

# Interaction Techniques for Navigation through and Manipulation of 2D and 3D Data

Dzmitry Aliakseyeu, Sriram Subramanian, Jean-Bernard Martens, Matthias Rauterberg

User Centered Engineering group, Faculty of Technology Management, Eindhoven University of Technology, The Netherlands

---

## Abstract

*In this article we present a working prototype incorporating some new interaction techniques for the navigation through and the manipulation of both 3D and 2D data. The prototype aims at professional applications like architectural design, surgical planning and geological exploration. Its design was influenced by the analysis of user requirements and by the requirement for a natural interface. The prototype permits the user to navigate through 3D and 2D data in order to explore the internal structure. 3D navigation is accomplished by means of a natural selection of dynamically updated cross-sections. The user can navigate through these selected cross-sections of 2D data and can perform manipulations with them. As a proof of concept we focus on two disciplines, viz. architectural design and surgical planning. Our usability evaluation of the interaction styles demonstrates that they are perceived as useful and enjoyable additions to existing techniques in these disciplines.*

## Keywords:

*Augmented Reality, 3D interface, 2D interaction technique, Natural User Interface, volume data Navigation, and manipulation.*

---

## 1. Introduction

In order to better understand the user-interface requirements imposed by different application domains, one often starts by identifying important and frequently occurring tasks. Tasks are defined here as coordinated or logical sequences of actions, and may be shared by different applications. Tasks can hence be broken down into elementary actions. Wuthrich [1] identifies three kinds of elementary actions: selection/grabbing, positioning with  $n$  Degrees of Freedom (DOF), and deforming. While it is possible to change the shape of an object by selecting a point within the object and repositioning that point, Wuthrich regards this to be a composite of two atomic actions (selecting and positioning) rather than the atomic action of deforming. In his view, deforming takes place when an input device (like the cube-based shape deformation interface [2]) enables direct manipulation of the object shape. A similar argument can be used to argue the need for adding a fourth kind of atomic action to the list, i.e., sketching/writing. While entering text can be done through a keyboard, this form of entering relies on the atomic action of selection. However, a more natural and direct way of creating text and symbols is by writing and sketching with a pen.

The important tasks in the application domains that we consider here are the navigation through and the manipulation of 3D and 2D data. The dimensionality of the data is defined here as the dimension of the smallest

space needed to completely represent it. Thus, 3D data needs at least a 3D space to fully represent it, while 2D data can be represented in both 2D and 3D space. In this paper we address some user-interface requirements for the four possible combinations of tasks and data dimensionality, i.e.,

- 2D Navigation (T1): This involves establishing the location and identity of structures in 2D space.
- 3D Navigation (T2): This involves establishing the location and identity of structures in 3D space.
- 2D Manipulation (T3): This involves creating, modifying and positioning items in 2D space.
- 3D Manipulation (T4): This involves creating, modifying and positioning items in 3D space.

The following are some typical applications where the above-mentioned tasks are encountered:

- Medical diagnostics and surgical planning [3] often require anatomical images of the human body. Examples of volumetric (or scientific) data in this context are Computer Tomography (CT) data, which are derived by a computer algorithm

from multiple 2-D images taken at many different orientations. CT exists for many different modalities: X-ray, Magnetic Resonance, Positron Emission, etc.

- In geology, seismic recordings over a complete area may be collected to obtain information on the subsurface structure. This data must subsequently be analyzed, for example to better understand the shape and position of different geological layers, or to plan interesting trajectories for test drills (in oil or gas exploration or mining). Often, the computer can assist by making segmentations of the data. However, the results of such segmentation are seldom completely trustworthy, and they have to be checked by a human operator, and corrected whenever necessary [4], [5].

Data is only as useful as the users' ability to navigate and interact with it. Thus, the search for natural ways of performing the above interaction tasks is considered very relevant. In this article we will present some new interaction techniques that have been implemented on a working prototype. They aim to improve the efficiency, effectiveness and pleasure of use while performing the above-mentioned tasks. The proposed prototype allows the user to work simultaneously with both 3D and 2D data. The next section discusses the design requirements that have served as boundary conditions in our search for suitable interaction techniques.

## 2. Design Requirement

Ideally, interactions with data should look and feel to the user like s/he is directly navigating and/or manipulating the data (at a certain level of abstraction), without the need to pay much attention to the interaction itself. Thus, at a phenomenological level, the interaction should be as *transparent* as possible [6]. This leads to an important design requirement of naturalness of the interaction technique, which can be derived from observations of how we interact with the real world. By the time we reach adulthood, we have perfected many manipulation and navigation skills to the point where we can perform many tasks without conscious attention. The interaction device becomes invisible. It is this level of naturalness and transparency that one aims to achieve in interacting with computers. When done well, the interaction shouldn't feel like a human-computer interaction anymore, but rather like a human-product or human-task interaction. The focus should be on interacting *through* a computer instead of interacting *with* a computer.

As a general requirement for naturalness, the atomic actions required to perform the interaction tasks of interest should match the atomic actions provided by the interaction device. When the task requires more atomic actions than the atomic actions provided by the interac-

tion device, the user interface will need to support a complex dialogue of composed interactions, making the interaction *unnatural*. The relative positioning in 2D of a mouse is, for example, not adequate for performing an absolute positioning task in 3D. Mouse actions have to be modified by keyboard control keys in order to accomplish this. The converse, where the interaction device has more DOF than required by the task, can also be inefficient and confusing, since in such a case the interaction device is not physically constrained to the same set of atomic actions as the task [7].

Two increasingly popular ways of designing more natural interfaces are using Virtual Reality (VR) and Augmented Reality (AR). A VR system positions the user in a graphical representation of a computer-generated model, with the intention of completely immersing the user in this virtual world. VR systems require the user to wear devices like head-mounted displays, head trackers, data gloves, etc. which intrude on the users personal space (as defined in [26]). These devices often seriously limit the ability of the user to interact in a social environment. Concerns have also been raised in both scientific and popular journals about possible harmful effects of such devices [10], [27]. In an AR system, the real world of the user is being augmented with graphical or virtual information in order to enhance the users real world. The realization that people are often communicating with their environment while performing a task is an important design aspect that is respected more in AR than in VR. Another important design principle of AR systems is that they try to make optimal use of the well-developed human skills of (two-handed) interaction with real objects. One of the key technological problems in VR, i.e., providing realistic haptic feedback, is circumvented in this way.

In order to incorporate our idea of naturalness, we adopt the following five design guidelines [8] for interaction devices:

- Two-handed interactions should be preferred over one-handed interaction. The users know where their hands are relative to one another and this can help to develop interactions which are less demanding of visual attention [1]. (DG1)
- Visual feedback is important for creating a feeling of presence or spatial awareness [8]. (DG2)
- The action and the perception space should coincide [9]. (DG3)
- Minimal use of intrusive devices, such as head-mounted displays, should be preferred [10]. (DG4)

- Wireless props are preferred over wired ones [11]. A prop is a physical device, which allows the user to interact with the data. An interface includes one or more props. (DG5)

In an effort to get a better grasp on the navigation and manipulation tasks of interest to us, we have adopted the following operational definitions:

- 2D Navigation (T1): This involves finding items in a 2D dataset (such as a collection of pictures). Finding is typically accomplished by positioning the viewpoint sufficiently closer to the item such that the resolution and the field of view (FOV) of the displayed image are such that identification becomes possible. 2D navigation is hence required in instances where not all items can be rendered simultaneously with the necessary resolution. Depending on whether the items are arranged in a 1D or 2D array, positioning requires either 1 or 2 DOF.
- 3D Navigation (T2): This is in principle very similar to 2D navigation. Practical problems arise in current interfaces because the positioning of the viewpoint has to be done in 3D, and because more parameters are involved in creating the displayed image. Especially in case of scientific data, such as the medical CT data or the geological seismic data mentioned before, deciding on an adequate display mechanism is a problem in itself. One possible approach is to map the scientific data to a surface model. The required segmentation not only reduces the available information, but may also be incorrect and lead to a biased or erroneous interpretation. The alternative that we propose here is to navigate by means of dynamically updated cross-sections. This implies that time makes up for the missing dimension. We prefer this approach because it preserves all information in the data. A rigid rectangular frame that can be positioned freely in space will be used to create such intersections. This frame will then have 6 DOF.
- 2D Manipulation (T3): This task demands the atomic actions of positioning (3 DOF: 2 translations and 1 rotation), selecting and sketching/writing in 2D. It is familiar from existing desktop interfaces.
- 3D Manipulation (T4): Users can perform precision manipulation tasks better in 2D than in 3D space. Based on this assumption it may be worthwhile to consider performing 3D manipulations through a series of 2D manipulations. For

example, a path in 3D could be created by letting the computer interconnect user-specified points in a number of intersection images.

In the next section we describe some of the previous research in developing interaction devices for tasks T1, T2, T3 and T4.

### 3. STATE-OF-THE-ART

Table 1 compares some of the different interaction devices (both commercial and research tools) found in the literature, with respect to: (1) the atomic actions (namely, selecting, sketching/writing, deforming and positioning), (2) the tasks (T1, T2, T3 and T4) they support and (3) the interaction design guidelines (DG1, DG2, DG3, DG4, and DG5) they meet.

*SILK.* The electronic sketching system SILK [12] is an informal sketching tool that combines paper-based sketching with computer support. With SILK, designers can quickly sketch an interface using an electronic pad and stylus. SILK recognizes widgets and other interface elements as the designer draws them.

*BUILD-IT.* BUILD-IT [13] is a computer-vision based multi-user planning tool. Two projectors are used to create a horizontal output medium on a table and a vertical one on a screen, respectively. Grouped around a table and employing tangible physical bricks, the users can select and manipulate virtual models within the scene that they are planning.

*Passive Interface Props.* Hinckley et al [14] proposed a two-handed interaction device for neurosurgical visualization using several props. The surgeon is provided with: (1) a head prop, in the shape of a small rubber sphere, for manipulating the view point (like zooming, rotation etc), (2) a cutting plane prop, used in combination with the head prop, for specifying the position and orientation of an arbitrary slice through the patients anatomy, and (3) a trajectory selection prop that allows to specify 3D vectors and points.

*Cubic Mouse.* The Cubic Mouse [15] consists of a cube-shaped case with three rods and control buttons. The rods are approximately orthogonal and correspond to the axes of a co-ordinate system. The device has an embedded tracker, which enables the users to control the viewpoint. Moving the rods can create slices through the volumetric data along canonical axes.

If we compare the atomic actions and the tasks supported by the interaction devices, it is clear that there is no interaction device that supports all the tasks in a natural way.

**Table 1.** Comparison of four different Interaction devices on the atomic actions they perform, the tasks they support and the guidelines they meet.

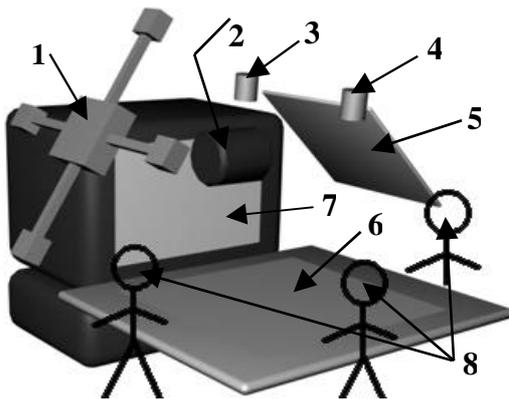
Interaction Device	Atomic actions supported	Task Supported				Interaction Device Requirements supported				
		T1	T2	T3	T4	DG1	DG2	DG3	DG4	DG5
SILK [12]	SW, S, P (2 DOF: 2 Translation)	Y	N	Y	N	N	Y	N / Y	Y	Y
BUILD-IT [13]	S, P(3DOF: 2 Translation, 1 Rotation)	Y	Y	½ Y	½ Y	Y	Y	Y	Y	Y
Passive Interface Prop [14]	S, P (4DOF:3 Translation, 1 Rotation)	N	Y	½ Y	Y	Y	Y	N	Y	N
Cubic Mouse [15]	S, P(6DOF: 3 Translation, 3 Rotation)	N	Y	N	N	Y	Y	N	Y	N

N = No; Y = Yes; ½Y = Supports task partially.  
S- Selecting, P- Positioning, SW- Sketching/ Writing.

We use the hardware setup (the Visual Interaction Platform) described in the next section in order to implement an interaction technique for performing the tasks T1, T2, T3 and T4 while trying to comply with the design requirements mentioned in the earlier section.

#### 4. Hardware Setup

The Visual Interaction Platform (VIP) [16] is an augmented reality platform that enables different natural interaction techniques such as writing, sketching, manipulating and navigating in 2D and 3D data. It is an extension of the BUILD-IT system [13].



**Figure1:** Visual Interaction Platform. 1-IS600 tracking device, 2-LCD Projector, 3-Camera, 4-Infrared light source, 5-Mirror, 6-Wacom Tablet (action-perception space), 7-Back projector (Communication Space), 8-Users

The VIP uses a LCD projector to create a large computer workspace on a horizontal surface. The horizontal surface contains a Wacom UltraPad A2® (<http://www.wacom.com>) digitizer tablet. The size of the projected image can be varied by changing the distance between the projector and the mirror (see fig. 1).

UltraPad can accurately record digital pen movements and allows the user to perform precise actions in the horizontal workspace, thus making it a combined action and perception space (henceforth action-perception space). Precise actions are necessary for handwriting, drawing and sketching. In addition, the platform is equipped with an ultrasonic wireless position (6 DOF: 3 translation and 3 rotation) tracking device of InterSense (IS-600 Mark 2, <http://www.isense.com>), which extends the 2D action-perception space to the 3D interaction space above the table.

The system also contains an infrared light source and a camera. This allows the system to track physical objects, such as small bricks, that are coated with infrared-reflecting material. The camera and the infrared light source are located above the table as shown in figure 1.

Apart from the horizontal action-perception space, the VIP has a second, vertically oriented, workspace (communication space). The communication space is equipped with a stereo back projection system, which allows 3D visualization. This optional second workspace is usually used to supply the user with more extensive visual feedback for increased spatial awareness, or to

communicate with remote participants. Back projection circumvents the problem of the user standing/ sitting in the way of the projector and the vertical screen

The platform is equipped with two computers. An Intergraph® workstation is used for visualization, while a Dell OptiPlex GX1 is used for gathering the data from IS600 and camera. Both run Microsoft windows operating system.

We use Microsoft Visual C++ as the software development tool. For visualization and image analysis we use OpenGL (<http://www.opengl.org>), VTK (<http://www.kitware.com>) and Intel OpenCV libraries (<http://www.intel.com>).

The hardware setup of the VIP allows realizing several props that support different atomic actions.

## 5. Props

The VIP has four basic props namely, the Brick Element (BEL), the digital pen in combination with the digitizer tablet, the Enhanced Paper Prop (EPP), and the Rigid Intersection Selection Prop (RISP).

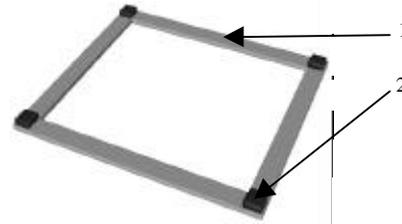
The BELS are physical objects coated with infrared reflecting material that can be tracked by the system. BELs are used for selecting and positioning objects on the action-perception space. An object is selected by placing a BEL on top of it. The object can be moved to the required position by dragging the BEL there. The BEL supports the atomic actions of selecting and positioning in 2D (3 DOF: 2 translation and 1 rotation).

The Wacom UltraPad consists of a tablet and a digital pen. The digital pen can be used like a conventional mouse or as a writing tool. In our current implementation the digital pen is used to make menu selections and 2D data (usually images) annotations. Thus, the digital pen supports the atomic actions of selecting and sketching/writing in 2D.

The EPP is a piece of real paper, which is placed on the action-perception space, The EPP contains infrared reflecting tags that allow the system to track it. Sketches made on the EPP, using the digital pen, can be traced into the computer. The visual feedback has two different modes. The sketch can either be projected onto the paper or can be created by a digital pen with an ink cartridge. The EPP supports the atomic action of positioning in 2D (3 DOF: 2 translation and 1 rotation).

The RISP (see Figure 2) is a rigid rectangular plastic frame (see 1 in Figure 2), which acts as a cross-section picker. The RISP is tracked using the IS600 tracking device. If the intended application involves interacting with 3D, then the user can use his/her non-dominant hand to move the (light-weight) RISP above the table,

thereby making the desired cross-section of the 3D data. Covering one of the IS600 trackers for at least ½ sec selects the current cross-section. The RISP hence supports the atomic action of positioning in 3D (6 DOF: 3 translation and 3 rotation) and selecting.

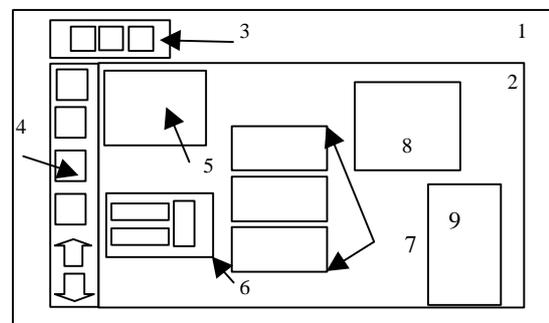


**Figure 2:** The Rigid Intersection Selection Prop (RISP). 1-Plastic Frame, 2 - IS600 trackers

## 6. Interaction technique

The layout of the action-perception space is dependent on whether the user is left-handed or right-handed. The layout described below is meant for a right handed person.

The layout (see Figure 3, see also Figure 4 and 5) consists of a linearly structured image database browser (see 3,4,5 in Figure 3), virtual paper (7), a floating toolbar (6), EPP (9) and a 3to2D window (8) in the action-perception space. The communication space is currently used to provide a surface rendering of the 3D model. This surface model is used to visualize the position of the RISP with respect to the 3D data.



**Figure 3:** The action-perception space layout. 1-Entire action-perception space, 2-Sketchable area (UltraPad), 3,4,5-Image database browser, 6-Floating toolbar, 7-Virtual Papers, 8 - 3to2D window, 9 - Enhanced Paper Prop.

The image database browser is located in the left margin of the action-perception space and consists of several images that the user has previously acquired (either through this interaction device or through other means). The browser contains an image database selector (3), image thumbnails from the selected database with two buttons for browsing (4), and a preview win-

dow (5), which shows a high-resolution version of the selected image.

A BEL positioned on the scrolling arrows and/or the thumbnails can be used to scroll through and select from the database. Here, the BEL is constrained to the atomic action of positioning with 1 DOF (1 translation). The second DOF is disabled once the BEL moves over the image data browser.

A copy of any image in the browser can be selected and dragged out of the browser into the working area, in order to become what we call a Virtual Paper (VP) [17]. Using the BEL, the user can orient and position the VP in the working area. The digital pen can be used to sketch and write on the VP. The VP has several properties like transparency level, size, sketching ink color and pen thickness. The action-perception space supports up to 10 VPs at any time.

To adjust the above-mentioned VP properties the user can use the movable semitransparent toolbar called the floating toolbar. The user can select and drag the menu toolbar by using a BEL. When any one of the four corners of the floating toolbar is moved within the boundaries of the VP, the floating toolbar gets attached to that VP, thereby enabling the user to change its properties. The properties can be changed using the digital pen.

The EPP is a combination of the real paper and the VP. The EPP can be positioned anywhere within the working area by dragging. Depending on the mode of use of the EPP, certain properties of the VP are inherited by the EPP as well.

The 3to2D window (see 4 in Figure 5) allows the user to interact with 3D data. The window is a preview window that allows the user to view different cross-sections of a 3D dataset. The absolute position of the 3D data is on the table above the 3to2D window. The user can navigate through the 3D data by moving the RISP above the table. If the user finds a certain cross-section interesting, s/he can capture the image by covering one of the IS600 trackers on the RISP. The selected cross-section image is then added to the image data browser. The floating toolbar can be positioned over the 3to2D window as well. In this case, sliders on the floating toolbar can be used to move the intersection plane along the canonical axes of the data in incremental steps.

The user is always provided with view of a surface model of the 3D data in the communication space. When the user lifts the RISP to interact with or navigate through the 3D data then the location of the RISP with respect to the 3D data is also shown.

Auditory feedback is provided to the user in situations where visual feedback is not clear or not possible. For example, when the user captures a cross-section, a sound is played to confirm to the user that his/her action has been successfully performed.

The platform supports collaborative work in a limited way. Several users can observe the interactions and use different BELs. Although different users may have separate digital pen, only one can be operated at a time. Currently there is only one RISP.

Note that the VIP does indeed conform to the guidelines for natural interfaces that were proposed earlier:

- two-handed interaction is possible (DG1)
- the platform provides extensive visual feedback through the communication space (DG2)
- the action and perception spaces coincide in the action-perception space (DG3)
- the users do not have to wear intrusive devices like head-mounted displays (DG4)
- there are no messy wires to hinder user movements (DG5)
- multiple users can collectively interact at the same time, using separate props, thereby promoting group work.

Additionally the use of props is easy to learn [17][18], since they rely on well-developed human skills for manipulating physical objects.

## 7. Application

In order to evaluate the usability of the interaction techniques, we considered applications in two different disciplines, namely architectural design and medical planning. The main reason for this selection has been the access to data from these domains and access to experts for evaluations. There is also a substantial difference in working practice of professionals in these domains. The emphasis in the medical domain is on the analysis of gathered data, while in the architectural domain it is on the creation of new data through analyzing existing data.

### 7.1. Architectural Design

#### 7.1.1. Problem Description

Pen and paper are traditional companions in many creative activities: "From mechanical engineering to the graphic art, designers comprehensively reject the use of computers in the early, conceptual/creative phases of designing... designers prepare to use paper and pencil"

[19]. Early architectural design is a typical example where traditional design tools such as sketching on paper still dominate over computer-assisted tools. In addition to the task and system requirements mentioned in previous sections, we must consider the following discipline specific requirements for computer assisted architectural design tools. They are based on our user studies (through questionnaires and brain storm session) and a literature review (see [19], [20], [21]).

The architect must be able to sketch, write, model or search for images, or any other information, in an easy, *intuitive* way. Intuitiveness in using the program is important in order to enable the architect to focus on the design problem and not on how to use the program. Intuitiveness can be provided to the architect by a tool that is natural to him [20].

Different kinds of *information* are useful to the architect. This can be information concerning the assignment, the situation of the building envisioned, the architect's vision of architecture, etc. It is hence important that a tool can handle *more representations* at the same time, meaning sketches, images, text and 3D models. Sketches are very powerful representations. They are vague on purpose: sketches can be often interpreted in many ways. An image can inspire the architect: sometimes it displays materials or atmosphere that the architect likes, or it displays a certain composition of proportions, which appeal to him/her in an esthetical way. Text is used as annotation, explanation or as keywords; sometimes it is used in schemes or in a description of the conceptual ideas [20].

### 7.1.2. Prototype Implementation

We created a prototype of the system for architectural design, which attempts to meet the user and system requirements mentioned above. Our prototype has the ability to integrate and inter-relate all textual and graphical data.

In the action-perception space (see Figure 4), the VP contains: previously made or scanned sketches; inspirational photos, for example, photos of work of other architects; images from previous projects; images with different building elements; photos of the site and surroundings; and/or other relevant material. These materials can be retrieved from the image database browser.

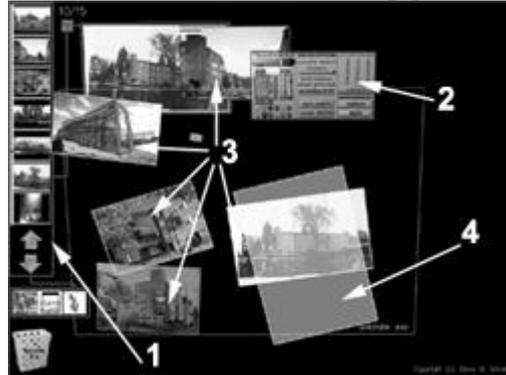
By moving the BEL with the non-dominant hand, the designer can change the position and orientation of a virtual paper. A pen in the dominant hand can be used to annotate or sketch on the virtual paper. This annotations or sketches can be saved or printed for future use.

Architect do not use opaque media very often - they use translucent media such as tracing paper and polyester film to copy, elaborate, annotate and evolve draw-

ings. Hence the system offers transparency to assist in re-drawing and over-drawing which is more important than editing and tweaking [21].

The EPP is also part of the prototype for the architectural design. A digital copy of the EPP content (sketches, annotations, etc.) can be printed or saved and used as virtual paper. The system can also add visual information that supports the design, since any virtual paper can be placed on top of the EPP.

The combination of real and virtual paper can help



**Figure 4** Action-Perception space for Architectural Design. 1-Image database browser, 2-Floating toolbar, 3-Virtual Papers, 4 – Enhanced Paper Prop.

in managing, storing and annotating images; managing, creating and editing sketches; and can assist in re-drawing and over-drawing. This interaction technique also preserves the naturalness of the traditional way of sketching.

## 7.2. Medical Planning

### 7.2.1. Problem Description

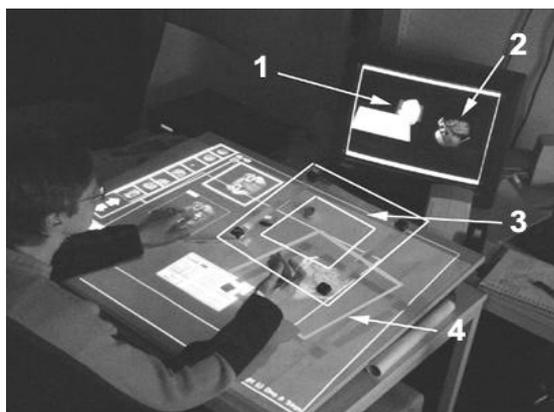
In surgical planning, surgeons may want to set out a trajectory in 3D in order to carry out a biopsy. Obviously this trajectory should avoid vital tissues. Traditionally neurosurgeons have planned surgery based on 2D slices acquired through MRI data. The 2D slices are also restricted to planes that are orthogonal to canonical axes through the patient's head. Although the MRI data are 3D volumetric data, the 2D slice presentations is an artifact of limited computer visualization and interaction technology. However, many surgical paths within the brain that are clinically useful are oblique to these views. Following such oblique trajectories has been risky, since it is difficult or impossible to produce appropriate visualization. This is why surgeons want to be able to see views at oblique angles and be able to relate them back to the more familiar canonical views [3].

From the above problem description, it is clear that the task requirements for neurosurgeons fit quite nicely

the task description mentioned in earlier sections. The surgeons need to navigate through the 3D volumetric data (in this case, the MRI brain model) and make selections of desired cross-sections and analyze them further to plan a 3D trajectory.

### 7.2.2. Prototype Implementation

In order to study the use of the platform for surgical planning, a prototype of the system was implemented using medical data as the volumetric data.



**Figure 5:** Action-Perception and Communication Space for Medical Planning. 1 – 3D model in the vertical projection space. 2 – Bitmap image of the slice, 3 – The RISIP, 4 – The 3to2D window

The 3D model (which is virtually resting on top of the workspace) can be positioned anywhere within the action-perception space by moving the resizable window 3to2D window (see 4 in Figure 5) to the desired location. The 3to2D window, as the name suggests, is a window to the 3D model through which the user can view the slices made by him/her. The user can dynamically view slices of the 3D model by moving the RISIP (see 3 in Figure 5) above the 3to2D window. If the user finds any of the cross-sections interesting and wishes to investigate it further, s/he can cover any of the (four) IS600 tracker discs (see Figure 2) for a short period of time (less than 1 sec) and immediately the bitmap is saved onto the hard disk of the computer. The image sliced then becomes available through the image database browser. Auditory feedback is provided to confirm that the user's action has been successfully performed. Any image in the image database browser can be used as a VP and examined as mentioned in the section on Interaction Technique

The communication space (vertical display area) provides a perspective view of the horizontal workspace with the surface rendered 3D model sitting in the location of the 3to2D window (see 1 in Figure 5). The 3D data requires some form of segmentation in order to map it into a surface model (Currently this segmentation is

based solely on the gray value). When the user moves the 3to2D window in the horizontal space the 3D model moves accordingly in the projection space. The RISIP is also visible in the projection space as a translucent plane moving through the 3D model. For user convenience the vertical projection space also provides the same bitmap that is shown in the 3to2D window (see 2 in Figure 5).

## 8. Evaluation

Due to the popularity of the platform as a demonstration tool we have had the opportunity of showing the prototype to many students (from different faculties within this university) and to some professionals from medical system and architecture companies. The platform had a clear "WOW" effect with all of them. Most observers were delighted when they saw the RISIP work. Almost all of them immediately requested to try using the Prop. The most common problem encountered was that the IS600 tracker loses the RISIP when the users block the (wireless) line of sight connection between the RISIP (i.e. tracker discs), and the tracker. But they learn quickly where to position the 3to2D window to work effectively. We plan to mount the IS600 trackers in a more suitable location in order to reduce this problem.

The prototype for architectural design was formally evaluated for different aspects of usability and usefulness. The experiment was performed in the form of a walkthrough guided by a scenario. The goal was to motivate the user to use all the functionality in the interface. The subjects were pleased with the system. They especially found that the use, functionality and limitations of the pen, the EPP and the VP are clear. They could easily learn how to work with these elements in less than 5 minutes. While developing the prototype we were quite uncertain of the need for the EPP, but the subjects (all architectural discipline experts) have confirmed the usefulness for the EPP.

The main problem encountered by most subjects was handling of the BEL. We noticed that some subjects tend to cover the top of the BEL while moving it, and failed to cover it completely when trying to remove it from the action-perception space. The use of the BEL demanded a conscious effort from the user to activate and deactivate it properly. We also compared the time taken by the subject to overdraw a simple sketch using both the system and the traditional way of using the transparent paper. There was no significant difference in performance times and the subjects preferred the system because of the flexibility it provides without the loss of speed.

## 9. Future Directions

We have planned formal experiments to evaluate the usability of the RISIP in performing the task T2 effectively and efficiently. This task of spatial navigation can be operationalized as the users ability to identify and locate objects within a 3D model. In a first experiment,

the users will have to browse through an artificially generated 3D dataset (for about 1 minute) and have to report the number of different objects that they see by filling in a questionnaire. They will then be shown 4-6 translucent versions of the 3D object (ranging from very similar to far from similar to the original 3D data) and they have to pick the one they think best describes the 3D data. This experiment will provide a measure of accuracy of performing task T2. In the second experiment we want to measure how precisely users can make cross-sections with the RISP. The users will be asked to create a cross-section that aligns with a white disc within the 3D model. If we measure error as the number of non-white pixels in the user generated slice then we expect the error to decrease with an increase in the thickness or diameter of the white disk. We are especially interested in the required thickness and diameter for (almost) flawless performance.

As a next step we intend to redesign some interface elements, such as the BEL, in order to incorporate the most important user feedback on the usability of the system. We will then perform formal usability evaluations, based on the ISO 9241-11 standard [22], of the system for effectiveness, efficiency and satisfaction. The evaluations will be done with discipline experts.

Finally, we intend to extend the functionality of the interaction technique to include spoken input. Spoken input is successful in situations where the user's hands and eyes are busy [23]. Furthermore, spoken input can increase the naturalness of the interaction technique [24]. There have been studies [25], which suggest that users prefer spoken input in any dialogue system as long as the speech recognition is reliable.

## 10. Conclusion

In this article we have presented new interaction techniques for navigating and manipulating both 3D and 2D data. Tools for architectural design and surgical planning were implemented on the prototype system. Informal evaluations by domain experts suggest that the system has added value in task performance. More formal experiments are planned to confirm this.

## References

1. Wuthrich, C.A., An Analysis and a Model of 3D Interaction Methods and Devices for Virtual Reality, *Proceedings of the Eurographics Workshop*, 1999, pp 18-29
2. Murakami, T., and Nakajima, N., Direct and intuitive input device for 3D shape deformation, *Proceedings of CHI*, 1994, pp. 465-470.
3. Hinckley, K., Pausch, R., Proffitt, D., and Kassell, N. (1998). Two-handed virtual manipulation. *ACM Transactions on CHI* 5, 260-302.
4. Johnson B.D. and Bacigalupo-Rose S. (1993). Three-dimensional data imaging in mine geology applications. *International Mining Geology Conference*, Kalgoorlie - Boulder, 5-8 July 1993, Australasian Institute of Mining and Metallurgy, Publication Number 5/93, 35-46.
5. Lin, C.R., Loftin, R.B. and Nelson Jr., H.R., Interaction with Geoscience Data in an Immersive Environment, *IEEE Virtual Reality 2000 Conference*, New Brunswick, USA, 55-62, March 2000
6. Winograd, T. and Flores, F. (1986). Understanding Computers and Cognition. Reading, MA, USA: Addison-Wesley Publishing Company.
7. Herndon, K.P.; Van Dam, A. and Gleicher, M., "Workshop on the challenges of 3D interaction", *SIGCHI Bulletin* 26, October 1994, pp. 1-9.
8. Subramanian, S. and IJsselsteijn, W.A. Survey and classification of spatial object manipulation techniques. *Proceedings of OZCHI 2000, Interfacing Reality in the New Millennium*, 330-337, December 2000
9. Smets, G.J.F., Stappers, P.J., Overbeeke, K.J., and Van der Mast, C., Designing in virtual reality: Perception-action coupling and affordances. In: Carr, K. and England, R. (eds.), *Simulated and Virtual Realities. Elements of Perception*. London: Taylor & Francis, 1995, pp. 189-208.
10. Seymour, J., Virtually real, really sick, *New Scientist* 1996, pp. 34-37.
11. Gribnau, M.W. Two-handed interaction in computer-supported 3D conceptual modelling, *Doctoral dissertation, Delft University of Technology, The Netherlands, 1999*.
12. Landay, J., and Myers, B. (2001), Sketching Interfaces: Toward more human interface design, *Computer*, March 2001, pp. 56-64
13. Rauterberg, M., Fjeld, M., Krueger, H., Bichsel, M., Leonhard, U. and Meier, M. (1997), BUILD-IT: A Computer Vision-based Interaction Technique for a Planning Tool, *Proceedings of HCI '97*, Berlin: Springer, pp. 303-314.

14. Hinckley, K., Pausch, R., Goble, J. and Kassell, N., Passive real-world interface props for neuro-surgical visualization, *Proceedings of CHI*, 1994, pp. 452-458.
15. Frohlich, B. and Plate, J. The cubic mouse: A new device for three dimensional input. *Proceedings of the CHI 2000*, 526-531.
16. Aliakseyeu D., Martens, J.-B., Subramanian, S., Vroubel, M., and Wesselink, W. Visual interaction platform. *Proceedings of Interact 2001*, Tokyo, Japan. July 2001, pp. 232-239
17. Vroubel, M., Markopoulos, P. and Bekker, M. Developing interaction styles to support informal communication at home, *Proceedings of IFAC Symposium on Analysis, Design and Evaluation of Human-Machine Systems*, 18-20 September 2001, Kassel, Germany
18. Greef, P and Ijsselsteijn, W. Social Presence in a photoshare tele-application, in *Proceedings PRESENCE 2000*.
19. Gross, M. and Yi-Luen Do, E., Ambiguous intentions: a paper-like interface for creative design. *Proceedings UIST'96 Symposium on User Interface Software and Technology*. Seattle, ACM Press 1996, pp. 183-192.
20. N.M. Segers, H.H. Achten, H.J.P. Timmermans, B. de Vries. *A comparison of computer-aided tools for architectural design*. - Timmermans, H.J.P. & Vries, B. de (eds.) *Design & Decision Support Systems in Architecture - Proceedings of the 5th International Conference*, August 22-25 2000, Nijkerk, pp. 325-340.
21. Trinder, M., The Computer's Role in Sketch Design: A Transparent Sketching Medium, in: Godfried Augenbroe, Charles Eastman (eds.), *Proceedings on the 8th International Conference on Computer Aided Architectural Design Futures*. Atlanta: Kluwer Academic Press, pp. 227-244
22. ISO/DIS 9241-11, Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11 Guidance on Usability, *International Organization of Standardization*, 1998.
23. Shneiderman, B. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Amsterdam: Addison-Wesley.
24. Kamm, C. User Interfaces for Voice application, In: D. B. Roe and J. G. Wilpon (Eds.): *Voice Communications between Humans and Machines*, pp. 422-441. Washington DC: National Academy Press.
25. Bilici, V., Krahmer, E., te Riele, S. and Veldhuis, R. Preferred Modalities in Dialogue Systems, *International Conference on Spoken Language Processing ICSLP 2000, Beijing (China) 2000*.
26. Cutting, J.E. and Vishton, P.M. (1995), "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth". In: *Epstein, W. and Rogers, S. (eds.), Perception of Space and Motion*. San Diego: Academic Press, pp. 69-117
27. Ijsselsteijn, W. A., de Ridder, H. and Vliegen, J. (1999), "Effects of Stereoscopic filming parameters and display duration on the subjective assessment of eye strain", *Proceedings of the SPIE 3957*, 12-22