Visual Perception of Spatial Subjects

Kurt R. S. Osterloh [1], Uwe Ewert [1]

[1] Federal Institute for Materials Research and Testing (BAM), VIII.3, 12200 Berlin, Germany, phone: +49 (0) 30 8104 3654, e-mail: kurt.osterloh@bam.de

Abstract

Principally, any imaging technology consists of two consecutive, though strictly separated processes: data acquisition and subsequent processing to generate an image that can be looked at, either on a monitor screen or printed on paper. Likewise, the physiological process of viewing can be separated into vision and perception, though these processes are much more overlapping. Understanding the appearance of a subject requires the entire sequence from receiving the information carried e.g. by photons up to an appropriate processing leading to the perception of the subject shown. As a consequence, the imagination of a subject is a result of both, technological and physiological processes. Whenever an evaluation of an image is critical, also the physiological part of the processing should be considered.

However, an image has two dimensions in the first place and reality is spatial, it has three dimensions. This problem has been tackled on a philosophical level at least since Platon's famous discussion on the shadow image in a dark cave. The mere practical point is which structural details can be perceived and what may remain undetected depending on the mode of presentation. This problem cannot be resolved without considering each single step of visual perception.

Physiologically, there are three "tools" available to understanding the spatial structure of a subject: binocular viewing, following the course of perspective projection and motion to collect multiple aspects. Artificially, an object may be cut in various ways to display the interior or covering parts could be made transparent within a model. Samples will be shown how certain details of a subject can be emphasised or hidden depending on the way of presentation. It needs to be discussed what might help to perceive the true spatial structure of a subject with all relevant details and what could be misleading.

Keywords: image (data) processing, image interpretation, physiological capabilities and constraints, visibility of certain features, perception of a subject

1. Introduction

It has been already aware to the ancient Greek philosophers that we perceive the surrounding reality as an image, which is two-dimensional in its nature though the reality is three-dimensional. In his famous work "Politeia" (commonly translated as "The State"), Platon discussed in the seventh volume extensively how shadows of figures are being perceived and interpreted as "reality" and how problems may raise when being confronted with the world "outside the cave". The question how to present the spatial reality on the dimensions of a canvas has been tackled by numerous artists, some of them even playing jokes with the presentation of space [1]. However, the human visual sense has mechanisms to realise the environment in all three dimensions. Understanding how they operate has contributed to find ways to presenting spatial objects visibly in all three dimensions.

The purpose of this contribution is to make aware of the three major and basic physiological mechanisms of perceiving space visually. Besides understanding perspective drawings, these are the binocular viewing and time resolved perception. This includes gathering views from various directions while moving around the specimen or turning it itself. None of these mechanisms operate independently; even the binocular and time resolved viewing cooperate as it can be demonstrated by the Pulfrich effect [2]. Though the primary sensor for light is located in the eye this site is more than a mere receptor. Interconnecting neurons care for both, high pass filtering for sharpening contours and data reduction, i.e. the information from more than 100 million receptor cells are forwarded to the brain by som 1 million ganglion cells. In contrast to technical systems where data acquistion and processing are rather separated the biological one has both steps more integrated along the whole data propagation and processing chain. Even more, the receptor properties of the retina are not equally distributed over the plane; high resolution and the perception of colours are allocated in the central areas whereas the peripheral ones are more sensitive to fast motion. The adaptation to the ambient light conditions is achieved on several levels from the retina to the thalamus. The following description of the mechanisms of human vision will be focussed on the three-dimensional perception, i.e. of spatial subjects.

2. Physiological Principles

2.1 Physiological "Data Processing"

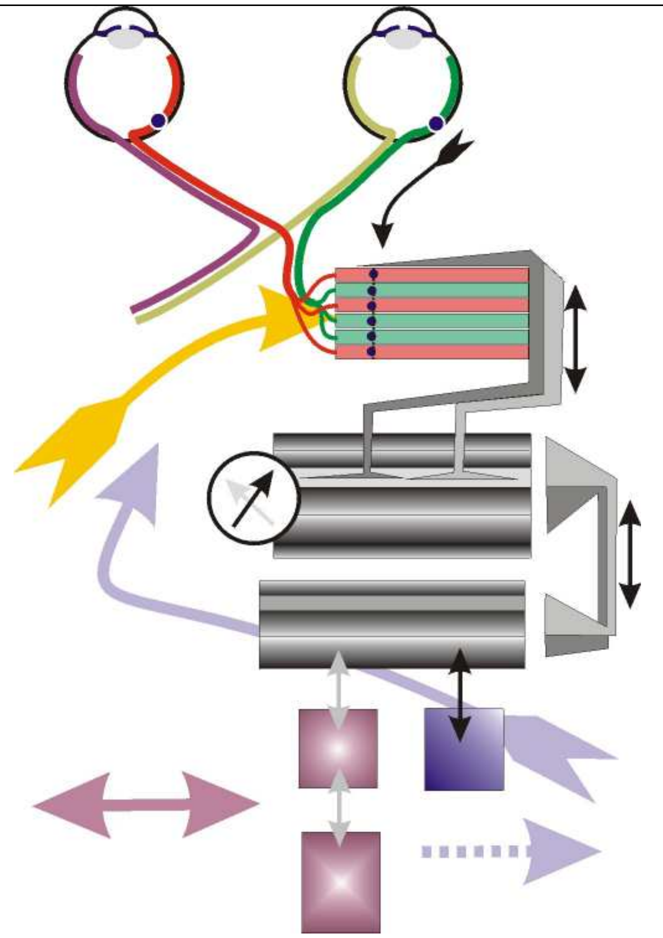
eye retina		adaptation, contrast, data reduction
thalamus lateral geniculate body		contrast, modulation, generation of a static image, association and imagination fed from memory
cortex primary visual cortex (V1) subsequent cortical areas (V2 …)		merging of disparate images, directions, perspective, motion (separate lane)
memory		shape, pattern, association

Figure 1: Allocation of physiological processing

The primary receptor of light, i.e. the information carried by photons, are the light sensitive cells of the retina of the human eye. From there on, propagation and processing of visual information are closely interwoven and operating bidirectionally from the thalamus on. The allocations to the corresponding anatomical structures are outlined in

Figure 1 with the parts of the brain involved in the left column and the specific functions in the right one. The optic nerves are partially crossing over before reaching the two lateral geniculate bodies of the thalamus. For the sake of simplicity, only one of the two symmetical halfs is depicted in Figure 1, the missing one has to be figured just mirrored on the left side.

The crossover of the optical nerves links the right half planes of both eyes totally to the right lateral geniculate body and, vice verse, the left ones to the left body (not shown in this figure). Within the thalamus, the images received from the eyes exist in a stack of copies. This is the site to generate a static image where all motions of the eyes are compensated. The black dot in the retinas of both eyes should represent a point in an image currently received by the eye. Propagated to the thalamus, this point will be found aligned in the alternately stacked copies from both eyes. Since the viewing directions of both eyes are differently tilted toward the central axis the images received by both retinas are slightly different, i.e. disparate (see below). Both disparate images are merged in the primary visual cortex, where also the perception of directions is allocated. Following the perception of perspectives, there is a separation into the recognition of shapes, patterns and the association with the memory on one side and following motions on the other side. The latter may have an influence on the attention, i.e., if we realise that something is passing by we may follow it with focussing our viewing into the direction of the object passed and even turn our head, if necessary. At this stage, visual perception has entered consciousness, we are knowingly aware of what we see. Before that, all processes compiling an image remain unconsciously. All processes such as adaptation, filtering, merging disparate images etc. work automatically without burdening our mind. Though association takes place deep in the memory it may have a feed back influence even up to the thalamus region, i.e. into the area of unconsciousness. So vision is a rather complex physiological process from receiving optically an image across several mechanisms including feed back loops finally to the understanding what is being seen.

2.2 Binocular Viewing: Disparate Images

The principles of binocular viewing are summarised in Figure 2. Since the retina is spherically shaped all points of an outer circle in focussing distance appear to be equally distant, i.e., there is no stereopsis. This circle is called the horopter (Figure 2 a). If a central point (red viewing axis) is shifted to the left (green axis), then it is shifted on the retina of both eyes in the same direction. This is different for any point in front of or behind the horopter. Any closer point causes an outward shift on the retinas and, vice verse, any more distant one an inward shift. This is the principle of binocular viewing generating disparate images. This can be pretended by presenting technically two images with the disparate features separately to each single eye. One classical way is to overlay such images with different colours and to view them with spectacles that have complementarily stained glasses. Since the eye has different sensors for red, green and blue light, the pair of colours used classically for this purpose is red/green (see below). In some optical systems the viewing axis of each eye is diverted in a way that it is viewing a separate image each, e.g. with optical prismas.

a	b	c

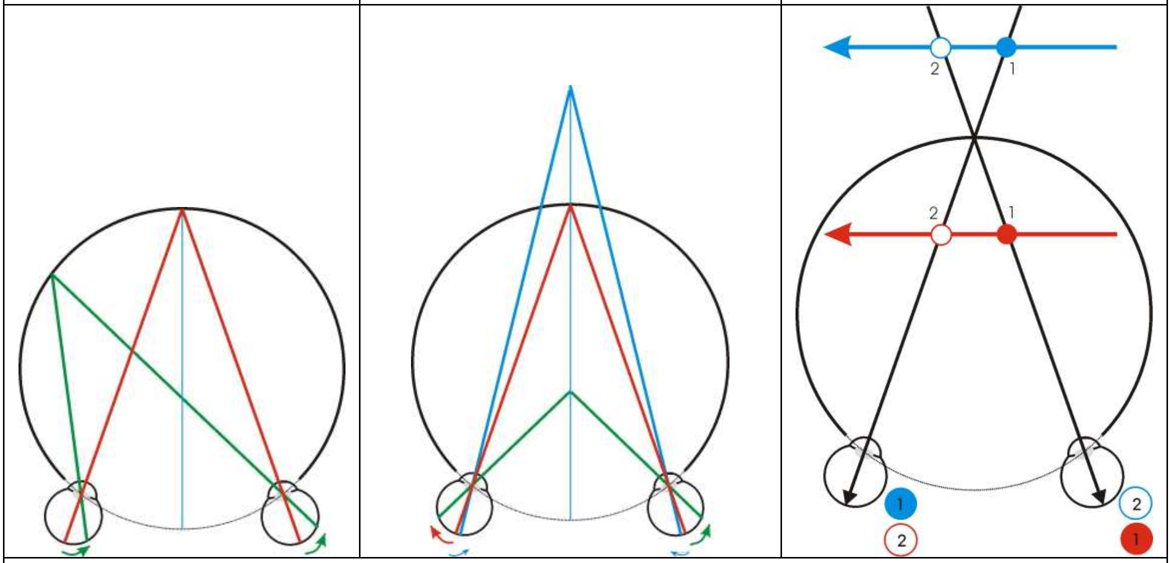


Figure 2: Binocular viewing:
a) horopter (curve with no stereopsis), **b)** disparate images, **c)** Pulfrich effect

2.3 Motion while keeping the direction of viewing

Another way of pretending space is by introducing a time shift between the two eyes when receiving an image of a moving object, commonly known as the Pulfrich effect [2]. This can be achieved by reducing the light intensity in one eye by a grey filter which has a retarding effect on the perception in the respective eye so the moving object is registered slightly later than in the bare eye. In an experiment, such a moving object can be presented as a pendulum that is restricted e.g. with two strings to move in one plane only. When watching this pendulum with one eye darkened by a grey filter, the object appears to leave the plane whenever passing through the centre of the oscillating motion which is physically not possible. Figure 2 c should help to understand this effect. Supposed that the real pendulum moves tangentially to the horopter from right to left, shading one eye with a grey filter causes the time delay as described. If the right eye is shaded, then the pendulum is registered with the left eye first as indicated with the blue symbols, i.e., the pendulum apparently leaves the restricted area to the rear. Vice verse, if the left eye is shaded, then the right one is the first one to register the pendulum so it apparently comes closer. This effect has been used for three-dimensional movie presentations only needing spectacles with a grey filter on one eye and nothing more than a usual screen display.

2.4 Motion to collect multiple aspects

Understanding an object fully in its spatial dimensions requires collecting aspects from all sides. We are either walking around a large object to see it from the side and from the back, if possible, or we are turning around a small one in our hands, to perceive it in its full spatial extension. Principally, this is a feature common with tomographic technologies where projections are collected systematically from any direction of viewing. While physiologically the result is intellectually an abstract imagination the goal of the technological computed tomography (CT) is the generation of a complete set of spatial density data. It requires a secondary set of tools to present these data visually, either as image sequences or as a perspective drawing. In spite of the common features, the main difference between CT and physiological perception remains evident. The technical process consists of clear-cut subsequent phases: data acquisition, processing and computing a three-dimensional set of data which is subsequently transformed into visible images. Physiologically, several mechanisms operate in parallel from the very beginning up to the imagination of a subject in its full three-dimensional extensions.

2.5 Perspective presentation

Apart from all binocular mechanisms there is the understanding of perspective with a horizon and vanishing points. This tool has been used extensively in the art of painting, and a plethora of examples can be found in nearly any gallery. Some artists like Escher challenged the illusion by perverting the perspectives within an image in a way so that e.g. down streaming water ends up in its origin or people are climbing a never ending staircase [1]. However, weird perspective presentations can also being produced photographically, as shown in Figure 3 (the object is real; the photo is not a fake). Particularly in the art of painting, light and shadows are used to underline the spatial feature of the presented scenery. Since this is a process within the cortex it would exceed the scope of this contribution to tackle this aspect more in detail in this context.

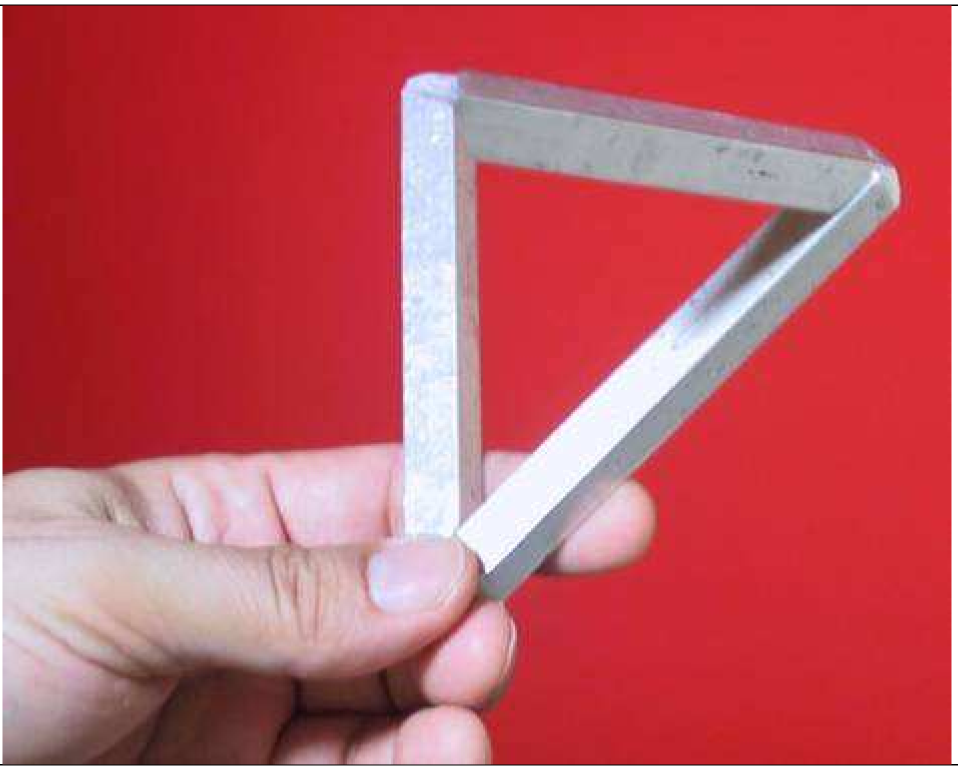


Figure 3: photo of an “impossible” triangle

3. A radiographic example

The practical example in Figure 4 demonstrates how to gain an impression how the details are arranged in the space of the interior of object shown. The specimen consists of a bottom, a top and two side plates, all screwed together, and is stuffed with parts of various densities which are not all visible form the outside (Figure 4 a). The first radiograph (Figure 4 b), taken with a nearly parallel X-ray beam (source in 3 m distance), shows a clear projection of the interior and, in particular, the top, bottom and one side plate with clear edges. Starting from here, there are three approaches to mediate the three-dimensionality of the object: to turn it (Figure 4 c), to introduce a vanishing point by a cone beam radiation geometry (reducing the distance from the source to 0.7 m, Figure 4 d) or by presenting disparate images (Figure 4 e).

Turning the object (Figure 4 c) reveals the nature of the large non-transparent part at the bottom, i.e. a battery pack. However, this is only one another of numerous possible projections. While the technical process of CT moves stepwise around the object collecting hundred and more single projections which are subsequently uses to reconstruct a data set representing the entire object. Upon completing this step, the data set can be presented in various forms. In physiology, a few looks from various sides already allow an assessment of the spatial construction of an object, i.e., an imagination is generated in the brain. When comparing the radiographs taken with the sources at different distances (Figure 4 b and d), it is quite obvious that the lateral areas of the image appear distorted due to the cone beam geometry. Moreover, the parts in the front, i.e. towards the source, appear larger. Both, the trapezoidal distortion and the shift in size of the parts within the projection image are typical features of perspective presentations. The vanishing point is the perpendicular projection of the focal point of the source onto the imaging plane; however, this is not always as obvious as in perspective drawings or paintings. The better way to generate an impression of space is to present disparate images separately to each single eye as explained above. Figure 4 e shows such a presentation where the two images are distinguished by their red and green colour. To obtain a spatial impression, this images needs to be viewed at with red-green spectacles.

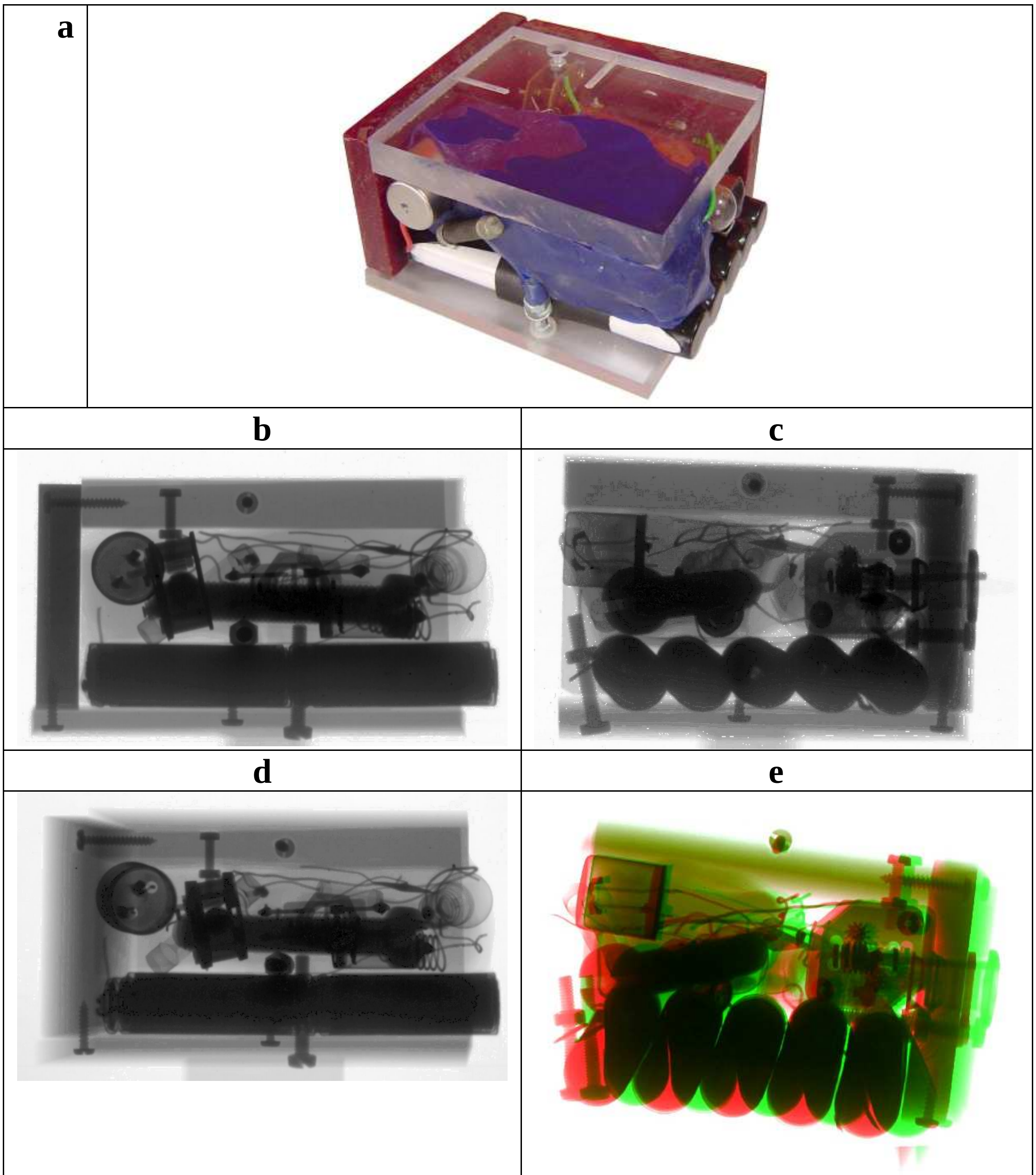


Figure 4: Radiography: space projected into a two-dimensional plane,

a) object containing details easily visualised radiographically

b) nearly parallel X-rays (source in 3 m distance)

c) same object irradiated from the right side

d) cone beam geometry of X-rays (source in 70 cm distance)

e) overlay of disparate images, to be viewed with red/green-spectacles

4. Conclusion

Various physiological mechanisms contribute to the perception of objects and understanding their spatial structure. In difference to technical systems, where the image information is acquired first and subsequently processed, acquisition, processing and data reduction are parallel processes already starting in the eye. To conceive a spatial structure, the predominant physiological mechanisms are: a) binocular viewing and merging the resulting disparate images, b) collecting motion sequences and c) understanding the perspective presentation of spatial objects on a two-dimensional canvas. The process of visual perception is allocated in different parts of the brain and remains unconscious in large parts. It becomes conscious later on in the cortex of the brain, particularly when it comes to the understanding of the course of perspective angles as well as the presence of a horizon and vanishing points. However, just a single viewing perspective can generate confusing illusions. Therefore, it may require several aspects to perceive a spatial subject correctly.

References

1. Maurits Cornelis Escher, 1898 – 1972,
2. A. Anzai, I. Ohzawa and R.D. Freeman, 'Joint-encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect', Nature Neuroscience, Vol 4, No 5, pp 513-518, 2001. (http://neurosci.nature.com)