Dynamic Visual Masking in a Virtual Environment

Jason Drummond* and Anthony Steed†

University College London

Abstract

Virtual environments (VEs) offer more than just simulation; they allow us to create perceptual effects not feasible in the real world. Arising from perception research, one technique that can be used is visual masking: under certain temporal and spatial conditions a masking stimulus can be used to prevent explicit, visual awareness of a target stimulus although the target may still be processed non-consciously. This positional paper describes on-going work into the use of visual masking within VEs. Our previous work has shown that it is possible to use static, visually masked objects within a VE to alter navigation and preference behaviour, without the participant being consciously aware of the stimuli. Building on this, it is intended to animate masked virtual faces to see if larger effects are possible.

Keywords: Virtual environments, visual masking, perception.

Index Terms: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality; J.4 [Social and Behavioral Sciences]: Psychology

1 Introduction

Visual masking describes a process whereby the presentation of one visual stimulus, the mask, effects the conscious perception of another, the target. One paradigm, known as structural pattern masking, uses simple graphic figures for both target and mask [1]. Paracontrast describes a stimulus onset asynchrony (SOA) whereby the mask precedes the target, metacontrast being the reverse of this. With common onset masking the target and mask appear simultaneously, the SOA being zero, but are differentiated by the length of exposure to the participant.

An example of a basic masking paradigm would be where a graphic figure, the target, is repeatedly flashed upon a screen, each time being replaced by another figure, the mask, for a longer duration. If the target is present for only 10-20 milliseconds each time and the mask for 200 milliseconds then the viewer should only be aware of the mask. This would be an instance of backwards (metacontrast) masking. It is worth noting that the term 'visual masking' is also found within computer graphics, applied to algorithms designed to hide shading defects. This is not the type of masking being discussed here.

In order to achieve a masking effect care must be taken over the form of the targets. Although information can be processed rapidly, in the order of milliseconds, there is a limit as to the complexity of the target. For example, sentences cannot be generally processed and indeed neither can long words [5]. However, although a face appears to be quite complex in form, at least in comparison to the basic figures used in many masking experiments, there appears to be neural systems devoted entirely to processing facial forms [8].

We know that people respond in similar ways to computer generated facial expressions and real expressions [6]. For example, amygdala activity was measured during facial emotion recognition tasks which used both computer generated and human faces. It was found that a 'robust amygdala activation was apparent in response to both' [9]. This suggests that, to some extent the brain treats both computer and human based emotional stimuli in the same way.

Work has been carried out using images of human faces as both targets and masks: Winkielman and Berridge [11], looked at how visually masked, emotionally valent, static facial images can affect behavioural responses. Participants experienced either a masked smiling, neutral or angry face, all masked by a neutral face. They were subsequently asked to drink a beverage offered and rate it for monetary value. Those who indicated that they were thirsty prior to the trial, *and* experienced a masked smiling face, tended to drink more and award a higher monetary value than those in the other conditions. Another study, Dimberg *et al* [3], again used masked, static emotive expressions but this time looked at facial muscle reactions via electromyography. It was found that appropriately matched facial muscles were activated as if the participant were mimicking the masked stimuli.

The facial images used in the above studies were planar as indeed targets and masks have tended to be traditionally in masking research. All of the visually masked faces were static. Little or no research appears to have been carried out using animated, masked objects such as non-planar, virtual faces found within VEs.

2 Visual Masking in Virtual Environments

When depth disparity is studied it tends to be of a form where both mask and target are planar objects and are presented on two different, or apparently different, depth planes. Often planar noise images such as random dot stereograms (RDS) are used. RDS can control for monocular perspective cues in a similar way to stereoscopic images that separate eye movement convergence from focus accommodation [10]. RDS depth cues tend to result in an inhibition of the masking process [7] which is known as 'binocular unmasking' [2]. However, there appears to be a little work which looks at objects with internal depth disparity [2] such as non-planar meshes with found within virtual environments.

A recent study showed that it was indeed possible to use non-planar, virtual objects as targets [4]. This showed that whilst people were not able to consciously perceive the masked objects in a VE they displayed an awareness of object position. The target objects had an internal depth disparity in line with the viewer, perpendicular to the planar mask. The masking effect appeared to work as long as the mask and the target were sat as closely as possible on similar depth planes. It was discovered that there was some leeway as to whether the mask sat just in front of, bisected or sat just behind the target. However, this may be dependent on the size of depth disparity in the target. Also it was found to be advantageous for the mask and target objects to carry the same texture or shader material.

*e-mail: j.drummond@cs.ucl.ac.uk
†e-mail: a.steed@cs.ucl.ac.uk

Another experiment demonstrated a small, affective effect using emotive faces as targets [4]. Similarly, these were non-planar, virtual objects with internal depth disparity, again the masks were planar. The expressions carried by the faces were happy, angry and neutral. An overall significant effect on participant choice behaviour was shown. Contrasts showed angry faces scoring significantly less than the other two, happy and neutral had no significant difference between them.

Both of these studies used static virtual objects as targets. From them we know that both masking and affect driven masking effects are possible in a VE using non-planar targets. However, although the effects were considered statistically significant they were not large. The second study in particular may have been hampered by use of static emotional expressions.

3 WORK PENDING

It is proposed to test animated, non-planar, masked expressive faces, using planar door objects as masks will remain as the masks, see Figure 1. An initial problem may be the way an animated target is 'read' by the brain: as a series of separate images or a fluid animation. Each frame of the animation will be shown in a longer time frame than would normally be the case for animation. For instance, if each target is shown for 15ms, the mask for 100ms and the between-stimulus gap is 85ms then the target will only appear at a rate of five frames per second (fps). Low frame rates can disrupt the illusion of fluid movement. However, even if we drop to around 1fps movement can still be perceived although it appears jerky, like the second hand on a watch. Therefore, animated masking may be possible.

Figure 1. Angry target and masking door

The VE will consist of a series of rooms ending in two doors, see Figure 2. Each door will lead to the next, similar room. In front of one of the doors will be a visible object, with a neutral surface texture. The other door will be a planar mask and so its visibility will oscillate rapidly, thereby masking an animated, facial expression, either angry or happy, within this second doorway. The 'flashing' mask should divert attention onto the target doorway but also navigating through the doorway will force participant attention in the target direction. Once through the flashing door (and therefore the masked face) the participant will rate the visible object's appeal. They will then move through into the next room and so on, until a series of rooms have been navigated and scores collected.

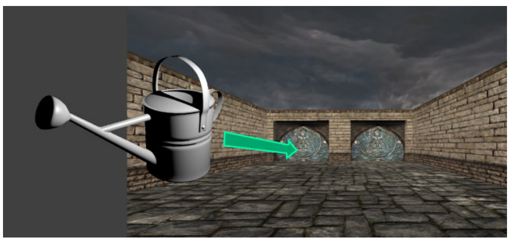

Figure 2. Visible object example and VE

The rating scores will be analysed to look for preference influence arising from the valence of the masked facial expression. It can be argue that preference ratings indicate some form of non-conscious understanding of the animated expressions. For example, visible objects presented alongside masked angry faces may engender a lower preference score than those placed alongside happy expressions.

A number of strategies can be employed to make it harder to remember a previous score given for a similar object that may be associated with a different expression. These should control any bias introduced by the participant trying to score consistently for the same object type.

Each visible object type would need to be associated with each expression at some point to prevent object preference biasing the results. A control condition could be run consisting of the same visible elements but lacking the masked faces. This again would ensure that object preference was not a confounding variable. Also, using the same neutral texture on all the visible objects would prevent a preference bias from texture/object interaction where a particular object may be seen as more attractive with a particular texture than a different object with that same texture.

4 CONCLUSION

Previous work has shown that visual masking is possible using non-planar targets in a VE. The next challenge will be to see if a dynamic, non-planar target paradigm can be made to work. This could open up a whole new set of possibilities, particularly through using masked avatar faces to drive affective behavior. For instance, it could be applied to training scenarios where there was a need to assess how a person might respond under implicit but heightened 'emotional pressure'. This may also help to increase the sometimes dampened affective responses found within VEs.

REFERENCES

[1] B.D. Breitmeyer & H. Öğmen. Visual Masking: Time Slices Through Conscious and Unconscious Vision. Oxford University Press, UK, 2006.

[2] S. De La Rosa, B. Schneider & G. Moraglia. Binocular unmasking is important for the detection, categorization and identification of noise-masked real-life objects. *Proc. XXIX International Congress of Psychology*, 2008.

[3] U. Dimberg, M. Thunberg, & K. Elmehed. Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*, 11: 86-89, January 2000.

[4] J.Drummond, N.Berthouze & A.J.Steed. Affective Reactions to Visually Masked Stimuli within a Virtual Environment, *Proceedings of CyberTherapy2011*, awaiting press.

[5] R.R. Hassin, J.S. Uleman & J.A. Bargh (editors) *The New Unconscious*. Oxford University Press, NY, 2005.

[6] E. Krumhuber, D. Cosker, A. Manstead, D. Marshall, & P.L. Rosin. Temporal dynamics of smiling: Human versus synthetic faces. *XIVth Conference of the International Society for Research on Emotions*, poster, July 2005.

[7] S. Lehmkuhle & R.Fox. Effect of depth separation on metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, 6: 605–621, November 1980.

[8] N.G. Martin. *Human Neuropsychology*. Prentice Hall, Europe, 1998.

[9] E. Moser, B. Derntl, S. Robinson, B. Fink, R.C. Gur & K. Grammer. Amygdala Activation at 3T in Response to Human and Avatar Facial Expressions of Emotions. *Journal of Neuroscience Methods*, 161: 126-133, March 2007.

[10] M. Slater, A.J. Steed & Y. Chrysanthou. *Computer Graphics and Virtual Environments: From Realism to Real Time*. Addison Wesley/Pearson Education, 2002.

[11] P. Winkielman & K.C. Berridge. Unconscious Emotion. *American Psychological Society*, 13: 3, 2004.