In: J.G. Verly (Ed.), *Enhanced and Synthetic Vision 2003*, SPIE-5081 (2003), pp. 58-65.
Bellingham, WA., USA: The International Society for Optical Engineering.

Additive and subtractive transparent depth displays

Frank L. Kooi[1] Alexander Toet
TNO Human Factors, PO Box 23, 3769ZG Soesterberg, The Netherlands

ABSTRACT

Image fusion is the generally preferred method to combine two or more images for visual display on a single screen. We demonstrate that perceptual image separation may be preferable over perceptual image fusion for the combined display of enhanced and synthetic imagery. In this context image separation refers to the simultaneous presentation of images on different depth planes of a single display. Image separation allows the user to recognize the source of the information that is displayed. This can be important because synthetic images are more liable to flaws. We have examined methods to optimize perceptual image separation. A true depth difference between enhanced and synthetic imagery works quite well. A standard stereoscopic display based on convergence is less suitable since the two images tend to interfere: the image behind is masked (occluded) by the image in front, which results in poor viewing comfort. This effect places 3D systems based on 3D glasses, as well as most autostereoscopic displays, at a serious disadvantage. A 3D display based on additive or subtractive transparency is acceptable: both the perceptual separation and the viewing comfort are good, but the color of objects depends on the color in the other depth layer(s). A combined additive and subtractive transparent display eliminates this disadvantage and is most suitable for the combined display of enhanced and synthetic imagery. We suggest that the development of such a display system is of a greater practical value than increasing the number of depth planes in autostereoscopic displays.

Keywords: image separation, image fusion, transparent displays, stereoscopic depth, situational awareness.

1. INTRODUCTION

The developments in display technology have resulted in the ability to visually present increasing amounts of data. One of the coming break-throughs is generally believed to be the introduction of "3-D displays": displays that create a true sense of depth. Although these types of displays already exist in a variety of forms, their application is still very limited. The two primary bottlenecks are the associated lack of viewing comfort and the interference with other tasks. In this paper we argue that the construction of "transparent depth" displays is an important new development because it does not suffer from either of these two drawbacks. We believe the time is near to introduce transparent depth displays in professional environments like the cockpit, command & control workstations, vehicles, and hand-held devices.

2. THE VISUAL SYSTEM: HOW 3-D VISION WORKS

Our two eyes in essence make two-dimensional images of the world, similar to photographs. The 3-D structure of the world around us therefore needs to be interpreted (mentally reconstructed) from the image pairs of the left and right eyes. This interpretation takes place in the visual part of the human brain, and not in the eyes themselves. 3-D vision can therefore be disturbed both by a problem with the eyes (unsuitable input imagery) or by a problem in the brain (erroneous interpretation).

2.1 The depth cues

The brain uses a number of *tricks* to make a 3-D interpretation of the 2-D input images from both eyes. These tricks are commonly referred to as "depth cues". Stating the process this way implies that the interpretation is not always correct. And indeed many if not most of the visual illusions are the result of an erroneous 3-D interpretation. Figure 1 shows an example of a visual depth illusion. The depth cues listed in Table I are essential for a thorough understanding of 3-D technology.

[1] Kooi@tm.tno.nl, www.tno.nl

Table 1. The four main categories of depth cues.

	Depth cue	Parameter unit
1	Stereopsis	Binocular disparity (deg)
2	Accommodation	Optical power (diopters)
3	(Motion) parallax	Relative position (deg)
4	Pictorial depth cues	—

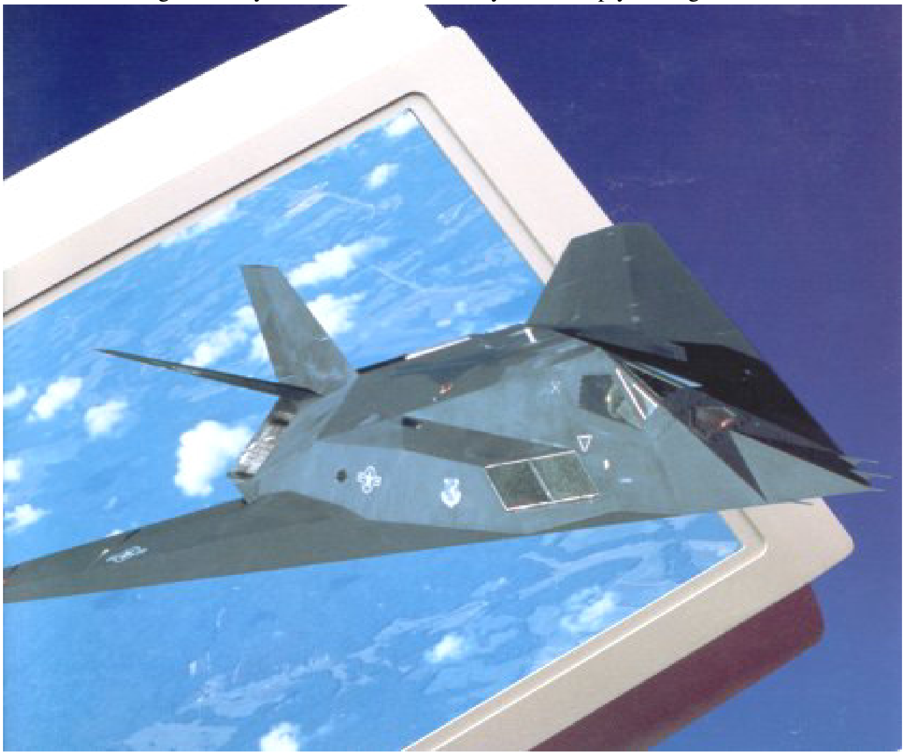

Figure 1. A depth illusion showing how easy it is to fool the visual system. Simply adding some cast shadows and occlusion to a 2D image already does the job. (From a Tektronix brochure.)

2.2 Convergence

Stereopsis is the result of viewing with two eyes rather than with one[1]. When an object that is close by is viewed, the eyes turn inward to fuse the object. This process is called convergence. Turning the eyes outward is called divergence. Vergence is neurologically coupled to accommodation. When converging, the eyes accommodate; when diverging the eyes relax the accommodation. The reverse is also true: when the eyes accommodate, they also tend to converge. This coupling is very convenient because it helps to prevent diplopia (seeing double) and blur. Objects that require convergence to fuse have a crossed disparity (the eyes need to cross). The visual system is able to convert this stereoscopic disparity to a sense of depth[1].

2.3 Accommodation

Accommodation is necessary because the eyes have a limited depth-of-focus. At any point in time only one distance is truly seen as sharp; everything in front and everything behind is blurred to some extent. While at first this may seem to be unfortunate, it in fact greatly helps to prevent visual attention from wavering all over the place. Figure 2 demonstrates this case. In the left image the foreground and background of the monitor are seen as sharp. In the right image, only the object of interest (the monitor) is seen as sharp as long as the two eyes are converged on it. It is much easier to view the display without being distracted by the foreground in this case.

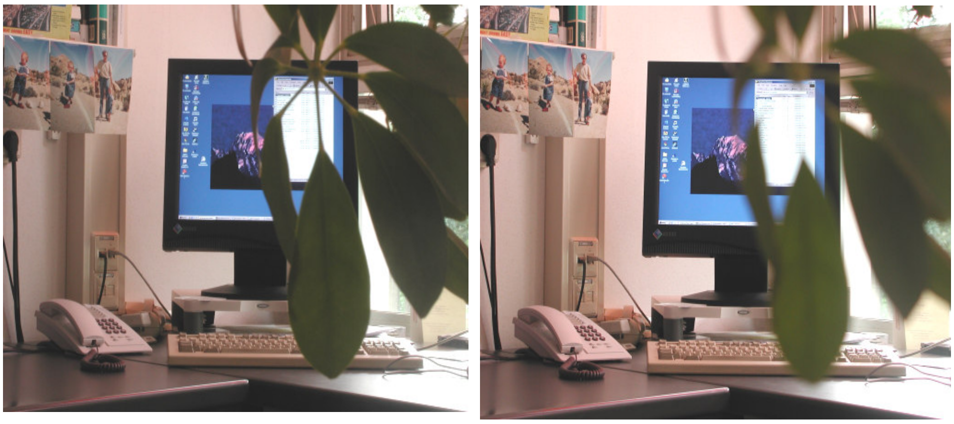

Figure 2. Illustration of the principle of accommodation. On the left: picture taken with a small aperture, resulting in a sharp foreground. On the right: same image taken with a wider aperture, resulting in a shallower depth-of-focus. In the right picture it should be easier to ignore the (fuzzy) leaves in the foreground.

2.4 (Motion) parallax

Moving the head sideways or up/down has two effects:

1. it provides a depth percept during the motion from the optic flow, and
2. it provides different points of view after the head motion is stopped.

The first (dynamic) effect is analogous to stereopsis. The second (static) effect is demonstrated in Figure 3 where part of the background objects can only be seen from the right point of view. Moving the head sideways is a natural part of our behaviour to get a better view.

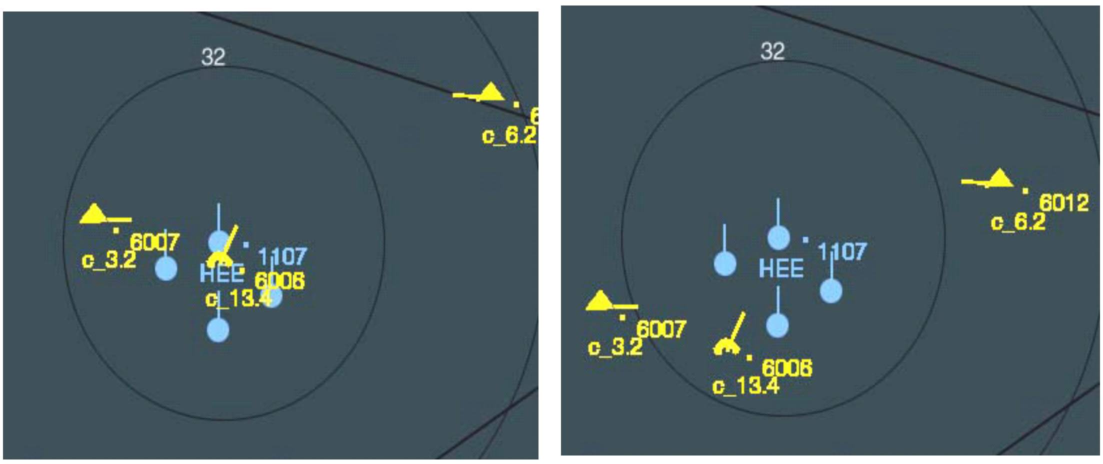

Figure 3. Illustration of parallax. On the left: the yellow foreground objects occlude the blue background objects. On the right: a (head) movement to the right and upwards makes all objects clearly visible, simulating the "de-cluttering function" of parallax.

2.5 Pictorial depth cues

Pictorial depth cues provide a depth perception when viewing the world with one eye closed and the other kept perfectly stationary. Pictorial depth cues are therefore also referred to as *monocular depth cues.* The depth that can be perceived in photographs is based exclusively on these cues. Examples of monocular depth cues are perspective, occlusion, and shading. The depth illusion shown in Figure 1 includes all three of these cues. While pictures can convey a strong sense of depth, it is also easy to find scenes whose 3-D layout cannot be perceived from a picture, whereas the 3-D structure immediately becomes apparent when stereo and/or parallax are available.

2.6 3D depth aids object detection

Scientific experiments have shown that depth is a potentially more powerful cue than color to help find an object[2]. A representative task is to find the RED O among RED X's and GREEN O's (color cue) or the in-front O among in-front X's and in-back O's (depth cue). When the number of "distractors" in the task that requires a color cue goes up, the search time goes up proportionally. This is called serial search[3]. When the number of distractors in the task that requires a depth cue goes up, the search time stays roughly constant. This is called parallel search and indicates that people are able to search within a single depth plane, while ignoring the other one(s). This ability provides a distinct advantage to an operator who knows where (in what depth plane) to find the desired information. Color coding does not provide the same ease of target detection and is in this sense inferior to depth coding [2,4-6].

3. THE TECHNOLOGY OF 3D DISPLAYS

Simply stated, 3-D technology adds the sense of depth by imitating one or more of the visual depth cues. Here we describe how the 3D technologies achieve this result. A good overview can be found at the following website: http://www.stereo3d.com/3dhome.htm. More detail on the human factors aspects can be found elsewhere [7].

3.1 Convergence

Convergence can be activated by presenting (slightly) different images to the left and right eyes. The left-right difference is called the stereoscopic disparity. The most common methods are shutter glasses, polarized glasses, Red/Green glasses, and Head-mounted displays. They share the common disadvantage of constraining the user. Most importantly, eye contact is disturbed, hampering communication with others. These devices are therefore not well suited for group activities.

To avoid the constraints imposed by wearing optics in front of the eyes, so-called "auto-stereoscopic 3-D displays" are being developed. The word "auto" signifies that the user does not need to wear an optical device. The optics are incorporated in the display, splitting the image in a left eye and a right eye component at the display instead. These optics are typically called "lenticular screens", and are glued to the flat-panel display. A lenticular screen consists of small lenses that bend the light from different display pixels in different directions. An inherent feature of auto-stereoscopic displays is that the head needs to be positioned at the right place. If for example the right eye is shifted 6 cm to the left, it will see the left eye image. Though solutions exist that allow some freedom of head movement, a price is paid in terms of an increase in cross-talk which reduces the viewing comfort[8] or in terms of added complexity in the form of a head tracker[9]. For a comparison of 3D methodologies on visual comfort see [10]. The disadvantages of the four main techniques can be summarized as follows:

1. Shutter glasses: low luminance, flicker in daylight environments
2. Polarised glasses: need to keep the head straight up
3. Red/Green glasses: no color vision, chromatic aberration, cross-talk
4. HMD's: image moves with the head, cables or weight

In the next chapter we will also describe a fifth technique (transparency).

3.2 Pictorial depth cues

Displays that contain symbols are oftentimes incompatible with the use of monocular depth cues because these cues tend to interfere with the clarity and standardization of the symbols.

3.3 Accommodation and parallax

The 3-D displays described above simulate the convergence depth cue but do not provide accommodation and parallax, which means that the depth percept is incomplete. Parallax can be added by tracking the head movements and adjusting the view point accordingly. However, even with a fairly powerful computer a time delay between head movement and image adjustment remains noticeable. Except for one prototype 3-D display in Oxford [11], the accommodation cue can only be added by imaging the scene at physically different distances. The most advanced system is the US Navy sponsored "volumetric display" which achieves the accommodation cue by imaging on a rotating drum [12]. However, its large volume (approximately 1 m^3) makes it unsuitable for the type of applications we have in mind.

A relatively simple way to include accommodation and parallax in the depth percept is to optically superimpose two or more image slices through the scene. Such a transparent display presents "true depth" in the sense that all depth cues are present except for occlusion[2]. An example transparent display is shown in Figure 4. A New Zealand Company[13] is the first to have marketed a compact transparent, 2-plane display. The display consists of two LCD filters, one placed in front of the other, making a subtractive transparent display. Recently a 20-plane transparent display[14] has come on the market. We believe these developments to be highly significant, as argued in the next section.

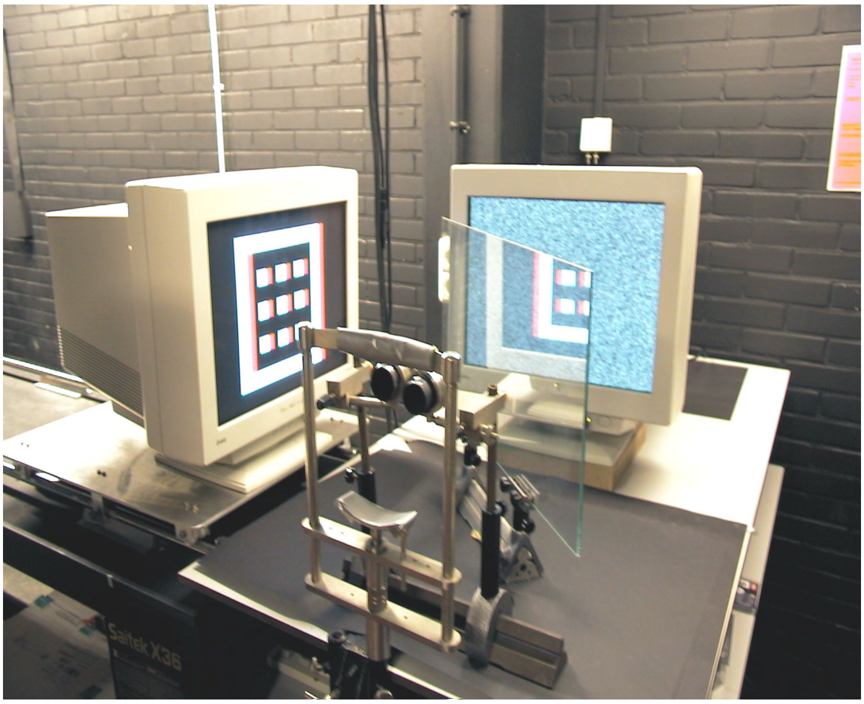

Figure 4. The experimental transparent 3-D setup at TNO Human factors. Two images are combined with a half-silvered mirror. Because the light from the two displays adds up, we call it an additive transparent display. The experiments examine the influence of various design parameters like amount of depth and scene content.

4. TRANSPARENT DEPTH DISPLAYS

4.1 Limited number of depth planes

So far not much research has been done on transparent displays. This is probably because the limited number of depth planes makes them unsuitable for the display of 3D pictures and videos. The technologies described above can in principle present as many depth planes as the number of pixels in the display. We believe however that for professional applications involving the display of symbols, 2 or 3 depth planes will be sufficient to provide a large operational advantage. By analogy, many of the "full color" cockpit displays by no means fully exploit their color gamut; often a display only contains the four primary colors. Similarly the information content of control displays oftentimes can be naturally divided into two (friend & foe) or three (above, below, & on the surface) layers. We therefore argue that the advantages of a transparent depth display will out-weight the disadvantage of the limited number of depth planes.

4.2 Optimal viewing comfort and depth perception

In the case of transparent depth displays the depth percept is truly extra. The user does not pay a price in terms of resolution, color, viewing angle, the need for special glasses, luminance, or viewing comfort as is the case with all other 3-D displays. Secondly, our present research shows that the depth percept "pops-out" immediately while the other types of 3-D displays require some amount of time for the depth percept to build up. Figure 5 shows how large the perceptual time delay is for unfavorable 3D stimuli. Thirdly, thanks to the parallax, occlusion of one object by an other can easily be eliminated by moving the head sideways or vertically (Figure 3). This is important if two objects are located at the same x,y coordinates but different heights. We therefore believe transparent planes to be very promising for the representation of computer generated data and symbols.

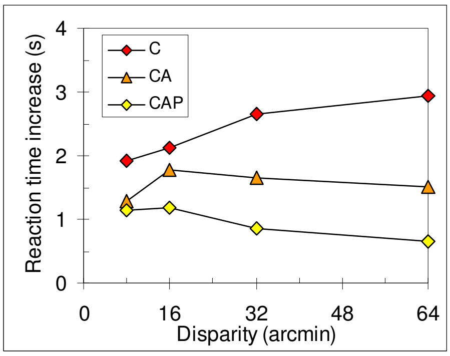

Figure 5. Data substantiating the claim that accommodation (A) and motion parallax (P) substantially influences the ease of depth perception when the depth gradient is large. Shows is the extra time required to perceive the depth relationship of two adjacent dots when a distracting object is added at a different depth. The horizontal axis contains the amount of depth difference. The increase in reaction time caused by the distractor is 1 to 2 seconds greater for the common type of 3D displays (C: Convergence cue only) than for transparent depth displays (CAP: accommodation and parallax as well as the convergence depth cue). These results imply that transparent depth displays are more natural to view than the standard 3D displays described in Chapter 3 and particularly suited for cluttered 3D imagery.

4.3 Subtractive and additive transparency

Another topic of current research at TNO Human Factors is design of the image content. We have shown that a transparent depth display, if designed wrong, can be very hard to fuse[15]. Secondly, the content of an additive transparent display (the front plane adds light: Figure 4) needs to be designed differently than a subtractive transparent display (the front plane removes light: e.g. Deepvideo[13]). The back layer of an additive display needs to be mostly dark, of a subtractive display mostly white. Otherwise the information in the other layer will not show up.

4.4 Occlusion

The 3D displays listed in Chapter 3 do fully show occlusion, the phenomenon that the front object ‘hides’ the back object. An additive 3D display however is not able to show occlusion and a subtractive 3D display only partially. In addition, color mixing leads to erroneous mixed colors: the color of the front object is influenced by the object behind and vice versa. For example, a yellow object on a purple background will appear as red on a subtractive 3D display (Figure 6). For any application involving warning symbols this is unacceptable. is at the expense of unwanted color mixing. We therefore argue that only an <u>additive + subtractive</u> transparent 3D display will support all the depth cues listed in Section 2. This may therefore well be the next major step forward in 3D technology.

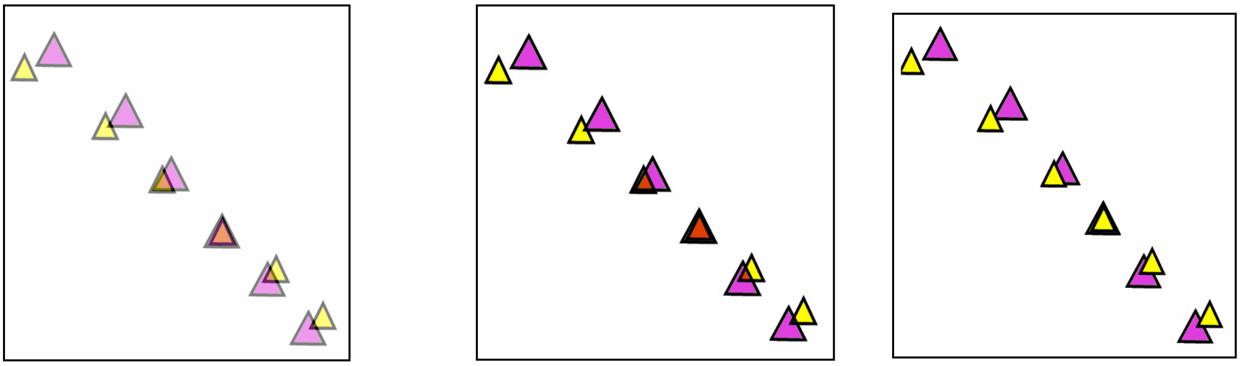

Figure 6. The effect of additive (LEFT) and subtractive (MIDDLE) transparency: the colors combine, easily causing confusion. In this example yellow and purple combine to orange and red respectively. The picture to the RIGHT shows occlusion, the yellow triangles positioned in front occluding the purple triangles.

4.5 Examples of potential applications

The positive outcome from our research combined with the developments in compact, transparent depth displays are promising for a number of applications that involve symbol-laden displays:
Command & control stations
Cockpit head-down displays
Compact mobile displays
We are presently testing the added value of 2-plane displays over existing display formats. If successful, a prototype 2-plane display may be tested on a Navy ship’s bridge or in an Air Force flight simulator.

5. CONCLUSION

We believe that the time of 3-D displays has arrived for professional applications. To start with, possibly not in the form of 3-D glasses, but by replacing flat-panel displays by 2-plane or 3-plane transparent displays. The development of a combined additive + subtractive transparent 3D display is, from a human factors point of view, the next major step forward.

6. REFERENCES

1. Howard, I.P. and Rogers, B.J., *Binocular vision and stereopsis* , Oxford University Press, Oxford, UK, 1995.
2. Nakayama, K. and Silverman, G.H., Serial and parallel processing of visual feature conjunctions, *Nature,* 320*(6059)* ,pp. 264-265, 1986.
3. Treisman, A. and Gelade, G., A feature integration theory of attention, *Cognitive Psychology,* 12 ,pp. 97-136, 1980.
4. Kooi, F.L., Toet, A., Tripathy, S.P. and Levi, D.M., The effect of similarity and duration on spatial interaction in peripheral vision, *Spatial Vision,* 8*(2)* ,pp. 255-279, 1994.
5. O'Toole, A.J. and Walker, C.L., On the preattentive accessibility of stereoscopic disparity: evidence from visual search, *Perception & Psychophysics,* 59*(2)* ,pp. 202-218, 1997.
6. Theeuwes, J., Atchley, P. and Kramer, A.F., Attentional control within three-dimensional space, *Journal of Experimental Psychology: Human Perception and Performance,* 24 ,pp. 1476-1485, 1998.
7. Pastoor, S., Human factors of 3D displays, *Displays,* 14, 1993.
8. Philips. *Philips 9 view auto-stereoscopic display*, Website: http://www.research.philips.com/generalinfo/special/3dlcd/index.htm(2003).
9. SeeReal Technologies d4d. Website: http://www.dresden3d.com(2003).
10. Kooi, F.L. and Lucassen, M.P., Visual comfort of binocular and 3-D displays, In: B.E. Rogowitz & T.N. Pappas (Ed.), *Human Vision and Electronic Imaging VI*, pp. 586-592, The International Society for Optical Engineering, Bellingham, WA, 2001.
11. Eagle, R., Paige, E., Rogers, B. and Sucharove, L., Accomodation cues reduce latencies for large disparity detection, In: pp. 2000.
12. Soltan, P., Lasher, M., Dahlke, W., McDonald, M. & Acantilado, N. *Improved second-generation 3-D volumetric display system,* (Report Spawar Technical Report 1763), Space and Naval Warfare Systems Center, San Diego, CA, (1998).
13. Deep Video Imaging Ltd. *Interactive dual plane imagery*, Website: http://www.Deepvideo.com(2003).
14. 3dMedia. *Multi-planer 3D volumetric display*, Website: http://www.3dmedia.com(2003).
15. Schoumans, N., Kooi, F.L. & Hogervorst, M.A. *Binocular fusion of transparant images,* (Report in preparation), TNO Human Factors, Soesterberg, The Netherlands, (2003).