Optic arrays and retinal images

In their recent *Perception* editorials on motion parallax and stereopsis, Barbara Gillam (2007) and Brian Rogers (2007) have argued that improved clarity is achieved by considering first the information in optic arrays. I offer the following remarks to this debate.

(i) Standard practice in computer vision when building a stereovision system is to begin with the projective geometry of the optic array for each eye. This starting point is perhaps why the confusions discussed by Gillam and Rogers are missing from the computer vision literature. Consider the following answer by John Mayhew when asked by Vincent Torre to comment on what “seems to be [a] controversy concerning the effect of eye rotations on the disparity field”. Neither Gillam nor Rogers, of course, need to be reminded of what Mayhew says, but readers less well versed in the stereopsis literature might find its clarity both helpful and relevant.

“The problem arises I think because the difference between the concepts of optic array and the measured disparity field is frequently confused. If we define the optic array to be the lines of sight through the nodal points of the eye and assume that the eye rotates around this point, then clearly the optic array cannot be changed by any rotation of the eye (and as is well known, this is why it is impossible to recover information about the three-dimensional structure of the scene from rotations alone). Similarly if we assume another eye, it too will have associated with it a different optic array, which cannot be changed by rotating that eye. Furthermore, the angle between the lines of sight which derive from the same point in the world will subtend at the nodal point of the eyes an angle (referred to as binocular parallax) that cannot be changed by the rotations of the eyes. This is a good thing, because it implies that there is information about the three-dimensional structure of the world which is in principle available to the stereo mechanism independently of which point in the world the eyes fixate.

The problems arise when one considers the issues involved in measuring binocular parallax (the angle between the corresponding lines of sight) of the point in the world. The physiological mechanism for this measures the retinal disparity, by which is generally meant the difference in retinal position in the left and right eyes of the two lines of sight derived from the same three-dimensional point. Of course it is now simple to see how moving the eyes from fixating one point to fixating another must affect the pattern of measured disparities on the retina. Subject to physiological abnormalities, every point in the world that the eyes fixate gives rise to a zero disparity measurement. Whereas in general, when the eyes fixate another different point, the previously fixated point will give rise to a measured retinal disparity which is a function of both the difference in three-dimensional positions of the points, and the directions and magnitudes of the rotations of the eyes which were needed to change fixation.” (These remarks follow Mayhew et al 1992, pages 325 – 326. My recollection is that Torre posed his question because he had been having discussions with champions of the proposal that it is best to confine geometric analyses of binocular vision to optic arrays, and that consequently Mayhew and Longuet-Higgins were barking up the wrong tree.)

These remarks illustrate my contention that the discipline of actually trying to build a vision system strongly encourages a clarity of analysis sometimes lacking from discussions in the psychological literature. Computer vision often gains insights from careful examinations of camera images, and/or simulations of them, noting the information that they do (and do not) contain.
(ii) Gibson changed his mind about what he wanted to be understood by his concept of the optic array. Specifically, and it has come as a recent surprise to me, (1) Gibson rejected the definition used by Gillam and Rogers, vide the following quotation (italics in the original):

“The heart of ecological optics is the concept of the ambient optic array at a point of observation. The ambient array is to be distinguished from the ambient light. The former constitutes stimulus information; the latter constitutes stimulus energy. Boynton thoroughly approves of the concept of ambient light energy coming to a point (I call it a ‘Boynton point’) but he is doubtful of the concept of a purely relational array or structure. He catches me up for having once defined it in terms of rays, and he is quite right to do so. The formula of a ‘dense interconnecting network of rays’ was a mistake; all I meant to imply by it was that the steady state of illumination in a living-space is projective. I now define an ambient optic array as a nested set of adjacent solid angles, not rays, each solid angle corresponding to one of the large faces or facets of the environment. The solid angles are separated by contours or contrasts. These contours I take to be mathematically definite, and to be independent of an observer. So defined, the array as such is invariant from noon to sunset.” (Gibson 1974, page 310)

It is clear from this that care is needed in using the term optic array. The meaning used by Gillam, Rogers, and Mayhew is based on the projective geometry of lines of sight to points in the scene. It is the meaning that Gibson came to reject, as we can see from this quotation if, as seems reasonable, we regard lines of sight as equivalent to rays of light for the purposes of this discussion (whether he subsequently changed his mind again I know not). Of course, as Gillam notes, the principles of the optic array so defined were developed by Renaissance artists for painting in perspective. Hence Gibson has no originality in that regard. Moreover, as far as I am aware, Gibson’s specification of the optic array as a nested set of solid angles, not rays, has played no part in clearing away confusions in the literatures of motion parallax and stereopsis. Indeed, I find it hard to imagine how it could do so. If this is correct then Gibson should be accorded no credit for the benefits that Gillam and Rogers note for beginning with the optic array, because they are not using that concept in the sense in which Gibson wanted his use of the term to be understood. Indeed, it would appear that the prefix ‘Gibson’s’ should be dropped when referring to optic arrays of light rays. Gibson does not own that concept, either in terms of originality or in terms of how he wanted his concept of the optic array to be understood.

(iii) Gillam explains why she thinks beginning with the optic array assists in explaining the concepts of crossed and uncrossed disparity. I would like to make a more radical suggestion. Let us abandon those terms altogether from current usage. They seem to me to add nothing to the terms ‘convergent’ and ‘divergent’ disparities, which are more natural and less confusing for students. I have found in my teaching that trying to explain the crossed/uncrossed distinction is a confusion-inducing distraction from getting students to grasp the elements of stereo geometry. Much better, I submit, is to concentrate on teaching the core concept of disparity vector fields. I find teaching the concepts of crossed/uncrossed disparity unhelpful in that regard.

(iv) My final point concerns different meanings of the term ‘retinal image’ that need to be distinguished when trying to make sense of the quotations of Helmholtz provided by Rogers. First, if the context is purely projective geometry then the image refers to

(1) I am indebted to our Editor-in-Chief for drawing my attention to the chapter from which this is taken. Gregory has discussed at length Gibson choosing to ‘deny the retinal image’. This was a phrase used by Gregory (2000) in which he discussed the mystery of Gibson’s blind spot on the retina (forgive my pun but our Editor-in-Chief loves them). Gibson did finally retreat from his extreme position of denying the retinal image (Gregory, personal communication), but the fact that he ever adopted it has perhaps been a source of confusion. Indeed, Gregory (2000) suggested that this could be a case of “the sometimes misleading power of top–down belief on observation”.

the projection of the optic array on to a surface that might be called an ‘imaginary retina’. This kind of retinal image suffers from no distortions, being a construct in the mathematical world of projective geometry. When considering images in this sense, we are in the domain of discourse that Marr (1982) called ‘computational theorising’. The issues there are to do with specifying a clearly stated vision problem and identifying constraints available for its solution. Such constraints can sometimes be found in geometry, as in Mayhew and Longuet-Higgins (1982) and Gillam and Lawergren (1983).

A second meaning of the term retinal image conceives it as the physical pattern of light rays arising from a real scene striking a real retina after passing through a real lens. This pattern of light will be subject to various kinds of distortions, such as spherical and chromatic aberrations.

A third usage of the term retinal image is to refer to the pattern of activity in rods and cones. Here the meaning is an ‘image’ encoded in the receptor mosaic. A similar sense of the word ‘image’ is commonplace in the computer vision literature when referring to pixels arrays in digitally encoded camera images.

These distinctions between different meanings of the term retinal image are helpful when thinking about Rogers’s comments on Helmholtz’s ideas on ‘celestial spheres’ and retinal images. Rogers regards Helmholtz as similar to Gibson in arguing against the dangers of “attaching too much importance to the particular characteristics of retinal images” (quote from Rogers’s penultimate paragraph).(2) However, this recommendation seems inappropriate if made about retinal images in the projective geometry sense. Rogers (personal communication) sees the optic array (of lines of sight) and its projection as an image as two sides of the same coin. I would agree, up to a point, that point being that the coordinate framework generally used to describe a retinal image is fovea centric and hence retinal images (unlike optic arrays) change with eye movements (vide the quotation from Mayhew). This fact is crucial when analysing how to extract useful scene properties from retinal disparities.

Rogers concludes his Editorial by saying: “Helmholtz ... was surely right in stressing the advantages of describing and thinking about the visual world in terms of his celestial sphere, which, like Gibson’s optic array, is uncontaminated by the distortions, imperfections, and philosophical confusions that belie our retinal images”. Presumably the distortions and imperfections referred to here concern ‘real’ retinal images, not ‘projective’ ones, using the terminology I introduced above. But who today is handicapped when thinking about retinal images, be they real or projective?

John Frisby
Department of Psychology, University of Sheffield, Sheffield S10 3TE, England, UK;
e-mail: j.p.frisby@sheffield.ac.uk

References
Gibson J J, 1974 “A note on ecological optics”, in Handbook of Perception volume 1 Historical and Philosophical Roots of Perception Eds E C Carterette, M P Friedman (New York: Academic Press) chapter 15, pp 309–312
Gillam B, 2007 “Guest editorial: Stereopsis and motion parallax” Perception 36 953–954
Gillam B, Lawergren B, 1983 “The induced effect, vertical disparity, and stereoscopic theory” Perception & Psychophysics 34 121–130
Gregory R L, 2000 “Adventures of a maverick”, in Psychology in Britain: Historical Essays and Personal Reflections Eds G C Bunn, G Richards, A D Lovie (Oxford: Wiley/Blackwell) pp 381–392

(2) Nowadays, all (good) teachers of introductory courses on visual perception emphasise that we do not ‘see’ our retinal images—they are only a data resource for perceptions. That truth can hardly be said often enough in introductory courses given the destructive pervasiveness of the photographic metaphor for seeing, and it seems to be the viewpoint that Helmholtz was railing against, judging from Rogers’s account.
Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)

Mayhew J E W, Longuet-Higgins H C, 1982 “A computational model of binocular depth perception” *Nature* 297: 376–379

Mayhew J E W, Zheng Y, Cornell S, 1992 “Use of stereopsis under variable camera geometry” *Philosophical Transactions of the Royal Society of London, Series B* 337: 315–326

Rogers B, 2007 “Guest editorial: Optic arrays and celestial spheres” *Perception* 36: 1269–1273