Detection and Processing of Visual Information in Three-Dimensional Space

ANTJE HERBON & MATTHIAS RÖTTING

Center of Human-Machine-Systems, Berlin University of Technology, Germany

Keywords: three-dimensional, depth, spatial interaction, spatial cognition, visual information presentation, visual space, attention, task-specificity

Abstract

This paper reviews literature on visual information presentation in depth. A theoretical model of 3-D spatial interactions by Previc (1998) is presented. Fundamental research on spatial attention and applied studies in the fields of aviation and automobile driving are described, which were concerned with the task-specificity of information presentation in space. Necessary research is identified and an experimental set-up for intended future studies is presented.

1. Introduction

When Human Machine Interaction (HMI) in virtual or augmented reality is concerned, frequently asked questions are: How are objects to be implemented? How do users view the environment? How will they interact, i.e. what kind of tools are to be used? The question of *where* in three-dimensional space, especially in depth, information should be presented or interaction should take place has been rather neglected, excepting aviation and automobile driving. In those fields of application, many researchers have been concerned with the utilization of either head-up displays (HUDs) or head-down displays (HDDs) for information presentation (e.g. Foyle, Sanford & McCann 1991, Horrey, Alexander & Wickens 2003, Liu & Wen 2004).

This paper will present evidence, which indicates that the matter of *where* information should be presented in depth, is worth investigating. First, a neuropsychologically based model of spatial interaction by Previc (1998) will be introduced. Secondly, results from empirical studies examining attentional issues in depth in both, laboratory and field settings, will be presented. Finally, open research questions are discussed and an experimental set-up for future studies is presented.

2. A theoretical model of 3-D spatial interactions

When considering information presentation in three-dimensional space, it is of importance to make assumptions on whether detection and processing of information differ according to location in space. Integrating neuropsychological work on three-dimensional spatial interaction, Previc (1998) suggested the division of the space surrounding us into four realms (see figure 1). Each realm has specific properties concerning neurological processing of stimuli and reactions, influencing human perception and behavior.

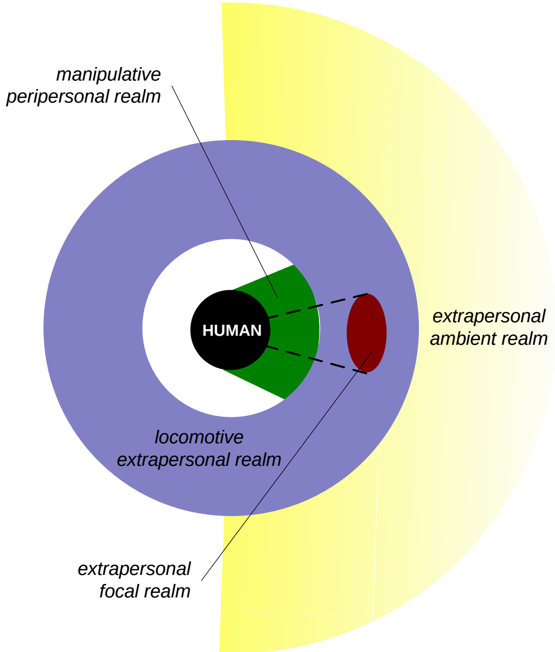

Figure 1: A theoretical model of 3-D spatial interactions, adapted from Previc (1998)

The first realm is the peripersonal or – extended by May (2006) – the manipulative peripersonal realm. It extends from zero to two meters into depth and is used for visual grasping and object manipulation. Its lateral extent is 60° (central). The peripersonal realm is to be understood as the realm inside hand-reaching distance, even though it has been shown to be able to expand with tool use, when the manipulable region expands therewith (Berti & Frassinetti 2000). As far as visual perception is concerned, this realm is specialized for global form, depth and motion. There is a lower-field bias, meaning that attention is more readily allocated to and manual reaction times are faster in the lower field (Sheliga, Craighero, Riggio & Rizzolatti 1997).

The second realm is the extrapersonal focal realm. It radially (i.e. in depth) extends from 0.2 meters to the distance at which a respective object is no longer resolvable, while the lateral extent is 20-30° (central)[1]. It is mainly used for visual search and object recognition, which entails that areas and objects of interest at a certain loca-

[1] Note that the red ellipse represents a focused object/position, while the dashed line depicts the lateral border of the realm.

tion inside this realm are focused for identification and classification. As far as visual perception is concerned, this realm is specialized for color and form, including high-resolution contour analysis. There is an upper field bias.

The third realm is the extrapersonal action or – as May (2006) called it – locomotive extrapersonal realm. It extends from two to 30 m or more in the full 360°. Positions in this realm can be reached by walking or other movements of the body. It is mainly used for navigation, target orientation (incl. motion-based actualization of object location) and scene memory. There is an upper field bias.

The fourth realm is the extrapersonal ambient realm. It extends up to a few kilometers and is mainly used for spatial orientation, postural control and locomotion beyond the locomotive extrapersonal realm. Its lateral extent is 180° and there is a lower field bias.

This model of three-dimensional space has been examined and supported mainly by neuropsychological studies (e.g. McCourt and Garlinghouse 2000; Weiss et al. 2000; see Halligan, Fink, Marshall and Vallar (2003) for a review). Studies in the field of cognitive psychology are rather scarce. However, literature from HMI-research, especially aviation and automobile driving reports many studies on information presentation in different depth-planes. There, focal and ambient vision are mainly differentiated and can alternatively be interpreted as peripersonal, focal extrapersonal or ambient extrapersonal realms most of the time. Some of those studies are being presented in the following.

3. Empirical Literature on Visual Attention in Depth

There are quite a number of laboratory studies which were concerned with the spread of attention in depth. Drive for most of this research was the question of whether attention allocation differs depending on the number of dimensions (two or three) that are concerned. However, just a few experiments directly examined allocation of attention in (and between) peripersonal and extrapersonal space.

3.1 Fundamental research on attention allocation in depth

It has been shown that attention does indeed spread in depth and is not limited to two-dimensional space: Atchley, Kramer, Andersen and Theeuwes (1997) found that reaction times were slower when subjects had to switch attention in x-, y-, and z-dimension (depth) than only in x- and y-dimensions. However, this effect was only apparent when distractors, i.e. increased perceptual load, were present.

Many researchers (e.g. Gawryszewski, Riggio, Rizzolatti & Umilta 1987; Andersen & Kramer 1993; Kimura, Miura, Doi & Yamamoto 2002) reported an asymmetry in the allocation of attention in depth where subjects were able to switch attention faster from far to near objects than from near to far objects. Arnott and Shedden (2000) found that this viewer-centered asymmetric depth gradient is dependent on perceptual load, i.e. it is not apparent when perceptual load is low because in this case a narrow attentional focus is not necessary.

As in two-dimensional space, attention in three-dimensional space can be either object- or space-based; a trade-off between the two possibilities is suggested (Atchley & Kramer 2001).

3.2 Peripersonal and extrapersonal space

Couyoumdjian, di Nocera and Ferlazzo (2003) conducted three experiments on allocation of attention within and between peripersonal and extrapersonal space. They presented four pairs of LED-cubes, one on the left and one on the right, at 40, 80, 120 and 160 cm from the observer and cued their onset either validly (at the same location) or invalidly (at a different location). They found that reaction times were significantly faster when invalidly cued and target locations were in the same realm than when subjects had to shift their attention across realm boarders, distances between cue light and target cube being equal. Results remained consistent when fixation point and target distances where manipulated also. These findings strongly suggest that – on top of the time needed to switch attention between two points in depth – there is an added cost when attention has to be switched between two perceptual realms.

In spite of this supportive evidence on perceptional issues, there have been contradictive results on the behavioral part of the model of 3-D space presented above. Schoumans, Kappers and Koenderink (2002) could not find any differences in a pointing task in 40 and 120 cm distance from their subjects. Moreover, they replicated systematic context-based errors in both distances. However, it would be of interest if using the pointer lead to an extension of the manipulative peripersonal realm and if effects were due to the fact that the two distances were actually part of one instead of two realms.

3.3 Applied studies in aviation and automobile driving

With the technical maturation of head-up displays (HUDs), research on information presentation in aviation and automobile driving has increased greatly. HUDs provide the opportunity of moving information from displays inside a vehicle to the windshield, resulting in reduced eyes-off-the-road time. Comparative research on HUDs and conventional head-down-displays (HDDs) has lead to the assumption that the optimal location of information presentation is task-specific.

Summarizing previous simulator studies, Horrey and Wickens (2004) suggested, that certain combinations of multiple tasks associated with operating a vehicle can be time-shared more efficiently than others. They proposed that the reason for this difference is that some tasks utilize focal, whereas other tasks utilize ambient vision. The multiple resources model (Wickens 2002) would then predict that two focal tasks will interfere and lead to poorer performance while a focal and an ambient task can be completed in parallel. Tasks utilizing ambient vision thereafter are lane-keeping and speed control, whereas hazard detection utilizes focal vision. This finding is very much in line with Previc's (1998) model, where the extrapersonal ambient realm is mainly used for spatial orientation and the extrapersonal focal realm is very sensitive for object recognition.

However, this generalization can be problematic. Liu and Wen (2004) for example showed in a goods delivery task with commercial vehicle operators, that reacting to urgent events and *speed keeping* require focal vision and do therefore interfere with a side-task that requires focal vision also. This discrepancy to Horrey's and Wickens' findings can be explained by the specifics of the presentation of the information that was needed in order to complete the task: Speedometer information was presented in form of a number on an HUD or an HDD and could therefore not be perceived using ambient vision, since this high-resolution form identification utilizes focal vision. Thus, speed-maintenance using information on the outside-world moving by (as in Horrey, Alexander & Wickens 2003) is very different from speed-maintenance as operationalized by Liu and Wen. The importance of this distinction was also apparent in a study conducted by Foyle, Sanford & McCann (1991), who had subjects complete a flight-task and presented altitude information necessary for an altitude-maintenance task in two different ways: as digital information on an HUD and as indirect source of information using sketched buildings super-imposed on the sides of the flight path. Results indicated that reading digits utilizes focal, while using the sketched buildings as height information utilizes ambient vision. In this study, altitude-maintenance was competitive to path-keeping, which most interestingly utilized focal vision. However, this is likely to be explained by the curvature of the path, because workload has been shown to modulate the allocation of visual resources (Horrey, Alexander & Wickens 2003).

Recently, Crawford and Neal (2006) have reviewed selected literature dealing with perceptual and cognitive issues associated with HUDs in aviation. They identified cognitive tunneling as one of the main attentional problems in the use of HUDs, which causes an impairment of the pilots' ability to detect events outside their vehicle because of their attention being captured by the information on the windshield. This is especially important when two tasks both depending on focal vision are competing for limited resources. Levy, Foyle and McCann (1998) found that linking HUD-symbology to the outside world (i.e. displaying information as if it were located on the flight path, for example, instead of on the windshield) can solve this problem by directing attention to a different depth. In their experiment, it did not matter where exactly on the path the symbology – in this case an analog gauge – was located. Performance on the main focal task was always better than in the traditional HUD-situation, where information was projected directly onto the windshield without additional depth information.

The empirical literature cited above gives strong evidence for an allocation of attention in depth. Task-type, which supposedly influences attention has been investigated in applied HMI-studies, but not yet in fundamental research. However, in order to generalize findings to other different fields where Augmented Reality (AR) technology is utilized, further research on task-specificity in attention allocation is needed. For example, AR is widely-used in production (e.g. Reiners, Stricker, Klinker & Müller 1998; Sarval, Baker & Filipovic 2005), has been utilized for navigation (e.g. Biocca, Tang, Owen & Fan 2006) and has found its way into Smart Homes (e.g. Hammond, Sharkey & Foster 1996; Intille 2002). The Tangible Media Group at MIT Media Lab have designed a number of applications for social interaction (Chang, Resner, Koerner, Wang & Ishii 2001; Bonanni, Vaucelle, Lieberman & Zuckerman 2006), sports (Ishii, Wisneski, Orbanes, Chun & Paradiso 1999), infotainment (Ishii & Ullmer 1997; Ishii 2004) and work (Ishii, Wisneski, Brave, Dahley, Gorbet, Ullmer & Yarin 1998). Tasks in these applications differ a lot, but might still be classified according to their requirements on information detection and processing.

4. Task-Specificity

Lacey and Lacey (1970) have described different stressors they used in physiological studies, which produced task-specific physiological response patterns. Mental arithmetic, reversed spelling, making up sentences and noxious stimulation lead to an increase in heart rate and heart rate variability, while attending to photoic flashes, white noise or a dramatic recitation resulted in a decrease of both parameters. The authors characterized the reaction to the two groups of stimuli as *rejection* and *intake of the environment*, respectively. In the first group with cardiac acceleration, tasks require internal cognitive elaboration (calculating, putting together letters and words) and in the case of the cold pressure test (noxious stimulation) simply the suppression of the unpleasant feeling. Perception of the environment is therefore not necessary and not wanted, thus the environment is *rejected*. In the second group with cardiac deceleration, tasks require attention on visual and auditory stimuli in the environment, i.e. the environment is *taken in*.

This classification can be compared to Norman's (1993) notion on experiential and reflective cognition. Tasks that are completed in experiential mode require environmental intake. A person has to perceive and/or react to stimuli from the outside world. Responses in this mode of cognition are automatic and not reflected upon. By contrast, reflective cognition utilizes conscious engagement in the task, e.g. planning or decision making. Once input from the outside world is attained, environmental intake is no longer needed but instead hinders and detracts from the problem solving process.

5. Intended Research

Task-specificity of attention allocation in depth is to be examined further. A main research goal is to draw implications as to where to present information in augmented and virtual realities independent of a specific application-field. Lacey's and Lacey's (1970) and Norman's (1993) notions on task classification will form a basis for experimental task design.

Previc's model on 3D-spatial-interaction will be used in order to define perceptual realms in an experimental set-up. It has been shown to be in line with previous experimental laboratory findings on allocation of attention in depth and can be used to interpret HMI-research results also. Even though cited applied-studies were not based on the model, they did give evidence for the division of the space surrounding us into different perceptual realms.

An experiment is currently being designed to investigate coherences between task-characteristics as described above (tasks with different requirements on visual perception as in the applied studies, reflective vs. experiential tasks) and perception of information in depth. As shown in figure 2, different tasks will be presented in two different depths, one on a transparent screen close to the observer and a second one on a far projection screen, several meters away from the observer.

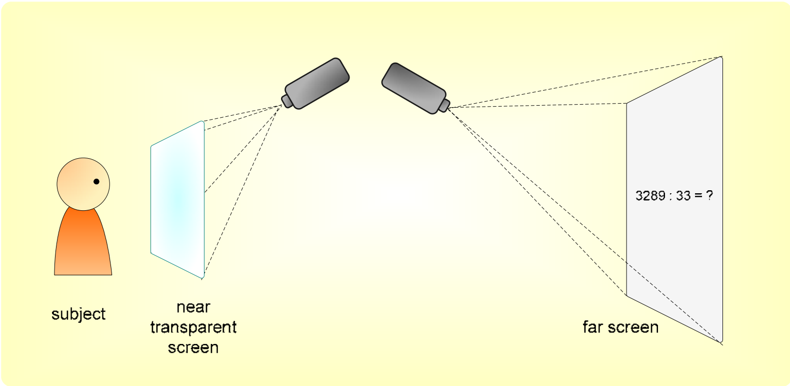

Figure 2: Intended experimental set-up

Suitable tasks will be identified and classified according to their requirements on visual perception and reflective thought. Elements from all groups of tasks will then be presented in both depths. Parameters to be inspected are reaction times, error rates and workload. Additional collection of physiological parameters, such as eye movement for investigation of attention allocation or stress-level indicators such as skin resistance or heart rate is also possible.

Acknowledgements

This research is supported by the German Research Foundation (DFG) as part of the Research Training Group 'Prospective Design of Human-Technology Interaction' (no. 1013).

References

Andersen, G.J. & Kramer, A.F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics*, 53 (6), 658-667.

Arnott, S.R. & Shedden, J.M. (2000). Attention switching in depth using random-dot autostereograms: Attention gradient asymmetries. *Perception & Psychophysics*, 62 (7), 1459-1473.

Atchley, P. & Kramer, A.F. (2001). Object and space-based attentional selection in three-dimensional space. *Visual Cognition*, 8 (1), 1-32.

Atchley, P.; Kramer, A.F.; Andersen, G.J. & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a „depth-aware“ attentional focus. *Psychonomic Bulletin & Review*, 4 (4), 524-529.

Berti, A. & Frassinetti, F. (2000). When Far becomes Near: Re-mapping of Space by Tool Use. *Journal of Cognitive Neuroscience*, 12 (3), 415-420

Bonanni, L.; Vaucelle, C.; Lieberman, J. & Zuckerman, O. (2006). PlayPals: Tangible Interfaces for Remote Communication and Play. *Proceedings of CHI '06*.

Chang, A.; Resner, B.; Koerner, B.; Wang, XC & Ishii, H. (2001). LumiTouch: An Emotional Communication Device. *Extended Abstracts of Conference on Human Factors in Computing Systems*.

Couyoumdjian, A.; di Nocera, F. & Ferlazzo, F. (2003). Functional representation of 3D space in endogenous attention shifts. *The Quarterly Journal of Experimental Psychology,* 56A (1), 155-183.

Crawford, J. & Neal, A. (2006). A Review of the Perceptual and Cognitive Issues Associated With the Use of Head-Up Displays in Commercial Aviation. *The International Journal of Aviation Psychology*, 16 (1), 1-19.

Foyle, D.C.; Sanford, B.D. & McCann, R.S. (1991). Attentional Issues in Superimposed Flight Symbology. *Proceedings of the Sixth International Symposium on Aviation Psychology*, 577-582.

Gawryszewski, L.D.G.; Riggio, L.; Rizzolatti, G. & Umilta, C. (1987). Movement of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia*, 25 (1A), 19-29.

Halligan, P.W.; Fink, G.R.; Marshall, J.C. & Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *TRENDS in Cognitive Sciences*, 7 (3), 125-133.

Hammond, J.C.; Sharkey, P.M. & Foster, G.T. (1996). Integrating augmented reality with home systems. *Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech.*

Horrey, W.J.; Alexander, A.L. & Wickens, C.D. (2003). Does Workload Modulate the Effect of In-Vehicle Display Location on Concurrent Driving and Side Task Performance? *DSC North America Proceedings*.

Horrey, W.J. & Wickens, C.D. (2004). Focal and Ambient Visual Contributions and Driver Visual Scanning in Lane Keeping and Hazard Detection. *Proceedings of the Human Factors and Ergonomics Society 48th annual meeting*.

Ishii, H. (2004). Bottles: A Transparent Interface as a Tribute to Mark Weiser. *IEICE Transactions on Information and Systems,* E87-D (6), 1299-1311.

Ishii, H. & Ullmer, B. (1997). Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Proceedings of CHI '97.*

Ishii, H.; Wisneski, C.; Brave, S.; Dahley, A.; Gorbet, M.; Ullmer, B. & Yarin, P. (1998). ambietROOM: Integrating Ambient Media with Architectural Space. *Proceedings of CHI '98*.

Ishii, H.; Wisneski, C.; Orbanes, J.; Chun, B. & Paradiso, J. (1999). PingPongPlus: Design of an Athletic-Tangible Interface for Computer Supported Cooperative Play. *Proceedings of CHI '99.*

Intille, S. S. (2002). Designing a Home of the Future. *IEEE Pervasive Computing*, 1, 76-82.

Kimura, T.; Miura, T.; Doi, S. & Yamamoto, Y. (2002). *Top-down and Bottom-up allocation of attention in three-dimensional space when observers are moving forward*. Technical Report on Attention and Cognition No. 16, http://L.u-tokyo.ac.jp/AandC/

Lacey, J.I. & Lacey, B.C. (1970). Some Autonomic-Central Nervous System Interrelationships. In: P. Black (ed.). *Physiological Correlates of Emotion*, 205-227. New York: Academic Press.

Liu, Y.-C. & Wen, M.-H. (2004). Comparison of head-up display (HUD) vs. Head-Down-Display (HDD): driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies*, 61, 679-697.

May, M. (2006). Raum. In: J. Funke & P.A. Frensch (eds.). *Handbuch der Allgemeinen Psychologie - Kognition*, 66-74. Göttingen: Hogrefe.

McCourt, M.E. & Garlinghouse, M. (2000). Asymmetries of visuospatial attention are modulated by viewing distance and visual field elevation: Pseudoneglect in peripersonal and extrapersonal space. *Cortex*, 36, 715-731.

Norman, D.A. (1993). *Things that make us Smart*. Reading, MA: Addison-Wesley Publishing Company.

Previc, F.H. (1998). The Neuropsychology of 3-D Space. *Psychological Bulletin*, 124 (2), 123-164.

Reiner, D.; Stricker, D.; Klinker, G. & Müller, S. (1998). Augmented Reality for Construction Tasks: Doorlock Assembly. *Proceedings of the 1st International Workshop on Augmented Reality (IWAR'98).*

Sarval, A.; Baker, C. & Filipovic, D. (2005). Head-Worn Display-Based Augmented Reality System for Manufacturing. *Proceedings of SPIE 2005*.

Schmidt, L.; Wiedenmaier, S.; Oehme, O. & Luczak, H. (2005). Benutzerzentrierte Gestaltung von Augmented Reality in der Produktion. In: C. Stary (ed.). *Mensch & Computer 2005: Kunst und Wissenschaft – Grenzüberschreitungen der interaktiven ART*, 51-60. München: Oldenbourg Verlag.

Shoumans, N.; Kappers, A.M.L. & Koenderink, J.J. (2002). Scale invariance in near space: pointing under influence of context. *Acta Psychologica*, 110, 63-81.

Sheliga, B.M.; Craighero, L.; Riggio, L. & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research*, 114, 339-351.

Weiss, P.H.; Marshall, J.C.; Wunderlich, G.; Tellmann, L.; Halligan, P.W.; Freund, H.-J.; Zilles, K. & Fink, G.R. (2000). *Brain*, 123, 2531-2541.

Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3 (2), 159-177.