Visual Detection with Hyperstereoscopic Video for Aerial Search and Rescue

Kit Man Cheung & Paul Milgram
Ergonomics in Teleoperation and Control (ETC) Laboratory
Department of Mechanical & Industrial Engineering
University of Toronto
Email: kit@etclab.mie.utoronto.ca, milgram@etclab.mie.utoronto.ca

This paper summarises the results of two laboratory studies conducted to explore the potential benefits of hyperstereoscopic video for aerial search and rescue (SAR). The studies were conducted using a single camera hyperstereoscopic video system (designed for aerial SAR) as the test platform. The first study was designed to evaluate potential benefits of hyperstereopsis provided by large viewpoint separation. It was found that individual differences play an important role in defining a possible "optimal" viewpoint separation for the system. The second study was conducted to test the effect of increased viewpoint separation on tolerance of vertical disparity. The results corroborated expectations that increased viewpoint separation would reduce vertical disparity tolerance. Guidelines are given for future modifications and system implementation.

INTRODUCTION

Aerial *search and rescue (SAR)* is an extremely important responsibility of the military in Canada, with SAR units responsible for co-ordinating and conducting search operations over a variety of environments. One such type of operation involves searching for downed aircraft over heavily forested terrain, where one of the key features used for detection of a possible crash site is the broken treetops in the foliage caused by the crash.

Currently, highly skilled SAR technicians conduct searches by naked eye and, when closer inspection of a particular area is needed, binoculars are used as an aid. An exploratory project was conducted to study the possibility of using a stereoscopic video system as a means to assist the SAR technicians in both on-line and off-line detection of terrain targets. More specifically, a video system was designed to explore some of the theoretical benefits of *hyperstereoscopic video*.

Under hyperstereoscopic viewing conditions, the separation of the two camera viewpoints is exaggerated (Diner & Fender, 1993). The key effect of increased separation is a corresponding increase in *horizontal (binocular) disparity* in the stereo pair presented to the viewer. In short, an increase in disparity should increase the relative depth separation perceived between objects in the scene. Under some conditions this increase may provide potential benefits for enhancing the detection of crash sites during SAR operations.

Binocular disparity can provide more than simple depth perception, however. In the literature, many studies have demonstrated advantages of stereoscopic displays over conventional monoscopic displays, most notably, as far as the present study goes, in terms of detection of barely distinguishable targets. In Merritt (1991), for example, *visual noise filtering* was identified as one of the potential benefits of stereo viewing, where, by presenting two different views of the same scene, it is possible for the brain to filter out uncorrelated noise and help detection of correlated target information presented to both eyes. In a comparative study of monoscopic versus stereoscopic displays, Smith & Pepper (1981) concluded that stereo displays are less affected by visual noise in the environment, especially under low contrast conditions. In another series of studies, Schneider and colleagues (Schneider, Moraglia et al., 1989; Schneider & Moraglia, 1994) identified an effect which they called *binocular unmasking*, which has been postulated to enhance a viewer's ability to detect camouflaged targets.

In light of the well documented advantages of stereo displays, it was reasoned that *hyper*stereoscopic camera settings should provide even further enhancement, by increasing the depth resolution of the display system and exaggerating the perceived relative depths between different objects. In one of the few published tests of this theory, Spain (1986) found that increases in interaxial separation are useful for enhancing target detection time and recognition. Extending this concept to the case of detecting breakage or holes in forest foliage, therefore, a properly designed hyperstereoscopic display should cause the gaps caused by breakage to appear deeper and thus facilitate this particular aspect of SAR.

Hypersterescopic Video Prototype

A prototype of a single camera hyperstereoscopic video system was implemented and two laboratory studies were conducted to evaluate the potential benefits provided by the prototype. The experiments, described in the following section, were conducted to evaluate users' performance using the hyperstereo prototype. For a more detailed account of both the technical design and the lab study, the reader is referred to Cheung (2000).

The system was designed with an emphasis on mechanical simplicity and flexibility for altering viewpoint separation. Unlike traditional stereoscopic video systems, however, the present prototype provided stereo video using *only one camera.* This was possible because of the motion of the search vehicle. That is, to generate the two images needed for a stereo pair, the camera can capture one image and store it. As time elapses, the aircraft (and the camera mounted to it) moves to a new location where another image is captured. Assuming that the vehicle does not alter its course drastically during this period, the two images captured can be used to create the necessary stereo pair.

In a sense, therefore, the two *effective cameras* are *separated in time*, rather than in space. The physical separation between the locations at which the two images are taken is referred to as the *effective camera separation*, and can be altered by adjusting the time delay between pairs of images stored in the buffer. This is achieved in our system, without the need for any additional mechanical hardware, simply by using the memory of a computer workstation as an image buffer.

Figure 1 illustrates how stereo pairs are generated in our system using the motion of the search aircraft. The software portion of this prototype was used for the laboratory studies described in the following.

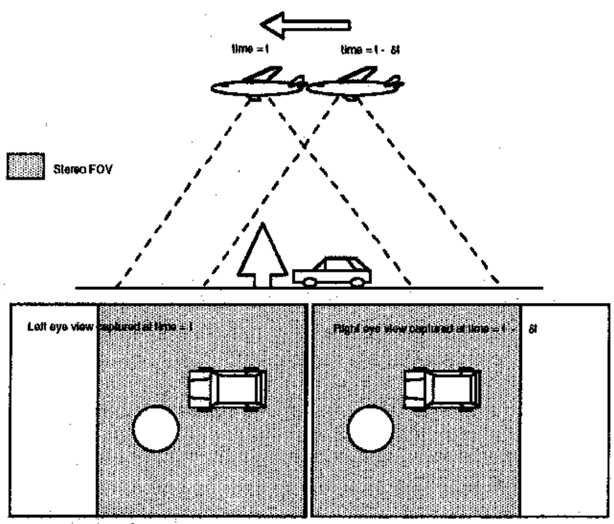

Figure 1 - Stereo pair created using the single camera hyper-stereoscopic video system.

Laboratory Study 1
Visual Detection using Hyperstereoscopic Video

The experiment described here was designed to test some of the theoretical advantages of hyperstereopsis with our prototype video system. Geometrically, a video system with a large camera separation should produce a larger horizontal disparity, and the resulting hyperstereo-scopic images should have a higher depth resolution due to exaggerated relative depth. Perceptually, however, it is not possible to maintain fusion for binocular disparities of indefinite magnitude since, as disparity increases, maintaining fusion becomes increasingly more difficult (Boff & Lincoln, 1988). The present experiment was thus designed to test not only the existence of the hypothetical advantages of using the hyperstereoscopic video system for detection, but also to investigate the limits of fusion with the system.

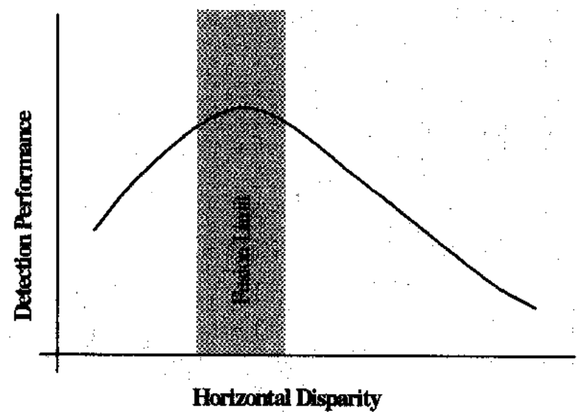

Figure 2 - Hypothetical detection performance as a function of horizontal disparity.

Expected Results:

The performance expected from the experiment is illustrated in Figure 2. Because increased horizontal (binocular) disparity is supposed to exaggerate depth differences, detection performance was expected to improve as horizontal disparity was increased. However, it was also expected that, as the binocular (horizontal) disparity approaches the fusion limit of the observer, detection performance should degrade quickly due to the difficulty in maintaining fusion – hence the U-shaped curve shown. Furthermore, since the fusion limit varies with individuals, the horizontal disparity for optimal performance may be different for each participant, as depicted by the shaded area in the figure.

Experimental Design

Because it was not possible to use real aerial video, the experiment was conducted using simulated video scenes as the input to the hyperstereo software. The video scenes were created using Alias PowerAnimator animation software, with a simplified terrain model.

A signal detection theory (SDT) type experiment was used for this evaluation. In the experimental task, participants were asked to detect a set (2-6) of disks floating above a flat surface as a (virtual) camera passed over the surface. Over the surface were also scattered many disks, placed there as distracters. The participants were asked to identify the target (by pointing at the screen) and to rate their confidence in the identification on a 5 point scale. A null response (i.e. the subject remained silent) while a target disk was on screen was considered a miss. A null response while only distracter disks were present was considered to be a correct rejection.

Eight participants took part in the within-subjects experiment. Each participant performed the detection task with four different 3-minute scenes at each of five different horizontal disparity levels (47.5, 95, 142.5, 190, and 285 minutes of arc).

Results and Discussion

Hit and False Alarm ratios were used to compute d' values for each of the different horizontal disparity (HD) levels. The mean d' values for all participants are shown in Figure 3 (with +/- 1 standard deviation).

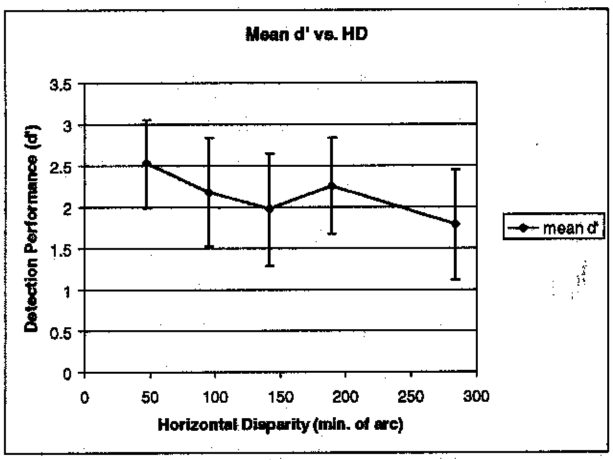

Figure 3 - Mean d' as a function of Horizontal Disparity.

The results obtained were not as unequivocal as portrayed in Figure 2. The average d' values indicate that performances levels for the lower four values of HD were very similar, with perhaps a slight dip in performance at the highest HD level. A one-way analysis of variance showed that there was no significant main effect of HD among the d' levels however.

Examining the individual performance data, it appears that the effect of individual differences in fusion limits may have been stronger than initially suspected. To better understand the experimental results, a second ANOVA was performed, this time with the data blocked by subject to give some insight into the significance of individual differences. Using a randomised block design, the second analysis showed that there was in fact a significant effect due to individual differences ($F_{(7, 28)} = 148$, $p < 0.01$). Furthermore, detection performance now showed a significant effect due to HD level ($F_{(4, 28)} = 4.45$, $p < 0.01$).

Study 1 Conclusions

Not shown in Fig. 3 is the fact that the performance contributing to the d' levels obtained derived from a Hit rates in the vicinity of 70%, with False Alarms at about 6%. Unfortunately, it was not possible (for technical reasons) to assess performance objectively for HD=0 (equivalent to monoscopic video). Therefore, although detection of the targets could by most accounts be considered to have been quite good, we are not able to make any valid claims about performance with the hyperstereo system in an absolute sense.

In a relative sense, the other objective of the experiment was to identify whether different levels of hyperstereopsis can maximise detection performance, as hypothesised in Fig. 2. The results showed that the postulated performance was found in some subjects, but for others the exaggerated stereo may have been disruptive. Furthermore, detection performance varied greatly across subjects, an apparent result of individual differences in adapting to different horizontal disparities. Since the fusion limit appeared to vary with individuals, it is possible that each participant may have exhibited peak performance at a different level of HD. (If this were the case, it would imply that we should not necessarily reject the hypothesis, but that it could be better tested if our system were to allow for finer control of HD levels.) One of the main practical implications of the experiment, therefore, is that, due to individual differences, it may be important under actual operational conditions to provide flexible control of horizontal disparity to the SAR personnel.

Laboratory Study 2
Tolerance of Vertical Disparity

The purpose of the second study was not to replicate the literature on vertical disparity tolerances but to explore the particular effect of horizontal disparity on the vertical disparity tolerance of observers within the context of the single camera stereo system. Since the system utilises forward vehicle motion for generating horizontal disparity (HD), any lateral movements (for example, due to cross winds) will translate into a vertical disparity (VD). In general, VD is disruptive in stereoscopic viewing (Boff & Lincoln, 1988). This study was thus designed to gather data to be used as a design guideline for controlling VD in future implementations.

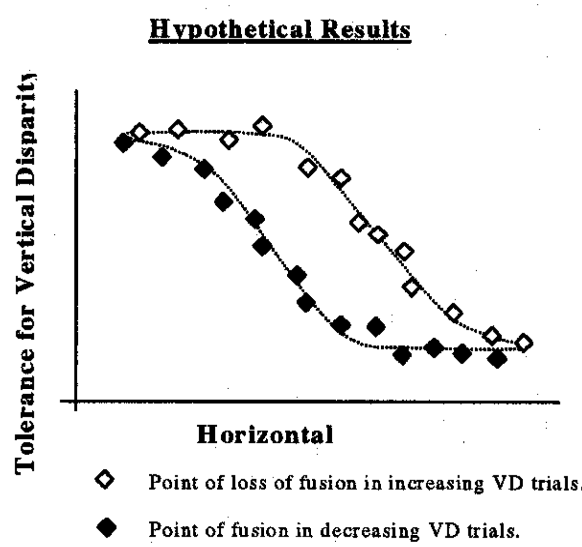

Figure 4 - Hypothetical result for Study 2.

Expected Results

Tolerance of vertical disparity (VD) was investigated using a method of limits experiment, consisting of trials of both decreasing and increasing values of VD. During trials, participants were asked to identify, respectively, onset and loss of the fusion state. Changes of state could be either from an unfused pair to a fused image (decreasing VD trial) or from fused to unfused (increasing VD trial). The participants also performed at different frame delays, thus introducing HD variation into the experiment. The expected performance is shown in Figure 4, where we see that, as HD increases, fusion should become less stable, expressed as a decreasing tolerance for VD. Opposite, but not identical, behaviour is expected in the opposite direction, resulting in a sort of hysteresis effect.

Experimental Design

Another simulated video was generated using computer animation as the stimulus for this study. The task this time was not to detect a target, however, but to identify changes in the state of binocular fusion. For this reason, the stimulus was simply a blue cross-shaped 3D pillar, which moved across a flat green surface in an arc, as shown in Figure 5. Note that movement along the arc (which resembles a ground trajectory for a constant airspeed in the presence of a cross wind) was the source of the independent variable, vertical disparity (VD).

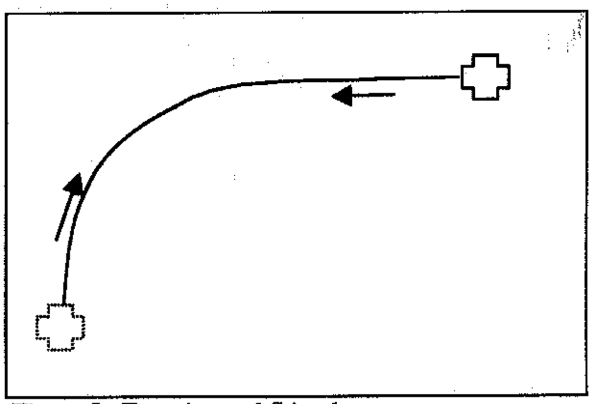

Figure 5 - Experimental Stimulus

Nine participants took part in the within subjects test, with each performing 4 trials at 4 different frame delay (HD level) settings.

Result

The threshold data obtained in the study are presented in Figure 6 as a scatter plot.

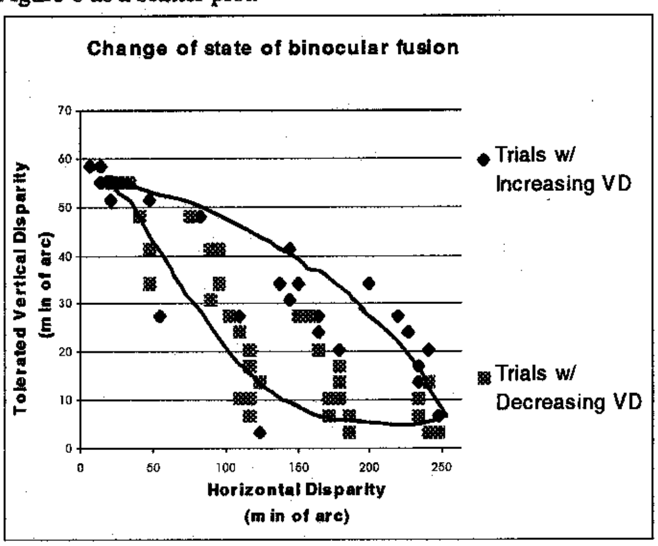

Figure 6 - Experimental Results

Study 2 Conclusions

The experimental evidence obtained suggests that our expectations of a reduction of VD tolerance with increasing HD, following an approximate hysteresis pattern, of VD were correct. Figure 6 suggests that, as a guideline, when operating with HD levels of 100 minutes of arc or more, tolerance for VD falls off fairly rapidly and VD should thus be kept to below about 10 minutes of arc. Such a conservative estimate should thus accommodate a reasonably wide range of operational disparity conditions.

ACKNOWLEDGEMENTS

This project was carried out under contract W7711-5-7263/001/TOS for, and with the generous support of, the Defence & Civil Institute of Environmental Medicine (DCIEM). The authors also thank Alias/Wavefront Inc. for their generous donation of the PowerAnimator software. Finally, we thank Dr. Julius Grodski, Bart Hall, Dr. Shi Yin, David Drascic and Rishi Bageshware for their invaluable assistance and contributions.

REFERENCES

Boff, KR & Lincoln, JE (eds) (1988). *Engineering Data Compendium: Human Perception and Performance*. AAMRL, Wright-Patterson AFB, OH, 1988.

Cheung, K.M. (2000). An investigation of potential benefits of hyper-stereoscopic video for aerial search and reconnaissance. Unpublished M.A.Sc. dissertation, University of Toronto.

Diner, DB & Fender, DH (1993). *Human Engineering in Stereoscopic Viewing Devices*. Plenum, New York 83-84.

Merritt, JO (1991). Evaluation of stereoscopic display benefits. *Introduction to Stereoscopic Displays & Applications - Short course notes*. L.Hodges, D. McAllister and J.O. Merritt. International Society for Optical Engineering (SPIE).

Schneider, B. & Moraglia, G. et al. (1989). Binocular unmasking: An analog to binaural unmasking? *Science* 243: 1479-1481.

Schneider, B. & Moraglia, G. (1994). Binocular vision enhances target detection by filtering the background. *Perception* 23(11): 1267-1286.

Smith, DC & Pepper, RL (1981). Stereo TV improves operator performance under degraded visibility conditions. *Optical Engineering,* July/August 1981, 20, 579-585.

Spain, EH (1986). *Effects of Extended Camera Baseline and Image Magnification on Target Detection Time and Target Recognition with a Stereoscopic TV System.* Naval Ocean Systems Center (NOSC), Tech Report 1100. San Diego, CA.