EYE TRACKING AS A USER BEHAVIOR REGISTRATION TOOL IN VIRTUAL ENVIRONMENTS

JAN DIJKSTRA, HARRY TIMMERMANS
Eindhoven University of Technology,
Faculty of Architecture, Building and Planning
P.O.Box 513, 5600 MB Eindhoven, The Netherlands

AND

WALTHER ROELEN
Eindhoven University of Technology,
CALIBRE BV
P.O.Box 513, 5600 MB Eindhoven, The Netherlands

Abstract. Registration of user behavior in a virtual environment is a particular aspect of an ongoing research project which aims to develop a conjoint analysis - virtual reality system. In this paper, the registration of user behavior by eye tracking techniques will be described. It will be advocated that eye-tracking techniques offer interesting possibilities for recording user behavior.

1. Introduction

The registration of user behavior in virtual environments is a particular aspect of an ongoing research project which aims to develop a conjoint analysis and virtual reality (CA&VR) system as part of a newly defined research program on the development of a design information system in virtual reality. This research program, called VR-DIS[1)], aims to develop a design system that can be used for interactive design and evaluation, using a VR interface. The concept of VR-DIS has a two meanings. From a technical point of view VR-DIS stands for Virtual Reality - Distributed Interactive Simulations. From the application point of view, VR-DIS stands for Virtual Reality - Design Information System.

1) VR-DIS is part of the Design and Decision Support System in Architecture and Urban Planning (DDSS) research program of the Eindhoven University of Technology.

CAADRIA '98 : Proceedings of The Third Conference on Computer Aided Architectural Design Research in Asia.
eds. T. Sasada, S. Yamaguchi, M. Morozumi, A. Kaga, and R. Homma
April 22-24, 1998. Osaka University, Osaka, Japan. Pp. 57-66

The research program will mainly focus on the development and application of VR-DIS technology as an integrated approach to support (design) decisions. The particular focus of this research program concerns the integration of information from different disciplines within a dynamic design process in combination with the application of the new possibilities that VR offers to present design information to the user.

VR as an User Interface will probably replace many of the existing techniques due to the possibilities which are available to communicate intuitively between man and machine. VR uses specific input and output devices that can stimulate all senses and capture all behavior of the user (Coomans and Timmermans, 1997). Especially in the architectural environment, the challenge lies in developing a new work environment in which the design process can take place.

The system architecture consists of interface functions, information storage and computer systems. An interface function may use one or more storage's and one or more systems. Figure 1 shows a part of the system architecture.

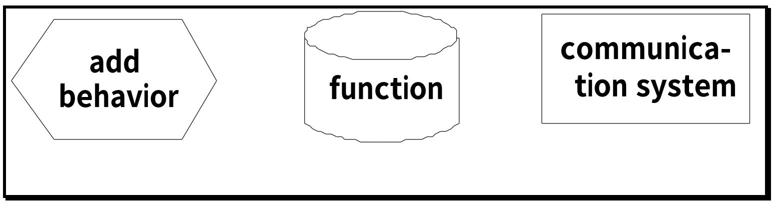

Figure 1. Interface function: add behavior.

User behavior may be recorded in different ways. In this paper, we focus on perceptual aspects of user behavior. The problem definition is how to record user perceptions of virtual environments, especially with respect to specific design concepts which are of interest to the designer/researcher.

The paper is organized as follows. First, we will discuss the possibility of registrating user behavior by using eye tracking techniques. Then, in section 3, we will discuss the basics of a CA&VR-system as a evaluation tool based on user perceptions. In section 4, a simple example illustrates the interpretation of the registration of user behavior by eye tracking techniques and the CA&VR-system. We will finish with a brief discussion.

2. Registration of user behavior by eye tracking

As visually oriented creatures, humans use their eyes intensively for a large variety of purposes, such as for reading, watching, gathering information to plan their actions, perceiving and learning new things, and exploring and navigating in environments. Generally, we do not realize how great an effort our eyes put into our perception process, and what immense amounts of information they process. We can concentrate our cognitive processes by operating on the concepts that surround us, leaving the intake and basic processing of optical information to our eyes and visual system. We observe the surroundings, which is also the most important role of our eyes. In this way, the eyes are used as INPUT -organs. But eyes can also operate as OUTPUT -organs; the output they are capable of producing is, on the face of it, direction. The eyes are pointed in one direction, thus indicating what is being focused upon (Glenstrup and Engell-Nielsen, 1995).

By eye tracking techniques, it is possible to monitor the orientation of eyes, and thus the direction of gaze. The ability to track the direction of gaze enhances the communication between the subject and the computer. In this section, we discuss successively eye tracking basics, the relation with architecture and registration of user behavior in virtual environments.

2.1. EYE TRACKING BASICS

Advances in communication have been made in terms of the communication from the computer to the user (e.g., windows systems, graphical presentation of data) as well as the communication from the user to the computer (e.g., keyboard, mouse, joystick). By tracking the direction of the gaze of the user, the amount of potential information transfer can be increased by using the information about what the user is looking at, and even designing objects specially intended for the user to look at. By monitoring the user, the computer can react to all kinds of gestures. In this way, what the user is really interested in becomes more transparent. As a positive consequence of eye-gaze interaction, handicapped people for example are allowed to concentrate on interacting with the data presented by the computer. That is a new way of regarding the computer not only as a tool that must be operated explicitly by commands by one that can also be operated visually.

Eye tracking works, but only under controlled conditions. There are two methods of eye tracking which are most commonly used: limbus tracking and video tracking. Limbus tracking works by illuminating the eye of the user with an invisible near-infrared LED. A photodetector is used to pick up the intensity of reflected light from the white, or sclera, portion of the eye. The dark regions of the eye are represented by the iris and pupil. The intensity of reflected light will vary with the eye position because of the varying proportion of the light and dark regions of the eye exposed to the detector.

Video tracking on the other hand works on the principle of Pupil Center/Corneal Reflection (PCR). The eye is illuminated with a near-infrared LED. A video camera collects images of the eye. From these images, a computer calculates the center of the pupil and the specular highlight of the LED which is referred to as the corneal reflection. Because the lens of the eye protrudes out in front of the sphere of the eyeball, the pupil and reflection move relative to each other as the eye shifts gaze. The computer uses the vector between the pupil center and corneal reflection to calculate the direction of the gaze. With either method, a calibration process is used to relate eyeball motion into point of gaze which relates to the real world.

On video tracking systems, in some cases the optics are fixed to the room and the system can directly measure eye line of gaze with respect to the room. This system (called *remote system*) is an eye tracking system which is not attached to the user being tracked. It may be attached to the floor by a tripod, or perhaps sitting on a table near a display which the user is looking at. In other cases the optics are fastened to the user's head or to a helmet or headband worn by the user and eye line of gaze is measured with respect to the head (*head mounted* system).

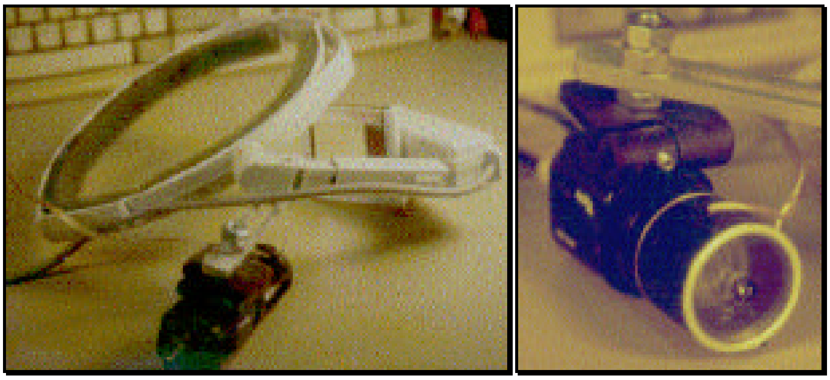

Figure 2. The Heads-Up unit, the miniature camera.

A head mounted system can measure no matter how the user turns his head or what he holds. A great deal of user freedom is possible although the user does have to wear some device, and the measured quantity is eye line of gaze with respect to the head. If you need to know the point of gaze on a stationary scene or object, either the head must be rigidly fixed or the position and orientation of the head must also be measured. The necessary head reference can be provided by a head mounted scene camera and/or by one of several head position detection systems.

2.2. EYE TRACKING TECHNIQUES AND ARCHITECTURE

The application of eye tracking techniques to the field of architecture has hitherto hardly been explored. Studies of eye movement have traditionally dealt with normal reading, reading disabilities, visual tracking and scanning. Studies which emphasize the influences of formal characteristics in visual patterns have been undertaken. Noton and Stark's (1971) studies pertain to the physiological and cognitive foundations of eye movement with a particular focus on scanpath patterns. The scanpath has been defined as a repetitive sequence of saccades and fixations, idiosyncratic to the viewer and to the picture.

The project 'Oculomotor Research in Architecture' by Weber (1996) represented an attempt to record how the visual experience of architecture is influenced by various formal-geometric characteristics. The study was undertaken with the use of computer-controlled video equipment measuring scan-path of the human eye during the perception of three dimensional architectural models according to the remote video tracking system. It shows how people look at buildings and architectural spaces; what elements of architectural form trigger the attention of the eye more strongly than others and how spaces are perceived when architectural elements are altered of replaced. One of the most important apparent findings of this initial research in relation to architectural issues appears to be that the perception of architectural forms and spaces is not based on the scanning of individual stimuli, such as contours, but on forms as a whole.

2.3. REGISTRATION OF USER BEHAVIOR AND VIRTUAL ENVIRONMENTS

In architectural and real estate simulations, it is interesting to get as realistically as possible impressions of a designed model by means of virtual reality. Consideration can be given to modeling autonomous objects and to the simulation of operations on objects. In contrast to conventional CAD systems, virtual reality systems provide support for activities which characterize the early conceptual phase of design, called 'prototyping of designs in virtual reality' (Coomans and Oxman, 1996). Virtual reality techniques can be used to create an interface that allows modeling in an intuitive way. Simulation-based design with VR enhances the capabilities of the design to manufacturing process. The product can be visualized, design changes can be made, and new concepts, without the traditional expense of prototyping, can be tested. Advances in VR techniques enable users to be immersed in new virtual environments and experience new designs, products or services. This aspect is of interest to get a better insight into user behavior, evaluating design concepts and support product testing. To what extent users are interested in specific design concepts can be based on measurement of visual attention by eye movements. The perception of a scene in a virtual environment involves a pattern of fixations, where the eye is held fairly still, and saccades, where the eye move to foveate a new part of the scene. Eye movements can be measured by eye tracking techniques. In our research we are particularly interested into user behavior concerning charming design concepts and/or effectiveness of design concepts.

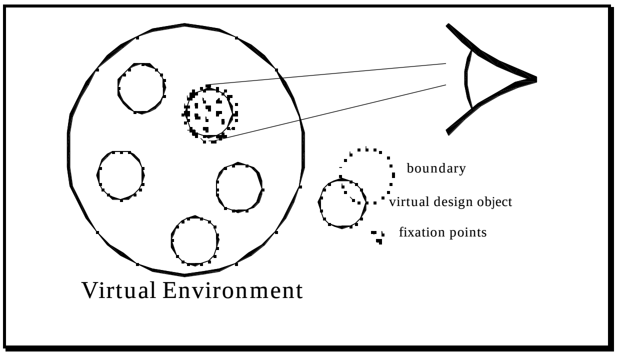

Figure 3. Visual attention field.

Eye movement measurement allow one to determine which areas of the scene in the virtual environment act as visual attention fields, on which the eye focuses and with longer duration of fixation. $FT_{o,1}$ is the measured fixation time on the visual attention field of the virtual design object. Herewith, fixation time is the elapsed time of fixing one's eye on the visual attention field, which is determined by the boundary (region) of the virtual design object. Considering a profile of *n* design objects of interest, the total fixation time is the summation of the fixation time of each virtual design object:

$FT_i = FT_{o,1} + \ldots\ldots + FT_{o,n}$; $FT_{i=}$total fixation of profile$_i$

We think about the head mounted eye tracking system as the most convenient system because the user travels through the virtual environment with less restrictions.

3. The CA&VR-system basics

If we are interested in specific design concepts namely their features with different levels of attribute values, the CA&VR-system provides a tool to evaluate users' perception to particular design alternatives. In this section, the basics of the CA&VR-system tool founded on conjoint analysis methodology will be described. Conjoint analysis or experimental choice analysis represents a widely applied methodology for measuring and analyzing consumer preferences (Carrol and Green, 1995). It is based on a number of paradigms in psychology, economics, marketing and urban planning that are concerned with the quantitative description of consumer preferences or value trade-offs (Timmermans, 1984; Louviere, 1988). Conjoint analysis involves the use of designed hypothetical choice situations to measure individual preferences and predict their choice in new situations. Conjoint experiments involve the design and analysis of hypothetical decision tasks. Alternatives are described by their main features, called attributes. Multiple hypothetical alternatives, called product profiles, are generated and presented to subjects (respondents, users, consumers), who are requested to express their degree of preference for these profiles. To some extent, judgments about presented product profiles are hypothetical. After all, the product profiles have been specified by the researcher, in contrast to observed behavior in real choice situations. Although conjoint analysis has been applied typically for measuring preferences and choices, the underlying principles can also be used for measuring perceptions. Experimental designs can still be used to vary to design parameters or solutions of interest. Rather than measuring preferences or choices, the interest in this case shifts to user perceptions.

Most studies of conjoint analysis have involved a verbal presentation format, although some studies use a pictorial presentation format. Consequently there may be a lack of realism with the traditional presentation formats, which might influence the choice behavior of interest. Especially if the interest concerns user perceptions a visual representation of attribute profiles becomes a (sine qua non) strict requirement. In the CA&VR concept, virtual reality depicts product profiles in a three-dimensional environment and allows subjects to interact with these product profiles. A product profile consists of a virtual environment model and dynamic objects representing the attributes with their levels. Both the virtual environment and objects model can be designed by 3-D graphical and virtual reality software. Advantages of a virtual environment include a more realistic 3D presentation and improved experience.

A concept of a CA&VR-system is developed in Dijkstra, Roelen and Timmermans (1996). It is motivated by the possibility to create or design 3D representations of environments, varied according to the principles of experimental designs thereby reliving the controlled observations of user reactions to design alternatives. By recording these observations in a particular way, various statistical and other methodologies can be applied to evaluate design alternatives or predict user reactions to possible future situations.

A simple illustration demonstrated the potential contribution of the virtual reality system; subjects experience two possible guidance alternatives and were asked to express the one they prefer (Dijkstra and Timmermans, 1997a). Another illustration considered virtual wayfinding as a dynamic decision making process. Indications like signs presented as different virtual objects in a virtual building could be tested in a fire drill simulation (Dijkstra and Timmermans, 1997b).

4. A simple illustration

As we have advocated in the previous section the application of the CA&VR system is not restricted to problems to preference and choice but can also be applied to perceptual issues. A design can be presented by a virtual environment model and dynamic virtual objects representing the different design aspects (called attributes) of interest with their respective levels. Each attribute level is a different state of the concerned virtual design object. We consider the functional design aspect of wayfinding with the emphasis on graphic components. This means we look at information that gives users an overview of where they are, and where the destination is. In this illustration the attribute ‘orientation’ involves the ability to perceive an overview of a given environment, which has two proposed levels: *floor plan* and *directory*. Also other graphic components can be considered, like ‘directional signage’ and ‘identification’. All these graphic components, representing attribute levels, will be presented as different virtual design objects in a virtual environment. A given situation of the graphic component attributes represents a product profile. In this case of virtual wayfinding, indications like signs presented as different virtual design objects could be tested for their suitability. The perception of virtual design objects in the virtual environment gives the necessary feedback. With a head mounted eye tracking system, this feedback will be given by measuring the duration of fixation on the visual attention fields of virtual design objects. As a perceptual task becomes more difficult, by definition time to perceive objects increases. In a perceptual task that involves scanning virtual design objects, this change is correlated with a change in one or more eye movement parameters. The parameters of eye movement involved in the perception of virtual design objects are duration and the location of eye fixations. Fixation time on attributes can be assumed to reflect the effectiveness of the concerned attributes. Randomization in experimental design helps to reduce undesirable effects of a subject’s expectations and strategies. We let subjects experience two profile alternatives and measure the fixation time on each attribute of each profile as well as the total fixation time of each profile. The total fixation time (FT) can be considered to be a measure of the computed level of interest or effectiveness. An experimental design can be used to vary the attributes of interest.

We are not interested in fixation at a certain point of gaze but in eye fixations at visual attention fields. Therefore, as a first try-out, we will simplify the head mounted eye tracking system by head tracking. The subject is wearing a headband with a head position sensing device, which transmits positional information to the scene of the virtual environment. Head motion will be measured, while navigating through a virtual building, which could be done with a joystick or 3D mouse looking to a great projection screen. The projected image of a 3D-cursor in virtual space should align visually with the real position of the 3D-input device that controls it (Hall, 1997). To refine the illustration, aspects to be considered are (i) testing design objects for their suitability by measuring fixation time involves the use of more virtual building alternatives, (ii) how accurately the eye position must be with respect to the head be measured, and how far the user can move his head, (iii) eye movements and therefore fixation time can be influenced by attention for the hand-eye coordination (Hall, 1997).

5. Discussion

VR systems allow the creation of interactive environments. A CA&VR-system tool can be used to observe user reactions and decision making in environments not yet existing. Such a system allows evaluation of an a priori building performance. Especially, it offers an opportunity for generalizing the findings beyond the actual environments that were incorporated in the virtual environments.

In the present paper, we have discussed the integration of the technology of eye tracking with the field of virtual environments. In a conjoint experiment, the objective of the task is defined by conjunctions of attributes. A profile, defined as the identification of virtual design objects representing the conjunctions of attributes, plays an important role. Attribute profiles are varied according to experimental design principles. Traditionally, in conjoint experiments, subjects are invited to express their preference for the experimentally varied profiles by rating or ranking them in terms of overall preferences or subjects may be asked to choose the profile they like best. By measurement of fixation time at visual attention fields of virtual design objects, such systems can also be developed for perceptual tasks. This means an extension to the conventional basis of conjoint analysis. We have discussed the role of eye tracking techniques with restrict to fixation time measurement in this context.

When testing this approach, we are conscious of the following aspects: (i) Subjects fulfilling the experimental task could exhibit different strategies in moving the eye and head. For instance, one subject will display almost exclusive use of head movement while another subject will display a more equitable share of eye and head movement; (ii) the duration of the eye fixation is closely related to the duration of cognitive processes, yet the two durations are not necessarily identical; (iii) the appearance of distance effects. By approaching the visual attention fields, the visual field increases and at that moment the fixation time decreases. We hope to empirically test relevance of the presented approach in the near future.

References

Carrol, J.D. and Green, P.E.: 1995, Psychometric Methods in Marketing Research;: Part I, Conjoint Analysis, *Journal of Marketing Research* **XXXII**, 385-391

Coomans, M.K.D. and Oxman, R.M.: 1996, Prototyping of Designs in Virtual Reality, *in* H.J.P. Timmermans (ed.), *3rd Design and Decision Support Systems in Architecture and Urban Planning Conference, Vol I: Architecture Proceedings*, pp. 13-25

Coomans, M.K.D. and Timmermans, H.J.P.: 1997, Towards a Taxonomy of Virtual Reality User Interfaces., *Proceedings of the International Conference on Information Visualisation (IV97)*, pp. 17-29

Dijkstra, J., Roelen, W.A.H. and Timmermans, H.J.P.: 1996, Conjoint Measurement in Virtual Environments: a Framework, *in* H.J.P. Timmermans (ed.), *3rd Design and Decision Support Systems in Architecture and Urban Planning Conference, Vol I: Architecture Proceedings*, pp. 59-71

Dijkstra, J. and Timmermans, H.J.P.: 1997a, Exploring the Possibilities of Conjoint Measurement as a Decision-making Tool for Virtual Wayfinding Environments, *Proceedings of The Second Conference on Computer Aided Architectural Design Research in Asia*, pp. 61-71

Dijkstra, J. and Timmermans, H.J.P.: 1997b, The Application of Conjoint Measurement as a Dynamic Decision Making Tool in a Virtual Reality Environment, *Proceedings of the 7th International Conference on Computer Aided Architectural Design Futures*, pp. 757-770

Glenstrup, A.J. and Engell-Nielsen, T.: 1995, *Eye Controlled Media: Present and Future State*, Thesis, University of Copenhagen, Copenhagen

Hall, T.W.: 1997, Hand-eye Coordination in Virtual Reality, using a Desktop Display, Stereo Glasses, and a 3-D Mouse, *Proceedings of The Second Conference on Computer Aided Architectural Design Research in Asia*, pp. 73-82

Louviere, J.J.: 1988, *Analyzing Decision Making: Metric Conjoint Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no.07-067. Beverly Hills: Sage Publications

Noton, D. and Stark, L.W.: 1971, Scanpaths in Eye Movements during Pattern Perception, *Science*, **171**, 308-311

Timmermans, H.J.P.: 1984, Decompositional Multiattribute Preference Models in Spatial Choice Analysis: A Review of Some Recent Developments, *Progress in Human Geography*, **8**, 189-221

Weber, R. 1996: Oculomotor Research in Architecture: A Proposal for a Theory of Visual Centers in Architectural Spaces bases on Eye-Movement Patterns, *Handout of paper presented at the 3rd Design and Decision Support Systems in Architecture and Urban Planning Conference*, Spa