Do Velocity Vectors Support Multiple Object Tracking?

Krista Oinonen¹, Lauri Oksama¹, Esa Rantanen² and Jukka Hyönä¹

¹University of Turku, Finland
²Rochester Institute of Technology, Rochester, NY, USA

We examined the effect of velocity vectors on human ability to extrapolate movement of multiple moving objects. This was achieved using an identity tracking task, in which objects (line-drawings of familiar objects) first move for 5 s, after which they temporarily disappear from view. In order to examine different aspects of vector information, we compared 4 velocity vector conditions: (1) no vectors were displayed; (2) only a history trail was displayed; (3) a condition where the end of the velocity vector pointed to exactly where the objects would reappear after the masking; (4) a constant heading length was displayed. Based on recent studies on Keane and Pylyshyn (2006) and Oinonen et al. (2009), we hypothesized that the condition, where exact location information is presented, would improve performance compared to baseline. The results supported this hypothesis by indicating that history trail or directional vector do not improve performance but only complete visual information about the future location will help to anticipate movement in multiple object tracking.

INTRODUCTION

In many real-life dynamic visual environments, such as traffic, sports, or air traffic control (ATC), it is important to anticipate in what direction critical objects are about to move and where they would be positioned in the near future. In order to aid to predict the targets’ trajectories into the future, velocity vectors (or, track velocity and direction vector lines) are common in ATC plan view displays (PVD). The velocity vectors achieve a very simple thing of making the dynamic characteristics of targets (i.e., aircraft) on controllers’ display visually available. Without such visualization the controller would be under a heavy cognitive burden having to rely on past experience on the target dynamics (memory) and integration of a large number of circumstantial factors to predict the targets’ trajectories into the future.

It appears that the putative benefits of visual representation of target dynamics were almost coincidental to the old radar technology and cathode ray tube (CRT) displays. Because the radar antenna continuously rotates and thus ‘sweeps’ the atmosphere for returns from solid objects, a target is only ‘illuminated’ for the moment the radar antenna is pointed at it; phosphorus coating of the CRT displays provided an afterglow of the radar return (Cole, 1985), forming a history trail for a moving target for past 3-5 positions (depending on the phosphorus decay rate). From such history trails controllers can glean targets’ direction, velocity (the distance between successive past positions), and rate of turn information, and extrapolate from these the future trajectories of the targets. In subsequent digital radar PVDs the history of the targets continued to be provided, together with the computed velocity vector projected ahead of the target (Nolan, 1998). However, remarkably little research has been published on the theoretical underpinnings of predictor displays in ATC.

But are vectors really helpful? Do they really help in movement extrapolation? We raise these questions as recent studies on tracking of multiple moving objects that temporarily disappear from the observer’s field of view have suggested that observers are not able to update or anticipate objects’ locations (Keane & Pylyshyn, 2006; Oinonen, Oksama & Hyönä, 2009). Our ability to extrapolate movement of multiple moving objects during interruption or occlusion hence appears to be quite poor.

According to Keane and Pylyshyn (2006) observers can only remember where the critical objects were located at the time of their disappearance or slightly before that. Even knowing the objects’ movement directions does not help observers to extrapolate and predict the future locations of multiple objects. According to the displacement hypothesis of Keane and Pylyshyn, predicting object’s location is the poorer the farther the objects move while invisible, causing bigger displacement from their last visible locations.

These results seem to cast a shadow on the intuitive idea that observers can easily anticipate multiple objects’ movement and benefit from velocity vector information. In particular, it follows from these studies that vectors including only directional information (heading or history) may not be beneficial at all, as they do not give any exact visual information on targets’ future locations. Cognitively demanding movement extrapolation process is still required. Instead, only vectors that provide exact information on targets’ future location may be truly helpful.

Because these experimental tasks are clearly analogous to those of air traffic controllers, where the occlusion of targets corresponds to the controller taking his or her eyes off the PVD to perform other tasks, an ATC task offers an attractive approach to deeper understanding of the contribution of velocity vectors to human multiple object tracking performance.

Past Research on Predictor Displays

Much of predictor display literature pertains to control engineering (e.g., Kelley, 1962; 1968) and active vehicular control with applications in nuclear submarines (Berbert & Kelley, 1962) and lunar landing (Fargel & Ullbrich, 1963). For a flavor of this body of literature and further examples of driving an automobile in traffic, a blind pedestrian using a cane or electronic obstacle detector, and remote manipulation of solid
Velocity Vectors in ATC

There is little empirical research on different predictor displays in ATC, despite their ubiquitous presence on most modern ATC systems. Perhaps this is due to the long evolution of predictor information on ATC displays and the recent research focus on the development of cockpit displays (i.e., CDTIs). Unfortunately, the results from CDTI research generalize poorly to ATC because the air traffic controller’s task is fundamentally different from the pilot’s. In all of the experiments reviewed in the previous section the task was to estimate at the most two targets’ relative position in the immediate future, or compare a vehicle’s state to some criterion value in direct vehicular control. However, while controllers need to estimate aircraft trajectories and judge potential conflicts between aircraft in a pair, they also must maintain situation awareness on a much larger scale, involving tens of aircraft in busy sectors. The research paradigm of tracking multiple moving objects that are temporarily masked appears uniquely suitable for examination of the effects of velocity vectors on tracking performance in a task that is closely analogous to that of air traffic controllers.

The present study

In the present study we were interested in the influence of velocity vectors on movement extrapolation. A version of multiple identity tracking task (MIT, see Oksama & Hyönnä, 2004, 2008) was employed, where at some point all moving targets temporarily disappear from the screen (Oinonen, Oksama, & Hyönnä, 2009). Participants were instructed to continue tracking the invisibly moving targets and extrapolate their movement. When the targets reappeared, the participants’ task was to indicate as fast as possible the current location of one auditorily probed target. In MIT, the moving elements all have distinct identities analogously to real-life tracking tasks; thus, the task required continuous updating of what–where-bindings. The task was designed to be a sensitive measure of immediate situation awareness of the observers.

In order to examine different features of vector information, we compared four velocity vector conditions (no vectors, only a history trail, velocity vectors corresponding to the masking duration, and heading vectors with constant length) over three target masking durations. Based on Keane and Pylyshyn (2006) and Oinonen et al. (2009) we expected that the...
vectors corresponding to masking duration, and thereby visually presenting exact target location into the future, would improve performance the most compared to the baseline.

METHOD

Participants

Twenty undergraduate psychology students (median age 24 years) from University of Turku, Finland, volunteered for the study.

Apparatus and Experimental Task

The stimuli were presented on a 19-inch Samsung SyncMaster monitor with a resolution of 1280 by 1024 pixels controlled by NVIDIA GeForce FX5200 card, AMD Sempron 2400+, 1.66 GHz, 1.0 GB of RAM computer. The experiment was programmed with the E-prime software (Schneider, Eschman, & Zuccolotto, 2002a, 2002b). The program that generated the motion sequences was written in Visual Basic.

The stimuli included 8 black-and-white line drawings (familiar objects and animals), chosen from the picture corpus of Snowgrass and Vanderwart (1980). The pictures were chosen so that their names in Finnish began with a consonant and comprised 5 letters and 2 syllables. Their pronunciation time was about 700 ms. The names were also matched with respect to their frequency in a written language corpus (Laine & Virtanen, 1999). Pictures were 75 x 39 pixels in size, and extended a visual angle of 1.9 x 1.0 degrees.

The objects’ movement directions were randomly chosen from among the four intermediate compass directions (north-east, south-east, south-west, and north-west). Object speed was 2.42 degrees/s. Objects were allowed to move with intersecting trajectories. The objects moved for 5 s. after which they disappeared from the display but continued to move. Participants were instructed to continue to track the invisibly moving targets and extrapolate their movement during masking (disappearance). After a given time the objects reappeared and the participants’ task was to immediately locate and indicate the auditorily designated target on the screen.

Design

Independent variables. We manipulated the masking duration (1680, 2520 and 3360 ms, resulting in the objects’ displacement during masking of approximately 4, 6, and 8 degrees of visual angle, respectively) and the number of targets to be tracked (2 and 3 out of the total of 8 objects). We also employed four velocity vector conditions: (1) in the baseline (B) condition, no vectors were displayed; (2) in the track (T) condition, only a history trail was displayed, indicating the objects’ movement direction but not their future positions; (3) in the congruent (C) condition the velocity vectors corresponded to the masking duration, that is, the end of the velocity vector pointed to exactly where the objects would reappear after the masking; (4) finally, in the heading (H) condition the objects’ velocity vector lengths were constant predicting the objects’ positions 2520 ms into the future. Note that this matched the congruent condition when the masking duration was 2520 ms, but with masking duration of 1680 ms the vector pointed farther and with 3360 ms masking duration closer than where the objects would reappear. This condition tested whether the length of the vector aided in anticipation of the targets’ movement. In all visual vector conditions (track, congruent and heading) the vectors were presented in all eight objects.

RESULTS

The results supported our hypothesis. The congruent condition where the vector length corresponded to the masking duration and thus predicted the target position upon its reappearance, resulted in the fastest identification (see Figure 1).

Before analysis, erroneous responses (1.3%) were removed from the data. A few response time outliers (0.6%) were also removed (z < -3.29 or z > 3.29). The remaining response time data were analyzed with a 2 (Set-Size) x 3 (Masking Duration) x 4 (Vector Condition) repeated measures analysis of variance (ANOVA). A Greenhouse-Geisser correction was applied to the p-values whenever needed.
The main effect of set-size was significant, $F(1, 19) = 233.53, p < .001$; RTs were lengthened as the set-size increased. The main effect of masking duration was significant as well, $F(2, 38) = 10.77, p < .001$; RTs were lengthened as a function of masking duration. The main effect of vector condition was also significant, $F(3, 57) = 4.49, p < .01$. Planned comparisons were conducted where the baseline condition was compared to other vector conditions. Comparisons indicated that the congruent vector condition differed from the baseline, $F(1, 19) = 6.51, p < .05$; RTs were faster in the congruent vector than in the baseline condition. The other vector conditions did not differ from baseline; track, $F(1, 19) = 1.68, p > .20$; heading, $F < 1$.

The interaction between set-size and masking duration was significant, $F(2, 28) = 5.84, p < .05$; with the effect of masking duration was more robust for set-size three than set-size two. The other interactions were not significant (Set-size x Vector Condition, $F(3, 57) = 1.87, p > .10$; Masking Duration x Vector Condition, $F < 1$; 3-way, $F < 1$).}

**DISCUSSION**

Our study examined the influence of velocity vectors on movement extrapolation when compared to the condition where no visual vectors were present. Surprisingly, neither the history trail nor the forward heading vector with constant length (not pointing to the future location in all trials) did improve performance. If anything, the history trail condition was slightly worse than the baseline condition (no velocity information at all).

Our results suggest that only explicit visual information about the future locations of invisibly moving targets will help to anticipate their locations after occlusion. On the other hand, movement history or directional information of targets’ trajectory does little to help on movement extrapolation with multiple targets. Our results are consistent with Keane and Pylyshyn’s (2006) displacement hypothesis. According to this hypothesis, when objects disappear from view, observers maintain the last visible locations of the target objects in the visual short term memory (VSTM). When the objects reappear, this VSTM representation is used as a reference in relocating the targets. This view is also supported by Oksama and Hyönä (2008), whose model of multiple identity tracking (MOMIT) suggests that in order to maintain an up-to-date representation of multiple moving targets focal attention is serially switched between the targets. When focal attention is directed to a target element its identity-location binding is refreshed and updated. And when a target becomes unattended, or in this case invisible, its last visible and attended location is stored in the VSTM. Therefore, relocating a recently unattended target relies on stored information on that target’s last position. Congruent vectors provide explicit location information of the targets’ position after occlusion, which information can then be stored in VSTM for future use. This is why congruent vectors are superior to other vector conditions in multiple identity tracking. This may also explain why the trail condition produced the worst performance. The trail may draw observers’ visual attention toward it and thus bias movement extrapolation in the wrong direction.

**REFERENCES**


