277
B237

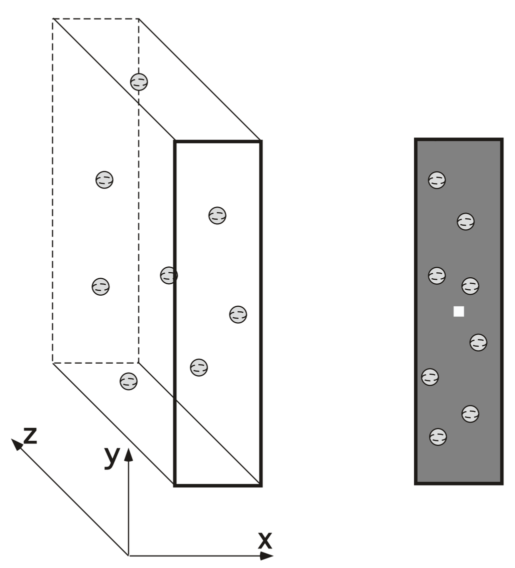
Measuring Attention During 3D Multielement Tracking

Erik Blaser, Zenon Pylyshyn, Fulvio Domini*

RuCCS, Rutgers University, New Jersey

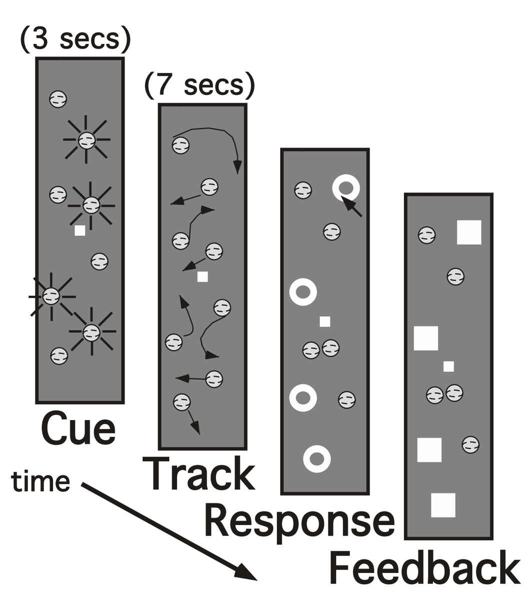
*University of California, Berkeley

What is 3D multielement tracking?



Observers are cued to track 4 out of 8 identical spheres as they move randomly in 3D. After tracking, the observer attempts to pick out the cued set.

After the response, the observer is given feedback and the next trial begins.



Experiments were run on an SGI, with disparity from LCD shutter glasses.

Each sphere subtended 1.0 deg when centered on fixation.

Speed and direction of spheres varied randomly and independently on all axes.

Depth cues included shading, occlusion, binocular disparity and those resulting from parallel projection.

Observers were instructed to maintain central fixation, and to do their best to track all 4 cued targets.

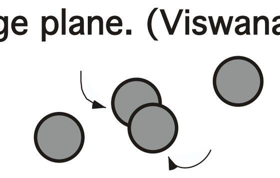
Performance is quantified as the number of spheres successfully tracked, after guess correction.

What is being studied?

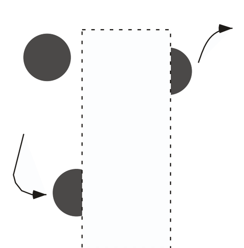
In 2D, it has been shown that observers can track up to 5 in 10 targets, under the proper viewing conditions. (Pylyshyn & Storm, 1988; Yantis, 1992)

Additionally, it has been shown that depth information influences tracking:

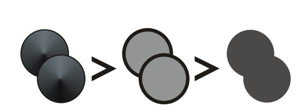
-Tracking is unhindered when objects are allowed to pass "over" one another in the image plane. (Viswanathan, 1998)



-Tracking is unhindered when objects are allowed to pass behind occluders, even invisible occluders. (Scholl, 1999)



-The addition of depth cues, such as shading, T-junctions, and binocular disparity boosts tracking performance. (Viswanathan, 1998, 1999)



These results suggest a tracking system which is fundamentally concerned with position in depth, both to resolve image plane ambiguitites, and as a source of location information.

Given this, our present research has two main goals:

-The first is to determine if depth information can be used at a high level to restrict tracking to a cued subvolume of 3D space.

-The second addresses the extent to which depth information is useful for tracking; that is, the relative "resolution" of attentional tracking along x, y, and z.

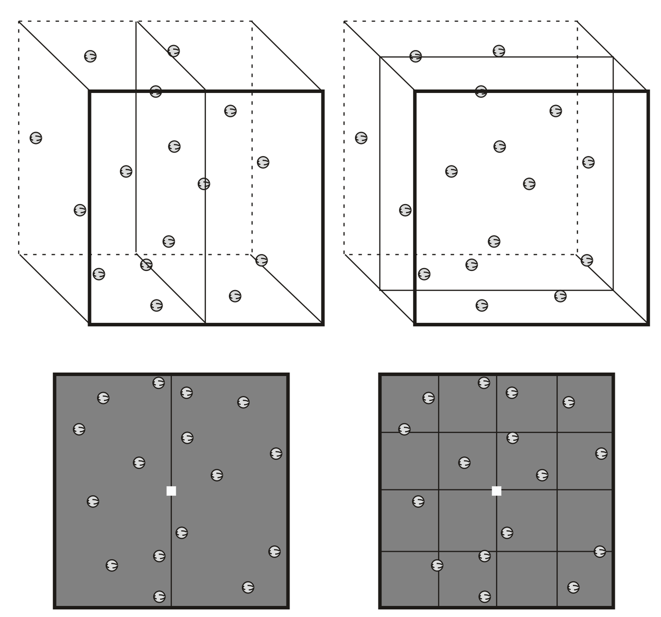
This second goal is approached by introducing a computational model of attentional tracking in 3D.

Can observers restrict tracking to a 3D subvolume?

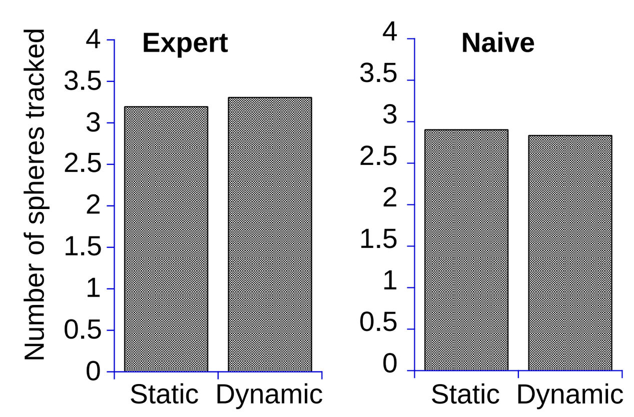
In this experiment, observers attempt to restrict tracking to a particular half-volume.

This half-volume is either the left or right, or the front or back of the full volume.

The region opposite to the tracking task is populated with another set of 8 distractors which are either dynamic or static, depending on the trial.

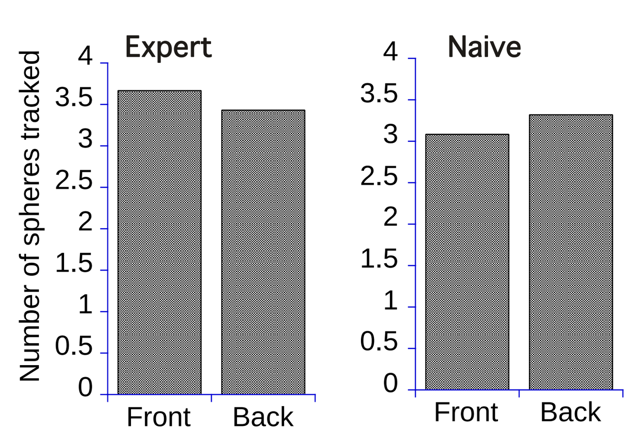


If tracking performance is unhurt by the dynamic distractors, then observers restricted tracking to the half-volume.



There is no significant difference.

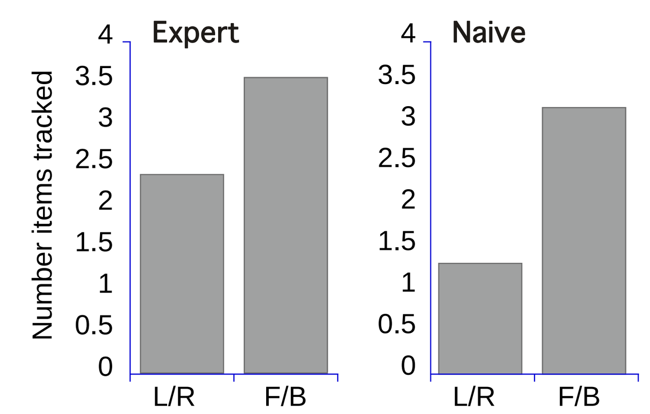
Secondly, if tracking performance is the same in the front half-volume as in the back, this indicates successful selection.



Again, there is no significant difference.

Clearly, observers can restrict tracking to a cued half-volume.

How does the shape of the subvolume affect tracking?

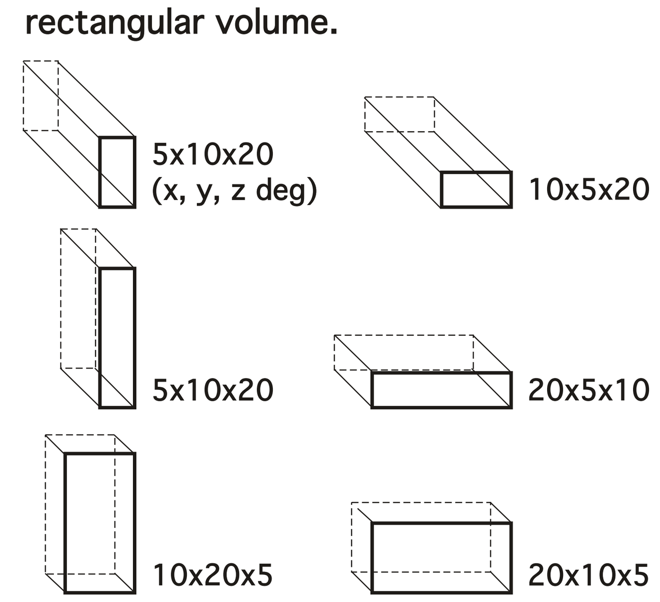


There is, however, a significant difference between the Left/Right tracking task versus Front/Back.

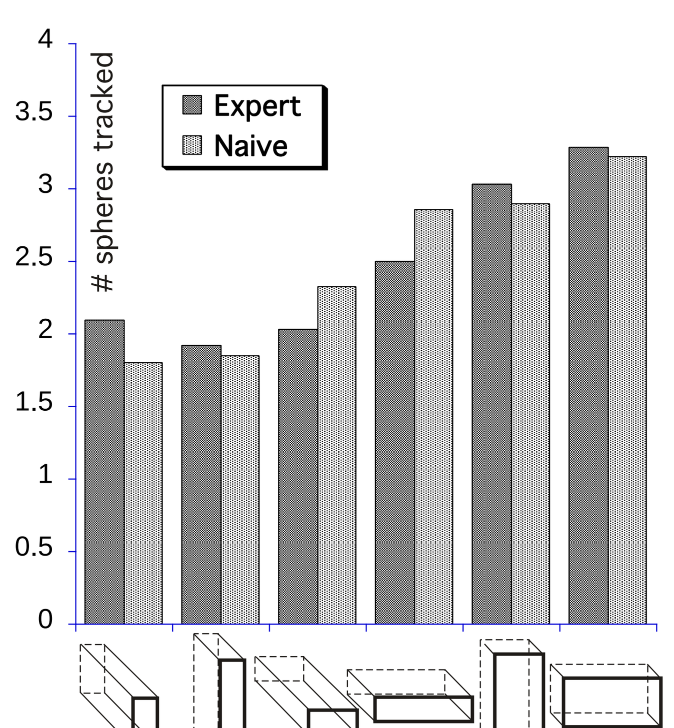
This suggests that, for the attentional tracking system, x, y and z information is not equivalent.

To study this further, Experiment 2 manipulated the relative size of the x, y, and z dimensions.

This was accomplished by simulating a change in the observer's viewpoint of a rectangular volume.



In 3D, all stimulus attributes, including relative speeds, distances, and densities are constant across conditions.



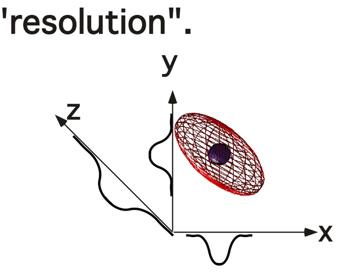
Again, volume shape significantly affects performance.

How can this be modeled?

-Objects are tracked by a system which codes location in 3D.

-The variability with which location is coded differs along x, y, and z.

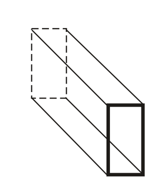
-It is possible to express the variability of location coding in x, y, z as the three variance parameters of a Gaussian distribution over potentially coded locations. In 3D, this forms an ellipsoidal Gaussian "cloud". In our terminology, higher variability means lower "resolution".



-The probability that two objects are "confused" (a new object is mistakenly tracked) is determined by the extent to which this new object has entered the "uncertainty cloud" of the original object.

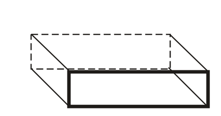
Importantly, what is varying from condition to condition in Experiment 2 is the likelihood that objects will be aligned along their x, y, or z axes.

For instance, in this condition,



objects are twice a likely to aligned along their z axes, than along y. And 4 times as likely to align along z than x.

While in this condition,



objects are most likely to align along x. Alignment in depth is less likely, and alignment along y is least likely.

Given this, performance depends on the interaction (i.e. compatibility) of the shape of the volume with the shape of the uncertainty cloud.

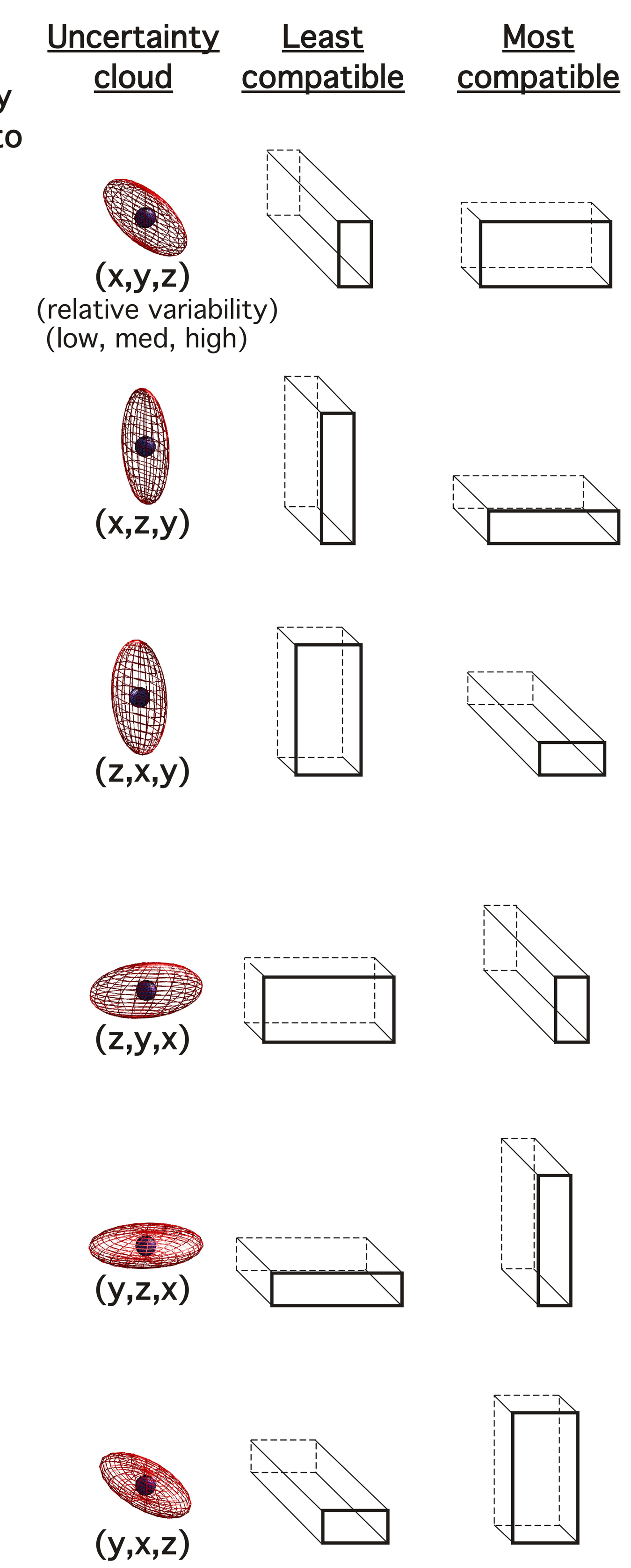
What does the model predict?

-A volume which leads to the most shallow intersections of distractors with the uncertainty clouds of targets yields the highest performance.

-A volume which leads to the deepest intersections yields lowest performance.

-In fact, all of the volumes can be rank ordered by the performance expected given a particular uncertainty cloud shape.

For instance, for any of the possible uncertainty cloud shapes specified by the relative variability along the three axes, it is possible to specify the most and least compatible volume shapes:

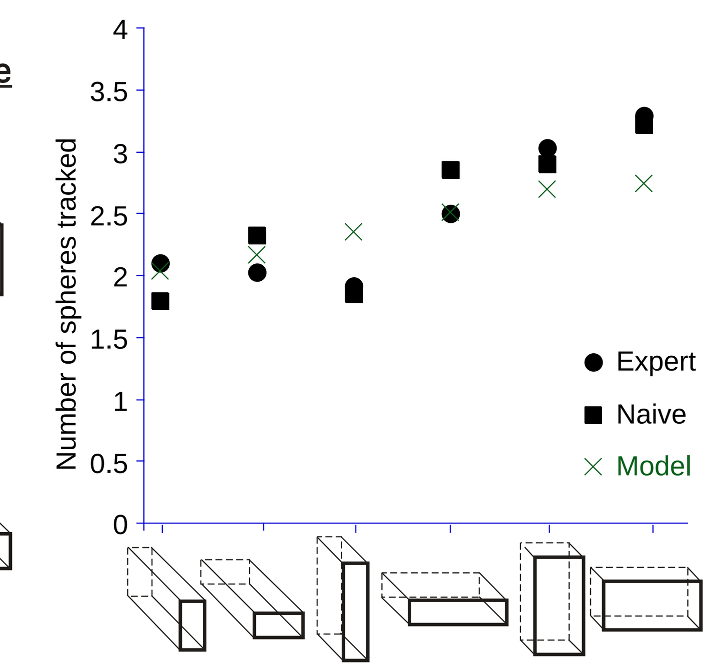


How is the model tested?

We fit the three variance parameters of the 3D Gaussian to determine which shape uncertainty cloud was most likely to have produced observers' data.

Once the model is fit, it imposes a rank ordering on the 6 volume shapes in terms of performance.

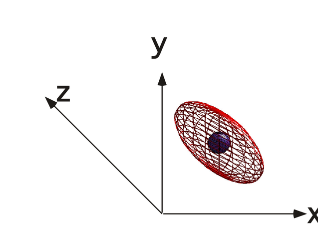
This ordering can be compared to the rank ordering imposed by observer's performance to determine how well the model fits the data.



The model's rank ordering of the conditions is nearly identical to that of the observers.

Conclusions

Qualitatively, this is the modeled shape of attention resolution during 3D multielement tracking.



Location coding is least variable (highest resolution) along x. Variability is higher along y, and higher yet (lowest resolution) along z.

In closing.

-Observers are readily able to restrict tracking to half of a 3D cube-volume.

-The simulated viewpoint of the observer with respect to this volume, however, strongly affects performance.

-We provided a computational model which captures this dependence on viewpoint (or equivalently, volume shape). The resulting model reveals that although depth information is useful for tracking, the resolution of attentional location coding along the z axis is worse (more highly variable) than along y (which in turn is worse than along x).

Future work

-Explore subvolume selection further by testing more complex, embedded subvolumes and space-curves.

-Specify not only the relative resolution of x, y and z, but also absolute resolution.

References

Pylyshyn, Z. W. & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision, 3, 179-197.

Scholl, B. J. & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. Cognitive Psychology, 38, 259-290.

Viswanathan, L. & Mingolla, E. (1998). Attention in depth: Disparity and occlusion cues facilitate multi-element visual tracking. Boston University Technical Report CAS/CNS-98-012.

Viswanathan, L. & Mingolla, E. (1999). Dynamics of attention in depth: Evidence from multi-element tracking. Boston University Technical Report CAS/CNS-TR-99-010.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. Cognitive Psychology, 24, 295-340.