The Use of Depth in Change Detection and Multiple Object Tracking

A. Dünser*,[1] and G. Mancero[2]

[1]HIT Lab NZ, University of Canterbury, New Zealand

[2]Middlesex University, Engineering and Information Sciences, London, UK

Abstract: Users require to quickly and reliably process information in visually complex scenarios. Detecting changes in visual displays and tracking multiple moving targets are important in tasks ranging from surveillance, air traffic control and process management to gaming. By integrating depth information into visual displays we may alleviate some of the difficulties users face in these tasks. In this paper, we report on a study investigating the allocation of attention in three-dimensional space and on the use of depth in the detection of visual changes and multiple object tracking. For this we developed a task combining change detection with multiple object tracking. Stimuli were presented on a Multi Layer Display that allows displaying information on different depth layers. We found that participants detected colour changes faster than changes in depth and there was no additional benefit in combining colour and depth change. They could track more simultaneously moving objects correctly when they were equally distributed across two depth layers. Increasing the complexity of the tracking task had less effect on performance in a concurrent change detection task when the tracking objects were distributed across two depth layers.

1. INTRODUCTION

Day to day experience is filled with situations like driving, sports, or even crossing the street, that call for sustained attention to multiple objects [1]. Change detection and object tracking are important tasks in fields such as air traffic control, emergency ambulance dispatch, engineering, and medicine, where operators have to deal with visually complex scenarios and quickly and reliably perceive relevant information. Numerous systems and processes rely heavily on visual displays to convey information, so failures in the detection of changing information may have implications for human-computer interface design [2]. In many of these systems operators might have to deal with and track multiple moving objects simultaneously. This is often combined with other concurrent tasks, which compete for the operator's attention.

In many real world tasks such as driving, flying or process management, people monitor multiple information sources over an extended period of time [3-5]. Experts learn to allocate their attention in an adaptive way, shifting attention between various channels with a frequency determined by their relative importance and bandwidth [4]. Even when people's scanning strategies are well tuned, it may still be useful to alert the driver, pilot, or operator to unexpected, infrequent, or high-priority events, interrupting the normal path of attentional scanning to ensure that important information is quickly encoded [4, 6]. In such instances, effective design requires cues to guide attention to crucial information in a bottom up manner [7]. Segregating information in depth is one way to allocate attention to different foci which might help to better deal with visually complex tasks.

*Address correspondence to this author at the HIT Lab NZ, University of Canterbury, New Zealand; Tel: +64 3 364 2349; Fax: +64 3 364 2095; E-mails: andreas.duenser@hitlabnz.org, a.duenser@gmx.net

The purpose of this research is to investigate the utility of a Multi Layer Display (MLD) for change detection and multiple object tracking tasks. We aim to determine whether visual depth enhances the detection of changes in visual displays and how the allocation of attention to different layers in depth influences multiple-element tracking performance. We discuss issues of attentional control, present Multi Layer Displays as one means of displaying actual depth information to the user, and discuss related work on the allocation of attention in 3D.

2. MULTI LAYER DISPLAYS

A Multi Layer Display (MLD) [8] is a device in which two LCD displays are stacked on top of each other, separated by a transparent layer. This allows visual information to be presented on two different physical layers. Information that is displayed on the back screen is visible through the front screen. Compared to other displays that use stereoscopic depth cues, MLDs show actual depth created by the physical separation of the two screens. Although human perception has been comprehensively studied with interfaces using stereoscopic depth or other simulated depth cues, few studies have explored such issues with actual depth displays.

When using MLD technology the user can perceive depth information without needing an additional apparatus such as glasses (e.g. polarised, coloured) or mirror setups to separate views. In workplace environments where the use of such additional equipment is either inconvenient or can hinder task execution, this can be an important factor. Thus MLD technology could be an option in such cases.

The visual affordances of a MLD [9] imply that the MLD has the potential to improve various aspects of visual information search and detection. However, only some of these affordances have been explored with formal user studies. In general, research suggests that depth information

1875-9343/09

can be valuable in complex search tasks, multiple object tracking, alerting or change detection tasks and divided attention tasks.

A Stroop test setup was used by Aboelsaadat and Balakrishnan [10] to study interference with one layer and two layer displays. For spatially overlapping stimuli they found that interference from the foreground stimuli on the perception of the background stimuli is higher with two layer displays. Performance degrades using two layer displays when the stimuli semantically compete for user attention. For non-spatially overlapping stimuli, they found performance depends on the assignment of stimuli to various layers, with the single layer display equalling or outperforming the two-layer display in all cases.

Prema *et al*. [11] developed and analysed new rendering techniques for the MLD. They did not find a general technique that works best for all applications but they gave some guidelines for producing effective scenes and enhancing perception. These guidelines include emphasizing important objects by displaying them on different layers, separating datasets across different layers, extruding objects across layers, transitioning objects smoothly between layers, and making use of the transparency of the front layer.

Depth information could also help to de-clutter displays. Wickens and Hollands [12] argue that in displays designed to facilitate parallel processing it is sometimes difficult to narrow the focus of attention and shut out unwanted inputs. However, depth cues can be used to assist with focusing on relevant information and they found that search times were shorter for targets separated in depth. The experiment of Hayes, Wong, and Moore [13] indicates that the MLD can offer benefits in helping users to focus on relevant information, and reducing visual clutter while still retaining all the information necessary for maintaining awareness of the overall situation.

Wong *et al*. [14] explored the effectiveness of using depth and alpha-blending to create varying levels of transparency and a sense of visual depth, while comparing objects presented on both layers of a MLD to a control condition using a Single Layer Display (SLD). They found that under easy task conditions there was no difference in response times for selecting targets between MLD and SLD conditions. However, in more complex conditions, such as the need to perform cognitively demanding comparisons, the MLD showed significant improvements over the SLD.

3. ATTENTIONAL CONTROL

3.1. Multiple Object Tracking and the Allocation of Attention in 3D

Multiple Object Tracking (MOT) tasks have been useful tools for studying the deployment of limited-capacity visual resources over time [15]. Pylyshyn and Storm [16] introduced the MOT paradigm and suggested that the elements must be tracked in parallel instead of using a serial process.

The MOT paradigm has also been used to investigate conditions under which depth may aid the allocation of attention when the visual system must simultaneously track a subset of identical moving objects. Several studies have explored the allocation of attention in 3D space, with apparently inconsistent results [17]. Some experiments showed that attention cannot be directed to locations in 3D space [18, 19], whereas other studies indicate that attention can be deployed to different depth locations [20-24].

Studies have shown that it is possible to focus attention on a particular depth plane defined by binocular disparity [20, 21, 24]. Nakayama and Silverman found conjunctive tasks combining stereoscopic disparity (20 arc min) with either colour or motion were qualitatively different and much easier than other conjunctive searches. They argued that the visual system can perform a parallel search in one depth plane without interference from target-like distractors in another depth plane [20]. Theeuwes *et al*. [24] found that directing attention to a particular depth plane is possible and that it provides processing speed benefits even when detecting a target defined by a single feature. When the colours of the target and distracters are identical attention can be captured from another depth plane. However, when the colors of the target and distracters are different directing attention to a particular depth plane can prevent attentional capture from another depth plane [24].

Viswanathan and Mingolla [25] showed that performance in a multi-element tracking task improves when attention is allocated across two depth planes instead of within a single depth plane. They used stereoscopic displays in which the disparity of the closest surface was -0.13 degrees, and the far surface was +0.13 degrees. Results showed that both the depth factor and the surface factor proved to have strong influence on performance in a multi-element tracking task. They concluded that it is possible to selectively attend to targets that move in depth as well as horizontally and vertically in the presence of identical distracters that move in a similar fashion.

A study by Bolia *et al*. [25] incorporated the multiple tracking paradigm in an experiment with a MLD that did not produce conclusive results. The authors compared depth and transparency in single and dual task conditions (a MOT and a relatively simple digit pair task). In the MLD condition stimuli were displayed on different depth planes, while in the SLD transparency was used. The authors did not find that depth had any effect on performance. They point out that multi-task experiments should be designed so that they are demanding enough in mental workload to avoid ceiling effects. Ideally tasks with varying difficulty should be considered.

3.2. Change Blindness

Change blindness refers to the failure to see large changes in objects and scenes when the changes are gradual [26] or happen during a visual disruption [27-30]. Studies in laboratory and real-world settings have found that people regularly fail to detect visual changes that take place within their field of view [27-31]. According to Rensink [30], change detection is the apprehension of change in the world around us, "it denotes not only proper detection (reporting on the existence of the change), but also identification (reporting what the change is) and localization (reporting where it is)" (p. 246).

An important distinction must be drawn between change and motion. According to Rensink, motion is defined as a variation referenced to location, whereas change is referenced to structure [30]. This distinction has important implications in the perceptual processes involved since motion is detected in the initial stages of visual processing whereas change is processed by more complex structures that must maintain spatiotemporal continuity [30].

Rensink [30] also makes a distinction between detection of dynamic and completed changes. The former refers to the perception of a change in progress, the transformation itself. The latter refers to the perception of a structural change at some point in the past. Presumably, dynamic changes are easier to detect because the signal caused by the change draws attention, facilitating detection [30].

Many researchers agree that focused attention is needed to see change [29, 32, 33]. Visual transients – detectable visual cues that signal a change in the environment over time- such as colour, flashing, highlighting, boxing, and motion have been well studied as cues to capture attention in visual displays [34-36], however the efficiency of those cues varies between situations.

At least three factors help to determine the likelihood with which a visual signal captures and holds attention: stimulus salience, stimulus newness and the observer's attentional set. Salient objects are those that differ substantially from their surrounding in some simple visual feature. According to Yantis [17], the appearance of a new perceptual object is an important perceptual event that has significant consequences for the deployment of attention. It is important to emphasize that new objects do not have absolute control over attention, but the visual system appears to be predisposed to attend to objects that require the creation of a new perceptual object representation. Finally, the ability of a visual signal to capture attention depends on the observer's readiness and goals. If the observer is not searching for a visual signal or is not ready for it, then it is more likely that the observer will miss it [4].

The role of depth for change detection was studied in [2] with the use of a MLD. With an inattentional blindness paradigm the authors studied the effect of depth in the detection of an unexpected event when attention was diverted to a primary task. Results showed that unexpected events in the front layer become more noticeable than at the rear layer. Detection was three times higher when the unexpected event was in the front layer and within 20 degrees of visual angle from the focus of attention. However, this study did not control for the condition in which all stimuli are presented in the front layer. The front layer generally produces a sharper image than the rear layer due to the Perspex layer in-between which might have influenced some of these results.

In a second study, the authors compared depth to colour as a cue to highlight changes. Changes were dynamic and no visual disruption was added. Overall the findings suggested that depth is not as efficient as colour for change detection. However, for changes that occurred outside the parafoveal region (over five and 30 degrees of visual angle), depth transients were more effective than colour-transients. Peripheral changes that were highlighted with a depth cue were detected with slightly longer latency than those at fixation, but with no less accuracy than the more central changes. The authors argued that depth should not be used as a sole cue to highlight changes, but could be used when colour and flashing are no longer appropriate [2].

3.3. Attention Models

"Attention seems to involve a perceptual resource that can both intentionally and automatically select – and be effortfully sustained on – particular stimuli or activities" [1, p50]. Previous research has attempted to resolve the issue of whether visual attention is solely spaced-based or object-based. The metaphor of the spotlight and the zoom-lens model can be adopted to characterize the space-based theory of attention. The spotlight model suggests that attention "illuminates" a restricted convex region of the image. It moves continuously from one location to the next as a spotlight sweeps across the surfaces that it illuminates as it moves [17, 37]. This is different from the "zoom-lens" model according to which the size but not the location of the attended region may change continuously [38]. This means that all visual information remaining within the beam will get processed even if it is unwanted information, therefore, causing a disruption on focus attention on the relevant information [4]. However, Multiple Object Tracking (MOT) studies have effectively ruled out the single-spotlight explanations *via* additional computational modelling which concluded that single-spotlight performance could never match actual human tracking abilities for those same trajectories [16]. Therefore, the underlying architecture of MOT "must involve parallel selection and tracking – perhaps including up to four separate loci of attention, which might then directly explain the fact that tracking suffers beyond this number of targets" [1].

Object-based attention models, on the other hand, argue that attention is preattentively directed to simple features on the display in addition to objects or regions [39, 40]. Hoffman and Mueller [41] concluded that object-based selection may operate on a representation that contains depth information. Much of the evidence on preattentive processing comes from visual search tasks, usually framed around Treisman's feature integration theory [42]. Pylyshyn's theory of visual indexing complements other theories of object-based attention by postulating a mechanism whereby pre-attentive object-based individuation, tracking and access are realized [43]. He argues that a limited number of spatial indexes (fingers of instantiation or FINSTs) can travel with a limited number of tracked objects [16]. Alvarez *et al.* [44] found that MOT tasks could be interrupted for brief intervals without being disrupted. They established that multiple object tracking and visual search do not continuously draw on the same attentional resources and that the two tasks must rely on some independent resources. Their findings support two possible conclusions: MOT and visual search require two different types of visual-spatial attention or one requires spatial-memory and the other requires attention.

Rensink's *Coherence Theory* [45] suggests that focused attention is needed to see change. Since only a small number of items can be attended at any time, most items in a scene will not have a stable, detailed representation. If attention has not been automatically directed to the change, change blindness occurs [45]. His theory assumes that object-based attention "is intimately involved with the formation of representational structures with spatiotemporal coherence" [46, p346]. Rensink [45] proposes a different view of attention, rather than being the main gateway of all visual perception; "attention is just one of several concurrent streams, namely the stream concerned with the conscious perception of coherent objects. The other streams do not rely on attention, and so can operate independently on it" [46, p. 66].

Such "non-attentional" streams are the systems that underlie motor actions such as reaching, grasping, and eye movement. These systems were discussed by Milner and Goodale [47] who suggested that visual perception and visual control of actions are governed by two different visual systems. This extended Ungerleider and Mishkin's work who found that there were two perception streams in the brain: the dorsal and the ventral stream [c.f., 48]. The dorsal stream determined *'where'* things were, the ventral stream determined *'what'* things were. Rensink argues that *what* requires attention while *where* does not [45]. Alvarez *et al*. established the possibility that *what* requires attention and *where* requires a separate type of attention [44]. If *where* does have an upper limit on capacity then it would be logical to describe this limitation in terms of some kind of attention, even if the *where* attentional mechanism is completely different from that use for *what*.

4. EXPERIMENT: MULTIPLE OBJECT TRACKING AND CHANGE DETECTION

In this experiment we investigated the deployment of attention in different depth planes in a dual task setup. Similar to Bolia *et al*. [25] we combined a MOT task with a change detection task in which targets changed in either colour or depth or both stimulus dimensions at the same time. We aimed at studying the utility of the MLD for these tasks and how visual depth enhances the detection of dynamic changes in visual displays as well as how the allocation of attention to different layers in depth influences multiple-object tracking performance.

4.1. Participants

Twenty university students (6 females and 14 males) aged between 18 and 31 years participated in the study. All participants had normal or corrected to normal vision.

4.2. Equipment

Stimuli were presented on a 17-inch MLD set at a resolution of 1024 x 768 pixels (per screen). The two LCD screens were physically separated by 7 mm. An X-keys USB programmable keyboard (20-keys) and a computer mouse were used as interaction devices.

4.3. Stimuli

The display contained 16 circles (radius = 10 pixels), of which either 4 or 6 were designated targets. After the targets flashed, all the items started moving in different random directions. Their trajectories were restricted so that they would not coincide with each other (in both layers) while they were moving. The movement phase lasted 12 seconds with a speed of 40 pixels/second.

The four targets for the change detection task had a radius of 12 pixels and were located in the four corners of the display. One out of these four objects changed for 300 msec. During the 12 seconds of each trial two or three changes occurred so that participants could not predict the exact number of changes. Participants had to press a button on the X-keys keyboard when they detected a change within 1000 msec. Keyboard presses after 1000 msec were counted as 'not detected'.

4.4. Design

The following conditions were tested:

1. MOT only
2. MOT with change detection task: objects changing colour (blue to red)
3. MOT with change detection task: objects changing depth (back to front layer)
4. MOT with change detection task: objects changing colour and depth

The MOT task changed pseudo randomly between one and two layers and four or six targets with conditions being equally represented within blocks of trials. In the one-layer case, all tracking items were restricted to lie on the same depth plane (the front layer). In the two-depth cases, targets and distracters were equally distributed between the two depth planes. Objects in the change detection task were located in the back layer. In case of change in depth, they 'jumped' to the front layer for 300 msec. Rensink [30] suggested that this approximately is the temporal range for detecting dynamic changes.

4.5. Procedure

The MOT task (Fig. **1**) showed a set of moving circles on the screen. Objects moved randomly around the screen using a momentum technique that ensured smooth movement. When two items approached each other too closely, their velocity vectors were set directly away from each other. In this way, no two objects ever overlapped, in one or two layers.

The participants were asked to track four or six targets from a total of 16 independently moving items. In each trial, the 16 items (drawn as identical blue circles) were initially shown in a static display. After 1s, the targets started flashing for 500 milliseconds (msec) and then all 16 objects began moving independently and randomly about the display. After 12 seconds of motion, the objects stopped moving, and the participant used the mouse to indicate (click) which of the 16 circles they believed the targets were (Fig. **1**).

In the change detection task the participants were required to monitor four objects in the corners of the display for changes in colour, depth, or colour and depth of distracter stimuli and respond (keystroke) as quickly as possible.

All trials were completed in one session that lasted approximately 45 minutes. The experiment included one practice block and one experimental block for each condition. The practice block included four trials, and the experimental block of 16 trials, resulting in 64 experimental trials per participant. The presentation order of the trials was randomized.

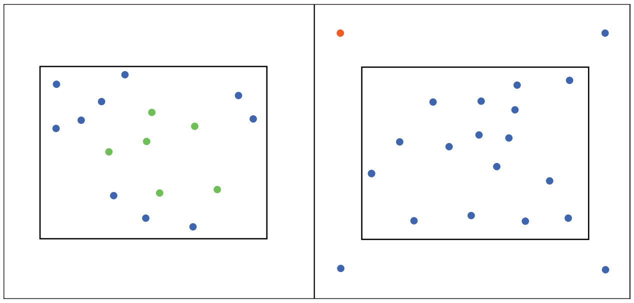

Fig. (1). Left: Selection of 6 objects after the MOT moving phase; Right: Example for MOT and change detection task (the object in the top left corner changes from blue / back layer to red / front layer).

After completing each trial block the participants were asked to fill out a questionnaire on how easy they thought the particular multiple object tracking condition was and how well they thought they performed (4-point Likert-scale). In addition the participants filled out six questions taken from NASA's TLX scale [49] (Mental demand, Physical demand, Temporal demand, Effort, Performance, and Frustration). After the condition without the change detection task (MOT only) the participants had to fill out separate questionnaires for the single and double layer conditions. In the end the participants were asked to rank the three conditions with the change detection task from 1 to 3 (1 being the best). Rankings were made for "Which condition was easiest?" and "In which condition did you perform best?"

4.6. Results

4.6.1. MOT Accuracy

The percentage of correctly tracked objects was used as a measure of the object tracking accuracy (Table **1**) and analysed with a 4 (conditions) x 2 (layers) x 2 (objects) within subjects ANOVA. All main effects were significant but none of the interactions.

Table 1. MOT: Percentage of Correctly Tracked Objects

	Layers				Total	
	Single		Double			
	Mean	SD	Mean	SD	Mean	SD
MOT only	82.79	13.45	84.31	13.02	83.55	13.18
Colour	74.81	16.52	79.27	16.29	77.04	16.46
Depth	74.41	18.71	73.58	17.93	73.99	18.21
Colour+Depth	73.13	17.41	75.16	17.86	74.14	17.55
Total	76.28	16.91	78.08	16.77		

Accuracy in the object tracking task was significantly different across the conditions ($F_{3,\ 57} = 14.37$, $p < .01$). Post Hoc analyses showed that participants achieved higher accuracy in the MOT only condition compared to conditions with change detection task. No difference in accuracy could be observed between conditions with the change detection task (Table **1**).

There also was a significant difference in object tracking accuracy between tracking objects in one or two layers ($F_{1,19} = 7.89$, $p = .01$). When the stimuli were distributed over two layers, there was a small but significant increase in the number of objects that participants could track correctly, independent of the number of objects that had to be tracked (78.08% correctly tracked objects in the double layer conditions compared to 76.26% in the single layer conditions).

Regarding the number of objects the participants had to track, accuracy was higher for simultaneously tracking 4 objects ($M = 83.88$) than for simultaneously tracking 6 objects ($M = 70.48$) ($F_{1,19} = 101.41$, $p < .01$).

4.6.2. Change Detection Task Performance

The reaction times for detecting changes in the change detection task were analysed. A 3 (change detection task conditions) x 2 (layers MOT task) x 2 (number of MOT objects) within subjects ANOVA was performed with change detection task reaction times as the dependent variable. The main effects for condition ($F_{2,\ 38} = 27.94$, $p < .01$) and objects ($F_{1,\ 19} = 18.57$, $p < .01$) were significant as well as the interaction layers x objects ($F_{1,19} = 6.53$, $p = .02$).

Table 2. Mean Correct Detection Times (Milliseconds)

	MOT Objects				Total	
	4		6			
	Mean	SD	Mean	SD	Mean	SD
Colour	590.89	109.87	629.38	124.62	610.13	118.33
Depth	699.89	138.17	749.48	162.33	724.69	151.84
Colour+Depth	563.45	93.09	591.99	96.46	577.72	95.28
Total	618.08	128.66	656.95	146.03		

Post Hoc analysis showed that in the change detection task, participants responded significantly faster ($p < .01$) to changes in colour (610.13 msec) or colour+depth (577.72 msec) compared to changes in depth only (699.89 msec). A closer look at the effect of objects shows that participants reacted faster to changes in the change detection task, when they had to track 4 objects in the MOT-task (see Table **2**).

The Post Hoc analysis for interaction of layers x objects (illustrated in Fig. **2**) showed that there was no effect of number of objects on change detection time when objects were distributed over two layers. However, change detection was slower when tracking six objects ($p < .01$).

The percentage of detected changes in the change detection task was analysed with a 3 (change detection task conditions) x

2 (layers MOT task) x 2 (number of MOT objects) within subjects ANOVA. Only the main effect of condition was significant ($F_{1.34,\ 25.53} = 14.54$, $p < .01$). Post Hoc analysis revealed that detection rates were lower for the depth only condition ($M = 83.14$) than for the colour ($M = 94.32$; $p = .01$) and colour+depth ($M = 96.74$; $p < .01$) conditions.

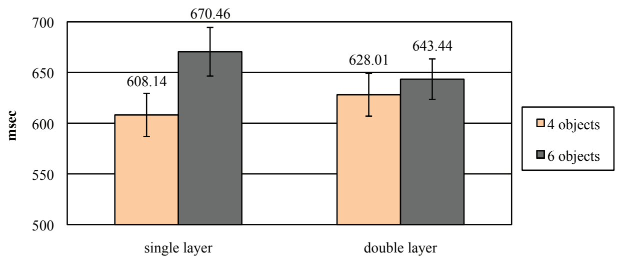

Fig. (2). Mean correct reaction times for detecting changes in the change detection task while performing the MOT task in single or double layer conditions tracking either four of six objects (error bars +/- SE).

4.6.3. Subjective Data

Ratings for Single and Double Layer MOT Trials. Participants' easiness and performance ratings of single and double layer MOT trials did not differ. However, comparing the answers on the TLX showed higher performance scale ratings for single layer trials ($p = .03$).

Table 3. User Ratings for Conditions (4-Point Likert Scales; Easy: 4 = Very Easy; Performance: 4 = Very Good)

	Easiness		Performance	
	Mean	**SD**	**Mean**	**SD**
MOT only	2.10	0.74	2.33	0.73
Colour	2.05	0.76	2.00	0.56
Depth	1.80	0.70	1.90	0.64
Colour+Depth	1.60	0.60	1.80	0.70

Ratings for Conditions. The perceived easiness differed significantly between the four conditions ($F_{3,57} = 2.95$, $p = .04$). Post hoc analyses showed that the condition with changes in colour+depth was not perceived as easy as the MOT only and colour conditions (see Table **3**). There was also a difference in the perceived performance ($F_{3,57} = 3.85$, $p = .01$). Post hoc analysis showed that performance ratings for the MOT only condition were higher than the ratings of the other conditions. Of the TLX scales, the performance scale was the only one that showed a significant difference between the conditions ($F_{3,57} = 4.23$, $p < .01$) with the MOT only condition having the highest scores.

Ranking. Of the conditions with change detection, the colour condition was ranked first for both easiness and performance, followed by the depth and the colour+depth conditions respectively.

5. DISCUSSION

We found that tracking performance was significantly more accurate when stimuli were distributed across the two depth layers. Independent of the number of objects to be tracked, participants were slightly more successful in tracking the objects correctly when the objects were in two distinct depth layers. While this finding replicates the results of Viswanathan and Mingolla [23] who found that performance in a MOT task improves as attention is allocated across two depth planes, with only 2% better tracking performance in the double layer condition, the improvement we found is rather small.

The tracking performance deteriorated under divided attention compared to MOT only trials. However, the tracking performance did not differ amongst the three conditions where MOT was combined with change detection tasks. Thus we only found a general effect of having to perform a concurrent change detection task on the MOT task, but the tracking performance was independent of the type of change detection.

In the change detection task the participants reacted faster to changes in colour or combined changes in colour+depth than to changes in depth only. The accuracy in detecting changes was lowest when the objects changed in depth only. Detection rates for changes in colour (blue to red) and colour+depth were higher than the depth-only changes and did not differ from each other. Overall this shows that colour changes are more effective than changes in depth while combining colour and depth is not any more advantageous than using colour only. Thus under divided attention (when participants had to concentrate on another relatively demanding task) changes in depth only seem to be not as noticeable as colour changes.

Earlier experiments have found that in a static display, when changes occur outside the parafoveal region, depth transients are more effective than colour-transients. However, in our study the target objects were located at the periphery, thus our results contradict these findings. Overall we found colour to be more effective than depth, and the combination of both cues did not improve detection rates.

We found an interesting interaction between the number of MOT objects and the number of layers. Results showed that if participants tracked objects from a single layer, the number of objects to be tracked had an impact on the concurrent change detection task. Change detection times increase significantly when a higher number of objects had to be tracked. In the double layer case the number of objects that had to be tracked did not affect change detection times. Hence increasing the complexity in the MOT task has less effect on change detection when the MOT objects were distributed over two depth layers. This might be due to lower mental workload for tracking objects in two depth layers. However, this assumption does not coincide with the participants' perceived workload. The MOT performance was rated higher for the single layer trials compared to the double layer trials.

In earlier studies [2] we also found advantages in arranging elements in different depth planes and that users were more accurate at change detection when the stimuli were distributed across two layers rather than confined to a single plane. Although participants were able to deploy their attention in both layers of the MLD, the capacity of tracking successfully up to four identical stimuli was not improved.

6. CONCLUSION

Many systems rely heavily on visual displays to convey information. However, research has shown that humans often have problems with managing visual attention. It is difficult to detect large visual changes and it is also hard to simultaneously monitor multiple sources of information. Effective design for visually complex user interfaces requires integrating effective cues such as visual transients to guide attention to crucial information.

This paper studied visual depth provided by a MLD as a tool for improving the management of visual attention. The issue of allocation of attention in different depth planes was investigated using a multiple object tracking (MOT) paradigm. We found that tracking moving objects from different depth layers can increase tracking accuracy. Thus dividing attention across two layers in depth appears to facilitate tracking multiple moving objects simultaneously. We also found that increasing the complexity of the MOT task has less effect on change detection performance when information is divided over two depth layers. This suggests that segregating stimuli in depth might lower mental workload and therefore could facilitate the design of multi-task environments with less interference across tasks.

Our research uncovered some interesting issues for the design and use of systems intended for monitoring and controlling multiple entities. When designing interfaces for monitoring and change detection tasks, designers should consider the applications of both space and object-based attention theories. Care should be taken to design interface tools in such a way that changes in the displayed information can easily be retrieved. Displays that allow the allocation of attention across depth layers such as the MLD could reduce the effort required to move attention between locations. This can allow quick access to and integration of information.

ACKNOWLEDGEMENTS

The study was conducted as part of a research project funded by PureDepth Inc. Thanks to Stephen Gaukrodger for his help in programming the application for the experiment.

REFERENCES

[1] Scholl BJ. What have we learned about attention from multiple object tracking (and vice versa)? In: Trick DDL, Ed. Computation, cognition, and Pylyshyn. Cambridge, MA: MIT Press 2009; pp. 49-78.

[2] Mancero G, Wong W. An evaluation of perceptual depth to enhance change detection. Human Factors and Ergonomics Society 52nd Annual Meeting 2008; pp. 338-42.

[3] Schriver AT, Morrow DG, Wickens CD, Talleur DA. Expertise differences in attentional strategies related to pilot decision making. Human Factors and Ergonomics Society. 52nd Annual Meeting; 2008; pp. 21-5.

[4] Wickens CD, McCarley JS. Applied attention theory. Boca Raton: CRC Press 2008.

[5] Werner S, Thies B. Is change blindness attenuated by domain-specific expertise? An expert-novices comparison in change detection in football images. Vis Cogn 2000; 7: 163-73.

[6] Nikolic MI, Orr JM, Sarter NB. Why pilots miss the green box: how display context undermines attention capture? Int J Aviat Psychol 2004; 14(1):39-52.

[7] Chu MM. Contextual cueing of visual attention. Trends Cogn Sci 2000; 4: 170-8.

[8] PureDepth. WHITE PAPER Multi-Layer Displays 2007; [Electronic Version]. [Retrieved: 4 Sept 2008]. Available from: http://www.puredepth.com/Docs_Downloads/HardwareWhitePaper_Feb_2007.pdf

[9] Wong W. Technology affordances of the ActualDepth(TM) multi-layered display: TBG Grant Ref. Num. DVIL 0201: Human-Computer Interaction of the ActualDepth(TM) 3-D Display Technology 2003.

[10] Aboelsaadat W, Balakrishnan R. An empirical comparison of transparency on one and two layer displays. HCI 2004 – The British HCI Conference, September 6-10, 2004; pp. 53-67.

[11] Prema V, Roberts G, Wuensche BC. 3D visualisation techniques for multi-layer display technology. IVCNZ '06; 27-29 November, 2006. Great Barrier Island, New Zealand, 2006; pp. 251-6.

[12] Wickens D, Hollands J. Engineering psychology and human performance. 3rd ed. Upper Saddle River, NJ: Prentice Hall 2000.

[13] Hayes J, Wong W, Moore A. Information layering to de-clutter displays for emergency ambulance dispatch. In: Grote G, Rizzo A, Wong W, Eds. 13th European Conference of Cognitive Ergonomics ECCE13; 2006 20-22 September 2006; ETHZ, Zürich, Switzerland: ETHZ Press 2006.

[14] Wong W, Joyekurun R, Mansour H, Amaldi P, Nees A, Villanueva R. OzCHI '05. 19th conference of the computer-human interaction special interest group (CHISIG) of Australia on computer-human interaction. Canberra, Australia: CHISIG Australia 2005; pp. 1-10.

[15] Place SS, Wolfe JM. Multiple visual object juggling. J Vision 2005; 5(8): 27.

[16] Pylyshyn ZW, Storm RW. Tracking multiple independent targets: evidence for a parallel tracking mechanism. Spat Vis 1988; 3(3): 1-19.

[17] Yantis S. Control of visual attention. In: Pashler H, Ed. Attention. East Sussex, UK: Psychology Press 1998; pp. 223-56.

[18] Iavecchia HP, Folk CL. Shifting visual attention in stereographic displays: a timecourse analysis. Hum Factors 1995; 36: 606-18.

[19] Ghiradelli TG, Folk CL. Spatial cuing in a stereoscopic display: evidence for a "depth-blind" attentional spotlight. Psychon Bull Rev 1996; 3: 81-6.

[20] Nakayama K, Silverman GH. Serial and parallel processing of visual feature conjunctions. Nature 1986; 320(20): 264-5.

[23] Viswanathan L, Mingolla E. Dynamics of attention in depth: evidence from multi-element tracking. Perception 2002; 31(12): 1415-37.

[21] Holliday IE, Braddick OJ. Pre-attentive detection of a target defined by stereoscopic slant. Perception 1991; 20: 355 - 62.

[22] Downing C, Pinker S. The spatial structure of visual attention. In: Posner MI, Martin OSM, Eds. Attention and performance XI. Hillsdale, NJ: Erlbaum 1985; pp. 171-87.

[23] Viswanathan L, Mingolla E. Dynamics of attention in depth: Evidence from multi-element tracking. Perception 2002 2002; 31(12): 1415-37.

[24] Theeuwes J, Atchley P, Kramer AF. Attentional control within 3-D space. J Exp Psychol Hum Percept Perform 1998; 24(5): 1476-85.

[25] Bolia RS, Nelson WT, Middendorf MS, Guilliams NM. Evaluating the utility of a multi-layered visual display for air battle managers. Human Factors and Ergonomics Society. 48th Annual Meeting; 20-24 September 2004, New Orleans, Louisana: Santa Monica, CA, 2004; pp. 21-5.

[26] Simons DJ, Franconeri S, Reimer R. Change blindness in the absence of a visual disruption. Perception 2000; 29: 1143-54.

[27] Rensink R, O'Regan JK, Clark J. To see or not to see: the need for attention to perceive changes in the scene. Psychol Sci 1997; 8: 368-73.

[28] Rensink R, O'Regan JK, Clark J. Change-blindness as a result of mudsplashes. Nature 1999; 398(6722): 34.

[29] O'Regan JK. Change Blindness 2000; [Retrieved November 7 2006]. Available from: http://nivea.psycho.univ-paris5.fr/ECS/ECS-CB.html

[30] Rensink R. Change detection. Ann Rev Psychol 2002; 53: 245-77.

[31] Simons DJ, Levin D. Change blindness. Trends Cogn Sci 1997; 1(7): 261-7.

[32] Rensink R. The dynamic representations of scenes. Vis Cogn 2000; 7: 345-76.

[33] Simons DJ. Current approaches to change blindness. Vis Cogn 2000; 7(1/2/3): 1-15.

[34] Fisher DL, Coury BG, Tengs TO, Duffy SA. Minimizing the time to search visual displays: the role of highlighting. Hum Factors 1989; 31(2): 167-82.

[35] Tullis T. Screen design. In: Helander M, Landauer TK, Prabhu PV, Eds. Handbook of Human-Computer Interaction. Massachussets, USA: Elsevier 1997; pp. 503-29.

[36] Bartram L, Ware C, Calvert T. Moticons: detection, distraction and task. Int J Hum Comput Stud 2003; 58(5): 515-45.

[37] Posner MI. Orienting of attention. Q J Exp Psychol 1980; 32: 3-25.

[38] Eriksen CW, Hoffman JE. The extent of processing of noise elements during selective coding from visual displays. Percept Psychophys 1973;14:155 - 60.

[39] Most SB, Simons DJ, Scholl BJ, Chabris C. Sustained Inattentional Blindness: The Role of Location in the Detection of Unexpected Dynamic Events 2000 December 2000 [Retrieved 2009 October 30]; Available from: http://psyche.cs.monash.edu.au/v6/psyche-6-14-most.html

[40] Duncan J. Selective attention and the organization of visual information. J Exp Psychol Gen 1984; 113: 501-17.

[41] Hoffman JE, Mueller S. An in depth look at visual attention. 35th Annual Meeting of the Psychonomic Society: St. Louis 1994.

[42] Treisman A, Gelade G. A feature-integration theory of attention. Cogn Psychol 1980; 12: 97-136.

[43] LePore E, Pylyshyn ZW. What is cognitive science? Oxford: Wiley-Blackwell 1999.

[44] Alvarez G, Horowitz TS, Arsenio H, DiMase JS, Wolfe JM. Do multielement visual tracking and visual search draw continuously on the same visual attention resources? J Exp Psychol Hum Percept Perform 2005; 31(4): 643 - 67.

[45] Rensink R. Internal *vs* external information in visual perception. Symposium on Smart Graphics. Hawthorne: NY, USA 2002; pp. 63-70.

[46] Rensink R. Visual search for change: a probe into the nature of attentional processing. In: Simons DJ, Ed. Change blindness and visual memory. Canada: Psychology Press 2000; pp. 345-77.

[47] Milner D, Goodale M. The Visual Brain in Action. 1998 [Retrieved 2009 November 7]; Available from: http://journal psyche.org/ojs-2.2/index.php/psyche/article/viewFile/2438/2367

[48] Goodale MA, Milner AD. Separate visual pathways for perception and action. Trends Neurosci 1992; 15(1): 20-5.

[49] Hart SG, Staveland LE. Development of a multi-dimensional workload rating scale: results of empirical and theoretical research. In: Hancock PA, Meshkati N, Eds. Human mental workload. Amsterdam, The Netherlands: Elsevier 1988; pp. 139-83.

Received: May 21, 2009 Revised: August 26, 2009 Accepted: September 10, 2009