knowledge: categorical and coordinative mental representations of objects in space. While the spatial information in the categorical representation combines declarative knowledge with topological relations between objects, like “the chair is to the left of the couch”, the coordinate representation specifies the exact (absolute) location in visual degrees, metres, etc. Assuming that the object in space that causes a mental representation can be defined as a landmark, these considerations allow us to discuss the role of landmarks when the user has to navigate, e.g., in an unknown terrain. A landmark is a combination of a prominent point in space

¹ In this article we concentrate only on interactive landmarks in space and ignore the existence of landmarks in the time domain.

and an object that carries additionally higher semantic information, and thus links the categorical and coordinate representations (Ware, 2000). This fact makes the landmark particularly suitable to support the user during spatial orientation and navigation tasks. Although every object in space has both representations, only landmarks are conspicuous enough to step into the foreground, attract the users' attention and invite them to interact.

In the article "The Nature of Landmarks for Real and Electronic Spaces", Sorrows *et al.* (1999) elaborate three global dimensions of landmarks: visual, structural and cognitive dimension in order to provide a coherent definition of landmarks for navigation in different contexts. In the visual dimension, a landmark provides the ability to contrast in the surrounding (prominence) and be memorable by the observer or user. The structural dimension represents a feature of landmarks that can be traced back to the structural property of the specific environment: road intersection, city plaza, but also hypertext index page, etc. Finally the cognitive dimension of a landmark refers to a contrast to the surrounding location due to typical meaning that may be personal and even have cultural or historical importance (the house where a famous person was born is a landmark amongst otherwise identical looking houses). Vinson (1999) created a more specific classification of spatial landmarks, e.g., path, edges, nodes and Landmarks following Lynch (1960), who found that people's cognitive maps generally contain these five types of elements. Vinson also proposed 13 guidelines for the design and placement of landmarks in virtual environments. In terms of the prior discussed aspects of landmarks, the structural dimension represents the spatial location of the landmark, while the cognitive dimension falls into the categorical representation (Kosslyn, 1987). Sorrows *et al.* state that the role of landmarks in physical space has been minimised to the structural dimension and that the discussion of landmarks in electronic space is narrowed down to the visual dimension. The authors come up with design issues of electronic and physical spaces, i.e., the reduction of links on top-level web pages or the support of difference between distance and local landmarks. Landmarks also have a link to semiotic: understood as a sign as in the user interfaces, a landmark can be an index, symbol or icon. A lighthouse is an index landmark for the coastline, while an arrow can be a symbol landmark and a door can be an icon landmark for the escape route (Gaver, 1991).

The decision of what a landmark is or not can be assumed to be intuitive, so that the perception of the landmark can be direct without learning. As described in Gibson's ecological approach to visual perception (Gibson, 1979), it can be assumed that landmarks have an affordance² that enables intuitive interaction with them and the environment. If a ship is approaching the coast, the first object that the sailors can see and indicating the coastline is the lighthouse that towers above everything else on the coast due to physical constraints of the perspective. In other words, the perception of the landmark lighthouse is based on the relationship of the physical conditions of perception and the seeking of the sailor of conspicuous objects on the horizon to determine the coastline. This perspective opens up the role of landmarks for user actions. Landmarks with an affordance signalise not only orientation and navigation hints to the observer, but also enable the user to take possibilities of action into account and plan further actions to reach a goal in the environment beyond the task of way finding. The famous example of the affordance of a door handle (Norman, 1988; Gaver, 1991) underlines this aspect: the door handle can be understood as a landmark on the overall flat door. The handle on the door is a conspicuous isolated object or point and in line with the definition, given above. If it is well mounted, it can be easily reached by the human hands. The aspect of affordance of landmarks adds to them the possibility to interact, consequently more tasks become conceivable beyond orientation and navigation that rely on the interaction with landmarks.

INTERACTIVE LANDMARKS

Landmarks that fulfil the definition given above are classified as primary landmarks. Primary landmarks usually play a passive role and support the observer or user in task like orientation and navigation. In addition, landmarks that provide possibilities of action that go beyond those of the primary landmarks and extend the context of the task will be called *interactive landmarks* from now on. Interactive landmarks provide access to additional explicit knowledge that is not directly apparent from the spatial domain representation of the object. Interactive landmarks facilitate change of inherent states of the object to which the landmark is referring. An example for an interactive landmark from the everyday environment is a free standing mailbox that enables the post man not only to find the destination address by street number and name but also to interact with the post box and to deliver the letter to the receiver. In addition, some mail boxes have signs indicate to the owner and receiver of mail whether mail has arrived or not. An other example is the area of the upper right corner of a standard application window on a computer display can also be understood as an interactive landmark because it allows the user to manipulate the window state, to close, minimise or maximise, respectively.

² Affordance is a word coined by J.J. Gibson (1979) to emphasis his view of the relationship between humans and their environment. Affordances of objects are offers to the human perception, and the perceived stimulus contains already the information about the usage of objects.

Interactive landmarks can be identified when the task of the user goes beyond orientation and navigation and when these tasks are subtasks to be carried out to achieve a goal of a higher-level task. One can argue that this extension enabled by interactive landmarks leads to an information rich environment (Bowman, 1999), i.e., an integration of additional information and possibilities of action at an appropriate spatial location for the user to recognise, memorise and interact. This can be compared also to earlier work on augmented reality systems, where additional task-relevant graphical information is displayed as an overlay to the real environment (Feiner, 1993). The supplementing part of an interactive landmark is also similar to the functions that are triggered by elements of a graphical user interface.

Implicit interaction with landmarks

While touching or clicking on an object can be understood as an explicit interaction of selection, more implicit interaction with object requires a method to decide whether a selection, for instance, is intended or not. For this problem, Jacob (1990) coined the term *Midas Touch Problem* in the context of eye-tracking interaction systems. He named it after the Greek character Midas who asked for the gift to turn everything into gold only by touching. Jacob emphasises the challenge to distinguish between an intended command only by looking at an object and the gazing at an object by chance. Therefore a discrimination method must be applied to transform an implicit into an explicit interaction. To turn the fact that the user is looking at an object into an associate explicit command, a fixation recognition algorithm was applied to the eye-tracking data. For the implicit interaction with interactive landmarks in virtual environments, a similar approach is suggested to guarantee that only the landmark will become selected that the user is intending to select. We consider two interaction methods that basically refer to the spatial distances between a user's representation, like a cursor, 'finger', etc. and a landmark in the environment. The basic implicit interaction methods with landmarks are:

- approaching landmark and
- pointing at landmarks.

The approaching method determines the proximity between the user's representation and a landmark, by calculating the L2-Norm (Euclidean norm) of the distance vector. Pointing at a landmark requires the orientation of the 'finger' of the user in order to determine the landmark nearest to the extension of the user's orientation (laser pointing metaphor). The latter method enables the interaction system to trigger a command without the user being in the vicinity of the target landmark and thus allows the user to act from a remote position. In both cases the resulting distance parameter can be used to discriminate whether the landmark is approached or pointed at or not. Additionally, a so-called *dwelling time* can be applied to include temporal aspects of the interaction. The implicit interaction method can be then understood as a gesture, a non-verbal way of communication by intentional movements of the human body. Additional processing has to be applied to work out the intention from the data of the movements. At some point the user's representation via a cursor, 'finger', tool etc. can also be understood as a special landmark representing an input object that is highly variable and a strong input character.

Controlling landmarks

Beyond the described basic implicit interaction of selecting a landmark, controlling a landmark is a possible option after the selection, while the selection of a landmark is a precondition to control it. To differentiate between selecting and controlling, the user has to either implicitly or explicitly change the mode of interaction. A typical interaction to change the mode is grabbing a landmark, comparable to the click and hold of the mouse button. Selecting and grabbing then enable the user to control, e.g., move the landmark in space. If the landmark is embedded in an object, for instance, the landmark is part of an object; then the interaction with the landmark can be interpreted so that the grabbing and moving has also an effect on the object. Whereas this interaction relies mostly on the experience with objects as we know it from our everyday environment, more abstract methods are conceivable. These methods utilise the spatial relation between the user's representation and a selected landmark and map them to other parameters than spatial parameters. A selected and grabbed landmark can be changed in its inherent state continuously due to the relative position of the user's pointer. If a landmark is embedded, e.g., in a virtual light bulb and the user has selected and grabbed this landmark, the rotation of the pointer can be used to control the brightness of the lamp instead of manipulating the position and orientation of the bulb. In the latter case a spatial attribute is controlled while in the former case a more abstract attribute of the brightness of the object is under control. The described forms of interaction can be also found partially in the two-dimensional graphical user interface on standard computer systems. A handle bar of a window, for instance, understood as a landmark with a strong structural characteristic (interpreting Sorrows *et al.*, 1999) allows controlling the position of the window on the screen. While other continuous interaction referring to more abstract parameters, e.g., the volume of the sound output is mediated by the use of an instrument or tool metaphor, like sliders and scrollbar (Beaudouin-Lafon, 2000).

The user groups of landmarks

As described earlier, landmarks realise a combination of a categorical and a coordinative mental representations of the user. But the integrating character of landmarks is not limited to the user interaction. Landmarks can also play a role for the different groups that are involved in the process of planning and creation of an interactive system: software engineers, simulation developer/designers, lesson authors, tutors or trainers and trainees. Landmarks have an outstanding communicative role because they integrate the different perspectives of groups so that cooperation, coordination and discussion is highly facilitated and supported. For example, the software engineer sees the landmark as a data object that receives messages from an input device and the designer of the model for the simulation treats a landmark as a 'usual' object that can be placed in the spatial scene. For both groups, the landmark is an entity and the pivot for communication about shape or function of the landmarks. For the users or trainees and the authors or tutors, landmarks are points that represent certain knowledge to which both groups are referring. In addition, this knowledge is independent from the actual position and context of the landmarks. Therefore, the author or trainer can easily employ landmarks to define and plan what the user can or has to do without referring to the actual spatial position of the landmark.

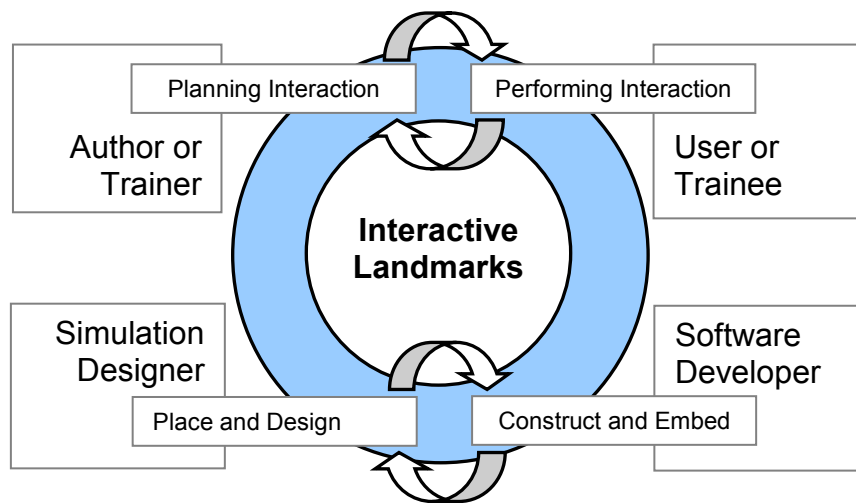


Figure 1: The concept of interactive landmarks integrates different perspectives of the groups that are involved in the process of planning and developing an interactive virtual environment. The interactive landmark is more than only an interface object for the user. It facilitates the communication amongst the groups that develop software, design the scene and plan the interaction. The diagram emphasises only the most prominent communication channels between the author and the user, and the designer and developer.

APPLYING INTERACTIVE LANDMARKS

We now discuss the important aspects of connecting a virtual environment with a knowledge-based reasoning system. We then will describe the application of interactive landmarks in a prototype of a surgical training environment. In a three dimensional virtual environment, objects are usually modelled with their physical properties. Then they are usually presented to the user via the visual, auditory and tactile modalities, with the goal to create the illusion of real objects. This approach enables users to interact in virtual environments with a simulation of objects in space, on the base of prior experience in a real space. For training applications, on which we are concentrating, it is agreed that virtual environments can help users build up their procedural knowledge before they get in contact with the real matter of the simulation for the first time (Cosman *et al.*, 2002). The fundamental difference in the structure of the representation of objects in a virtual environment and those in a knowledge base reasoning system are twofold: on the one hand, representations of objects in a knowledge base usually do not have spatial attributes, merely providing spatial information at less precision and accuracy than in a virtual environment. The relations of objects in a knowledge base are usually conceptual relations while those in a virtual environment are usually spatial relations. On the other hand, the knowledge retrieved from a knowledge base system refers to the so-called explicit or declarative knowledge, in contrast to tacit or procedural knowledge that is typically perceived in virtual environments and that often cannot be made explicit (Magill, 1993). From the point of human knowledge, these two types of knowledge are complementary, and the extension of one by the other could bring new benefits to a resulting interactive application (information-rich virtual environments: Bowman, 1999). An additional advantage for a training application is expected, because explicit advice, instruction or feedback that can then be provided to the users to guide them during learning in virtual environments and to support their decision-making.

A more technologically oriented difference between the two representations becomes apparent when an existing virtual environment is enhanced or supported by a knowledge base. A framework to develop applications for virtual environments is optimised for flexible and efficient rendering of a complex spatial model. Furthermore, instead of a procedural description, e.g., computer graphic function libraries, a declarative description in a so-called scene-graph representation is usually provided, using a high-level description language (VRML, Virtual Reality Modelling Language: Carey and Bell, 1997) with additional script nodes to incorporate user specific functionality. In contrast, a standard representation in a knowledge base uses a tree-like structure of nodes or terms that are in a conceptual relation to each other (e.g., KL-ONE: Brachman, 1985) as opposed to the spatial relation of the scene-graph. Knowledge can be acquired directly through the access of the definition of a node but also indirectly by deriving from the concept tree. Instead, the procedural knowledge is obtained from of a virtual environment while the user is interacting with the simulation. The corresponding component to the scene-graph of the virtual environment is a domain model of the objects in the knowledge base. Thus, the technological difference consists of the way applications are developed and the representation language they employ. This is driven by the unequal constraints of the quality of the individual performance system, e.g., high performance real-time processing for image rendering and data interoperability for flexible knowledge management.

The proposed concept of interactive landmarks helps bridge the gap in the structure of the representation of objects. As a pivot, the interactive landmarks can be easily represented in both the scene-graph and the knowledge base. As an add-on object to the simulation, an interactive landmark enables reasoning about the spatial situation in the virtual environment. From the viewpoint of the reasoning system, the environment consists only of collection of distinct dynamic points that are associated to instances in the knowledge base. This approach overcomes the structural barrier, and inter-process communication facilities can be used to guarantee consistency between these two representations. The fact that the complexity of a set of objects in a spatial scene is reduced to the most important points of the interaction makes the proposed concept of interactive landmarks flexible and computationally efficient.

Computer-supported surgical training

In the area of medical applications and especially in the surgical procedures, the technology of haptic virtual environments, computer assisted surgery and medical robots is expected to have a major impact on health care in the next decade (McCloy and Stone, 2001; Tendick *et al.*, 2002). 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 (Cosman *et al.*, 2002), 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) (McCloy and Stone, 2001), as well as skills, in case of suturing and burring (Moody *et al.*, 2003; Agus *et al.*, 2003). Although a skilfully performed operation is 75% decision making and only 25% dexterity (Darzi, 1999), 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 (Magill, 1993), and that is established only during practice, e.g., during sporting activities.

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 (Cosman *et al.*, 2002). The introduction of computer-based medical simulations for surgical training is founded on a haptic virtual environment that provides new direction for training motor skills. 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 initial prototype we developed addresses the issue of how skill acquisition can be made more effective in such an environment providing contextual information. It is expected that this information underpins decision making and leads to improvements in the learning curve.

Prototype setup

To experiment with and to demonstrate the concept of interactive landmarks, we built a vertical prototype that involves all important components of a virtual environment to train surgical skills. The core system architecture of the prototype is a haptic environment or Haptic Workbench (HWB, Stevenson, 1999) and the Virtual Document Planner (VPD, Moore and Paris, 1993) as a knowledge-based reasoning system. These systems are

heterogeneous in the structure of their representation of objects, as described above, and were chosen to be connected by the concept of interactive landmarks because of their reliable and acclaimed achievements.

The HWB consists of a PHANTOM haptic force feedback device in a hand-immersive display. This setup guarantees coherent spatial representation of haptic and visual objects (Poston and Serra, 1996; Stevenson *et al.*, 1999). The interactive scene is perceived fully three-dimensionally by using shutter glasses for stereo vision. An object oriented application programming interface (Reachin API, 2003) enables to development a simulation using a scene-graph description in VRML. This description combines haptic and visual aspects of a scene in a single scene-graph representation. In this scene-graph, the model is described as a hierarchical tree of objects. The approach enables the grouping and clustering of spatial related objects and changes within the tree structure in an efficient way. So-called transformation nodes allow designers of a scene to perform basic operations like translation and rotation of complete branches within the tree. The spatial properties of nodes in the tree are inherited or propagated to the next nodes in the scene-graph, and this procedure produces benefits in developing time of the scene as well as in real-time rendering performance (Reachin API, 2003). To experiment with our proposed concept of interactive landmarks, we used the Reachin API to build dedicated software objects that could be easily referred in the VRML scene-graph description. These objects continuously calculate their distance to the user's representation. This information is refined by a central object and passed as the state of interaction to the VDP using inter-process communication facilities.

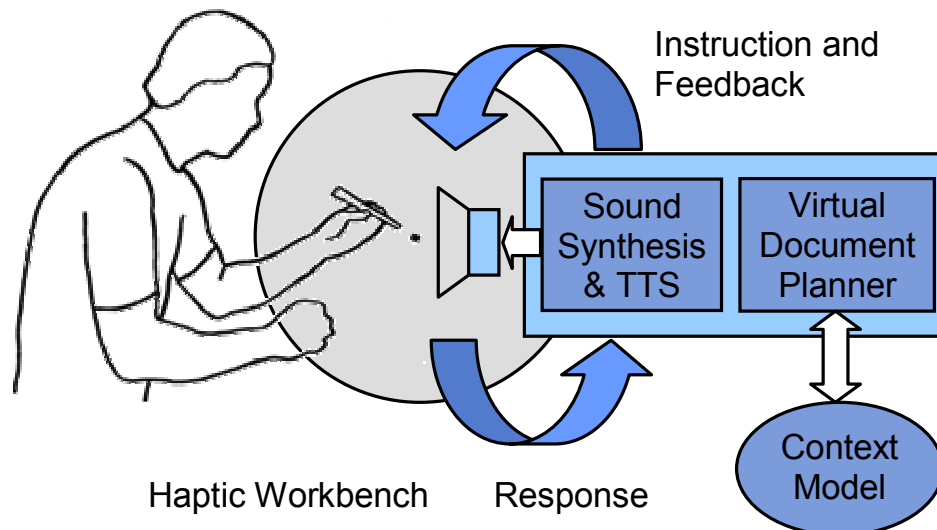


Figure 2: The prototype for surgical training with a Haptic Workbench and a Virtual Document Planner. The user or trainee interacts with the simulation in the virtual haptic environment. The interaction is guided by instructions and feedback that are generated by the Virtual Document Planner as the result of the actual user response and the current state of the context model. Aside the hand-immersive visual and haptic output, the instructions and feedback can also be delivered via text-to-speech (TTS) or non-speech audio.

To incorporate knowledge-based reasoning at that point, we pass the information about the interaction to a Virtual Document Planner (VDP) (Moore and Paris, 1993; Paris, 2002), which plans and presents tailored instructions and appropriate feedback to the trainees in a coherent manner. The VDP is extended for the described prototype to cope with the demand of the close interaction loop, i.e., rapid turn-taking of alternating instructions/feedback and user response (see figure 2). The reasoning is done on the basis of the user response and several knowledge bases that model, amongst others things, the domain, the user and the discourse history. The resulting system not merely provides visual and haptic feedback, but also gives instructions and feedback to the user taken the current context into account. These instructions and feedback can be presented with different feature and using different modalities, like text output or graphical output on the computer screen, text-to-speech (TTS) output using a Speech SDK (2003) or non-speech audio (Pure Data, 2003; Müller-Tomfelde, 2004). The resulting system can be classified as a multimodal system, following the definition from (Nigay and Coutaz, 1993) and hence offers the advantages for interaction described by Robertson (1998).

The interactive landmarks are realised as a VRML node for the Reachin environment. They have the appearance of a spherical object and can be controlled by the VDP to be visible or invisible, tangible or intangible depending on the actual context and demand in the application. Interactive landmarks also have a label that basically indicates the name of the landmark. The label is linked to a spherical representation of the landmark by a thin line that depicts the correspondence. The appearance of the landmark label can be also controlled by the VDP.

Interaction and training scenario

One of the basic tasks of a surgical trainee is to identify landmarks in the area of operation. During the identification process, the trainee builds up his or her mental model of the spatial relation of the anatomic structures in the present situation (Wilhelm *et al.*, 2002). The initial training scenario for the prototype is split into two lessons. In the first lesson, the trainees are asked to orient themselves in a scene that consists of some abstract objects that have one or more landmarks. While probing an object with a tool, the landmark is highlighted, and its label is provided, giving the name of the associated object. The orientation lesson lasts until all landmarks have been touched by the trainee. This allows the system to conclude that the trainee knows about the associated objects. In a second lesson, this recently acquired knowledge is tested. The trainee is asked to identify each anatomical landmark in a question and answer manner. In case of multiple wrong identifications/answers, the reasoning system skips the current landmark and continues testing knowledge about other landmarks. In contrast to an exercise supervised by an experienced surgeon who can observe and check the trainee's actions, the method in computer-based training changes: the trainee has to interact explicitly with landmarks to enable the system to reason about the interaction and to provide appropriate feedback.

To enable reasoning about selected landmarks or, about the correct or false identification of a landmark during the lessons, the distances between the landmarks and the tool tip are calculated and updated continuously. This allows the determination of the closest of all landmarks to the current position of the tool tip (landmarks are indicated with L and distances with d , see figure 3a). In order not to process absolute distances, the measurements are set into relation to the furthest landmark and therefore correspond to values of relative proximity. In addition, the use of the relative proximity demands more precise interaction from the trainee in an area with dense positioned landmarks than in areas where only a few landmarks have been placed. While the user holds the tool toward a landmark and remains close to it, the implicit interaction is transformed into an explicit interaction event. 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 (landmark L_1 has the minimal distance d_1 , see in figure 3a). Finally, the resulting interaction event is passed to be examined and further processed by the reasoning system (VDP) to generate the appropriate feedback or a next instruction.

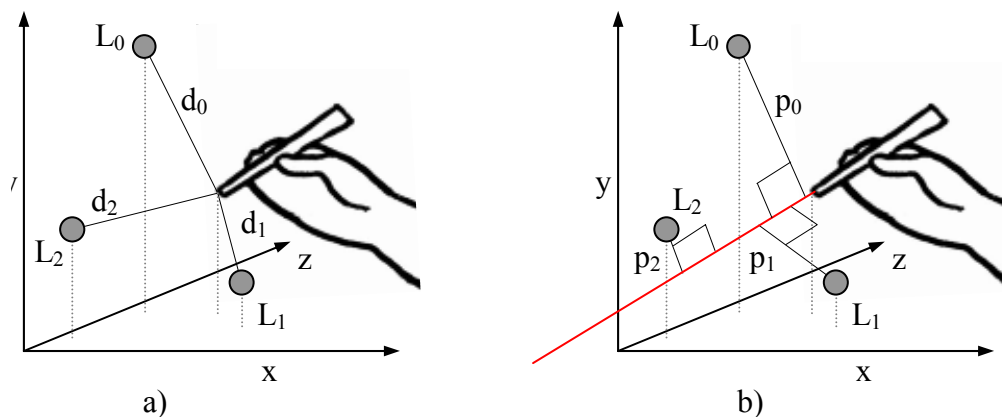


Figure 3. a) Approaching interactive landmarks with a tool: the distances d between the tool tip and landmarks L in the three dimensional space of a haptic environment. 3.b) Pointing with the tool at a landmark: the distances p of the landmarks according to the direction, in which the tool is pointing (red line). The grey, dotted and vertical lines represent the heights of the landmarks over the xz -plane.

Pointing at a landmark is the alternative way to select a landmark and thus refer to a location in space. 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 (red line in figure 3b) has to be close to the landmark to select it; in brief, the user must point at the landmark (laser pointer metaphor). The distance vectors are orthogonal to the stylus extension and lead to the distances p between the extended stylus and the landmarks (landmark L_2 has the minimal distance p_2 , see figure 3b). The output values are set into relation to the furthest of all landmarks, and, as a result, a proximity factor in percentage and the unique id of the landmark is passed to the VDP as a user response. For both interaction methods, different thresholds in the relative proximity and dwell time are applied to guarantee maximum interaction possibilities. The authors are currently working on the support of more continuous interaction to enable parameter manipulation, as described above, and to assess the performance in metrics of movements, e.g., 'economy' and 'purposefulness' or 'fluidity', like proposed in (Satava, 2003).

CONCLUSION AND FUTURE WORK

In this paper we presented the concept of interactive landmarks and discussed a vertical prototype to demonstrate and proof the concept in the context of training surgical skills. The prototype consists of a haptic virtual environment called Haptic Workbench and a knowledge-based reasoning system that are coupled together by interactive landmarks to bridge the gap in the structure of the representation of objects between both systems. The experience with the prototype is encouraging because of the simplicity and transparency of the usage of interactive landmarks by different groups that are involved in the realisation and usage of the system. It became clear that the information about the user interactions with these landmarks is extensive enough to reason about and to probably assess the user's interaction. The core constraint is the consistency between a landmark's reference in the scene-graph and the knowledge base. We modified the Virtual Document Planner (VDP) to be able to reason about the user's actions and to provide individual tailored information and add new benefits to the training session. The VDP will be employed in future also to plan, organise and maintain the curriculum of trainees. More sophisticated and complex interactions are conceivable anticipating the use of interactive landmarks in more complex training scenarios on the Haptic Workbench. Formal evaluations have to be planned, and experiments conducted to verify the first promising experience with the prototype and to demonstrate the effectiveness of incorporating information delivery in interactive haptic virtual environments.

REFERENCE

- Agus, M., Giachetti, A., Gobbetti, E., Zanetti G., and Zorocolo A. (2003) Real-Time Haptic and Visual Simulation of Bone Dissection, *Presence* 12(1): 110-122.
- Aylett, R. and Cavazza, M. (2001) Intelligent Virtual Environments: a State-of-the-Art Report, In *Eurographics 2001*, STAR Reports Volume.
- Beaudouin-Lafon, M. (2000) Instrumental Interaction: An Interaction Model for Designing Post-WIMP User Interfaces, In *Proceedings of the CHI2000 Conference*, ACM, CHI Letters 2(1): 446-453.
- Bowman, D., Wineman, J., Hodges, L., and Allison, D. (1999) The Educational Value of an Information-Rich Virtual Environment, *Presence* 8(3): 317-331.
- Brachman, R. and Schmolze, J. (1985) An Overview of the KL-ONE Knowledge Representation system, *Cognitive Science* 9: 171-216.
- Carey, R. and Bell, G. (1997) *The Annotated VRML 2.0 Reference Manual*, Addison-Wesley Developers Press.
- Cosman, P.H., Cregan, P.C., Martin, C.J., and Cartmill, J.A. (2002) Virtual Reality Simulators: Current Status in Acquisition and Assessment of Surgical Skills, *ANZ Journal of Surgery* 72: 30-34.
- Darzi A., Smith, S., and Taffinder, N. (1999) Assessing operative skill, *British Medical Journal* 318: 888-889.
- Feiner, S., MacIntyre, B., and Seligmann, D. (1993) Knowledge-Based Augmented Reality, *Communications of the ACM* 36(7): 53-62.
- Gaver, W.W. (1991) Technology Affordances, In *Proceedings of the CHI'91 Conference*, ACM, pp. 79-84.
- Gibson, J.J. (1979) *The ecological approach to visual perception*, Boston: Houghton Mifflin.
- Hyperdictionary (2004) URL <http://www.hyperdictionary.com>, Accessed 4 March 2004.
- Jacob, R.J. (1990) What You Look at is What You Get: Eye Movement-Based Interaction Techniques, In *Proceedings of the CHI'90 Conference*, ACM, pp. 11-18.
- Jung, B., Latoschik, M., and Wachsmuth, I. (1989) Knowledge-Based Assembly Simulation for Virtual Prototype Modeling, In *Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society, Vol. 4*, IEEE, pp. 2152-2157.
- Kosslyn, S.M. (1987) Seeing and Imaging in the Cerebral Hemisphere: A Computational Approach, *Psychology Review* 94: 148-175.
- Luck, M. and Aylett, R. (2000) Applying Artificial Intelligence to Virtual Reality: Intelligent Virtual Environment, *Applied Artificial Intelligence* 14(1): 3-23.
- MacCloy, R. and Stone R. (2001) Virtual Reality in Surgery and Medicine, *British Medical Journal*, 323: 912-915.
- Magill, R.A. (1993) *Motor Learning: Concepts and Applications*, 4th ed. Madison, Brown and Benchmark.
- Merriam-Webster Online (2004) URL <http://www.merriam-webster.com/>, Accessed 4 March 2004.

- Moody, L., Baber, C., Arantis, T.N., and Elliott, M. (2003) Objective Metrics for the Evaluation of Simple Surgical Skills in Real and Virtual Domains, *Presence* 12(2): 207-221.
- Moore, J. and Paris, C. (1993) Planning Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information, In *Computational Linguistics* 19(4): 651-694.
- Müller-Tomfelde, C. (2004) Interaction Sound Feedback in a Haptic Virtual Environment to Improve Motor Skill Acquisition, In *Proceeding of the 10th International Conference on Auditory Display (ICAD04)*.
- Nigay, L. and Coutaz, J. (1993) A Design Space for Multimodal Systems Concurrent Processing and Data Fusion, In *Proceedings of the INTERCHI'93 Conference*, ACM, Amsterdam, pp. 172- 178.
- Norman, D.A. (1988) *The psychology of everyday things*, New York: Basic Books.
- Oxford English Dictionary (2004) URL <http://dictionary.oed.com/>, Accessed 4 March 2004.
- Paris, C. (2002) Information Delivery for Tourism, In *IEEE Intelligent Systems*, Intelligent Systems for Tourism: Trends and Controversies, S. Staab and H. Werthner (Eds), pp. 53–66.
- Phantom Sensable (2003) URL <http://www.sensable.com/>, Accessed 1 Sept 2003.
- Poston, T. and Serra, L. (1996) Dextrous virtual work, *Communications of the ACM*, 39(5): 37-45.
- Pure Data (2003) URL <http://www.pure-data.org/>, Accessed 23 July 2003.
- Reachin API Programmers Guide (2003) URL <http://www.reachin.se/>, Accessed 8 Aug 2003.
- Robertson, G.G. (1998) URL <http://research.microsoft.com/~ggr/pui98.htm>, Accessed 10 April 2002.
- Satava, R.M., Cuschieri, A., and Hamdorf, J. (2003) Metrics for objective assessment, *Surgical Endoscopy* 17: 220–226.
- Sorrows, M.E. and Hirtle, S.C. (1999) The Nature of Landmarks for Real and Electronic Spaces, In *Spatial Information Theory, Cognitive and Computational Foundations of Geographic Information Science (COSIT '99)*, C. Freksa and D.M. Mark (eds.), Springer-Verlag, Vol. 1661: 37-50.
- Speech SDK (2003) URL <http://www.microsoft.com/speech/download/sdk51/>, Microsoft, Accessed 2 Nov. 2003.
- Stevenson, D., Smith, K., Veldkamp, P., McLaughlin J., Gunn, C., and Dixon, M. (1999) Haptic Workbench: A Multisensory Virtual Environment, *The Engineering Reality of Virtual Reality, Electronic Imaging '99*, San Jose.
- Tendick, F., Downes, M., Gokekin, T., Cavusoglu, M., Feygin, D., Wu, X., Eyal, R., Hegarty M., and Way, L. (2002) A Virtual Environment Testbed for Training Laparoscopic Surgical Skills, *Presence* 9(3): 236-255.
- Vinson, N.G. (1999) Design Guidelines for Landmarks to Support Navigation in Virtual Environments, In *Proceedings of the CHI'99 Conference*, ACM, pp. 278-285.
- Ware, C. (2000) *Information Visualization: Perception for Design*, Morgan Kaufman, New York.
- West, A. and Hubbard, R. (1998) Research Challenges for Systems Supporting Collaborative Virtual Environments, In *Proceedings Collaborative Virtual Environments CVE'98*, Manchester, p. 11.
- Wilhelm, D., Ogan, K., Roehrborn, C., Cadeddu J., and Pearle, M. (2002) Assessment of Basic Endoscopic Performance Using a Virtual Reality Simulator, *American College of Surgeons*, 195(5): 675-681.

ACKNOWLEDGEMENTS

The authors would like to thank Clara Teoh for her valuable contribution on extending the VDP. Thanks also to Chris Gunn and Matthew Hutchins for their valuable help on the Haptic Workbench. This work is being supported by a CSIRO Corporate Fellowship and the CSIRO ICT Centre.

COPYRIGHT

C. Müller-Tomfelde, C. Paris and D. Stevenson © 2004. The authors assign to OZCHI and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to OZCHI to publish this document in full in the Conference Papers and Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.