

visual perception

# Visual perception and the holographic image

Michael Dalton

[mdalton@mac.com](mailto:mdalton@mac.com)



## **About the author**

Michael is a physicist who became interested in holography around 1983. He was co-founder and Technical Business Director of Voxel—a company which developed and patented the concept of multiple exposure Digital Holograms (Voxgrams) for medical applications. He was awarded US patent number 6123733 for the simulation of Voxgrams on desktop computers. He has a strong interest in visual perception.

In a period where the term “Holographic” can be applied to just about anything, even a lipstick (L’Oreal’s Glam Shine), I thought it an good time to go back to first principles and to reflect on what is distinctive about a holographic image and in what when it is an appropriate solution to a visual problem. If we consider the relevant features of the viewers visual system we may be able to construct a better holographic image or three-dimensional imaging system.

A viewer’s visual perception of the world is based on a complex relationship between the eye and the brain. The eye consists of a physical optical system e.g. a lens, and light sensors. The brain contains the neurological system used to interpret the information passed from the eye through the visual cortex to the brain. The parts of the visual system of particular interest to us are the 3D depth perception mechanisms, or depth cues. Whilst depth perception is a complex sum of all of these depth cues, it can be divided into two broad categories: physiological and psychological.

## **Physiological depth cues**

These are considered the strongest depth cues because they are dependent on physical changes in the eye’s optical system. The first three (accommodation, motion parallax, and convergence) can be considered monocular depth cues, since they can be used by those with sight in only one eye.

- Accommodation – This depth cue is dependent on the control of the shape of the eyes by the ciliary muscles to bring the image formed on the retina into a sharp focus.
- Motion parallax – Of two objects moving at the same speed, that which is further from the viewer projects an image which moves across the eye’s retina more slowly than the one which is closer. Therefore the further objects will appear to move more slowly than those which are

closer. Correspondingly, if the viewer is moving and the objects are stationary, then the objects which are closer will appear to move faster than those at a greater distance.

- Convergence – To bring an object into focus in the most sensitive portion of the eye, the fovea, objects which are closer require the eyeballs to twist towards each other more (larger convergence angle), than those at a greater distance (smaller convergence angle).
- Binocular disparity – This is a complex binocular depth cue, which assumes that the image projected onto the area of the back of each eye ball, is focused on the same object, and hence is intimately related to the convergence cue. It relies upon the correlation, and hence disparity or lack of correlation, between the two images perceived by the brain. Points that are farther away will appear to have a greater disparity or separation between each perceived image, than those closer to each other. The brain interprets this disparity as being closer, or further to the viewer than the point of convergence.

## Psychological depth cues

These depth cues are not derived from physical changes in the viewers eyes, but rather from a higher level interpretation of the images passed from the eyes to the visual cortex in the brain. These depth cues result from our brain's previous knowledge of how 3D objects should appear in the real world, but, as the renaissance artists in the 15th and 16th centuries realized, can also be triggered by two-dimensional image representations, such as canvases, to create the illusion of depth. They are characterized by being relative measures of distance; they require comparison with other objects in a scene before they can be used to infer their relative distances from the observer. These depth cues, when separated from their natural (physiological) counterparts in the 2D world, can become ambiguous as many illusionists and artists, such as Escher, have shown to great effect.

- Stereopsis – One of the strongest psychological depth cues, for those with binocular vision is stereopsis, which can be considered as part of the binocular disparity cue, but which can be triggered by pairs of two-dimensional images or stereo displays such as anaglyphs.
- Occlusion – An object which is nearer to the viewer can partially or completely obscure those further away.
- Linear perspective – As an object moves away from the eye the image projected onto the back of the eye decreases in size and so we perceive them as becoming smaller.

- Aerial perspective – The further an object is, the more atmosphere there is between us and the object. The atmosphere causes light to be scattered, and has a tendency to scatter blue light more. Thus as objects recede in distance their contrast will decrease and they will appear bluer. This is most apparent when looking at distant objects such as mountain ranges, but the effect can also be seen in a smoke-filled room.
- Size – An object that is closer to us will appear larger than when it is moved farther away. If we recognise the object and can determine its relative size in the scene, we can judge the relative depths of objects in the scene.
- Shading/shadows – The shading of an object, the way that light falls on and is reflected by it, can give clues to its orientation, which can be used to relate it to other objects in a scene. When used with other psychological depth cues it can help to sort the relative depths of objects in scene. If an object casts a shadow onto another, then we can determine that it is closer to the light source – an interpretation which can also aid in sorting the relative distances of objects.

## Ranges of depth cues

Now let us consider over what distance ranges these depth cues are most effective. In this way we can determine which depth cues are most relevant in particular situations.

### Near range depth cues (15 cm–1.5 m)

These depth cues are most important when dealing with objects at arm's length, such as a surgeon operating with a scalpel or a wine grower picking a grape from a vine. They are implicitly used in everyday tasks which require visual feedback to the brain's motor control system, such as when we pick up an object. Of the psychological cues, occlusion (think of threading a needle) and stereopsis are most useful at this near range.

### Medium range depth cues (1.5–150 m)

Here recognising an object (is it a tiger or a cat?) and the rate that it is moving (is that car going to knock me down?) are important visual questions. The importance of the physiological depth cues starts to diminish with increasing depth and we become more reliant on the psychological depth cues.

## Long range depth cues (150 m–15 km)

Here the physiological depth cues dominate and at great distances the occlusion and aerial perspective are the only cues which have much effect.

**Table 1.** Importance of depth cues at different distances

| Range                | Near | Medium | Far |
|----------------------|------|--------|-----|
| <b>Physiological</b> |      |        |     |
| Accommodation        | ••   | •      |     |
| Motion Parallax      | •    | •      |     |
| Convergence          | ••   | •      |     |
| Binocular Disparity  | ••   | •      |     |
| <b>Psychological</b> |      |        |     |
| Stereopsis           | •    | •      | •   |
| Occlusion            | •    | •      | •   |
| Linear Perspective   | •    | •      | •   |
| Aerial Perspective   |      | •      | •   |
| Size                 | •    | •      | •   |
| Shading              | •    | •      | •   |

When we are considering the creation of an image or visual display system, by looking at these depth cues, we can construct a display which works to complement our visual perceptive system. To underline the importance of taking these into account, let us first the effect of situations which cause conflicting inputs to our perception system:

Travel sickness – This is a condition caused by the conflict between sensory inputs to the brain. If I am reading while I am being rocked about by the motion of the car or boat in which I am travelling, my visual system tells me that the book I am reading is stationary, but my inner ear tells me I am moving. This conflict between the visual system and the detection of motion from the inner ear can lead to feelings of nausea. By looking out the window instead of reading, the two perception systems both report the same information to the brain and the feeling of nausea will recede.

The complement to this type of conflict can be found in the works of the OpArt or Kinetic artists. Here the viewer is stationary but may see motion in the visual field caused by the interaction of high contrast edges at various angles and frequencies. A more serious occurrence of this type of perceptive conflict may be found in people with ‘scotopic sensitivity syndrome’ where high-contrast repeated patterns such as window blinds or even text on a page can cause feelings of nausea. Once solution being investigated by researchers is to find a particular coloured filter which minimises the effect on the viewer.

## Holographic systems

First let us consider the “gold standard” of holography – A full-aperture hologram made at a single wavelength and replayed with the same reference beam.

A full-aperture hologram of an object can trigger all the depth cues that the real object would. There is nothing to suggest that the shape of the object is not in fact “real” other than it’s monochromatic. Just as in the real world, the physiological and psychological depth cues work together and do not conflict each other. Any distortions in the reconstruction system tend to be minimized near the plane of the holographic film, so objects tend to be placed as close as possible to the film plane. Because of this proximity to the viewer, the physiological cues are the most important.

Therefore full aperture holograms are best suited to those displays where the viewer is going to come close to the display and be within a distance that they could interact with the object. Such examples would be in museums or exhibitions. Placing the hologram far from the viewer reduces the effect of the physiological cues and hence reduces its three-dimensional impact on the viewer – one may just as effectively use a video display or stereo display system (e.g. lenticular) to get the viewer’s attention. Unfortunately, the type of objects that can be used in this form of full-aperture holography is restricted by the size of holographic film, and by the need for the object to be motionless for the duration of the holographic exposure.

However there are circumstances where it is not practical to make a full-aperture hologram of the object, due to its size or susceptibility to motion. Here we may resort to stereo displays, which use pairs of two-dimensional images to trigger only the binocular disparity from the physiological and the psychological depth cues. Such examples can be found in lenticular displays, anaglyphs and stereograms. Stereograms use holography to create a series of virtual slits through which the viewer can see stereo pairs of images, without the need for glasses. Since the 2D images used in the recording of these stereograms can be generated from any photographic or other scanning technique, or from computer generated simulations or designs, any size of object can be used to create them.

There is, however, a problem with stereo display systems. When used to display near-range objects, the physiological depth cues of accommodation, and convergence both report to the visual system that they are looking at two flat images at a single depth. However the other depth cues, especially binocular disparity (and hence stereopsis) are all reporting differing depths, and this conflict can lead to nausea if the viewer is exposed to such displays for an extended length of time. Therefore, I would not consider using such a system for a surgical applications for example, where the surgeon may be operating for periods of 10 or more hours on a patient at arms length!

This conflict between physiological and psychological depth cues for stereo displays can be tolerated at near ranges for short periods of time. If the display has to be viewed for longer periods of time, e.g. movies, then we can limit objects appearing in the near range, where the physiological depth cues will report conflict, and instead keep them at the medium and long ranges where the physiological depth cues are weak. One other technique for reducing this conflict is to restrict the range of depths in the scene so that the disparity between the depth cues is minimized.

## Conclusion

I have briefly explored the relationship between the physiological, and psychological depth cues and how these consideration can be relevant when creating holographic images of scenes which consist of near, medium and far ranges. In a future article I will explore how an understanding of other aspects of the viewer's perception system, can be useful when constructing a holographic image or another 3D imaging system.