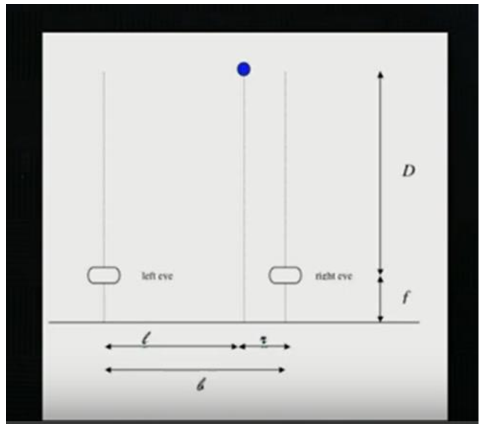
Theo Shin
Problem Set for Module 3

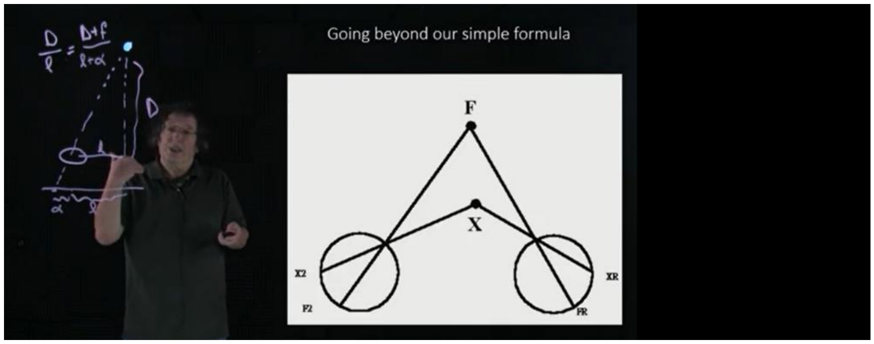
1. DISPARITY

A)



Since the distance of a point from the plane of the eyes is inversely proportional to disparity, and the sum of alpha and beta is equal to the disparity, when determining the distance of a point very far, the disparity will be proportionally small. The resolution of the retina is limited because the eyes don’t uniquely point forward and can swivel which can provide more information depending on how much they turn inward at a point, which takes into account the resolution not being constant. Without a smooth continuous surface, the projection of an object is directed toward the fovea which has a higher resolution. When dealing with objects very far, the computation D = fb/(a + b) isn’t very helpful since the disparity (a + b) approaches zero (f and b remain constant), meaning binocular vision grows ineffective for objects at a very far distance. When the distance is extremely large, the disparity becomes so tiny it goes beyond the retina’s resolution. Example from the lecture would be the difficulty in determining the size of the moon (“moon illusion”) where the view of the moon appears much larger when it is closer to the horizon compared to directly above in the center of the sky.

B)



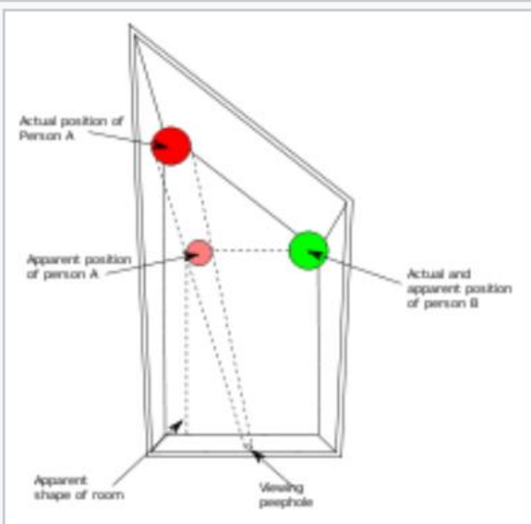
Again, looking at the formula D = fb/(a + b), as the (D)istance becomes extremely small (closer to 0) the denominator (disparity) becomes larger. When objects become extremely close to the retina, it becomes extremely hard to focus using binocular vision. An example would be holding a pencil and slowly moving the pencil closer to the eyes. At a certain point (for me, about six inches from the bridge of the nose) I am no longer able to have both eyes focused on the pencil and I begin seeing a less focused double image (which was predicted since the higher the disparity the more different the view is from each eye). This would have to do with how much the retina is being forced to swivel inwards. Observing the image above, when the 'X' becomes too close, 'X2' and 'XR' are unable to simultaneously align with 'X' due to the limitations of how much the eyes can rotate inwards. With the retina being a finite size and the distance from the lens to the retina being a certain distance (roughly 17mm according to a Washington.edu article), then another problem when objects are extremely close are retinal "blind spots." When an object is extremely close, I can still sense how close an object is. So, the problem with objects being extremely close is not the same as when objects are very far away as in part A. Beyond a certain point the eyes are unable to focus on the image.

2. GEONS FOR OBJECT RECOGNITION AND A FAMOUS ILLUSION

A) The rough estimate for calculating the number of recognizable distinct two-geon objects is 100,000. Begin by deriving the number of geons with varying sizes (20 geons x 5 sizes = 100) and since we are selecting 2 geons, we can multiply 100 * 100 which equals 10,000 just as Prof. Eisenberg demonstrated in his Object Recognition lecture. Lastly, with 10 orientations or qualitative relations we compute 10,000 * 10 to get our final estimate of 100,000 2-geon objects.

B) After introducing a third geon with the same number of varieties and sizes, we can multiply 100 * 100 * 100 (100^3; “100” to represent each geon) which gives us a result of 1,000,000. Just as in part a, with 10 allowed relations we can compute 1,000,000 * 10 to get our final estimate of 10,000,000 3-geon objects or roughly 100 times greater than the 2-geon estimate.

C) From the object recognition lecture, the evolutionary techniques by which we recognize geons are not exactly foolproof, i.e. assuming lines that appear parallel are actually parallel in the real world. This is seen in the ‘Ames Room’ visual illusion. Looking through a peephole, the Ames room appears to be an ordinary rectangular cuboid, with all dimensions and angles appearing “normal” due to the floor, walls, ceilings forming consistent parallel lines.

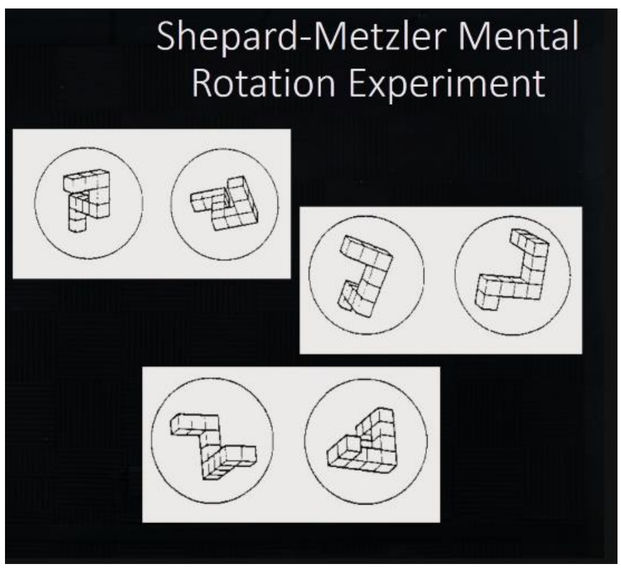


A diagram of the true and apparent position of a person in an Ames room, and the shape of that room

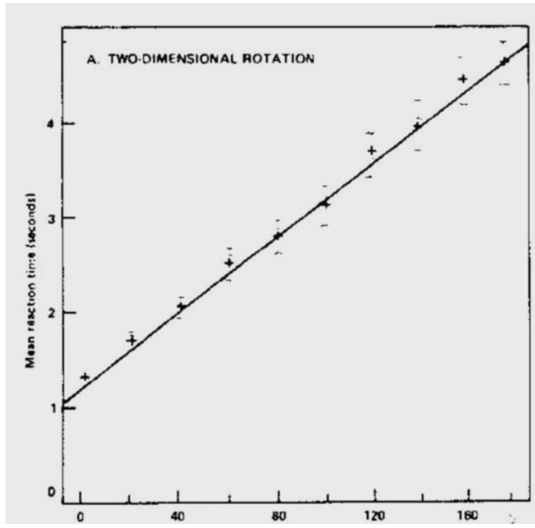
The true shape of the room is, however, that of a six-sided convex polyhedron so that one corner of the room is in fact further from an observer than the other. The reasons the 'Ames Room' works is by forcing the use of a peephole since from any other angle the true shape of the room would be revealed and to use only one eye when observing the room otherwise with binocular vision one could obtain sufficient information about the room that something was "off" (Also, according to wikipedia, strategic lighting, patterns of walls and floors, were also manipulated to increase the realism effect). By observing parallel lines and "Y-corners" throughout the room, one would incorrectly assume the room is just a rectangular room with 90-degree angled corners, which isn't the case for the 'Ames Room.' Another assumption we make when recognizing geons are lines that are occluded in some regions but appear continuous (seen in the "impossible doghouse" visual by Walter Wick).

3.

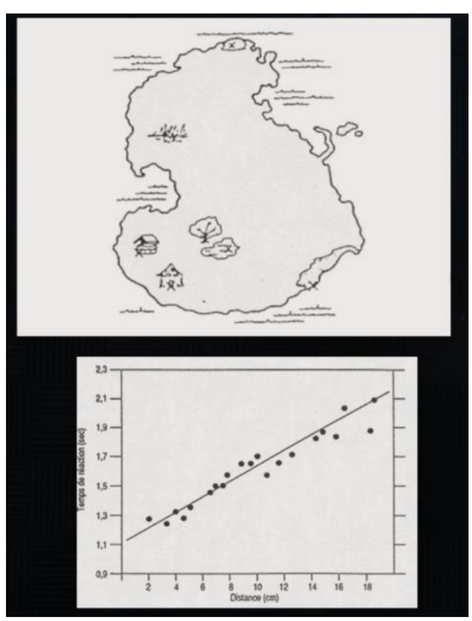
A)



In the Shepard-Metzler mental rotation experiment, subjects are tested on whether and how long an object could be rotated to coincide into its rotated equivalent (Seen above, the top and bottom pairings can be rotated into its equivalent while the middle pairing cannot). The surprisingly result was when the subjects were asked to press a button as quickly as possible when deciding whether one object could be rotated to coincide into the other object and the two objects were the same but with just a rotation difference, the time taken to answer was linearly related to the angular distance (demonstrated in the graph below).

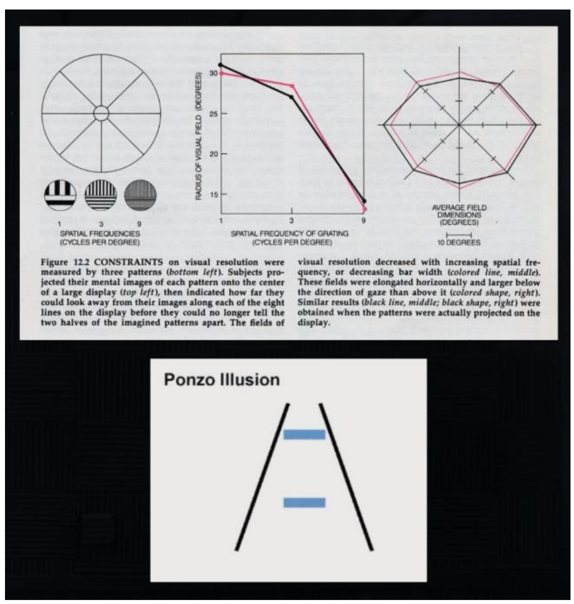


The linear function of angular discrepancy happens to fall on a straight line which according to the lecture is an enormous rarity whenever a psychological experiment results in an almost exactly linear manner. It's as if subjects are mentally taking an object and rotating it a standard pace (angular velocity) which then naturally it takes them longer to rotate the object 90 degrees compared to say 30 or 40 degrees, just as it would in the physical world. A similar mental imagery phenomenon as discussed in lecture was the Kosslyn, Ball, and Reiser experiment where subjects are shown a sketch of an island and asked to stare at it long enough to form a strong mental image of the island. The results of the experiment resulted again in a close approximation to a linear function when comparing the amount of time it took for subjects to shift their focus from various objects on the island to the Euclidean distance between those objects on the island map.



B) For the ambiguous figure (rabbit/duck) mental imagery experiment, the surprising result wasn't just that a subject could only see a rabbit or a duck, but when asked to mentally visualize a duck when they've only previously identified a rabbit, the subjects were unable to. However, when subjects were asked to draw the mental image out on paper, they were more inclined to identify the opposite animal and reinterpret the image once it had been drawn out. This is counterevidence to mental imagery stating that staring at a photograph products similar results. I'm personally trying to understand the reasons behind this phenomenon. I don't have a difficult time identifying both animals (most likely because I've seen that photo or similar photos at several stages in my life) but I suppose the way the mental images are etched into the mind, it becomes harder for the brain to identify it as anything else than the way it was conjured up. As opposed to when it's drawn on paper, we have some mental distance from the object and can more objectively see that it could easily be a rabbit just as much as a duck. A similar experiment was represented in the "Star of David" experiment where subjects had no difficulty identifying the triangles composed within the star image, but had great difficulty in conjuring up the parallelograms. In both instances, securing a mental image is not the same as looking at its picture counterpart.

C)

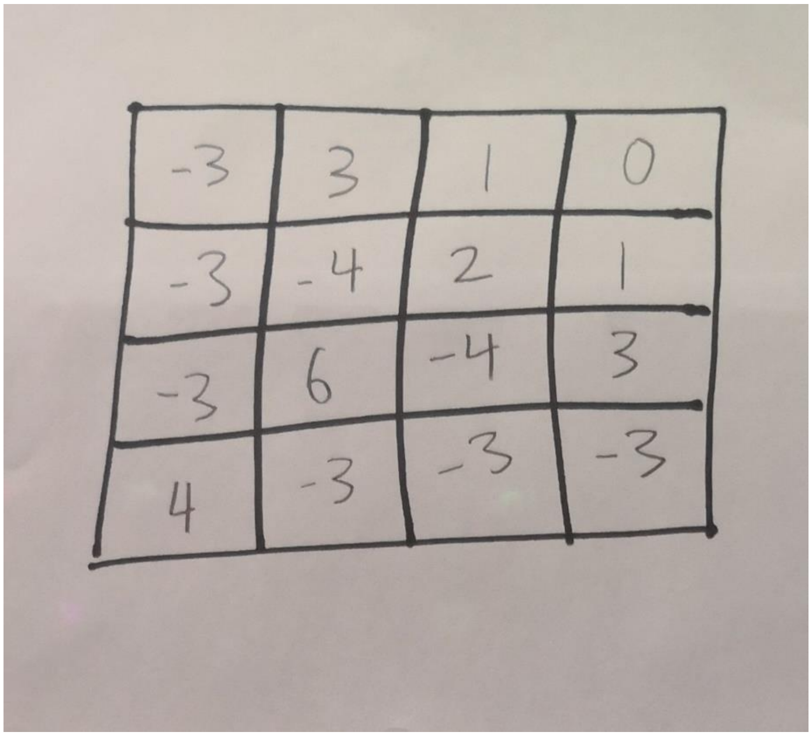


For the Finke experiment, subjects are presented with a variety of grids ranging from course, to medium, to fine. Subjects are then asked to place the varying grids onto the center of the target while shifting their gaze to one of the spokes of the wheel (radii) towards the edge of the circle and report at what point along the radius does the central image blur (subject can no longer distinguish between vertical and horizonal stripes). The results showed distances were reported longer for the course grained grids and shorter for the fine grid or as the spatial frequency increases, the distance to blur the image along the radii decreases. The surprising phenomenon derived from the Finke experiment was that by observing the two graphs on the top right of the picture (with red and black lines corresponding to mental imagery and looking at the actual picture, respectively) the results were pretty much the same or very similar whether the subjects were looking at the actual perceived picture or a conjured mental image. Also, from the graph on the far right, it was determined the blur distance around the target is greater when moving horizontally which coincides with vision being sharper horizontally rather than vertically. This suggests there must be a strong visual component to mental imagery and it can't be reduced entirely to symbols. Another example of this phenomena are the "Ponzo illusion" and "Muller-Lyer illusion." In both experiments, when the subject was looking at the actual picture, they believed one line to be longer (the y-line for Muller-Lyer or the top blue line for Ponzo). When asked to hold a strong mental image of either experiment, again, the subject imagined those lines to be longer when in reality, they were the same length.

D) For the McCollough Effect, it was determined after subjects were asked to stare for a good amount of time at a green patch and a red patch and then shown plain grids with vertical and horizontal stripes, the subjects reported seeing a pink background where the vertical stripes were associated with the green patch and a greenish background for the horizontal grid. The result is surprising given the result is not a retinal effect (or the fatiguing of the cones), but instead deals with the primary visual cortex where the colorless gratings appeared colored contingent upon the orientation of the gratings. It was also noted in the lecture that when subjects were asked to stare at just a block of red or green and then to form mental images of the black vertical or horizontal lines, the McCollough Effect still occurred, though weaker. From these experiments, mental imagery overlaps with many important aspects of human vision, but isn't quite the same. The significance of the varying results of these experiments is how human thinking factors into mental imagery and moving forward, how to computationally understand mental imagery for future implementation.

4. DISCRETE CONVOLUTION

A)



-3	3	1	0
-3	-4	2	1
-3	6	-4	3
4	-3	-3	-3

B)

-1	2	0	1
3	-2	3	0
-3	3	-2	2
5	-3	3	-1