Chapter 7

**DEPTH AND SIZE PERCEPTION**

Our visual systems represent the world on flat, 2D retinae but we must perceive a 3D world. The solution the visual system uses is the cue approach to depth perception. These cues are divided into three groups: oculomotor cues, monocular cues, and binocular cues. Oculomotor cues include accommodation and convergence. Monocular cues include occlusion, relative height, relative size, familiar size, linear perspective, texture gradients, atmospheric perspective, shadows and shading, and motion cues. Binocular cues compare the images from each retina. In the area where both eyes see the same part of the world we have binocular vision, or stereopsis. Stereopsis arises from binocular disparity, which occurs because our eyes are in different locations in our head, giving slightly different views of the world. Corresponding points refer to when points on the left and right retinae coincide if the two retinae were superimposed, whereas noncorresponding points refers to when they do not coincide. The horopter is the region in space where the two images from an object fall on corresponding locations on the two retinae. Disparity information is used to determine if an image is in front or behind the horopter. The correspondence problem is the problem of determining which image in one eye matches the correct image in the other eye. To study stereopsis, stereograms and anaglyphs are used to create a pop-out phenomenon that we see as depth. Random-dot stereograms show that object recognition is not necessary for disparity cues, suggesting that correspondence precedes object recognition.

Size and depth are deeply intertwined. Size-distance invariance refers to the phenomenon that perceived size depends on perceived distance and, often, vice versa. Size constancy is the perception of an object as having a fixed size, despite the change in the size of the visual angle that accompanies the change in distance. By manipulating depth cues, it is possible to see depth relations that are not really present in an image, as seen in visual illusions. Classic visual illusions of size and depth include the Müller-Lyer illusion, the Ponzo illusion, the Ames Room illusion, and the moon illusion.

Introduction

* Our visual system must re-create the 3D world using an essentially flat, 2D retina.
  + The solution that our visual system uses is known as the **cue approach to depth perception.**
  + That is, we must infer the third dimension from other cues our visual system provides.
  + These cues include oculomotor cues such as accommodation and convergence, monocular cues such as occlusion, relative height, and relative size, and binocular cues from comparing images from each retina.

Oculomotor Cues

Accommodation

* **Accommodation** is the process of adjusting the lens of the eye so that one can see both near and far objects easily.
* To focus on more distant objects, the ciliary muscles that contract the lens are relaxed.
* To focus on more near objects, the ciliary muscles are contracted.
* These contractions and relaxations take place automatically, but they can be sensed.
* It is the sensing of these movements that gives information about depth.

Vergence (or Convergence)

* **Vergence** (also known as convergence) happens when the eyes bend inward to see a near object and bend outward (diverge) when looking at a more distant object.
* Like accommodation, this process is automatic but the movements are sensed, giving us information about relative depth.
* Vergence is probably a more useful cue than accommodation, providing the visual system with reliable depth information to about 2m in length.

Monocular Depth Cues

* **Monocular depth cues** are cues that provide information about depth and distance but can be inferred from just a single retina.
* Monocular cues include **pictorial cues,** from which we can judge depth from static or nonmoving pictures, and **movement-based cues**, in which moving objects let us make inferences about depth.

Occlusion (or Interposition)

* **Occlusion** happens when one object partially hides or obstructs the view of a second object.
* Occlusion provides information about relative, not absolute, distance.

Relative Height

* **Relative height** means that objects closer to the horizon are seen as more distant.
* Objects below the horizon are seen as more near the viewer if they are closer to the bottom of the visual scene but objects above the horizon are seen as more near if they are closer to the top of the visual scene.

Relative Size

* **Relative size** means the more distant an object, the smaller the image will be on the retina.
* Thus, if there are two identical objects, the one that is closer will have the larger image on the retina.
* However, the object that is further away does not look abnormally small due to the mechanism call size constancy.

Familiar Size

* **Familiar size** means we judge distance based on existing knowledge of the sizes of objects.
* If we know that a particular object is smaller than another object, but it is taking up more space on the retina, we assume the smaller object is closer and the larger object is farther away.

Linear Perspective

* **Linear perspective** is the pictorial depth cue that arises from the fact that parallel lines appear to converge as they recede into the distance.
* Though parallel lines never meet in actuality, perceptually they do at the edge of the horizon.
* Parallel lines seem farther apart close up, because they take up more space in the image on the retina.
* Linear perspective is an important technique when painting a 3D scene.

Texture Gradients

* **Texture gradients** occur because textures become finer and smoother as they recede in distance.
* Common elements that are evenly spaced in an image appear more close together in the distance than they do in the foreground.

Atmospheric Perspective

* **Atmospheric perspective** is a pictorial depth cue that arises from the fact that objects in the distance appear blurred and tinged with blue.
* Because the atmosphere scatters light, more distant objects appear blue.

Shadows and Shading

* **Shadows** are a depth cue arising because an object is in front of its shadow, and the angle of the shadow can provide some information about how far the object is in front of the background.
* Objects in shadow must be farther from the light than objects that are not in shadow.

Motion Cues

* Objects moving at different speeds can reveal information about relative distance.
* **Motion parallax** is a monocular depth cue arising from the motion of a person in the environment.
  + Faster moving objects are closer to us; slower moving objects are farther away.
* **Deletion** is the gradual occlusion of a moving object as it passes behind another whereas **accretion**  is the gradual reappearance of a moving object as it emerges from behind another object.
  + The object that is deleted and later accreted is the object that is farther away than the object we can see continuously.
* **Optic flow** is the motion depth cue that refers to the relative motions of objects as a person moves forward or backward.
  + Like with motion parallax, closer objects appear to move more quickly and farther away objects appear to move more slowly.
  + Optic flow is often used to convey depth in movies

Binocular Cues to Depth

* Because humans have two eyes that see the world from slightly different angles, we have an important cue to depth.
* In the area where both eyes see the same part of the world, we have binocular vision.
* **Stereopsis** is the sense of depth that we perceive from the visual system’s processing of the comparison of the two different images from each retina.

Binocular Disparity

* Disparity occurs because our two eyes are in different locations in our head and thus have slightly different views of the world.

Corresponding and Noncorresponding Points

* *Corresponding points*refers to a situation in which a point on the left retina and a point on the right retina would coincide if the two retinae were superimposed.
* *Noncorresponding points*, in contrast, refers to a situation in which a point on the left retina and a point on the right retina would not coincide if the two retinae were superimposed.
* The **horopter** is the region in space where the two images from an object fall on corresponding locations on the two retinae.
  + Objects that lie along the horopter are perceived as single unified objects when viewed with both eyes.
* Images from the two eyes can also fuse for objects that fall inside **Panum’s area of fusion**, which is the region of small disparity around the horopter.
  + Outside Panum’s area, we see double images, or **diplopia.**
* Points closer to us than the horopter have crossed disparity and points farther away have uncrossed disparity.
  + **Crossed disparity** refers to the direction of disparity for objects in front of the horopter (the image in the left eye is to the right of the image of the object in the right eye).
  + **Uncrossed disparity** refers to the direction of disparity for objects that are behind the horopter (the image of the object in the left eye is to the left of the image of the object of the right eye).
  + Points along the horopter have **zero disparity** (i.e., retinal images fall along corresponding points).
* Retinal disparity can give us information about depth distances as small as 4mm at a distance of 5m.

The Correspondence Problem

* In some complex scenes, matching images to the left and right eyes may be difficult because real scenes often have complex textures, similar objects in motion, and other variations.
* The **correspondence problem** is the problem of determining which image in one eye matches the correct image in the other eye.

Stereograms

* In the 19th century, Charles Wheatstone invented the stereoscope, which is a small instrument that presents one image to one eye and a second image to the other eye.
  + The pictures are images slightly offset from each other in order to replicate crossed and uncrossed disparity, giving the viewer a single 3D image.
  + Because the images are so similar, they are combined by the visual system and give the illusion of depth.
* Some individuals can examine stereograms without using a stereoscope through a process known as free fusion.
* The anaglyph is another form of stereogram which is generally used in older 3D movies.
  + Two photographs are taken about 6cm apart and one photograph is printed in a shade of one color, such as blue, and the other photograph is printed in a shade of another color, such as red.
  + The two photographs are integrated into a common image.
  + When viewed with color-coded anaglyph glasses, each of the two images goes to one eye, allowing for the stereoscopic image to emerge.

Random-Dot Stereograms

* **Random-dot stereograms** are stereograms in which the images consist of a randomly arranged set of black and white dots.
  + The left and right eye images are arranged identically, except that a portion of the dots is moved to the left or right in one image to create either crossed or uncrossed disparity.
  + This makes us perceive that part of the image is either in front or behind the rest of the dots.
* Because people see depth information in random-dot stereograms, we know that correspondence precedes object recognition.

The Anatomy and Physiology of Binocular Perception

* Binocular cells in V1 have two receptive fields, one for each eye.
* These cells are usually similar with respect to their preferred orientation and motion sensitivity, suggesting that their main function is to match the images coming to each eye.
* Different binocular cells are tuned to different disparities.

Development of Stereopsis

* Research suggests that newborns are blind to binocular depth information and continue to show no stereopsis until 4 months, when stereopsis develops rapidly.
* Preferential looking tasks show that infants start becoming sensitive to disparity at about 5 months.

Size Perception

* **Size-distance invariance** refers to the relation between perceived size and perceived distance.
  + The perceived size of an object depends on its perceived distance, and the perceived distance of an object may depend on its perceived size.
* **Visual angles**, the angles of objects relative to one’s eyes, are important to studying size-distance invariance.
  + If lines were drawn from the top and bottom of an object to the eyes, we would have the visual angle of the object.
  + Smaller objects close up can have the same visual angles as larger objects farther away, like the moon and the sun.
* Depth provides information that allows us to appropriately scale the size of an object.

Size Constancy

* **Size constancy** is the perception of an object as having a fixed size, despite the change in the size of the visual angle that accompanies the change in distance.
* We tend to see an object as the same size regardless of the size of its image on the retinae.

Visual Illusions of Size and Depth

* By manipulating depth cues, it is possible to see depth relations that are not really present in an image.

The Müller-Lyer Illusion

* In this illusion, although both lines are exactly the same length, the left line in Figure 7.30 is seen as longer than the right line due to the smaller lines that split off from the main line.
* One viewpoint for why we see this illusion is that it is a result of misapplied size constancy.
* Another viewpoint is that the overall image of the left line is longer and, thus, its components are also perceived as longer.

The Ponzo Illusion

* In this illusion, the cow closer to the top of the page in Figure 7.32 looks bigger than the cow on the bottom of the page.
* We see this because there are several cues in the photograph that give normal indications of depth, such as the parallel lines of the side of the road, the texture gradient of flowers and grasses, and familiar size cues of the trees.

The Ames Room Illusion

* In this illusion, two people of normal size are put into corners of a room with trapezoidal walls and windows.
* When viewed through a peephole, the trapezoids are converted into squares by our visual systems so that we perceive the room as normal.
* Given that we perceive a normal room, we must therefore infer that the people are abnormally short or tall.

The Moon Illusion

* In this illusion found in natural phenomena, a full moon on the horizon is perceived as being larger than when it is higher up in the sky, though its size and distance from Earth remain constant.
* Because the horizon is perceived as being farther away than the zenith overhead, an object that takes up the same amount of space on our retina is perceived to be larger.

***In Depth: Stereopsis and Sports: Do We Need Binocular Vision in Sports?***

* Much of the research on stereopsis in sports has been done regarding a baseball batter.
  + Major league batters performed better on a stereopic task than minor league batters and that performance predicting batting average, suggesting that baseball batters do use stereo cues (Hofeldt & Hoefle, 1993).
  + Filters impairing the amount of light entering only one eye dramatically impaired batters’ ability in a batting cage, whereas filters impairing both eyes did not impair ability (Hofeldt et al., 1996).
  + These results suggest that binocular information is important for baseball batting.
* Binocular cues are also used in catching balls and shooting foul shots in basketball.