

Issues and Challenges of 3D User Interfaces: Effects of Distraction

Leslie Klein
kleinl@in.tum.de

In time critical tasks like when driving a car or in emergency management, 3D user interfaces provide an attractive approach to support the user. Upcoming interaction techniques enable faster and more intuitive interaction in comparison to standard interfaces. As 3D User Interfaces gain more capabilities, new challenges and problems arise. Unlike in 2d human-machine interaction, where standards are set, 3D postulates new metaphors to make usability simple, intuitive and effective. Furthermore in time critical tasks users have to focus primarily on their main task, they cannot lend their full attention to the user interface. Hence common requirements and rules for interface design need to be revised and extended. First an overview about 2D and 3D Interfaces in general will be given. This includes the definition of terms, as well as listing the main components like input, output and interaction techniques of 2D and 3D user interfaces. Followed by the implication of constraints in the application domain of time-critical tasks is taken under consideration. In addition, potential and hazardous effects like perceptual tunneling and cognitive capture are mentioned.

1 Introduction

This report starts with a survey about user interfaces in general, comprising their task within human-machine interaction, their main components like input devices, output devices and interaction techniques and furthermore the extension to 3D user interfaces. These include spatial input devices which allow gestural input and 3D output modalities, like head-mounted displays or spatial audio systems. In our context of time-critical tasks only output devices for the visual, auditory and haptic channel will be mentioned. In the next section interaction techniques, methods used to accomplish a given task via the interface, will be introduced. Categorization of these techniques into structures and interaction into general terms provide a more complete understanding of the tasks and thereby a helpful scientific framework for interaction design. In addition this section ends with the main questions which have to be considered by designing user interfaces.

Afterwards the focus moves to time-critical tasks and the special constraints to user interfaces within these tasks.

The third major section exposes that as a consequence of new technologies also new challenges and problems arise. Within this section psychological phenomena like perceptual tunneling, cognitive capture, change blindness will be explained in matters of using 3D user interfaces in the time-critical task driving. These phenomena refer to a potential situation where the driver is distracted by the interface from his main task driving.

2 User Interfaces

User Interfaces are the medium through which human users can communicate with computers or systems. They translate a user's action into a representation the computer can understand and act upon and the other way around. Input and output devices are physical tools to provide this interaction with the system, so they are important components in building user interfaces. The way interaction takes place is implemented by various interaction techniques. In traditional human-machine interaction the user interacts with the computer environment from outside via interfaces. This interaction can be compounded into a continuous interaction-circuit shown in Figure 1. The user is perceiving the actual system state and transfers his goals via the physical input devices like mouse or keyboard to the system. Task of the user interface is to translate users action into signals the system can understand. So, the system perceives the users goals and acts upon. Further on the interface transcribes the system goals into a human-understandable representation and transfers these via the output devices to the user.

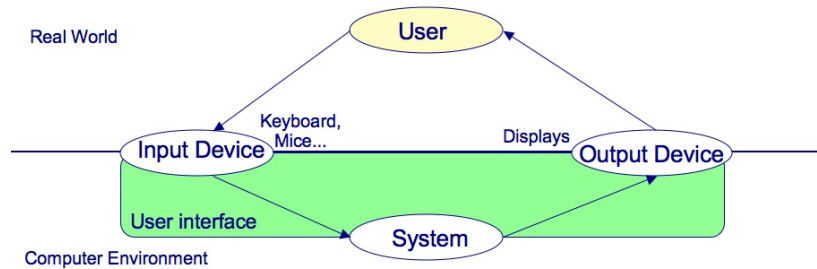


Figure 1: 2D User Interface

But in the course of time upcoming technologies extended the standard parts of interfaces, so human-computer interaction in which the user tasks are performed directly in a 3D spatial context are enabled. This means the user interacts *within* the computer environment. Hence, in terms of 3D interaction the user is not interacting through an interface, he can actually be seen as a part of the interface. (Figure 2)

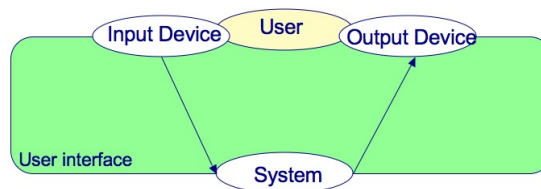


Figure 2: 3D User Interface

This section covers an introduction to the three main components of user interfaces. First input devices, followed by output devices and further on interaction techniques will be explained.

2.1 Input Devices

Input devices are physical components used to transfer user goals to commands or information the computer can perceive. In distinction to interaction techniques, input devices are physical tools performing the aims of interaction. In other words they are tools through which the user is able to control the system or the program. Many different ways of interaction can be mapped onto any given device. Clicking a mouse for example has a wide range of interaction possibilities depending on its context.

Input devices can be characterized in *discrete* and *continuous* input devices, depending on types of events they generate. Discrete input like pressing a button generates a single signal whereas tracking devices generate a continuous stream of events. Keyboards and mice are traditional two-dimensional input devices and are still prevalent, but new technologies enable input in 3D spatial context. Trackers are 3D input devices whereby position and orientation of moving objects are traced as streams of events. Thereby only parts of the users body like a hand with a Dataglove (Figure 3 left and right) or the eyes or the complete body (Figure 3 middle) can be tracked. Using datagloves as input devices for interaction with the virtual environment seems to be an intuitive approach since direct hand manipulation is a major interaction modality in natural physical environments. Though in comparison to the real world some cues are missing so it can be much more difficult for users to understand and act within 3D virtual environment. Therefore the choice of input is linked to the output devices, because the output is in some way a feedback for the user and therefore essential for the understanding of interaction within spatial context. Datagloves can be both, input and output device. To provide a natural, efficient and appropriate mapping between interaction techniques and hardware the designer of user interfaces must have a proper understanding of the advantages and limitations of the devices they use.

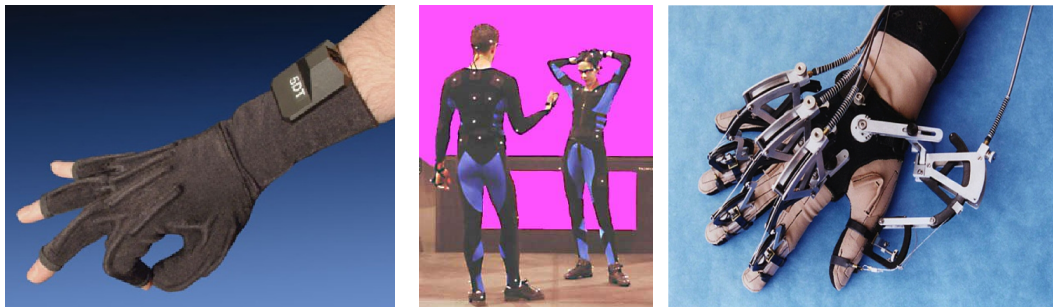


Figure 3: Tracking Devices

Spatial input modalities like gestures and speech gain more capabilities for interaction. Speech is a approach if the user has no free hand for interacting. Furthermore combination of various input devices enable multimodal interaction.

2.2 Output Devices

Output devices in user interfaces, commonly described as displays, are the physical components to present information generated by the computer and perceived by the user. The output information can be some kind of feedback for the user. In matters pertaining to human-computer interaction they can be characterized in visual, auditory, haptic, tactile and olfactory displays.

2.2.1 Visual Output Devices

Because a various range of information can be perceived by humans via the visual channel, visual displays are the most common ones. Fully-immersive displays occlude the real world with a graphical representation of a virtual world. Semi-immersive displays allow the user to see both the physical and the virtual world. In addition for aspects like mobility and the field of view different approaches exist. Depending on which target should be pursued a big variety of display types can be chosen. Surround-screen displays (like CAVE) consist of 3 up to 6 rear projection planes, arranged like are cube where the user is inside so a stereoscopic view is provided. Head-mounted displays (HMD) are displays with lenses embedded in a helmet, glasses or visor. By displaying an offset image to each eye, the device can be used to show stereoscopic images. A combination with head-tracking devices enables the user by turning his head to have a look around in the virtual world. This HMD technology can be used as a fully-immersive display or as well as semi-immersive, so that the reality can be augmented by virtual images. Virtual retinal displays include eye-tracking devices in such a way as to enable projecting information by a laser direct into the eye. Head-up displays (HUD) (Figure 4) are transparent displays that presents data without restricting the user's view. In aircrafts or automobiles they are used to present either symbolic information, like speed-limits or in combination with a head-tracking system to display conformal information.



Figure 4: Head-Up Displays: C-130J, U.S. Air Force and BMW

2.2.2 Auditory Output Devices

Auditory displays provide an attractive second communication channel, informational or alarm signals can be transferred acoustical, so cognitive load can thereby be reduced. Auditory signals can be used guiding users attention to a specific direction to support locating objects in environment or as substitution for another missing sensory modality like pressing

a button. Accordingly main interface goals are the generation of 3D sounds and sonification, transforming information into sounds. Furthermore ambient acoustical effects like birds chirping can provide a sense of realism in virtual environments.

2.2.3 Haptic Output Devices

Haptic and tactile displays are important devices to provide a sense of force, a sense of touch or both. With this device, attractive effects can be integrated to user interfaces, because feeling and touching makes the virtual environments much more realistic. Haptic devices can be used to catch the user's attention or give feedback or informational hints. To transfer haptic cues, output devices have to be physical connected to the user in some kind. This is provided either by way of ground-referenced tools, like flight-simulators and treadmills to provide moving or walking or by way of body-referenced tools, so the haptic devices are mounted on the user's body, whereby much more freedom to users motion is given.

2.2.4 Multimodal Output Devices

Multimodal displays are the combination of different device types, so the individual weakness of each can be vanquished. In real life the interplay of senses plays an essential role for humans to become aware of their environment. Humans perceive visual, acoustical and haptic cues simultaneously, therefore it is rewarding to combine diverse device types to provide a deeper sense of realism.

2.3 Interaction Techniques

The quality of 3D UI depends not only on output also on the range of possibilities of interaction. Thus the goal of user interface design is to make the user's interaction experience as simple and intuitive. Interaction techniques are methods used to accomplish a given task via the interface. In this context hardware and software components have to be considered. Interaction techniques are strongly related to the goals and tasks of a system, so strength and weakness of each is depending of the particular requirements and therefore no standards are established. Categorization of spatial interaction in general terms provides guiding principles for interaction design. To make human-machine interaction as natural and intuitive as possible *metaphors* are used. Metaphors are easily understandable mental models to allow the user to apply everyday knowledge in the virtual environment.

2.3.1 Manipulation and Selection

Manipulation tasks are fundamental for physical and virtual environments so it is indispensable to provide techniques for selection, acquiring a particular object, for positioning, moving an object and for rotating, changing orientation of an object. Several approaches exist, but their effectiveness is depending on a multitude of variables like application goals, object size, distance between user and object and state of the user. A virtual hand is an intuitive way of interaction because direct hand manipulation is a major interaction modality in natural physical environment. But only objects within a close area around the user can be picked up. To overcome this problem pointers, ray- or cone-casting techniques provide an easy, distance independent way for selection of objects. When a vector or ray, emanating from the virtual hand, intersects a virtual object it can be chosen by using a trigger, like voice command or

button. Such an approach provides only poor positioning and rotating technique, objects can only be manipulated in radial movements around the user and rotated only about the axis, defined by the pointing vector. In addition the problem of occlusion occurs. Objects behind another cannot be selected and accordingly not manipulated.



Figure 5: Selection with Cone-Casting

An alternative to the techniques mentioned before is selection with two vectors emanating from each hand of the user. The object where both vectors are crossing can be selected, so also occluded ones can be chosen.

2.3.2 Navigation

Navigation is a fundamental human task in physical environment, but as well in synthetic environment the user is faced with this task, navigating within a virtual world of computer games or navigating the world wide web via the browser are examples. To provide efficient movement, the user has to know where he is in relationship to other objects in his surrounding. He needs support for spatial awareness. Navigation can be subdivided into travel and way finding. Travel is the motor component of navigation, the task of performing the action that moves one from a current location to a desired direction or a target location. Way finding is the cognitive process of defining a path through an environment, thereby spatial knowledge has to be used and acquired to built up mental maps. Usage and acquisition of spatial knowledge can be supported by natural and artificial cues. Way finding support includes visual motion, real-motion and auditory cues.

2.3.3 System Control

System control is a users task to access computer system functionality by issuing commands. Commands are used to request system to perform particular function, to change the mode of interaction and to change the system state. The user issues what the system should do, but leaves it up to system how to perform the given command. Control widgets like windows, icons, menus and pointers (WIMP) are typical two-dimensional interaction styles. They are often adapted to 3D user interfaces because they are well-known, but are not effective in all situations. Inter alia a problem is, where to place such widgets in a virtual environment. Controlling the system by gestures would be an alternative, but is only useful when they are related to well-defined natural gestures or in combination with other modalities.

2.3.4 Symbolic Input

Symbolic input is the task in which the user communicates symbolic information, like text, numbers and other symbols or marks to the system. Objects, ideas and concepts are represented by abstract symbols and even the language is an abstraction of information, so humans use symbolic communication in everyday tasks. Using computers symbolic input tasks like writing emails, composing documents are used constantly. While the importance of this task are clear for 2D interfaces it is not as obvious for 3D user interfaces. As multiple users share a 3D environment, they commonly require facilities for interchanging information. So possibilities for communication or adding labels, questions, suggestions or others to objects in virtual or augmented environments have to be given. Therefore in 3D user interfaces symbolic input is indispensable, to provide communication, annotations or labeling and markup some information.

Mentioned at the beginning of this section the design of 3D user interfaces is depending on the requirements of the particular application and therefore the following questions have to be considered in the design of user interfaces:

- Which tasks need to be supported by the user interfaces?
- Which interaction techniques are appropriate?
- Which input devices map these techniques best?

3 Time Critical Tasks

To cope with tasks humans have to manage the complexity of information they perceive and react in an adequate manner. This activity is a complex and continuous circuit, shown in Figure 6. Information has to be perceived by all senses, processed by the brain and transcribed into the next step of interaction. In time-critical tasks, like driving or in emergency cases the procedure of this circuit has to be accomplished in a very short range of time.

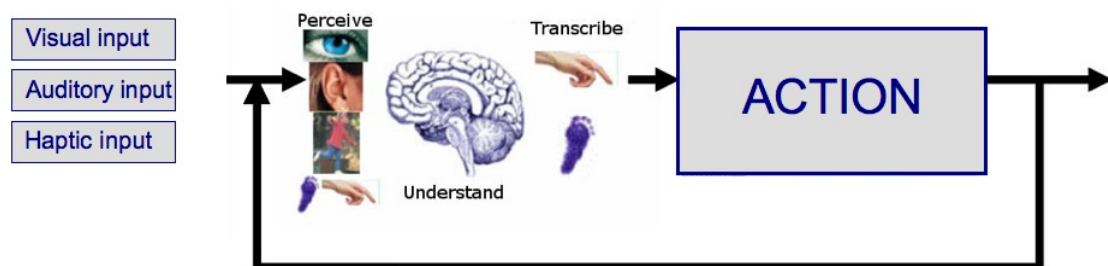


Figure 6: Interactivity Circuit

For instance driving a car is a highly complex tasks which needs a significant amount of a driver's attention. Accordingly users have to focus primarily on their main task, supplementary systems should first and foremost be as less distractive as possible. For the reason that all information can be potential distractive, strong constraints in designing user interfaces have

to be taken under consideration: To reduce the impact of distraction input and interaction should be intuitive, with no learning phase and should require only little visual attention. Furthermore to improve safety supplementary systems must be interruptible. Output devices should not distract cognitive skills.

4 Challenges and Problems

As 3D user interfaces gain new possibilities, new interaction techniques enable faster interaction in comparison to standard interfaces but as a consequence thereof also new challenges and problems arise. Embedding visual information into a HUD integrated into a car's cockpit is an example for using 3D user interfaces in time-critical tasks. This on the one hand can provide support but on the other hand can lead to distractive psychological phenomena. In this section these phenomena like perceptual tunneling, cognitive capture, change blindness will be explained in terms of the time-critical task driving.

4.1 Information Overload

Information overload refers to the state having too much information at the same time, so the user is not able to perceive the important information. For instance in car integrated Head-Up displays show significant information, such as speed limits or safe distance information to the driver (see Figure 7). In matters of keeping the driver informed large amounts of diverse informational signs can be displayed, but the effect of information overload should thereby stay in mind.



Figure 7: Information Overload in HUD

4.2 Change Blindness

Change blindness refers to a failure detecting what should be an obvious change. In human-computer interaction it occurs when more than one change is happening at a display at the

same time or the user's attention is distracted. The user has to memorize the state before the change and in comparison has to detect the new state to recognize that a change already had happened. An attempt to reduce the possibility to miss an important change is to make changes explicit.

4.3 Occlusion and Depth Perception

HUD integrated into car cockpit provides not only the possibility to present symbolic information, furthermore virtual objects like navigational arrows (Figure 8 left) can be displayed in their exact depth, so they seem to be aligned on the street and thus pretending to be like an real object integrated in the environment. This approach is under great research, because such a navigation system is more intuitive, the driver does not have to mentally transform the bird's eye perspective like in common systems into his perspective or does not need to interpret distance information, so cognitive load can thereby be reduced. Furthermore the driver does not need to turn his head away of the street-scenery. But additionally new problems arise. If there is leading traffic the arrow cannot be seen until the leading car has left the position. To avoid this, the arrow can be displayed closer to the own car (Figure 8 right), but even if the arrow is semi-transparent some areas of the surrounding field are occluded. This increases the risk to miss an important event in the traffic and thus to cause a crash. And furthermore the depth cue perception is reversed by such a visual presentation, so it is irritating and in some kind distracting for the driver.

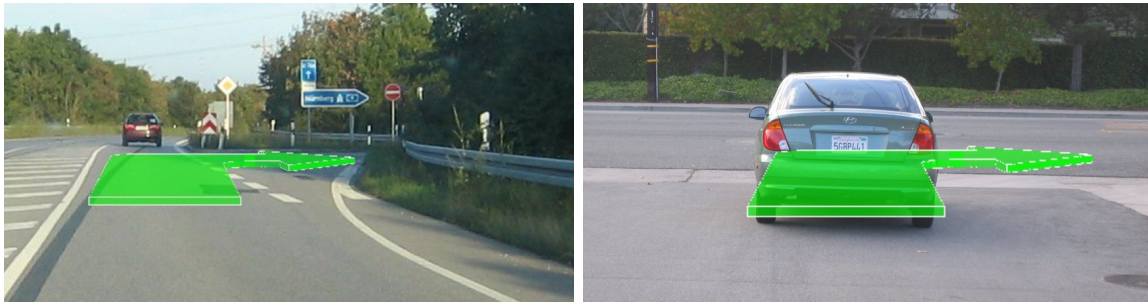


Figure 8: Occlusion

4.4 Perceptual Tunneling

Beside information overload displaying information, especially animated warning schemes into a car's HUD can lead to perceptual tunneling. This phenomenon is first being recognized in aviation and refers to a state where the human is focused a specific stimulus like a flashing warning signal and neglects to pay attention to other important tasks or informations like in this context driving, flying or the traffic. The human is only visual attracted by a specific stimulus, but does not spent cognitive load on the stimulus, he is only starring at something without thinking about it. This effect is known in everyday situation like for example being with friends in a bar where a TV is turned on and one has to stare at the TV without recognizing what he is looking at. But the visual attraction can easily change to a cognitive attraction and thereby lead to the next mentioned phenomenon cognitive capture.

4.5 Cognitive Capture

Cognitive capture refers to the situation where the driver may be totally "lost in thought" and thereby could lead to loss of situational awareness. This effect is known in everyday situations like for example a person in a cinema is looking out for a free seat and is thereof lost in thought about this task, so that he does not see a friend who is waving to him. Whereby in this case perceptual tunneling does not cause important consequences whereas in time-critical tasks this can cause hazardous effects, like increasing the risk of a crash.

5 Summary and Outlook

3D user interfaces gain more capabilities and thereby enable faster and more intuitive interaction in comparison to standard interfaces. New devices, new techniques and new metaphors on the one hand promising new possibilities for human-machine interaction and especially for time-critical tasks, but on the other hand as a consequence thereof no standards are set and new challenges and problems arise. So great care has to be taken by designing 3D user interfaces. The design is an interdisciplinary field of study of research, because not only technical aspects as well the physical and the psychological situation of the user have to be taken into account. Therefore knowledge in perception, cognition, linguistics, human factors, ethnography, graphic design and others is very important. Further on, to avoid cognitive phenomena usability tests are indispensable.

References

- [1] D.A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev *3D User Interfaces: Theory and Practice* Addison Wesley, 2004.
- [2] Marcus Tönnis, Verena Broy, and Gudrun Klinker *A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers* In Proceedings of the 1st IEEE Symposium on 3D User Interfaces (3D UI), March 2006.
- [3] Marcus Tönnis, and Gudrun Klinker *Effective Control of a Car Drivers Attention for Visual and Acoustic Guidance towards the Direction of Imminent Dangers* In Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR), October 2006.
- [4] Marcus Tönnis, Christian Lange, Gudrun Klinker, and Heiner Bubb *Transfer von Flugschlauchanzeigen in das Head-Up Display von Kraftfahrzeugen (Transfer of Flight Tunnel Visualizations into the Head-Up Display of Cars)* In Proceedings of the VDI-Gemeinschaftstagung Integrierte Sicherheit und Fahrerassistenzsysteme, October 2006.
- [5] Paula J. Durlach *Change Blindness: What you dont see is what you dont get* U.S. Army Research Institute for the Behavioral and Social Sciences; 12350 Research Parkway, Orlando, FL, 32826; 2005
- [6] Shamus P. Smith, Jonathan Hart, *Evaluating Distributed Cognitive Resources for Wayfinding in a Desktop Virtual Environment* 3D User Interfaces (3DUI'06), 2006.