When are Motion Trajectories Perceived Correctly?

Amanda Alvarez, David Hoffman, Martin S. Banks

Abstract—Recent studies have argued that trajectories of motion-in-depth stimuli are perceived quite inaccurately [1]. Here we investigate the perception of motion trajectories in the real world and on displays. The different cue availabilities in these viewing environments makes this an interesting topic for both vision and visualization research. The perceptual errors in judging motions on displays can be caused by the absence of appropriate changes in image size (looming), the absence of appropriate changes in focus cues (blur and accommodation), and the use of an estimation procedure that is subject to the response-mapping problem. To better understand the causes of these reported errors, we manipulated the availability of disparity, looming, and focus cues, and used a response measure that should not be subject to the mapping problem. Observers viewed stimuli that moved in elliptical paths in depth and judged whether the path was too compressed or stretched to be circular. Real world motion of a LED served as a condition in which all trajectory cues were consistent. In a second condition, an equivalent stimulus was shown on a computer display; in this case focus cues specified a frontoparallel path while other cues were veridical. In a third condition, the computer-displayed stimulus had constant angular size; in this case, looming and focus cues specified a frontoparallel path. As cues to depth were removed observers perceived circular paths as increasingly compressed in depth. Accurate percepts occur only when all cues specify the same path. Specific measures to recover cues to motion may be necessary for visualizations to use motion as an effective visual variable.

Keywords—Motion-in-depth, displays, depth perception, motion trajectory, stereopsis, animation, information visualization.

◆

1 INTRODUCTION

This work addresses a question of experimental and practical interest in the field bridging vision science and visualization. This very general question asks: How does perception of real world stimuli differ from the perception of virtual, or computer displayed stimuli? The import of this question has perhaps been underappreciated in the visualization community, or the knowledge from the vision science literature has not been communicated. The answer to this question lies in the differing availability of visual cues in the two situations, real and virtual.

We are interested in the question of the salient visual cues for motion perception. Investigating the cue space is particularly relevant in the case of information visualization displays, since most flat displays provide a cue-inconsistent viewing environment. This problem becomes even more exacerbated in specialized viewing environments, such as immersive visualization. In general, the more specialized or restrictive the display is, the more potential there is for salient visual cues to be weakened compared to a real world viewing situation. Here we consider real world and virtual viewing situations and their associated cue spaces.

Animation and motion in visualizations can be compelling and informative, especially for signalling and integration of information [3] and for displaying transitions between states [2] and orienting the viewer. Motion is known to be a pre-attentive visual cue, and unlike other cues such as shape or color, it does not require fixation to be effective. Moreover, since motion embodies spatiotemporal change, it is well suited to signalling changes over time. The availability of visual cues to depth, structure, and motion trajectory, however, is reduced on computer displays. Cues are also weakened in experimental settings, where observers are found to make errors in judging motion trajectories [1]. Effective visualizations require correct perception of motion trajectories, which in turn requires that the visual cues are veridical and consistent. We looked at how removing cues from the viewing situation affects the perceived motion.

In this paper we investigate the visual cues used by human observers to judge motion trajectories. We will first discuss previous treatments of this and related questions within perception and visualization frameworks. Then we present the experiments conducted to assess the use of cues in real world and virtual motion viewing. We find that motions trajectories are only perceived correctly when the motion cues that specify them are in agreement.

2 RELATED WORK

The primary motivation for the present study comes from a recent paper by Harris & Drga [1]. They suggest that observers use visual direction (the angle between an object's location and the direction the observer is facing) instead of readily available binocular information to judge how an object has moved. The authors had participants view linear trajectories coming towards them, varying from 0 degrees (moving directly towards the observer) out to 20 degrees; see Figure 1A for the viewing parameters. Visual direction α and trajectory angle θ were manipulated by changing the x and z components of motion. In the first experiment x was varied and z was held constant. For all observers, the estimated angle θ' increased as θ increased; performance was in the correct direction, if somewhat overestimating θ. When z was varied and x was held constant, displaying the same θ angles, observers' estimates were no longer an increasing function of θ. For physical angles of θ that were quite different, observers gave the same estimate of motion direction. The surprising conclusion of this paper is that the pattern of errors seems to indicate that observers are using α and a constant estimate of viewing distance to make their motion judgments. As the authors point out, using the visual direction angle for motion judgments is disadvantageous because it does not specify unique motions, leading to the observed errors.

- *Amanda Alvarez <alvareza@berkeley.edu> and David Hoffman: Vision Science Program, University of California, Berkeley.*
- *Martin S. Banks: Vision Science Program, Department of Psychology, Helen Wills Neuroscience Institute. University of California, Berkeley.*

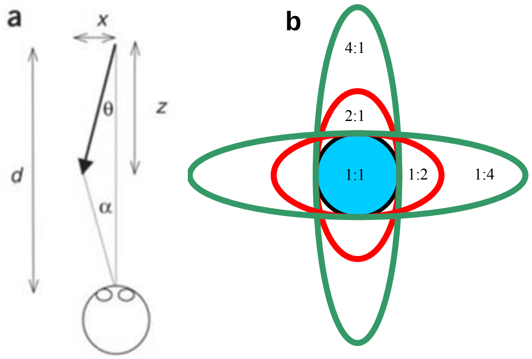

Figure 1. A. The stimulus geometry of [1], described in Section 2. The observer is facing straight ahead with viewing distance *d*. The visual direction of the target is α, its trajectory is θ. Reproduced from Figure 1A in [1]. B. A top view of the stimulus geometry in the elliptical motion experiments. Aspects ratios displayed were roughly within the bounds of the red ellipses.

We believe there a number of potential confounds in this study. The stimuli used were sufficiently small to not provide effective size cues; looming, a potentially important cue to depth, was not present. The appropriate changes in focus cues were also not present when stimuli were presented on displays. Although the experiments were reproduced with real 3D motion, the same concern of lacking looming cues still applies. Moreover, the experiments described above used an estimation procedure that is subject to the response-mapping problem. Subjects were asked to move a wooden pointer to reproduce the perceived 3D motion of the target. The mapping between the perceived angle and the angle of the pointer setting is unknown; that is, it is not known how the percept is translated into muscle movements, and with what accuracy. Since the percept is the interesting element, using a pointer is a needlessly ambiguous way to approach the problem. We have addressed this problem here by using a different response measure.

Our task seeks to addresses the limitations of this study. The observed errors in motion trajectory judgements in [1] were hypothesized to occur because viewers were not taking the viewing distance into account, and were only using two ‘snapshots’ of the motion trajectory to make their judgments. Unlike the linear motion task, our motion has no start and end points, so the task cannot be done based on perceptual ‘snapshots’.

Motion has also been investigated in visualization, though the emphasis has naturally been on usability, aesthetics, and coarse scale perceptual and cognitive judgments instead of the fine scale of basic visual processes. Tversky and colleagues [4] hold that animation in data graphics may not be effective because of a lack of equivalence between animations and their static counterparts. Animation may be particularly ineffective, the authors contend, because it violates principles of good graphics, congruence and apprehension. These are high level concepts that require elements of a visualization, such as its motion, to mirror the underlying concepts and be easily understood. Part of apprehension and congruence, we would argue, arises from correct perception of the basic stimuli, which draws back to the thesis of this paper: motion cannot be perceived correctly unless the cues specifying the motion agree.

It has been noted that apprehension of a data graphic can be greatly aided by using predictable motions and simple transitions [2]. Bartram [3] names periodicity as a guiding principle for the use of motion in visualizations, especially for signalling and grouping. Although Bertamini & Proffitt [5] contend that translations and divergence are easier to understand than rotations, the guidelines provided by both Heer & Robertson [2] and Bartram [2] indicate that the type of stimulus used here – an elliptical motion – may be well suited not only for psychophysics, but for representing a generic, periodic, predictable and easily apprehended motion path.

3 METHODS

For the pilot data described here, three observers with normal or corrected to normal vision participated. All were experienced psychophysical observers.

Real motion stimuli were displayed using a custom built elliptical orbiter. Elliptical paths were created by varying the sizes of two adjoining radii (see Figure 2 for an illustration). The properties of these motion paths are described in Section 3.1. In both the real and virtual cases, subjects viewed elliptical motion paths head on, that is, there was only motion in the *x* and *z* directions, not in the *y* direction. The task in both the real and virtual motion conditions was to judge whether the motion path was stretched or compressed compared to a circular path. A circle is used as the standard for a forced choice procedure, which allows a more direct measure of the percept. We have had success with similar forced choice procedures for judgments against other internal standards, such as that for a 90 degree open book hinge.

No fixed reference point was present during stimulus presentation, and no explicit instructions to fixate or track the stimulus were given. Each trial consisted of one clockwise revolution of the target.

3.1 Real elliptical motion

An ellipse is typically defined by its foci (two fixed points) and the distances from any point on the curve surrounding these foci. Specifically, an ellipse is the locus of points where the sum of the distances from any point on the curve to the foci is constant. We chose to define the real elliptical motion by two adjoined radii moving in opposite directions, as shown in Figure 2. The figure also gives an idea of the basic design of the real motion apparatus, the elliptical orbiter.

The size of the ellipse was varied by changing the radii of the two circles. We used radii of 4, 5, 6, and 7 cm for the large or main radius and varied the small radius to achieve the desired aspect ratios. Values of the small radius varied between ± 2.5 cm, negative values generating wide aspect ratios of less than 1. The small radius varied in increments of 0.25 cm. We used the method of constant stimuli in our real motion presentations. In each block of trials, the main radius remained constant while the small radius was changed after every other trial. The small radii within a block were randomized, as was the order of presentation of the main radius blocks. For the real motion conditions, subject DMH completed 432 trials, subject JOS completed 832 trials, and subject AAA completed 640 trials. Real motion is referred to as condition 1 in the results.

The real motion mechanism was manually powered and operated. This necessarily led to uncontrolled elements in the stimulus presentation. For example, the presentation contained accelerations and decelerations not present on the computer display stimuli. Measures were taken to make the presentation of real motion stimuli as consistent as possible across trials. As the results show, any differences in speed or length of the travelled path induced by manual operation had little adverse effect on observers’ judgements. There is also no reason to believe that confounds introduced by manual operation of the elliptical orbiter improved performance.

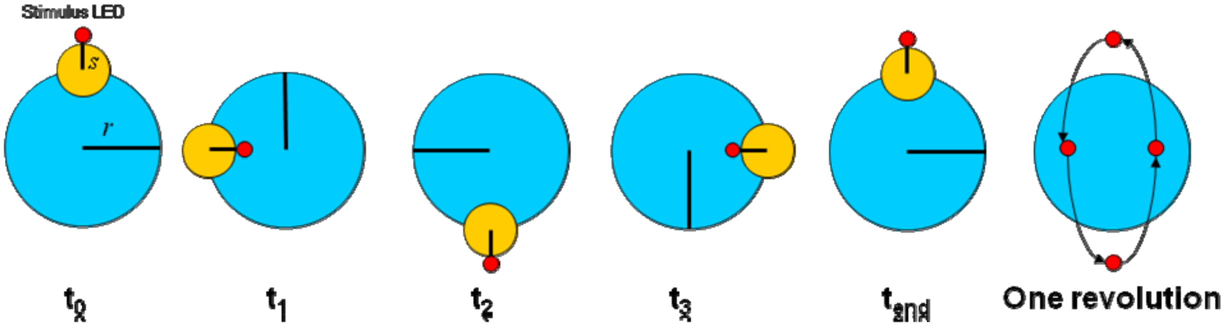

Figure 2. Top view of a real elliptical motion path produced by two adjoining circles of different radii. *s* = small radius, *r* = main radius

3.2 Computer displayed elliptical motion

For the computer displayed motion, stimuli were shown on a IBM T221 LCD display viewed through half-silvered mirrors; these create a volumetric display with three focal planes, of which only one was used for the present study. The resolution was 1920 x 1200 pixels (half the resolution of the monitor) and the refresh rate was 44 Hz. Stimuli were generated using OpenGL and C++. On the display, the main radii used were 4, 5, and 6 cm, and the small radii were between 0.2 and 3 cm. These sizes were used to provide maximum comparability with the real motion trials, though the largest main radius (7 cm) could not be used due to the limited lateral field of view of the volumetric display and the resulting clipping of the stimulus. The stimulus consisted of a red sphere with a 2.5 cm radius. During a trial, the stimulus appeared, was stationary for 900 msec, moved in one clockwise revolution for 1700 msec, returned to its starting position, remained visible for the same delay period (900 msec), and then disappeared. The response was given once the screen blanked out, after which a new trial would begin. No feedback was given.

Six adaptive staircases were used to control the presentation of stimuli on the display. 12 reversals of the staircase were required, with a maximum of 25 trials per staircase. Three of the staircases displayed stimuli with appropriate looming information while the other three did not. The first three thus comprise experiment 2A, the latter three experiment 2B.

3.3 Cues and conditions

In this study, we are interested in three visual cues to motion-in-depth:

- Disparity, the informative difference between the two eyes' images
- Focus cues: blur of the retinal image and accommodation to bring the image into focus
- Looming, or appropriate size change as an object moves in depth

In real world viewing, these cues are veridical and congruent, and specify the same motion path for a traveling object. On a flat display, focus cues are weakened. All light is emitted from one plane, instead of a volume like in real space. If a target moving in depth on a display does not change size appropriately, this further indicates that the path of motion must be frontoparallel or flat, instead of in depth. Thus, in condition 1, all cues specify a motion path in depth. In condition 2A, focus cues specify a frontoparallel path while other cues are veridical. In condition 2B, focus cues and a constant angular size of the target indicate a frontoparallel path. All conditions were viewed both monocularly and binocularly.

$$\text{Aspect ratio} = \frac{\text{main radius} + \text{small radius}}{\text{main radius} - \text{small radius}}$$

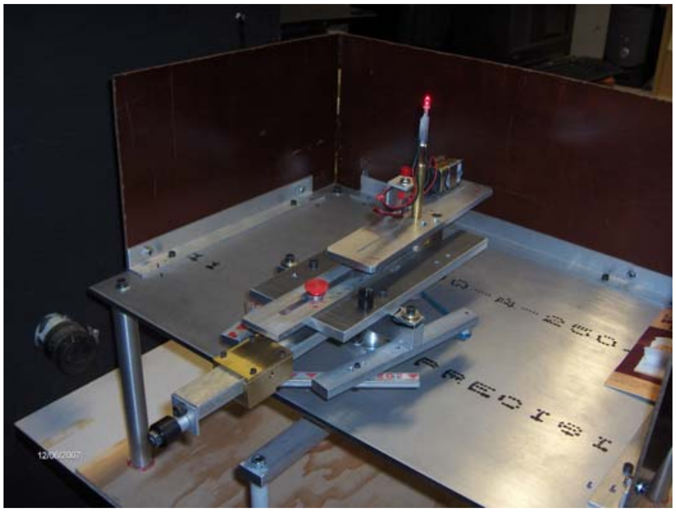

Figure 3A. Initial view of the real world motion apparatus (elliptical orbiter). The experimenter turns the crank under the mechanism. The subject sits on the other side of the panel and can only see the LED.

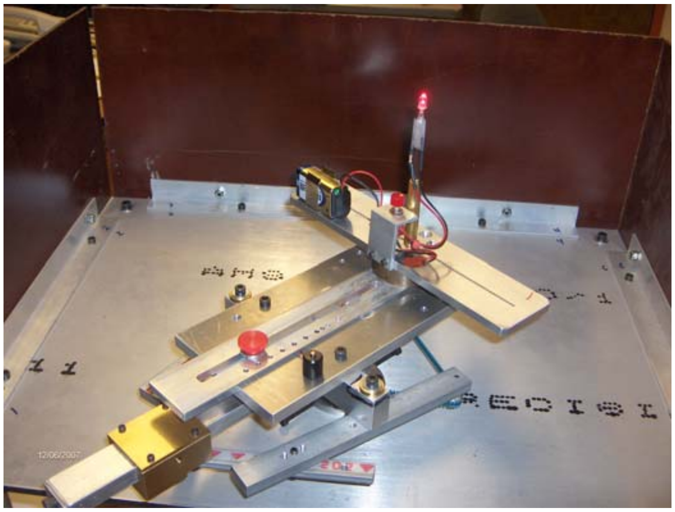

Figure 3B. View of the real world motion apparatus after rotation. The LED is at a small radius of 0 cm (the center of the top plate), so this path is a circle.

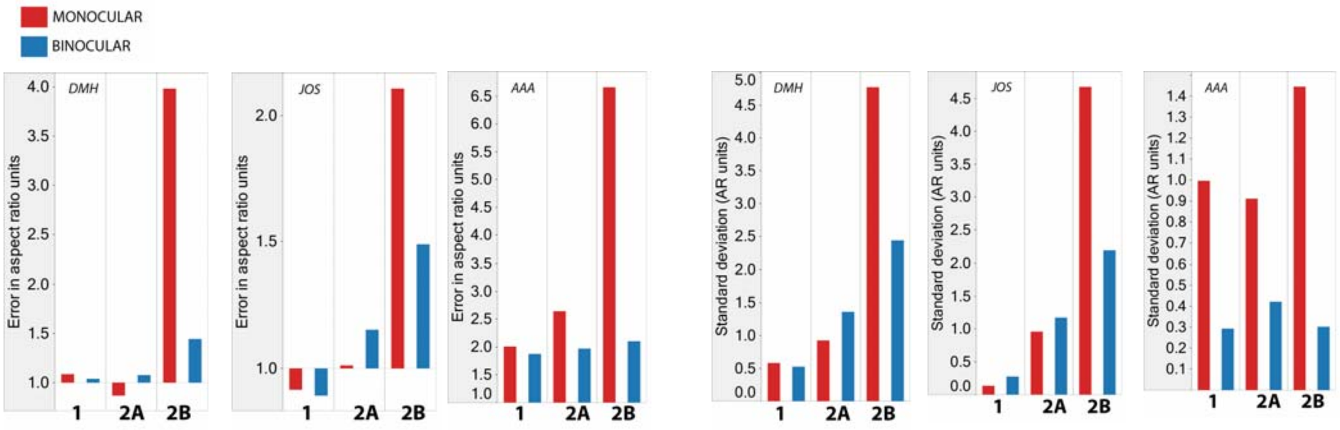

Figure 4. Results for three subjects from three conditions: 1 Real motion, 2A Display motion with looming, 2B Display motion without looming. Left three graphs show errors. 1.0 on the error ordinate indicates a percept of aspect ratio 1, i.e. no error. Right three graphs show standard deviations of same three subjects.

4 RESULTS

The results from the three conditions – real motion, display motion with looming, and display motion without looming – appear in Figure 4. These are averages across the different path sizes used, and across multiple runs. The ordinate of the bar charts is error in aspect ratio units, with the baseline of 1.0 representing a percept of a circular path with aspect ratio 1. Error is plotted for monocular and binocular trials in each condition, and for each subject. Positive error values indicate a percept of a path more stretched than a circle; negative error values indicate a percept wider or more compressed than a circle.

These bar charts, although they are based on the pilot data of only three subjects, already are indicative of certain trends. Accuracy was very good in the real motion case (condition 1). With computer displayed targets with appropriate looming (2A), errors were slightly greater than for the comparable real motion conditions. When looming was removed (2B), responses with binocular viewing were less accurate than in the other binocular conditions, and responses with monocular viewing were very inaccurate. (The inaccuracy displayed by subject AAA is likely due to having less data than for the other two subjects).

All three subjects displayed one pattern in particular: For targets of constant angular size (2B), the monocular condition was nearly impossible to do. Because the motion path appears flat in these cases, subjects continued to respond that the path was ‘wider’ than a circle, and so the staircase continued to display ever more ‘stretched’ motion paths. We expect that for subjects DMH and AAA, aspect ratios of roughly 4.0 and 6.5 do not represent the true percept of a circle in those conditions; rather, the staircase hit a ceiling. The idea of perceptual stretching and compression is also illustrated in Figure 1B. To observer DMH, for example, the green ellipse with aspect ratio 4:1 appeared circular when looming cues were not present. This is perceptual depth compression.

Not surprisingly, binocular viewing greatly improved accuracy. This would indicate that, contrary to the results of [1], binocular information serves a role as a useful and important cue for judging motion-in-depth trajectories. It is highly likely that subjects tracked the stimulus during its revolution. Therefore, there is good reason to believe that eye movement signals and not disparity are contributing to the percept.

5 CONCLUSION

Experimental situations and visualization environments lack the cue consistency of real world motion stimuli. The data show that observers are most accurate in their percepts for real motion. With the computer-displayed target and appropriate looming, responses with binocular and monocular viewing were somewhat less accurate for the corresponding real world motion. With fixed angular size, responses with binocular viewing were less accurate than the other binocular conditions and responses with monocular viewing were very inaccurate. As cues to depth were removed observers increasingly judged circular paths to be compressed in depth. Motion trajectories are perceived correctly only when all the cues specify the same path.

There are a great many additional manipulations that we intend to pursue in this framework. These include the use of three focal planes of the volumetric display and manipulation of focus cues to introduce a viewing situation on the display that more closely resembles natural viewing. We also plan to change the target size to assess the role of looming and introduce a fixation point to investigate the role of absolute versus relative disparity.

Once these conditions have been replicated on the volumetric display, this work may be particularly relevant to the use of 3D displays for visualization. Specialized visualization environments require specific measures to re-introduce the cue richness of real world viewing, and the case of motion serves to underscore how important these cues are for correct perception.

ACKNOWLEDGEMENTS

Ralph Bartholomew of the Optometry School machine shop built the elliptical orbiter. This work was supported in part by a grant from NIH.

REFERENCES

[1] J. Harris and V. Drga. Using visual direction in three-dimension motion perception. Nature Neuroscience, 8(2):229-233, February 2005.

[2] J. Heer and G. Robertson. Animated transitions in statistical data graphics. Proceedings of IEEE Information Visualization, 2007.

[3] L. Bartram. Enhancing Information Visualization with Motion. PhD thesis, 2001.

[4] B. Tversky, J. Morrison, M. Betancourt. Animation: Can it facilitate? International Journal of Human-Computer Studies, 57:247-262, 2002.

[5] M. Bertamini, D. Proffitt. Hierarchical Motion Organization in Random Dot Configurations. Journal of Experimental Psychology: Human Perception and Performance, 26(4):1371-1386, 2000.