COMPREHENDING STEREOPSIS

*Abhishek Saini,**Jyotsna Shukla
*Assistant Professor, Department of Physiology, S.M.S. Medical College, Jaipur, Rajasthan, India.
**Senior Professor, Department of Physiology, S.M.S. Medical College, Jaipur, Rajasthan, India

Abstract: Consider this: Our retinas are presented with tiny, distorted, inverted two dimensional images of the surrounding world and from that our mind ("the operational virtues of our brain") is capable of constructing a vivid and rich three dimensional representation, which seems nothing short of astonishing. It is logical to believe that the brain must deploy enormous amounts of computational operations onto the incoming stimuli which carries bits and pieces of information for accomplishing its goal of reciprocating the correlations of the outside world and forming reliable representations inside the egocentric mind. The goal and the end result of all these complex computations are not directed merely towards rendering the representation of the outside world as such and in its totality but to transform the available information into non retinal coordinates useful for guiding action hence unearthing the hidden components like logic, relevance and survival benefits of such representations alongside. Thus in a way, what we see is not just what is present in the outside world and what can be physically measured, but in fact, it is much beyond a simple stimulus registration mechanism so much so that most of the components of our experience of the outside world are "basically constructed internally".

One of the primary component entangled in the perception of surrounding world in all its glory, richness and vividness is the capacity to perceive the objects in three dimensions as well as the capacity to perceive a 'tangible' empty space between the objects.The mechanisms underlying these computational achievements have been far from being clear and still remain one of the unsolved mysteries of the science of vision despite exponential gains in the prowess of various disciplines like experimental electrophysiology, computational neurosciences, physiology, anatomy, psychology, etc. One of the reason precluding the development of a scientifically acceptable theory for stereopsis is its entanglement with the perceptual domains which are themselves very hard to address via prevalent objective assessment methods. On the other hand, due to its habitual appreciation and easy demonstrability, the phenomena of stereopsis poses itself as an appealing and suitable model system subserving understanding of the link between neuronal activity and perception.[1]Therefore many attempts have been made by researchers and theorists across multiple disciplines for putting forward an explanation towards this unique qualia of vision, called stereopsis and the present article is aimed towards addressing the 'phenomenology' of stereopsis as it stands till date.

Keywords: Stereopsis, Binocular disparity, Disparity Tuning curve

Corresponding Author: Dr. Abhishek Saini, Assistant Professor, Department of Physiology, S.M.S. Medical College, Jaipur,Rajasthan, India, Email: drabhi16@gmail.com, Phone: 9468651414

When viewed from both eyes, the outside world is projected as a pair of two dimensional images on to the two retinas that are subsequently processed by the brain leading to a vivid three dimensional visual percept that is qualitatively different from the perception of depth and 3-dimensionality observed monocularly or while viewing a picture.[2, 3] This characteristic and peculiar visual experience that is obtained on the basis of information derived from the two eyes is termed as "stereopsis" whose etymology is derived from the Greek words for "solid" ("stereos") and "sight"("opsis").[4] In other words, stereopsis is the distinctive experience of tangible solid forms and the immersive surrounding space and is bestowed upon due to binocular viewing wherein the horizontally separated eyes view the the outside world from different vantage points resulting in positional differences in the corresponding retinal images.

Monocular v/s Binocular depth perception

In general, our perception of depth and three dimensionality is dependent on many visual cues out of which some are monocular while other are binocular. Out of the many monocular ways to decipher depth, the most useful way is to move the head in order to sample multiple views over time, producing motion parallax wherein the objects near to the point of fixation move in a direction opposite to that of head motion and those farther away move in the same direction as that of head.[5,6] Some of the other monocular cues are: Relative size, Linear perspective, Aerial perspective, Shadows and Shading, Occlusion, Blur, Refection and Illumination, phenomena of Size Constancy and Size Contrast etc. Despite the presence of many intricate ways of depth perception by monocular cues, the qualitative experience of perceiving the depth via stereopsis is remarkably distinctive and very different both aesthetically as well as perceptually. The quotes from Susan Barry, a neuroscientist who recovered her binocular vision in late adulthood (Barry, 2009, 2011)[2,3] portrays the difference quite eloquently...

“[I saw] palpable volume[s] of empty space . . . I could see, not just infer, the volume of space between tree limbs . . . the sink faucet reached out toward me . . . the grape was rounder and more solid than any grape I had ever seen . . . Objects seemed more solid, vibrant, and real. (pp. 94–132)”......ref

As Susan Barry acquired the phenomena of binocular vision only in adulthood, she came out to be vividly aware of the direct sense of three-dimensionality of the surrounding world, a sense that most of us take for granted. Although her capacity to infer depth was present through her monocular repertoire, her real sense of an immersive surrounding space only became overt on arrival of her binocular vision. It is clear from her descriptions as well as from other experiences, that seeing the world twice, using both eyes, from two slightly different perspectives, makes a number of important visual computations that result in an altogether unique subjective experience of perceiving the world in three dimensions termed distinctively as Stereopsis.

Hence, the term stereopsis should be used wherein the visual dimension of depth is perceived via the differences in images projected onto to the retinas of the two eyes - ‘so- called binocular disparity’ by virtue of binocular vision. In the ensuing discussion we will first discuss the basic geometrical considerations of binocular disparity and then progressing towards discussing the mechanisms by which the brain encode and compute the disparity information for generating the visual perception of stereopsis.

Basic Geometrical considerations

Due to the presence of horizontal separation between eyes, the two eyes view the same visual scene from two slightly different vantage points. Whenever the eyes look at a particular feature in space, the two eyes try to fixate on that so that the image of the object is projected on to the fovea (the tiny retinal region of highest visual acuity) of both the eyes. Apart from this point of fixation, there is also a plane called “geometric horopter” on which any point will project onto the two retinas in such a fashion, that the projection image in each eye will be equidistant from the respective fovea. These points of projection are called as ‘corresponding points’. Also, to an observer, all the points on the geometric horopter will appear to lie roughly at the same depth.

The parts of the object that lie either in front or behind this plane, will lead to projections on to the retinas that will not be corresponding points and thus there will be a disparity in the position of such projections. This disparity is called as binocular disparity and it lays the foundation for development of stereopsis.

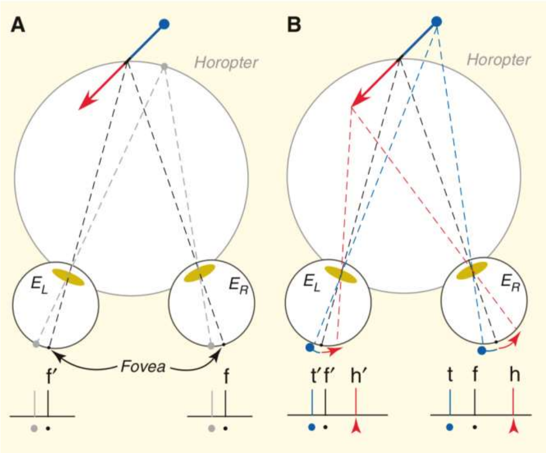

Figure 1: The geometry of stereopsis (adapted from Tyler 2004)[7]

An example: In the above figure (figure:1), consider two eyes fixating on a black dot on an arrow resulting in formation of images on to the two fovea. These points on the fovea are called corresponding retinal points and the images so formed are called corresponding retinal images. If someone start from this position and focus at a new point along the grey circle, the projections from this new point will move same distance on each retina hence resulting in corresponding points again. The locus of all such points whose projections falls on corresponding retinal points will constitute the geometric horopter (the grey circle in figure 1A). Keeping both eyes fixed on the black dot on the arrow, as previously, the head of the arrow which is placed in front of the horopter, will result in retinal projections that are not equidistant from the two fovea hence resulting in binocular disparity which in the present case can be mathematically represented as -[(h-f)-(h`-f`)]. It is customary to denote these disparities as crossed disparities and these are assigned negative values. Conversely, the binocular disparity [(t-f)-(t`-f`)] produced by tail of the arrow, which is placed behind the horopter, (Figure 1 B) is referred to as uncrossed disparity and positive values are assigned to it. At this point it seems worth mentioning, that there is a relatively narrow range of allowed disparities that could lead towards generation of stereopsis because image fusion is also necessitated for normal binocular vision. This range of admissible separation of the two retinal images producing a single visual percept can be traced back to its corresponding area in visual space called Panum's area beyond which fusion of the two images is not possible hence resulting in double vision wherein the observer is aware of each retinal image separately. However, despite the subjective experience of double vision, the individual will still be able to extract meaningful depth information upto binocular disparities of many magnitudes depending on the size of the feature.

The so produced binocular disparity is dependent on the depth of the object from the Horopter as well as on the distance of the object from eyes. The binocular disparity generated out of a given depth difference varies as the inverse square of the viewing distance, hence, at close viewing distances, exceedingly small depth differences can be discriminated. While for large viewing distances, the larger relative depth dimensions are needed to produce accurate depth perception via stereopsis.[4,8-10]

Historical Background

Stereopsis was first explained by Charles Wheatstone in 1838, demonstrating an illusion of depth from presentation of two flat pictures that had a slight horizontal shift with respect to each other. During the Victorian times, stereoscope was quite popular after the invention of prism stereoscope by David Brewster. Peter Ludvig Panum, Ewald Hering, Adelbert Ames Jr., Kenneth N. Ogle were among the many researchers who explored the realm of stereopsis in many different ways but were mainly concerned with explaining stereopsis empirically. In the same lines, in 1960s Bela Julesz[11] invented random-dot stereograms wherein pair of images of random dots which, when viewed with the aid of a stereoscope, or with the eyes focused on a point in front of or behind the images, produced a sensation of depth, with objects appearing to be in front of or behind the display level. Studies aimed towards explaining the neurophysiological basis of stereopsis started in late 1960s with the work of Horace Barlow, Colin Blakemore, Jack Pettigrew and others who published the first reports of neurons demonstrating the property of disparity selectivity in the primary visual cortex (V1, or area 17) in anaesthetised cats.[12,13] These authors proposed the presence of subgroups of neurons in the layers of primary visual cortex which demonstrate the capability of signalling different ranges of binocular disparities and ultimately encoding a range of object positions in depth. In the 1980s, Gian Poggio and others found neurons in the area V2 of the brains of monkeys that responded to the depth of random-dot stereograms.[14,15] .They also classified and established four basic classes of neurons (discussed below) which differed in their optimal response towards objects placed at various depths across the horopter.

Physiological Basis of Stereopsis

"Encoding the binocular disparity" As the basic root cause for stereopsis are the binocular disparities, it is quite imperative that the processing element should be endowed with a capacity to compare the signals from both retinas simultaneously. It is in the primary visual cortex (V1) where single neurons receiving inputs from both the eyes are present making them ideal candidate for performing computational operations necessary for stereopsis. Besides integration of the signals, these neurons are unique because they are sensitive to specific differences between the receptive fields in the two eyes (or to the degree of binocular disparity).Their individual sensitivity profile is specific towards a specific range of disparity, showing higher rates of discharges for some particular disparities and lower rates for other. Such a modulated firing rate pattern can be visualised graphically in form of a 'disparity tuning curve'. These curves are prepared by plotting the firing/spike rate of the V1 neurons evoked by visual stimuli of differing binocular disparities on the Y axis and the stimulus binocular disparities on the X axis in terms of degrees of disparities. These curves are categorised on the basis of allocation of peak firing response. Accordingly 4 typical binocular disparity selective neurons can be classified namely: Tuned Excitatory Neurons, Tuned Inhibitory Neurons, Crossed Disparity Selective Neurons and Uncrossed Disparity Selective Neurons. (See figure 2.)

a) Tuned Excitatory neurons (Figure 2A): these neurons demonstrate maximum firing rates when the disparity is near zero (or no disparity) and minimum response when the disparity is higher on either side of the zero.
b) Tuned Inhibitory neurons (Figure 2B): these neurons are suppressed by a small range of disparities around zero and appear to be an inverted version of the tuned excitatory cells
c) Crossed Disparity selective neurons (Figure 2C): these neurons have their response peaks at crossed disparities (produced by objects in front of horopter)
d) Uncrossed Disparity selective neurons (Figure 2D): these neurons have their response peaks at uncrossed disparities (produced by objects behind the horopter)

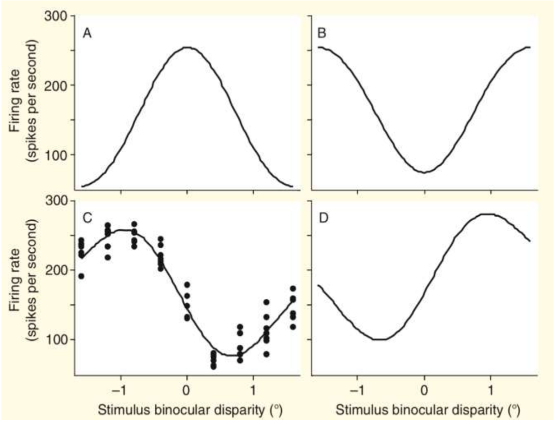

Figure:2 Spike rates evoked by different binocular disparities in different specialised neurons (adapted form Ponce and Born 2008) [16]

Though these 4 classes of neurones have been classed typically, there is a continuum of various neurons with different tuning curve characteristics spanning the gamut from near-, to zero-,to far- tuned neurons. Therefore for any disparity within the admissible range there is a specific cell in the V1 area of the cortex that is specific for it.[16]

Basis of disparity tuning characteristics

The primary visual area (V1) contains many cells in its recipient layers which are called - simple cells, exhibit a property of firing action potentials in response to a small spot of light appearing in some parts of their receptive fields (ON regions), but are inhibited when the same spot is illuminated in other regions of the receptive field (OFF regions)[17]. David Hubel and Torsten Weisel[17] revealed that most simple cells have receptive fields comprising elongated regions of ON excitation and flanked by elongated regions of OFF inhibition. The receptive field profile of such kind is explained by a mathematical function called Gabor function which is the product of a Gaussian sensitivity profile and a sinusoid. It has been shown that V1 simple cells are selective for specific stimulus position and orientation, hence binocular disparity can be coded by introducing small differences between the left and right receptive fields. It can be done via two ways: [18,19]

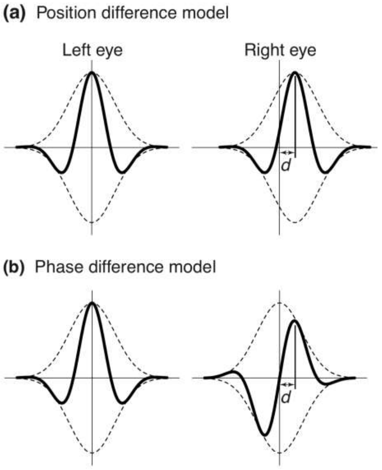

Figure 3: Disparity encoding models, Solid curves are Gabor Functions that represent the receptive field profiles of an idealised binocular simple cell and the dashed curves show the Gaussian envelopes of the profiles. The horizontal axis represent retinal position while the vertical axis represent sensitivity to a luminous increment. (Adapted from DeAngelis, G.C. et al. 1995)[20]

a) **Position disparity** (Figure 3a): the relative positions of excitatory and inhibitory regions in the left and right receptive fields are similar but there is a spatial offset between the retinal positions of each receptive field

b) **Phase disparity**(Figure 3b): in this case both left and right receptive fields share same retinal positions but the relative positions of excitatory and inhibitory regions within the receptive fields are different.

Position disparity versus Phase disparity encoding

As explained above (depicted in figure 3), position disparity between the two eyes or phase disparity represent the two possible ways that could encode a wide range of binocular disparities. The first testament towards a position disparity encoding theory came from the study of Hubel and Wiesel[17] who demonstrated that the spatial arrangements for excitatory and inhibitory regions were the same for the two receptive fields of binocular cells.A study done by Maske et al. gave similar insights wherein they reported that the receptive field profile of the two eyes were remarkably similar.[21] In a study on binocular neurons of the visual Wulst of the barn owl, Wagner and Frost[22] reported that the neurons exhibited a characteristic disparity in their responses to a stimulus consisting of dichoptic sine wave stimuli with variable spatial frequency. In contrast to these studies, Freeman & colleagues and DeAngelis et al. demonstrated that simple cells in the cat do exhibit interocular phase differences.[20,23-25] A hybrid type of model, supporting the coexistence of both position/phase encoding by the V1 cells was propounded by DeAngelis et al. and Anzai et al.[20,24,26,27]

According to theposition disparity encoding model, the range of disparities that can be encoded does not necessarily depend on the spatial frequency of the visual image whereas in the phase model, the range of disparities encoded by a population of neurons is inversely proportional to the spatial frequency.

Apart from these models which presume that V1 neurons are selective for detecting the absolute disparity of visual features on the retina, there have been studies which demonstrated that humans are most sensitive to the relative disparities between multiple visual targets.[28] It has also been postulated that there are some neurons in the area V2 which are responsible for signalling the relative disparity.[29]

Complex cells and other higher brain areas

Another type of orientation selective cell found in the primary visual cortex is called Complex cell. These cells are sensitive to optimally oriented lines anywhere in their receptive fields without showing a difference in firing pattern on the basis of different locations (ON or OFF regions)within the respective receptive field. It has been argued, that positional invariance of complex cells have arose due to its reception of feed from multiple simple cells.

A more intricate ‘energy model’ has been proposed to explain the behaviour of complex cells.[16] This model envisages disparity sensitive complex cell as a recipient of input from multiple simple cells having same binocular disparity and orientation preference but different relative position of peak excitatory domain within each receptive field . The disparity energy model accounts well for the observed diversity of disparity tuning curves which are postulated as a possible mechanism for providing a neural tool kit for stereopsis.[16] There had also been studies which postulated, that, by having a knowledge of vergence and horizontal disparity cues, the visual system is capable of knowing the visual distance of the object of interest and compute egocentric depth from this data.

Apart from the obvious roles of V1, the areas V2 and MT are also likely to be involved in processing stages for elaborating the disparity signals. It should also be noted that disparity selective neurons are also found in a number of visual cortical areas in primates including V1, V2, V3, MT, MST and IT.[16]

Despite of all the above seemingly plausible explanations for elucidating the neural mechanistic correlates of stereopsis, there have been many unresolved issues warranting the need of further enlightenment. Most of the theories proffered to explain the phenomena of stereopsis relied on a ‘bottom up’ approach of the stimulus processing but on many occasions, a ‘top-down’ modulation of neural processing seems a better explanation of the observed phenomena of stereopsis under certain circumstances.

Therefore a clear understanding of the neural processing subserving the phenomena of stereopsis still remains a goal to be achieved but many in-roads have been made in recent times taking advantage of the remarkable growth in technology and techniques enabling quantitative assessment of many perceptual attributes, a thing that alluded us in the past.

Conclusion

Ever since the original descriptions of disparity selective neurons in the late 1960s, most physiological research studies have revolved around measurement of the disparity tuning functions of cortical neurons for propounding an explanation for the observed phenomenon of stereopsis. Although, there are many shortcomings towards this approach but still many of the postulates proffered on the basis of these researches are increasingly been affirmed by new age objective scientific methods. With the convergence of traditional electro physiological methods with psychophysical and computational approaches, a new era of understanding the neural correlates and mechanism of stereopsis have begun to emerge promising a development of concrete understanding of the neural basis of stereopsis.

Abbreviations: MT- Middle temporal area MST- Medial superior temporal area IT-Inferior temporal gurus

No conflict of interest

References

1. Roe, A. W., Parker, A. J., Born, R. T. and DeAngelis, G. C. (2007). Disparity channels in early vision. J. Neurosci. 27, 11820-11831
2. Barry, S. (2009). Fixing my gaze. New York, NY: Basic Books.
3. Barry, S. (2011, October 15). Eyes on the brain: Teaching cross-eyed people to see in 3D. Psychology Today. Retrieved from http://www.psychologytoday.com
4. Howard IP, Rogers BJ (1995). *Binocular vision and stereopsis*. New York: Oxford University Press
5. Rogers, B. and Graham, M. (1982) Similarities between motion parallaxand stereopsis in human depth perception. Vis. Res. 22, 261–270
6. Ferris, S. H. (1972). Motion parallax and absolute distance. Journal of experimental psychology, 95(2), 258--63
7. Tyler CW, Likova LT, Wade AR. Properties of object processing in lateral occipital cortex. Journal of Vision. 2004 Jun 1;4(8):91-.
8. Julesz, B. (1971) Foundations of Cyclopean Perception, University of Chicago Press
9. Mayhew, J. (1982) The interpretation of stereo-disparity information: the computation of surface orientation and depth. Perception 11, 387–403
10. Bishop, P.O. (1989) Vertical disparity, egocentric distance and stereoscopic depth constancy: a new interpretation. Proc. R. Soc. London Ser. B 237, 445–469
11. Julesz, B. (1960). "Binocular depth perception of computer-generated images". *Bell System Technical Journal*. **39** (5): 1125–1163.

12. Barlow, H.B. et al. (1967) The neural mechanism of binocular depthdiscrimination. J. Physiol. 193, 327–342
13. Pettigrew, J.D. et al. (1968) Binocular interaction on single units in catstriate cortex: simultaneous stimulation by single moving slit withreceptive fields in correspondence. Exp. Brain Res. 6, 391–410
14. Poggio, G.F. and Fischer, B. (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey.J. Neurophysiol. 40, 1392–1405
15. Poggio, G.F. et al. (1988) Stereoscopic mechanisms in monkey visualcortex: binocular correlation and disparity selectivity. J. Neurosci. 8,4531–4550
16. Ponce CR, Born RT. Stereopsis. Current Biology. 2008 Sep 23;18(18):R845-50
17. Hubel, D.H. and Wiesel, T.N. (1962) Receptive fields, binocular interaction,and functional architecture in the cat‘s visual cortex. J. Physiol. 160,106–154
18. Freeman, R.D. and Ohzawa, I. (1990) On the neurophysiologicalorganization of binocular vision. Vis. Res. 30, 1661–1676
19. Nomura, M. et al. (1990) A binocular model for the simple cell. Biol.Cybern. 63, 237–242
20. DeAngelis, G.C. et al. (1995) Neuronal mechanisms underlying stereopsis: How do simple cells in the visual cortex encode binocular disparity?Perception 24, 3–31
21. Maske, R. et al. (1984) Binocular simple cells for local stereopsis:comparison of receptive field organizations for the two eyes. Vis. Res.24, 1921–1929
22. Wagner, H. and Frost, B. (1993) Disparity-sensitive cells in the owl havea characteristicdisparity. Nature 364, 796–798
23. Freeman, R.D. and Ohzawa, I. (1990) On the neurophysiologicalorganization of binocular vision. Vis. Res. 30, 1661–1676
24. DeAngelis, G.C. et al. (1991) Depth is encoded in the visual cortex by aspecialised receptive field structure. Nature 352, 156–159
25. Ohzawa, I. et al. (1996) Encoding of binocular disparity by simple cellsin the cat‘s visual cortex. J. Neurophysiol. 75, 1779–1805
26. Anzai, A. et al. (1999) Neural mechanisms for encoding binocular disparity:receptive field position versus phase. J. Neurophysiol. 82, 874–890
27. Anzai, A. et al. (1997) Neural mechanisms underlying binocular fusion and stereopsis: position versus phase. Proc. Natl. Acad. Sci. U. S. A. 94,5438–5443
28. Westheimer, G. (1979) Cooperative neural processes involved instereoscopic acuity. Exp. Brain Res. 36, 585–597
29. Thomas, O.M. et al. (1999) The responses of binocular V2 neurons torelative and absolute disparities. Soc. Neurosci. Abstr. 25, 2060