printed in:
MONTPELLIER INFORMATIQUE '95
4th International Conference
INTERFACE TO REAL & VIRTUAL WORLDS
Proceedings, Montpellier – France, June 26-30, 1995
ISBN: 2-2-910085-15-5
EC2&Cie – 9, rue Denis Poisson – 75017 PARIS – France
Tel: (33-1) 40 68 12 12 – Fax (33-1) 40 68 12 11

A DESIGN CONCEPT FOR N-DIMENSIONAL USER INTERFACES

Matthias Rauterberg (1) and Kornél Szabó (2)

(1) Usability Laboratory
Work and Organizational Psychology Unit
Swiss Federal Institute of Technology (ETH)
Zurich, Switzerland

(2) MultiMedia Laboratory
Department of Computer Science
University of Zurich
Switzerland

Abstract: We describe a classification concept, that allows us to discriminate four different domains: the dimension of the internal world model, presentation effects based on different techniques, perception mechanism, and the conceptualization of the world's dimensions in the user's mental model. This classification concept is applied to visual depth cues and acoustical signals. The user's perceptual feeling of being immersed in the context of virtual interfaces can be achieved by different presentation effects based on several techniques.⁄A major problem in designing n-D user interfaces is the fact, that designers do not have any metrics or benchmarks for applying the optimal software and hardware setting given a certain task. Different experiments for the visual information processing channel show, that, for example, depth cues for monocular perception are often sufficient versus binocular depth cues (often called stereo). Experiments show that combining different presentation and inter–action techniques in a synergetic way may give further advantages for the effectiveness of n-D user interfaces.

KEYWORDS: classification concept, user interface design, virtual reality, interactive 3D simulation, human computer interaction, empirical validation.

INTRODUCTION

Progress in information visualization, human-computer interaction, sensing technologies, integrated multimedia platforms and increased computing power extend traditional user interface metaphors towards multi-dimensional and multi-modal user interface concepts. As pointed out in [11] and [24], there is an increasing interest for spatial displays and instruments. There is a need for formal and empirical user studies to verify and expand on n-D user-interface concepts. The classification of 2-D, 2.5-D, or 3-D is not sufficiently when we take modern interface types into account (e.g., multimedia, multimodality, virtual reality). If we design more than one input and output channel, then we must pay attention to inter and intra modal dependencies [7], [16].

We need a deeper understanding in information representation technologies to produce optimal conceptual effects and new metaphors for fully n-D computer generated user-interfaces. Such n-D user-interfaces may allow the designer an effective use of design space and may help in better visualizing complex relationships between information units, and in consequence better understanding of complex data. Also the taxonomies used in traditional information processing design techniques have to be over thought regarding new potentials offered by modern simulation technologies.

DESIGN CONCEPT

As a first approach, to work out the parameters for immersion and their effectiveness on human perception, we point out a design concept, that allows us to discriminate four different domains. We found for the designer's side: (1) The computer internal presentation of scenes or objects, typically characterized by internal data structures or topological dimensions, and (2) presentation effects based on different techniques to visualize and to make the objects heard. We found for the user's side: (3) Human visual perception mechanisms of the presented objects and environments, and (4) the conceptualization of the world's dimensions in the user's mental model (see Figure 1 and 2).

Different transformation processes relate these dimensions to each other by [3], [13], [15]. The user's perceptual feeling of being immersed in the context of new multi-dimensional user interfaces can be achieved by different techniques. On the user's side psychological and physiological factors affect the success of viewing and interpreting.

Therefore we have to notice, that some people for example can not see stereo and a proportion of males is red-green color blind. This fact leads to a very important question: Is binocular disparity necessary for future n-D metaphors to support the user with n-D mental models? The results of empirical investigations presented in this paper try to give first answers to the above question. This is a first step to propose a metric suite to n-D environments with a maximum of immersion.

VALIDATION OF THE CONCEPT

Human perception consists of several channels: visual sense, auditory sense, smell, taste, skin sensation, kinesthesis and equilibratory senses. The visual and auditory sense uses information coming from the kinesthesis and equilibratory sense. Before we describe the visual and the auditory channel on a more detailed level, we will introduce the function of kinesthesis and equilibratory senses.

Our ordinary vocabulary lacks a word for the sensory system that informs us of the position and movement of parts of the body. Kinesthesis is the muscle, tendon, and joint sense (e.g., force perception). Without kinesthesis we would have great difficulty in maintaining posture, walking, climbing, and in controlling voluntary movements such as reaching, grasping, and manipulating (see for example [8]).

Cooperating with kinesthesis are the equilibratory senses, which deal with total body position in relation to gravity and with motion of the body as a whole. The relation of bodily parts to one another and to external objects is the responsibility of kinesthesis; the orientation of the body in space is the responsibility of the equilibratory senses. The equilibratory senses also signal accelerated motion in a straight line, but sometimes they produce illusions that distort the true path of motion. These illusions occur in flying, because of changes in speed and the banking and climbing of the plane. For example, when a plane is increasing its speed gradually, a blindfolded subject may feel sure that the plane is climbing; if its speed is decreasing gradually, he/she may feel equally sure that it is diving.

How important and strong is the influence of the kinesthesis and equilibratory senses of an actively moved subject to his/her visual and auditory information processing? This influence was demonstrated and proved by Held and Freedman [9]. Disturbances of this correlation can be the reason for vertigo and dizziness. On the other side, it is also possible to recover the three-dimensional structure from motion of observed objects. If a shadow is cast by rotating wire shape onto a screen, a passive viewer can readily perceive the shape of the structure behind the screen from the dynamic shadow pattern [29]; when static, the screen looks like a random collection of lines.

Visual signals

When the observer or an object moves, certain higher-order characteristics of the optic array remain invariant while others change. These invariants over time specify the layout of the environment precisely. The observer perceives simple by 'picking up' these invariances [15].

Definition of basic terms:

Light: The visible portion of the electromagnetic spectrum.

Color: The different qualities as seen when sunlight is sent through a prism (chromatic colors), and additionally black and white (achromatic colors).

Contrast and Sensitivity: The retina is designed to keep the response range of the visual system in register with ambient illumination, thus enabling the retina to form high-contrast neural images over a broad range of light conditions (e.g., dark adaptation). Two types of neural signals can be encoded by the visual system: a change-sensitive, transient signal as well as a level-sensitive (sustained) signal.

Feature extraction: The visual system has line, edge, and corner detectors. Other units analyze the difference between a coarse and a fine pattern. Texture perception includes higher-level processes. Surface perception is not only defined by its edges, but more importantly by its lightness, color, or texture.

Control Mechanisms in Localization: To put objects into the primary attention focus (e.g., the central region of the retina, called fovea), a human makes extremely rapid and coordinated movements of the head, eyes, and often even the limbs. For example, to hold the palm behind the ear conch brings sound into the auditory perception focus. There are at least two coordinating centers, one handling principally sensory input and the second concerned mainly with motor coordination.

Visual perception and looking:

To determine the point of visual attention, several studies measured eye movements. There are much unsolved problems to correlate eye movements with higher psychological processes. But, 'eyes as output' are one of the best empirical sources. Kahneman ([10] pp. 64-65) distinguishes three types of eye movements:

(1) *Spontaneous looking*, which is governed by collative features of signals (novelty, complexity, incongruity). Responses to such signals are 'enduring dispositions,' rooted in the innate tendency to respond to contours, and toward moving objects.

(2) *Task-relevant looking* is viewed as an allocation problem. It is a characteristic of the eye in that it has sharp vision at its center or fovea, while peripheral vision is increasingly less distinct on outwards. Parafoveal vision is very sensitive to movements. Sharp vision occurs in sequential glances. The problem of where next to look is resolved through the interaction of task constraints and the visual context.

(3) Looking is a function of the *changing orientation of thought*. Eye movements of this type seem to reflect the overall transitions between stages of thought, even when the location, where the human is looking, cannot possibly offer any 'new' information. The eye movements during thought seem to be related to the balance of activity between the two hemispheres, the rate of mental activity generally.

All these three types of visual information gathering strategies are one reason that we have a closed loop between the 'perception mechanisms' and the 'concept of the world' (see Figure 1 and 2).

Experimental investigations:
We use our design concept (Figure 1) to describe empirical results in a more detailed way.

Ware et al. [26] could show, that motion perspective is more important than stereo. The best combination is head coupling with stereo. Monocular depth cues only based on vde:{1.3, 1.4, 1.5} scored low in error performance. The first step of user improvement (33%) could be reached adding vde:{2.4}. The monocular depth cues with additional head coupling (vde:{4.1}) improved user performance by 83%. The greatest improvement (94%) was measured with vde:{1.3, 1.4, 1.5, 2.4, 4.1}.

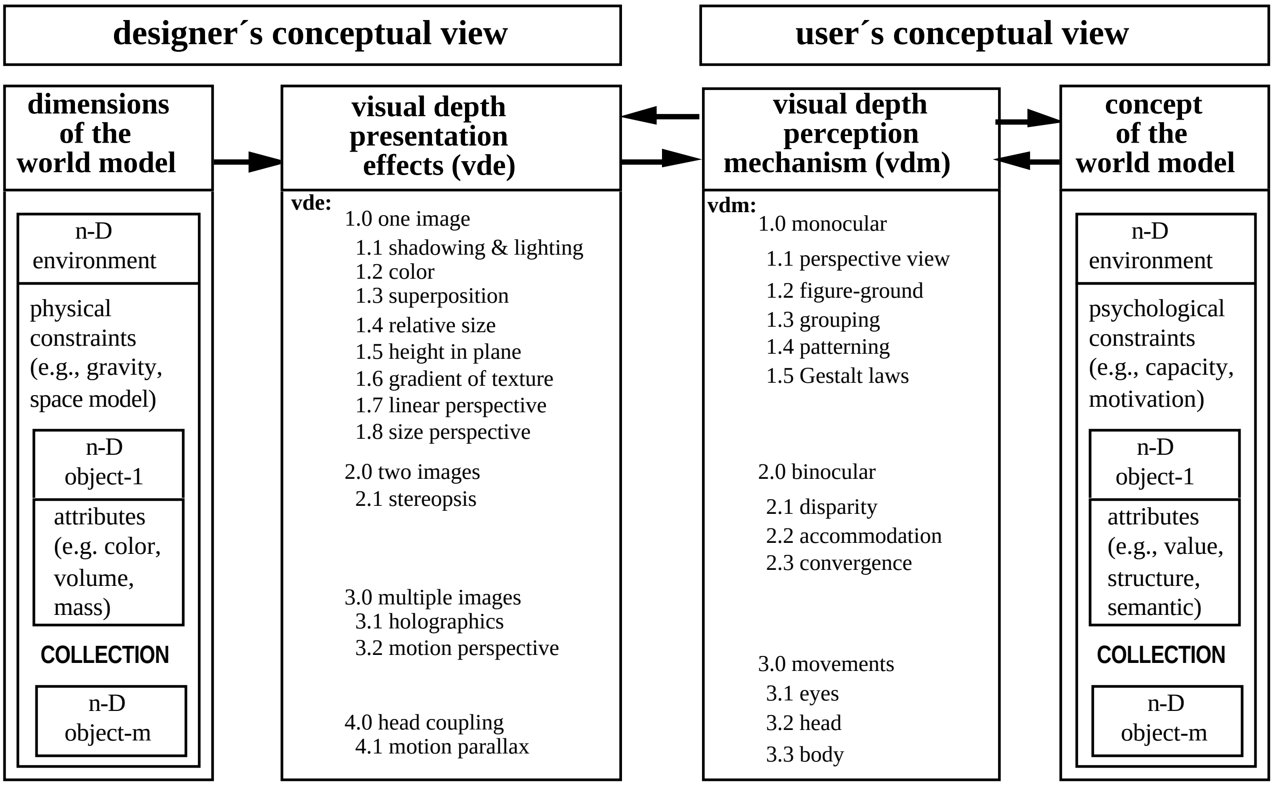

Figure 1. A design concept for the visual channel for an n-D user interface (see also [23]).

In one of our own experiments we could not proof the hypothesis, that 3-D perception based on anaglyph (vde:{1.4, 1.5, 2.1}) is better than a 2.5-D presentation only with monocular depth cues (vde:{1.4, 1.5}). For a more detailed discussion of this empirical investigation see [22].

Pfeffer et al. [18] could show the following improvements of distance estimations for near distance areas (6–8 m): 18% for vde:{1.3, 1.5, 1.6, 1.7, 1.8} compared with vde:{1.3, 1.5, 1.7, 1.8}; 26% for vde:{1.2, 1.3, 1.5, 1.7, 1.8} compared with vde:{1.3, 1.5, 1.7, 1.8}; 35% for vde:{1.3, 1.5, 1.7, 1.8, 2.4} compared with vde:{1.3, 1.5, 1.7, 1.8}.

On the other hand, Neisser [15] could show that the quality of perception is strongly influenced by movements (vdm:{3.1, 3.2, 3.3}). To demonstrate the importance of motion for visual perception the following experimental setting was used by Cornsweet [1]: A tiny slide projector was mounted on a contact lens attached to the cornea. A slide is projected onto a screen, and the eye wearing the lens looks at the image. Since the lens and the projector move with the eye, the image presented to the retina is stabilized; that is, the retinal image impinges on the same retinal receptors regardless of eye movements. When the projector is first turned on, the subject sees the projected figure with normal, or slightly better than normal, visual acuity. Within a few seconds the image begins to fade and within a minute disappears altogether. Changes in illumination on receptors are necessary for us to perceive objects (vde:{2.4, 3.4, 4.1} or vdm:{3.1, 3.2, 3.3})!

Presentation techniques:
Depth information within a virtual environment can be achieved by different techniques. In our model, we look at the amount of computed images or viewpoints to present a virtual scene to the user. One-image techniques with only one viewpoint -- such as superposition or size perspective -- are most common to produce depth cues. Superposition means that near objects overlap objects fare away from the user's viewpoint. Size perspective means that objects nearer are bigger than those further away.

With new visual display technologies, stereoscopic viewing can be generated. The two images -- to produce the stereo effect -- are either presented simultaneously (head mounted displays; red-green glasses) or consecutively (polarized glasses; LCD shutter glasses).

Acoustical signals

Auditory perception is a temporally extended activity. Sound waves exist only in time. The listener continuously develops more or less specific readiness for what will come next, based on information he has already picked up.

The perception of auditory signal patterns in everyday life can come in very different forms: a car driving by on the street, a dripping faucet, the confusion of voices from a crowd of people, opera music, a plane flying by, the buzz of a travel alarm clock, the beeping of a wristwatch, etc.

All of these acoustical signals are divided into four categories: speech, music, sound and noise; sometimes noises and sounds are heard and grouped together. All of these categories are described sufficiently in physical terms through the mixing and superposition of different pitches, frequencies, volumes and sound duration.

The audio channel can be divided into two components: auditory verbal (spoken words) and auditory non-verbal (sound effects and music). The terms that we use later in this paper are defined as follows.

Definition of basic terms:

Audio signal pattern: description of all perceptible audio signals.

Speech: The description of all audio signals that have describable grammar structures.

Music: Complex audio signal pattern that has rhythmic describable structures.

Tone: Simple audio signal pattern with rhythmic describable structure.

Everyday sounds: Acoustical signals that have a complex, non periodic structure (e.g., noise).

Everyday sound perception:

One of the essential differences between these categories, however, lies in their semantics: Speech serves primarily to convey information, while music and noise can have a pleasant or an unpleasant influence on the emotions. For musicians and other people who are intimate with this area, music and noise have a comparable semantic and informative character as speech does for the normal citizen. Besides from music and noise, the listener is interested next in the possibility of undisturbed, context-free perception.

In contrast to music and noise, everyday sounds have a self standing characteristic; they are extremely context sensitive and event related [4]. Through the physical interaction of different objects in 3-D space sounds of everyday life are created and through propagation they become audible through the air. In comparison with music and noises, the semantically relevant dimension of sound lies not with the characteristic quality of the auditory signal pattern itself, but rather with the quality of the sound producing event as it respects the concerned object [14].

This difference leads us to the conclusion that sounds are interpreted differently than music based upon their quality. When listening to music we are primarily interested in the effect of music on us; while when hearing everyday sounds we are interested in the quality of the sound producing object and the accompanying circumstances (e.g., surrounding conditions, events, etc.). Of course, music can be heard from the perspective of every day use; in this case the listener pays attention to the nature and tune of the instrument in use, to the tempo, to the acoustic, to the place of performance, etc. This method of listening to music is dependent upon the listener's knowledge of this domain field; only someone who is experienced with music will be able to extract all the various aspects from a piece of music.

The average adult is, for the interpretation of everyday sounds, an expert with a large degree of knowledge from experience. This knowledge allows one to evaluate everyday sounds according to the following criteria for relevant information:

1) Information about the physical occurrence: we hear, if the fallen glass clinks or breaks.
2) Information about hidden events: when knocking on the wall, we can hear if it is hollow.
3) Information about dynamic changes: when filling a wine glass, we can hear when it is full and runs over.
4) Information about abnormal conditions: we hear, when the car engine ceases to function properly and runs irregularly.
5) Information about occurrences outside of the visual field: the sound of footsteps behind us 'tells' us if someone is approaching.

Listening to everyday sounds is based upon the perception of events and not upon the perception of sounds in and of themselves. This fact becomes clear in the following example:

Illustrative example:

A pen dropped upon a piece of paper from a height of about 15 cm created a different sound than when it is dropped upon the hard surface of a desk. An altogether different sound is created when a rubber eraser is dropped upon the paper or, respectively, on the desk.

The sound created in each case of the previous example is neither a characteristic of any of the participating objects (pen, rubber eraser, sheet of paper, desk surface) nor a characteristic of the occurrence 'dropped' itself. The four different sounds in the examples are, with an observation that holds true to the reality of the situation solely determined by their respective interaction and environmental conditions. Everyday sounds are therefore due to a lack of better descriptive possibilities, often described through the underlying occurrence.

We describe -- for example -- every impact sound as the result of one or several interactions between one or several solid objects at a certain place and in a certain environment. The attributes of every interaction influence the generated sound. Simultaneously, the participating objects, which take part in the sound generation process, can consist of different physical conditions (states of aggregation), materials as well as their configurations. All relevant attributes have an influence on the generated sound. The hearing of sounds in everyday life is based on the perception of events caused by an interaction and not on the perception of sounds as such.

Every sound is also a result of one or more interactions between two or more objects in a definite place and in definite surroundings and can be defined as the following [20]:

Sound = f(Interaction(s), objects, surroundings, place, time)

Every interaction possesses attributes that have an influence on the produced sound. At the same time the shared objects can participate in the production of sound from different aggregate conditions, materials, and even their configuration. The configurations of these materials possess attributes that also can have an influence on the produced sound.

Presentation techniques:

Echo, reverberation, phase differences, overtones, sound conduction routes, masking effects, inter-aural intensity differences, position and orientation of listener's head are important factors in the perception of auditory information. Especially the modeling and consideration of the audio source position and head related parameters (head position and orientation, anatomy of the pinna as a linear filter), intensity differences, temporal or phase differences between arriving signals play a significant role in spatial sound perception and generation using head related transfer functions (HRTF). Timbre is affected by the envelope of the sound signal (the rate of amplitude modulations).

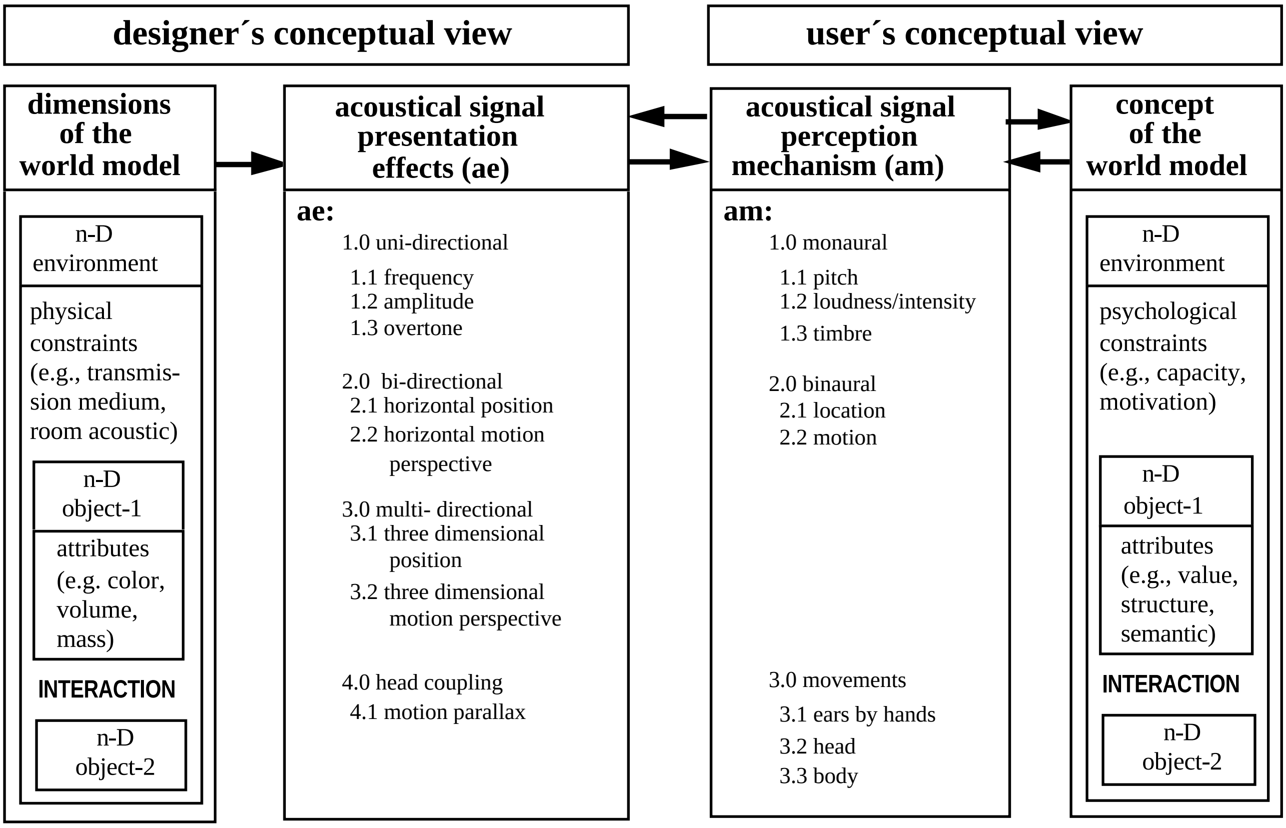

Figure 2. A design concept for the auditory channel for an n-D user interface.

Experimental investigations:
We use our design concept (Figure 2) to describe empirical results in a more detailed way.

Stevens and Newman [25] presented data on sound localization in a free field. The explanation maintains that non transient tones are localized through two sets of cues: for low frequencies, temporal cues are dominant; for higher frequencies, intensive cues are dominant (ae:{1.1, 1.2}). In the midrange, neither cue is effective and localization errors occur. For complex sounds like noise and impulses, both cues operate simultaneously.

However, the cues of Stevens and Newman [25] are not effective in localizing sound in a non-free field. In this case, head movements provide cue information (ae:{1.1, 1.2} and am:{3.2}) [27]; without movement, the time of arrival of the initial transients in the wave front provides a reliable cue of location when compared to the second (echo) transient [28]. The research on lateralization -- 'locating' the sound image within one's head when sound is presented dichotically with earphones -- has indicated that time and intensity cues are both operable in a complementary manner (ae:{1.1, 1.2, 3.1} and am:{2.1}).

We carried out an experiment to estimate the effect of sound feedback of hidden events [19], [21]. Eight students of computer science operated a process simulation program of an assembly line with computer numeric controlled (CNC) robots. Relevant information of disturbances and machine breakdowns was given only in a visual (vdt:{1.0} and vdm:{2.0}), and in visual and audible form (vdt:{1.0}, vdm: {2.0}, ae:{2.1} and am:{2.1}). The results indicate, that the additional sound feedback of hidden events improves significantly the operator performance and increases positively some mood aspects of the users.

Comparison of visual and acoustical signals

The textual representation of information is of most use when the user is familiar with the domain area and can demonstrate much experience and knowledge in that domain area. In comparison, more concrete (visual and auditory) representations of information that the user can query are of most use when the domain area is new and unknown [12]. By comparing audio signal patterns with visual signal patterns the different advantages of each can be shown (see Table 1).

Table 1. Guidelines for determining whether to adopt the visual or the audio channel in presenting information (see also [2] and [7]).

Use visual presentation if:	**Use auditory presentation if:**
1. Message is complex.	1. Message is simple.
2. Message is long.	2. Message is short.
3. Message will be referred to later.	3. Message will not be referred to later.
4. Message deals with location in space.	4. Message deals with events in time.
5. Message does not call for immediate action.	5. Message calls for immediate action.
6. Auditory system of the user is overburdened.	6. Visual system of the user is overburdened.
7. Receiving location is to noisy.	7. Receiving location is too bright or dark-adaptation integrity is necessary.
8. User's task allows him to remain in one position.	8. User's task requires him to move around continually.

Sounds and music can be used to improve the user's understanding of visual predecessors or can stand alone as independent sources of information. (For example: sounds as diagnostic support applied with the direction of a process simulation [7].)

The parallel use of different media and the resulting parallel distribution of information, for example by simultaneously showing a predecessor through a concrete representation and its explanation through audio distribution, leads to a denser sharing of information. In this case, the user can dedicate his attention solely to the visual information, which has parallel audio support. This reduces the need to change the textual or other visual delivery and prevents the overflow of visual information [5], [19].

The redundancy of information represented visually and auditory, as long as the representation of the information is realistically formed, is sensed not as disturbing, but instead it demands and increases information reception. It is important that with simultaneous information representation, that the information is harmonized together and that the different media are well synchronized [7].

CONCLUSIONS

Multiplicity of information is surely used in the act of perceiving. The mental schemata that accept information and direct the search for more of it are not visual or auditory or tactile, but perceptual [15]. The interdepencies between different perception channels are often neglected [7].

We could show that the differentiation of presentation effects on the output channel is helpful to describe and to categorize several empirical findings in a straight forward manner. For the description of empirical studies it would be of extreme value, if the test conditions would be classified in the context of our design concept.

The next step is to extend our design concept with other senses (tactile, etc.). It is not so easy to go this way , because we observed a substantial lack of empirical research in this area.

REFERENCES

[1] Cornsweet, T., Visual perception. New York: Academic, 1970.

[2] Deatherage, B., Auditory and other sensory forms of information presentation. in: Van Cott, H. & Kinkade, R. (Eds.) Human Engineering Guide to Equipment Design. Washington: U.S. Government Printing Office, 1972.

[3] Drösler, J., Das beidäugige Raumsehen. in: W. Metzger (Hrsg.) Handbuch der Psychologie, 1. Band: Allgemeine Psychologie, I. Der Aufbau des Erkennens, 1.Halbband: Wahrnehmung und Bewußtsein (S. 590-615) Göttingen: Hogrefe, 1974.

[4] Gaver, W., Auditory Icons: using sound in computer interfaces. *Human Computer Interaction,* 2: 167-177, 1986.

[5] Gaver, W., Smith, R., O'Shea, T., Effective sounds in complex systems: the ARKola simulation. in: S. Robertson, G. Olson & J. Olson (eds.), Reaching through technology CHI'91. (pp. 85-90), Reading MA: Addison-Wesley, 1991.

[6] Green, M., Virtual Reality User Interfaces: Tools and Techniques. in: T.S. Chua & T.L. Kunii (eds.) Computer Graphics International ´90 (pp. 51-68) Tokyo: Springer, 1990.

[7] Hartman, F., Single and Multiple Channel Communication: a Review of research and a Proposed Model. *Audio-visual Communication Review* 9: 235-262, 1961.

[8] Hashimoto, H., Kunii, Y., Buss, M., Harashima, F., Dynamic Force simulator for force feedback human-machine interaction. in: IEEE Virual Reality Annual International Symposium, (pp. 209-215), Piscataway: IEEE Service Center, 1993.

[9] Held, R., Freedman, S., Plasticity in Human Sensorimotor Control. *Science* 142(3591): 455-462, 1963.

[10] Kahneman, D., Attention and Effort. Englewood Cliffs: Prentice-Hall, 1973.

[11] Marcus, A., Graphic Design for Electronic Documents and User Interfaces. New York: ACM, 1992.

[12] Marmolin, H., Multimedia from the perspective of psychology. in: L. Kjelldahl (ed.) Multimedia: Systems, Interaction, and Applications. (pp. 39-52), Berlin, Heidelberg: Springer, 1992.

[13] Metzger, W., Das einäugige Tiefensehen. in: W. Metzger (Eds.) Handbuch der Psychologie, 1. Band: Allgemeine Psychologie I. Der Aufbau des Erkennens, 1.Halbband: Wahrnehmung und Bewußtsein (pp. 556-589) Göttingen: Hogrefe, 1974.

[14] Mountford, S., Gaver, W., Talking and listening to computers. in: B. Laurel & S. Mountford (eds.) The Art of Human-Computer Interface Design. (pp. 319-334), Reading MA: Addison-Wesley, 1990.

[15] Neisser, U., Cognition and Reality. San Francisco: Freeman, 1976.

[16] Ochsman, R., Chapanis, A., The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. *International Journal of Man-Machine Studies* 6: 579-619, 1974.

[17] Petigrew, J. D., The Neurophysiology of Binocular Vision. in: R. Held & W. Richards (eds.) Recent progress in perception (pp. 55-66) San Francisco: Freeman, 1976.

[18] Pfeffer, D., Christidis, A., Klärner, E., Untersuchung des Beitrages der Stereoskopie zur Tiefenwahrnehmumg. in: R. Möller (ed.) 2. Workshop Sichtsysteme (pp. 17-32). Berlin: Springer, 1991.

[19] Rauterberg, M., Styger, E., Positive effects of sound feedback during the operation of a plant simulator. in: B. Blumenthal, J. Gornostaev & C. Unger (Eds.) Human Computer Interaction. (Lecture Notes in Computer Science, Vol. 876, pp. 35-44), Berlin: Springer, 1994.

[20] Rauterberg, M., Motavalli, M., Darvishi, A., Schauer, H., Automatic sound generation for spherical objects hitting straight beams based on physical models. in: T. Ottmann & I. Tomek (Eds.) Educational Multimedia and Hypermedia. (Proceedings ED-MEDIA'94, pp. 468-473), Charlottesville: Association for the Advancement of Computing in Education, 1994.

[21] Rauterberg, M., Styger, E., Baumgartner A., Jenny, A. & de Lange, M., Additional Sound Feedback in Man-Computer Interaction: empirical investigations. in: A. Grieco, G. Molteni, E. Occhipinti & B. Piccoli (Eds.) Book of Short Papers of 4th International Conference on Work with Display Units. (Volume 3, pp. D18-D21), University of Milan: Institute of Occupational Health, 1994.

[22] Rauterberg, M., Kühni, M., Styger, E., Spiess P., 3-D versus 2 1/2-D: eine experimentelle Vergleichsstudie. Technical Report VR-2-93. Institut für Arbeitspsychologie, Zürich: Eidgenössische Technische Hochschule, 1993.

[23] Rhind, W. D., Spatial Data Handling in the Geosciences. in: A. K. Turner (ed.) Three-Dimensional Modeling with Geoscientific Information Systems (pp. 13-27). Dodrecht: Academic, 1992.

[24] Robertson, G. G., Card, S. K., Mackinlay, J. D., Information Visualization Using 3D Interactive Animation. *Communications of the ACM* 36(4): 57-71, 1993.

[25] Stevens, S., Newman, E., The localization of actual sources of sound. *American Journal of Psychology* 48:297-306, 1936.

[26] Ware, C., Arthur, K., Booth, K., Fish tank virtual reality. in: S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel & T. White (eds.) Conference on Human Factors in Computing Systems INTERCHI '93 (pp. 37-42). New York: ACM, 1993.

[27] Wallach, H., The Role of head movements and vestibular and visual cues in sound localization. J*ournal of Experimental Psychology* 27: 339-368, 1940.

[28] Wallach, H., Newman, E., Rosenzweig, M., The precedence effect in sound localization. *American Journal of Psychology* 52: 315-336, 1949.

[29] Wallach, H., O'Connell, D., The kinetic depth effect. *Journal of Experimental Psychology* 45: 205-217, 1953.