Transparent Visual Presentation: Image Information and Ambiguity.

Faye Oosterhoff & Huib de Ridder
Delft University of Technology.
Sub-Faculty of Industrial Design Engineering.
Jaffalaan 9
Delft, 2628 BX The Netherlands
+31 15.278.3908
f.h.oosterhoff@io.tudelft.nl

ABSTRACT

Presenting images in transparent overlay probably introduces ambiguity in figure-ground perception. We are interested in the factors that influence this ambiguity. A monocular see-through Head Mounted Display (HMD) presenting fixed focus images is used to create extended reality. We have investigated image visibility in relation to four virtual focus distances between 33 cm and infinite. The HMD presented transparent virtual images overlaying background images. Background images were presented with a 17-inch monitor at 1 m distance. The two images were perceived as one. An effective stimulus comprised one 'landolt-c' within a circular configuration of seven hoops. A checkerboard was used as an affective stimulus. Stimuli were gaussian blurred. A trial consisted of a fixation cross, the combined stimuli and random noise subsequently. Correlating fixation either to effective or affective stimulus disbanded attention and task. Stimulus and presentation characteristics were randomised. Subjects' score correct on 'landolt-c' orientation was obtained by means of a 2AFC paradigm. Spatial frequency was the visual component both stimuli had in common. Results showed that paradigm and stimuli were functional for visibility testing. The virtual image showed ambiguity by proximity of the background. Fixation at a different depth decreased visibility, although presentation behind the background yielded mostly constant results. We suggest investigating visual priorities and context for use of visual information, to avoid unnecessary ambiguity.

Keywords:
ambiguity, extended reality, head mounted display, psychophysics, transparency, ubiquitous communication.

INTRODUCTION

The visual system is extensively investigated since Helmholtz (1909), giving insights in its workings. Our visual sensory system can recognise many different kinds of information. This information is detected with two eyes that together have the ability to extract stereo information. We can detect movements and relative speed, colours, hues, centre and surround and spatial aspects in three dimensions. We derive patterns from it, making up our visual world. It makes us able to separate foreground from background and detect objects. [Gibson 1979, Grossberg 1994, vd Schaaf & v Hateren 1996, Swets 1988.] Our visual world does not only show visual elements, but it has a relation to our behaviour. Objects can be discriminated by their appearance, but also show functionality. Functionality means that the visual element represents a context to the observer. This context is often recognised as a meaning. Seemingly opposite to that, is the fact that our world is constantly changing as well as our sensory system. Nevertheless, the world as we know it remains fairly constant. The visual system and other sensory systems are constantly calibrated on each other by correlating input from the same source. [Howard 1982] This means that perception is in constant interaction with the world that is perceived.

Fig 1. *A photograph and an abstract representation of Monsieur Vollard by Pablo Picasso. (painting is dated 1909) Pictorial similarity is still recognisable, although a close up will only show unrecognisable features.*

Media like displays, camera's, microphones and speakers make it possible to perceive and interact with parts of our world beyond our physical reach. Currently, techniques are being developed to do this while moving around, thus creating ubiquitous communication. [Mann 1996] Images presenting a synthetic world or parts of the real world, secluded from the natural environment are something we are quite used to. However, extending the natural environment with such information may create a problem. Unexpected objects and their characteristics could influence our own behaviour and the understanding of our visual world. Added information should not create ambiguity. It could affect calibration of our sensory system. [Howard 1982] Therefore, development of user interfaces for ubiquitous communication calls for insights in the human sensory system and human behaviour. Artists have been studying and expressing ambiguous images early last century. They experimented with abstract forms of information in images, resulting to surrealism, magic realism, cubism and a lot of other ways to express ambiguity in an image. The purpose of presenting such images, quite often was estrangement while isolating a particular essence of the image. (See figure 1.) Creating such an estrangement in our environment on a regular basis could change our perception of the world. Scientific studies of ambiguous visual information are based on what we know about the visual system. Well known ambiguity is found for reduced cue stimuli, like the Necker-cube or the vase-faces image. (See figure 2.) We would like to know for which situation visual information is misjudged when a virtual image is added to a real image. Presenting images in transparent overlay is likely to create an ambiguous situation. [Leeuwenberg 1978, Metelli 1974] A reflection quite often transfigures images and in case of transparency, one of the images may be too vague to recognise. So, in a natural situation we tend to avoid transparent or reflected information rather than use it. For transparent presentation both sources of information are visible at the same retinal position and time. This means that the observer has to make a choice, unless the image is not ambiguous. Distinguishing figures from a visual context is a basic ability the observer needs to detect objects. If both figure and ground have the same visual aspect to discriminate on, they are likely to influence each other on visibility. For this study we concentrate on spatial frequencies and the ability to discriminate objects from them, while presented in transparent overlay. Investigating human physiological abilities, especially on sensory perception is usually done with a psychophysical paradigm. It gives an answer to the influence of added information in our visual habitat and it can show what is filtered out from the original information.

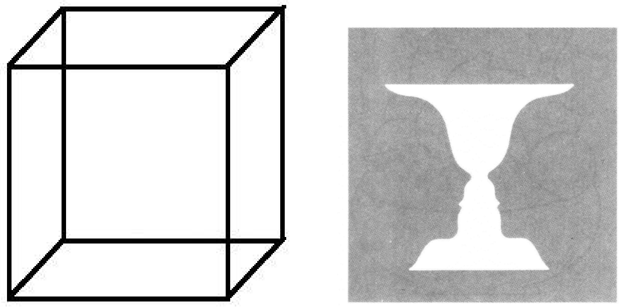


Fig 2. *Two different images showing the same kind of ambiguity. It isn't clear what is to be regarded as fore- or background. For the Necker cube it depends on where you looking. For the vase-faces figure it depends on preconception.*

Realising a set up.

We can investigate a virtual image as part of our visual environment. With a generic representation we can focus on the parameters we are interested in. In the research described below, we investigate human ability to deal with spatial frequency and transparent images. A background image is presented with a transparent overlay. By changing virtual depth of focus blurring can dissipate spatial frequencies. Such a set up is realised by using a monitor as background and looking at it through a semitransparent mirror that reflects the image of a small cathode ray tube (CRT) precisely as overlay. It allows the user to view displayed images and the environment simultaneously. [Melzer & Moffit 1997] [See fig 3.] Depth of focus can be changed by a set of lenses between CRT and mirror. However, two problems will occur with such a set up. The use of a set of lenses will yield aberrations in the image, and the brightness/contrast ratio of both monitors may differ. For our purpose, either an adjustment or description of these factors is done. Aberrations in the lenses are established by comparing 20 concentric circles of the virtual image with the same image printed on paper for five depths of focus. It shows variation up to 10%. The amount of aberration also varies with focal distance. The monitor's brightness and contrast were set at 50% and maximum respectively. We adjusted the HMD's brightness and contrast compared to the monitor's settings. The adjustment is achieved by showing opposite greyscale images viewed in overlay. [See fig 4.] For 50% visibility for both images, it should equal towards grey banners.

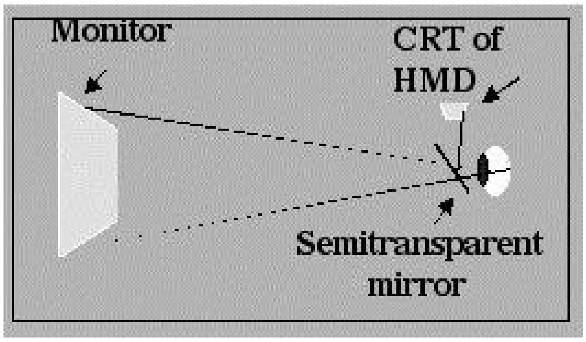


Fig 3. *Schematization representing a head mounted display set up. Lenses between mirror and crt would enable changing focus.*

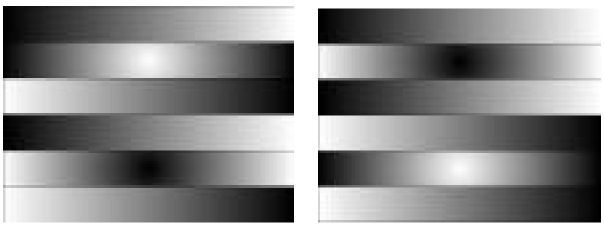


Fig 4. *Grayscale images for adjustment of transparent image display to 50%. The four central bands will middle to a gray field.*

Stimulus characteristics

The stimulus is contrived specifically to discriminate on spatial frequency. The stimulus a person sees and on which to perform the task are actually two stimuli. One stimulus is effective to the task (effective stimulus) and the other one is affective to the task (affective stimulus). They are both used for overlay or background presentation. The combined stimulus is called test stimulus. The affective stimulus shows a single spatial frequency in horizontal and vertical direction. To obtain such a stimulus, we've used checkerboards with field sizes of four, six and eight pixels. The effective stimulus is based on the 'c' shapes like used by optometrists, called landolt-c. It is assembled from two circles with the same centre. One circle has half the radius of the other. On a white background, the inner circle is white and the outer black. A rectangular white gap, in line of the radius, gives a 'c'-shape. The gap's width is half the size of the large circle's radius. The landolt-c is also used without the gap, forming a hoop. The sizes of the gap equal the checkerboards' field size. Stimuli are gaussian blurred over half the field or gap size to avoid hyper-acuity. [Badcock & Westheimer 1985] An overlay of both images results in addition or subtraction of luminance, according to image characteristics. For both images being based on spatial frequency they are quite likely to create ambiguity. Pilots showed 100% correct score for the effective stimulus only. With our experiments we want to investigate the workings of the visual system as well as the characteristics of our set up.

THE EXPERIMENT

The experiment is designed to test the functionality of monocular transparent presentation. We test human ability to use transparent, static visual information for various parameter settings of our set up and stimulus. Visibility scores are obtained by means of a two alternative forced choice (2AFC) paradigm. Not all stimulus effects will be reported in this paper. We will only look at effects of stimulus size and depth of focus of the virtual image.

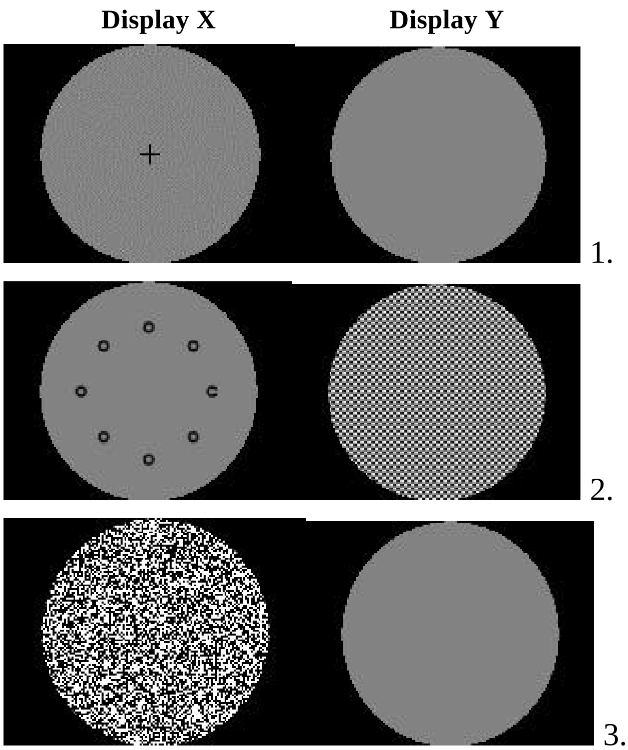
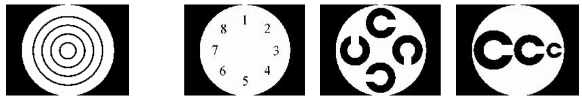


Fig 5. *A trial sequence (1,2 and 3), showing a possible configuration of images displayed to a subject. Images by display X and Y are seen in overlay by a subject. Display X and Y can be either Monitor or HMD.*

Table 1: *possible varieties in the effective stimulus.*



Eccentricity (pixels from center)	Position (deg)	Orientation	Gap size (pixels)
47	0	Top	4
94	45	Right	6
141	90	Bottom	8
188	135	Left	
	180		
	225		
	270		
	315		

Method

Apparatus

Virtual overlay images were projected by the HOPROS™ (Delft sensor systems). The mirror reflected 20% of light between 450 and 700 nm. The CRT source used phosphor P42. By means of an adjustment ring, the focus distance of the HMD image could be adjusted between 33 cm and infinity. The monitor displaying the background image was a 17-inch iijama vision master pro screen, model MT-9017T at 5000K display mode. Two 24-bit video cards, set at 640 by 480 pixels and 256 Grays depth were configured to present 1280 by 480 sized pictures at once, divided over two screens.

Stimulus

The elements that formed the stimulus are described in the section about stimulus characteristics. Both effective and affective stimuli were presented within a circle of 480 pixels diameter, in the centre of the screen. Outside this circle, the image was kept black, thus creating an aperture. A configuration of seven hoops and a landolt-c placed evenly on a hypothetical circle from the centre of the screen and beginning at the top, made up the effective stimulus. The landolt-c was given two orientations horizontally and two vertically. The imaginative circle was given four possible sizes. This way the following configurations were achieved. [See table 1.]

Sizes weren’t mixed per stimulus. Having 4 eccentricities, 8 positions, 4 orientation and 3 sizes, gives 384 possible images. The affective stimulus counted only 3 varieties for field size 4, 6 and 8 pixels. A configuration of 256 images was chosen such, that the properties we investigated were evenly distributed. The trial showed an image with fixation cross, the test stimulus and a 50% noise image. [See fig 5.] The test stimulus was shown for a < 100ms period. Fixation and noise image were only shown on one of the displays per session. Sessions were also randomised. All images except the noise and the affective stimulus had 50% grey backgrounds, keeping light emission at about 50%. Effective and affective stimuli were fully randomised in separate lists and then joined one on one. Randomisation took place for display assignment of either stimulus. 256 Stimuli had to be judged in a session. Two sessions were used to obtain data for parameter settings. Scores for ‘Horizontal’ and ‘Vertical’ are put together to avoid bias resulting from the task.

Procedure

The HMD was mounted on a chin rest to fixate the virtual image as an overlay on the image from the monitor. The monitor was set at 100 cm from the position of the subject’s eye. Due to HMD focus variation, this was an approximate distance. It had to be moved less than ± 7 cm to get a fitting overlay image for all distances. Tested parameter settings were HMD focus distance and the position of fixation. Four HMD focus distances were tested: 33 cm, 100 cm, 166 cm and infinite.

Subjects

The subjects were members of the department and trained for use of the stimuli and task. Two female subjects were aged around thirty. The male subject was in his forties. All subjects had their vision corrected to normal.

Task

The subject’s task was to look at the fixation cross. After the stimulus was presented, the judgement ‘horizontal’ or ‘vertical’ had to be entered by pushing the keys [J] or [N] respectively. During this period the noise image was shown.

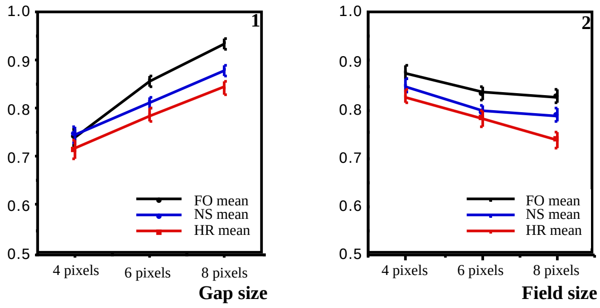


Fig 6. *Picture 1 shows correct scores for judgement on effective stimulus’s visibility at different sizes. Stimulus size is represented by landolt-c gap size. Picture 2 shows the effect on correct scores by different field sizes of the affective stimulus. Scores of subjects are shown separately. (n = 360)*

Results

Judgement on effective stimuli shows an increase in correct score for increase of stimulus size. (See figure 6.) Judgement on effective stimuli shows an increase in correct score for increase of stimulus size. The affective stimulus decreases correct score for increasing field size. For both kinds of stimuli, differences can be found between subjects’ individual scores but trends are approximately the same.

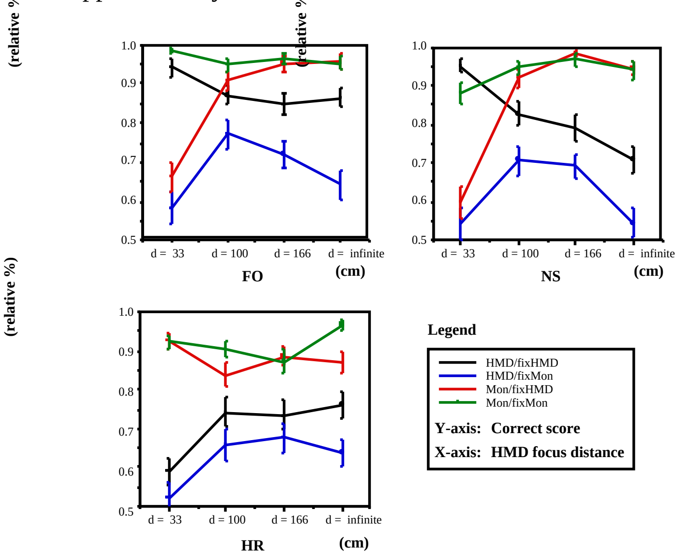


Fig 7. *The pictures show correct scores on visibility of landolt-c at different HMD focus distances by subjects FO, NS and HR separately. Separate lines represent different parameter settings. In the legend it shows the format: ‘effective stimulus on display X/ fixation on display Y’. X and Y can be HMD or Monitor (Mon). (n =70)*

We can look at figure 7 in two parts. Scores for d = 33cm show two groups of scores. Scores are near chance level for FO and NS when fixation and the effective stimulus are displayed on separate monitors. For the same subjects scores are at 90% (0.9 relative percent) or higher when fixation and effective stimulus are displayed on the same monitor. So scores depend on fixation. Subject HR shows a different pattern. Scores for effective stimulus at HMD are at chance level. When displayed on the monitor scores larger than 90% are found. Scores seem indifferent to where fixation is displayed. Second part are scores for d = 100cm, 166cm and infinity. Presentation of the effective stimulus on the monitor shows high scores for all subjects. Generally, scores for HR are slightly lower (80-95%) except at infinity (< 95%). For all subjects, scores are lowest (<80-50%) when the effective stimulus is displayed on the HMD and fixation was placed on the monitor. FO and NS show a decrease up to 10% for increase of distance. All scores are about 10% higher if fixation is placed on the HMD as well. For these settings, only scores by NS decrease for increasing distance. The other subjects show invariable scores. Now we can join the two parts in the results we’ve just observed. Putting fixation on the HMD and the effective stimulus on either monitor shows a cross over for scores by FO and NS between d = 33cm and d= 100cm. They do not for HR.

Discussion

Increase in correct score for increase of effective stimulus size means that we indeed measure visibility, which was the objective of stimulus choice and presentation method. Increasing field size of the affective stimulus resulting in decreasing visibility shows there to be an effect. It mainly suggests being a result of occlusion, although we also expected an influence of equal spatial frequencies. It would be interesting to investigate the influence of spatial frequency on a more comprehensive scope, using a set up with less aberration. FO and NS are slightly better at recognising the stimulus and between them, FO again is better. This could also be a learning effect, but there are too many varieties in the stimulus to determine this. Nevertheless, equal trends and comparable errors suggest it to be general difference in subject ability to recognise the stimulus. Results in figure 7 showed that scores crossed for the situation where fixation is displayed on the HMD, as long as the virtual image is presented in front of the monitor. Further studies, not presented here have corroborated this finding and shown it to be a gradual process. Anchoring primary attention by fixation at a different depth will blur the effective stimulus. Visibility of the virtual image appears to be influenced by proximity of the monitor's image, but not the other way around. The same effect could be expected behind the monitor, but scores remain approximately the same. It may be that effects like influence by monitor's image and the increasing distance cancel each other out. Only when fixation is put on the monitor and the effective stimulus has to be judged while being behind it, visibility is clearly decreased. Focussing on short distances will become harder with age. That is most probably the reason why results by HR do not seem to agree. HR complained about not having a sharp image when viewing the fixation cross at close distance. Closest comfortable reading distance for this subject was about 45 cm. This explains the different results at 33 cm. Scores are generally lower on the HMD when its focus depth is the same or higher than the monitor's. To some extent it can be due to less quality of the HMD image. Nevertheless, subjects said not to be able to establish a source. Placing fixation on the monitor shows decreased visibility of the HMD image for increasing difference in focus distance, reaching chance level for HR and NS. For focus distance close to the distance of the monitor visibility is good, reaching 80% for FO. These results can be explained by the increasing blur of the effective stimulus, when fixation distance is getting larger. Visibility on the monitor is not influenced significantly by affective stimuli on the HMD in case virtual focus distances are equal or higher. This is irrespective of fixation being on either display. It suggests the monitor's image either to be less ambiguous or just more visible. If focus distance of the HMD is at 33cm and shows the fixation cross, visibility of the effective stimulus on the monitor is at chance level. This shows that very close up fixation will diminish visibility a lot for further away images. This effect will decrease with increasing distance from the observer.

Conclusion

The effective stimulus works for visibility testing within the used paradigm. Results for the affective stimulus show that it works as intended. If we look at variations for focus distance, we see conditions that affect visibility scores. The first condition is when the task is presented as a virtual image. Apparently it is influenced by the proximity of the monitor's image making the objects to discriminate on more ambiguous, thus yielding lower scores. This doesn't work the other way around. Presentation of the effective stimulus behind the monitor where is fixated on reduces visibility with increasing distance. For other situations at or behind the monitor visibility remains constant. A third condition is viewing images close up. It depends on the ability to accommodate close up, but when doing so distant images get less influential in drawing attention.

Engaging transparent presentation and further study

The experiment describes a method to investigate image visibility with transparent overlay. The objective is to explain results and set up characteristics for an extended environment situation. Our set up is just a small part of a realistic situation that occurs when using virtual visual information in our natural environment. This set up can be compared to the situation where someone is walking around and suddenly is presented textual information, as a virtual and transparent image at a particular distance. The other situation is that someone is reading text and something in the environment suddenly draws attention. Apparently visibility may sometimes be close to zero for either situation. A third possible situation is that someone tries to read a text against interfering background texture, creating ambiguity. Of course conclusions from our experiment correlating to these situations are based on bestowed image information with reduced cue stimuli. We also saw that the virtual image was affected more by presence of the monitor image than the other way around. This suggests that the virtual presentation is more easily getting ambiguous, although the same kind of image information was used. Display quality was responsible for it to some extent, but results suggest a role of depth order as well. For the virtual image, we can say that dealing with static visual information behind focus of attention makes it less important. It may be that transparent visual information is regarded less interesting than non-transparent information, although subjects could not distinguish the source of the images. Still, we can conclude that the visual system designates priorities to particular visual information. These priorities seem to be context dependent. [Edgar & Reeves 1997] For particular kinds of information, we may be able to determine and manipulate context. If we want to design a ubiquitous visual presentation system that gives information with the least ambiguity, we have to use the right format and context for presentation. Knowledge on how to present visual information based on visibility, information content and context would be interesting. In the introduction we have stated that ambiguous presentation may even change the way we perceive our world. To avoid ambiguity having such consequences, extended reality must be adapted to the observer and its environment. [Feiner, MacIntire & Seligman 1993, Starner, Shawney & Pentland 1998] This study has shown some elements to be of importance. We want to investigate this further by looking at the spatial frequencies that are used for the experiment. To do this, we have to equal display properties. Another step is to add information to our stimulus. It would be interesting to learn the effect of colour on contrast and spatial frequency.

ACKNOWLEDGMENTS

This work was supported by the Netherlands Organisation for Scientific Research (N.W.O.) under project DEL55.3671: Mobile Multimedia Communication. This multi-disciplinary project aims to investigate various aspects of communication between mobile users using multimedia.

Bibliography

Badcock, D.R. & Westheimer, G. (1985). Spatial location and hyperacuity: The center surround location contribution function has two substrates. Vision Research, 25, 1259-1267.

Edgar, G.K. & Reeves C.A., (1997). Visual Accommodation and Virtual Images: Do Attentional Factors Mediate the Interacting Effects of Perceived Distance, Mental Workload, and Stimulus Presentation Modality? Human Factors, Vol. 39 (3), 374-381.

Feiner, MacIntire and Seligman., (1993) Knowledge-based Augmented Reality. Communications of the ACM, 36 (7).

Gibson, J.J., (1979) The ecological approach to visual perception. Boston: Houghton Mifflin.

Grossberg, S., (1994). 3D Vision and figure-ground separation by visual cortex. Perception & Psychophysics, 55, 48-120.

Helmholtz, H. von. (1909) Physiological Optics. Dover, New York. (1962 English translation by Southall J.P.C.)

Howard, I.P. (1982). Human Visual Orientation. Chichester, New York, Brisbane, Toronto: John Wiley & Sons.

Leeuwenberg, E.L.J., (1978). Quantification of certain visual pattern properties: Salience, Transparency, Similarity. E.L.J. Leeuwenberg and H.F.J.M. Buffart (eds), Formal theories of visual perception, Chichester, New York: Wiley.

Mann, S., (1996). “Smart Clothing.” TR 366, M.I.T. Media Lab Perceptual Computing Section, Cambridge, Ma.

Melzer, J.E. and Moffit, K. (1997). Head Mounted Displays, Designing for the user. Optical and Electro-optical Engineering Series, Fischer R,E and Smith W.J (ed.), MacGraw-Hill.

Metelli, F. (1974) The Perception of Transparency. Scientific American, Vol. 230 (4), 91-98.

van der Schaaf, A. & van Hateren, J.H. (1996). Modeling the Power Spectra of Natural Images: Statistics and Information. Vision Research, Vol. 36 (17), 2759-2770.

Starner T., Shawney, N. & Pentland A. (1998). Visual Context Awareness in Wearable Computing. Proceedings of the International Symposium on Wearable Computing. Pittsburgh, Pennsylvania, 19-20 October, 50-57.

Swets, J., (1988). Signal Detection and Recognition by Human Observers. Contemporary Readings, New York, Wiley.