

On Alternative Approaches to 3D Image Perception: Monoscopic 3D Techniques

Barry G. Blundell

Received: 27 January 2015 / Revised: 31 March 2015 / Accepted: 1 April 2015 / Published online: 28 April 2015
© 3D Research Center, Kwangwoon University and Springer-Verlag Berlin Heidelberg 2015

Abstract In the eighteenth century, techniques that enabled a strong sense of 3D perception to be experienced without recourse to binocular disparities (arising from the spatial separation of the eyes) underpinned the first significant commercial sales of 3D viewing devices and associated content. However following the advent of stereoscopic techniques in the nineteenth century, 3D image depiction has become inextricably linked to binocular parallax and outside the vision science and arts communities relatively little attention has been directed towards earlier approaches. Here we introduce relevant concepts and terminology and consider a number of techniques and optical devices that enable 3D perception to be experienced on the basis of planar images rendered from a single vantage point. Subsequently we allude to possible mechanisms for non-binocular parallax based 3D perception. Particular attention is given to reviewing areas likely to be thought-provoking to those involved in 3D display development, spatial visualization, HCI, and other related areas of interdisciplinary research.

Electronic supplementary material The online version of this article (doi:[10.1007/s13319-015-0047-6](https://doi.org/10.1007/s13319-015-0047-6)) contains supplementary material, which is available to authorized users.

B. G. Blundell (✉)
School of Computer and Mathematical Sciences,
Auckland University of Technology,
Private Bag 92006, Auckland 1142, New Zealand
e-mail: barrygblundell@gmail.com

Graphical Abstract



Keywords Stereopsis · 3D perception · Paradoxical monocular stereopsis · Monocular 3D · Plastic effect · Stereoscopy by default

1 Introduction

The basic function of any 3D display is the presentation to the human visual system of stimuli that evoke a vivid sensation of depth whereby objects, structures

and other entities comprising an image scene appear to coexist within a spatial framework that mimics key visual attributes we routinely ascribe to our physical surroundings. The potential benefits of 3D systems are diverse and vary according to content, method of presentation and application context. They may, for example, enhance realism and immersion, facilitate the interpretation of spatial/geometrical relationships (both static and kinetic), provide greater insight into surface profiles/textures, and better support spatial interaction techniques.

The visualization of our surroundings is achieved against a backdrop of rapid saccadic eye movements, via the numerous attributes (pictorial cues) associated with the retinal images, and with the support of the oculomotor and parallax cues [10]. In addition, prior experience and familiarity usually play important roles such that the perceptual experience is often strongly influenced by our visual expectations [47].

The ensuing vivid externalized qualitative impression of our local spatial surroundings is generally referred to as ‘stereopsis’ [from the Greek *stereos* (solid) and *opsis* (sight)]. Drawing on and adapting Vishwanath [52], we loosely define ‘conventional stereopsis’ as the externalized visual impression of tangible form, order and immersion that is set in a 3D framework of pervasive negative space and which is derived from binocular observation of our spatial surroundings, or by appropriate viewing of stereoscopic images. In this context the term ‘negative space’ is assumed to denote translucent permeable space in which physical entities generally appear to coexist and which enhances the tangibility of their spatial separation. The crucial role of this space is well recognized by artists such as Claude Monet who wrote: ‘*I want the unobtainable. Other artists paint a bridge, a house, a boat and that’s the end. They are finished. I want to paint the air which surrounds the bridge, the house, the boat, the beauty of the air in which these objects are located, and that is nothing short of impossible*’ (quoted in Barry [4]).

In recent years there has been an increasing tendency in literature to assume that 3D perception is exclusively based on spatial (geometric) disparities in the two retinal images. For example, Patterson [44] suggests that ‘*Stereopsis refers to the perception of depth based on binocular disparity, a cue that derives from the existence of horizontally separated eyes.*’ In fact the assumption that binocular vision (and

associated retinal disparities) is a fundamental requirement for vivid depth perception is a misnomer and well before the advent of stereoscopic imaging it was recognized that the spatial realism of perspective paintings, engravings and projected images could often be enhanced via certain viewing strategies and optical arrangements.

Interest in such approaches can be readily traced back to the early fifteenth century and was particularly evident during the eighteenth century—a period in which 3D first gained popularity and some commercial success. Subsequently Wheatstone’s demonstration of stereoscopic techniques in the first half of the nineteenth century [55] led to a revival of interest in 3D and quickly formed the foundations of a flourishing 3D industry. Koenderink et al. [36] observe that: ‘*Historically, a sharp caesura occurs with the invention of the stereoscope...any depth not based on [binocular] stereopsis was suspect.*’

Although mainstream attention has focused on stereoscopic methods, the delivery of 3D via non-binocular parallax based techniques has attracted steady interest—particularly within the vision science and arts communities. A review of ensuing research outputs to date coupled with first hand experimental observations firmly demonstrates that under appropriate conditions vivid spatial perception can be experienced in the absence of binocular parallax. Furthermore, it is evident that the 3D perceptual experience is multifaceted and can be catalyzed (or frustrated) by various stimuli. On balance it appears that 3D spatial perception (in contrast to 2D ‘Flatland’)¹ represents a natural result that the visual system will generally work towards achieving. Thus under appropriate conditions (relating to image content, composition and method of presentation) the visual system has the natural capacity to deliver a strong impression of 3D form, emergence and

¹ Here we assume that a 2D perceptual experience is one in which the image entities *that are the focus of attention* reside, or appear to reside, within two spatial dimensions—the negative space associated with the third dimension (normal to the depiction surface) being absent. However, such a perceptual experience is generally set within an overarching 3D framework. This arises because, for example, in most situations stimuli comprising a 2D image scene collectively continue to appear to be located at a finite distance. For near field viewing scenarios, this perceived distance is reduced (but not eliminated) when one eye is shaded.

recession from single, well focused, images rendered from a particular instantaneous vantage point. Here, such images are termed ‘monoscopic’ and are assumed to be created on (or projected onto) planar or curved surfaces and encapsulate only the pictorial cues to depth. It is understood that the surface does not sweep out a third dimension over time. The term is used for convenience and we acknowledge that the etymology of ‘stereoscopic’ contains nothing implying duality.

Within the 3D display technology community this area has attracted relatively little attention, although monocular approaches are occasionally reported. For example, Ohmori et al. [43] describe a 3D display implemented by directing image content into the tracked eye in such a way as to stimulate spatial perception by initiating changes in focal length of the lens. Thus depth perception is underpinned by the accommodation cue (although this may also stimulate changes in vergence).

Here we coin the expression ‘stereopsis by default’ (s-bd) when referring to a perceptual experience in which attributes associated with conventional stereopsis are derived from monocular or binocular viewing of single, or identical pairs of monoscopic images comprising static or kinetic content. This is based on a term used by Gabor [24]: ‘*Anybody can check “stereoscopy by default” in the cinema by closing one eye and forming a tube with the hand around the other to obscure the frame of the screen.*’

As illustrated in Fig. 1, in the case of stereoscopic images the depth map may be deduced solely from spatial (geometric) disparities in the left and right views—although in many natural situations the depth map may not be fully complete (see Sect. 2). In contrast when dealing with static monoscopic images, the qualitative depth map is less readily available but may often be inferred from various pictorial cues supported by prior experience, familiarity and expectations. On occasion, the ensuing visual experience may include erroneous content—suggesting the formulation of false perceptual ‘solutions’.

Binocular cues cannot be assumed to yield a comprehensive depth map to every point in an image scene, nor can we assume that all available sources of data are combined in a way that produces a single unified spatial representation able to support all activities requiring quantitative judgements. One hypothesis is to consider data as being used to provide a plurality of spatial

representations [26]. In this scenario, it is most likely that the presence or absence of spatial retinal disparities impacts in various ways on individual representations. This in turn suggests that often subtle changes are likely to occur when binocular parallax is suppressed.

This paper focuses on several methodologies which have the ability to support s-bd. These share the common characteristic of not being reliant on spatial disparities in the retinal images formed in the two eyes, and which directly relate to the positions of objects within 3D space. In addition, we particularly focus on qualitative aspects of the visual experience and pay less attention to complex issues relating to quantitative depth judgements.

In the next section we introduce various concepts and terminology relating to relevant aspects of spatial perception. In addition we consider characteristics of s-bd and in Sect. 3 summarise exemplar viewing techniques. Section 4 focuses on s-bd as a perceptual experience and subsequently in Sect. 5 possible underlying mechanisms are considered with particular emphasis on the disruption of the oculomotor cues to depth, reduction of the monocular field of view, observer immersion, geometrical considerations, and subsidiary awareness. Here we also briefly allude to several indicative controlled studies. Finally in Sect. 6 we provide summary discussion and consider the ongoing relevance of s-bd.

2 Background and Terminology

The locations, dimensions, geometric forms and kinetic characteristics of entities comprising a

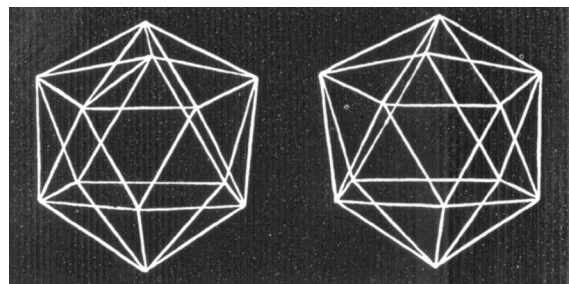


Fig. 1 An early hand-drawn stereogram. When fused, perception of three-dimensionality is derived solely from geometric disparities in the two views. However, in some instances it is possible to perceive three-dimensionality from single or identical pairs of monoscopic images—that is without recourse to the binocular parallax cue to depth

physical scene are assumed to be faithfully defined within an isotropic and homogeneous Euclidean ‘3D physical space’. When light emanating from objects in this space is cast onto a static planar surface using, for example, a suitably positioned converging lens or pinhole, a ‘2D perspective space’ image is formed. In this case, the physical size and true geometrical characteristics of physical image components are modified on the basis of a linear compression of space with distance from the imaging element. Here we coin the term ‘visually perceived space’ (vps) when referring in a general way to a person’s externalized egocentric perceptual experience of 3D physical space. In vps, conventional stereopsis is a continuously variable quantity—its role gradually diminishes with distance from the observer.

Thus major differences invariably exist between the geometrical characteristics of 3D physical space and vps. In addition, visual sensations routinely associated with the physical world and which relate to the wavelength and intensity of radiation have no physical basis and are unique to vps. In this latter respect, Johann Gottlieb Fichte (1762–1814) writes: ‘*Der erleuchtete, durchsichtige, durchgreifbare und durchdringliche Raum, das reinste Bild meines Wissens, wird nicht gesehen, sondern angeschaut, und in ihm wird mein Sehen selbst angeschaut. Das Licht ist nicht ausser mir, sondern in mir, und ich selbst bin das Licht*’² [23].

The viewing of monoscopic images under appropriate conditions can give rise to the three retinal disparity scenarios indicated in Fig. 2. Each has the capability of stimulating s-bd.

The difference in the overall luminance of the images cast onto the two retinae (here, termed ‘luminance disparity’) enables some forms of kinetic monoscopic image to be perceived as yielding s-bd. This was recognized by Carl Pulfrich in the early twentieth century and is widely referred to as the Pulfrich effect [12, 29]. It can be readily experienced by viewing appropriate image content³ with a neutral density filter placed before the dominant eye.

² ‘*Translucent penetrable space, pervious to sight and thrust, the purest image of my awareness, is not seen but intuited and in my seeing itself is intuited. The light is not without but within me, and I myself am the light*’ translation from Petic [45].

³ See example of the Pulfrich effect: https://www.youtube.com/watch?v=1mnWI_u_zBg. Accessed 20 March 2015.

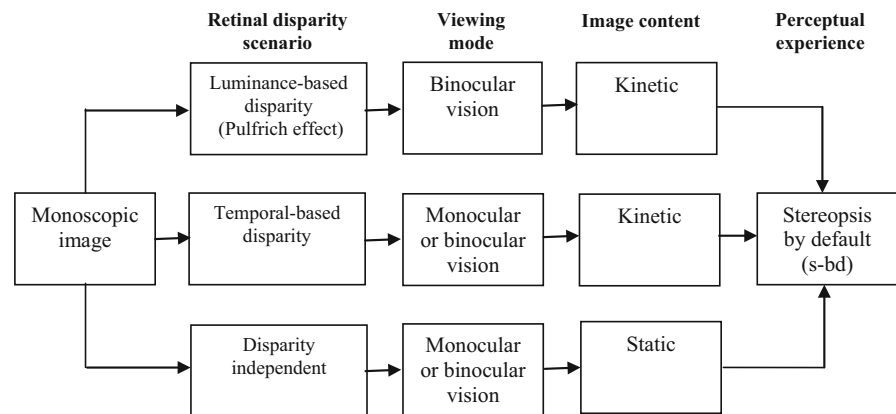
When an observer moves relative to static 3D physical space (comprising, for example, near and mid-field objects) dynamic changes in the retinal image ensue. Those directly relating to the spatial composition and form of the scene are assumed to yield ‘temporal disparity’. This embraces the motion parallax cue (relating to dynamic changes in the relative positioning of image components), changes in the geometrical form of entities and in both shadows and shading. Temporal disparities may also arise when kinetic physical world and monoscopic scenes are viewed from a fixed vantage point and when pan/zoom operations are applied to a physical or virtual camera used in the generation of a monoscopic scene.

Although dependent on temporal characteristics (which play a crucial role), viewing conditions and image content, temporal disparity provides a powerful means of facilitating s-bd. In contrast, dynamic retinal changes resulting from a shift in vantage point relative to a monoscopic image (in the absence of head tracking and associated image update) cannot convey information relating to the virtual depth and form of components comprising an image scene. Consequently, such changes are not considered to represent temporal disparity and in fact reinforce perception of a ‘2D pictorial space’ in vps. The temporal disparity mechanism can be effective for both monocular and binocular viewing.

The disparity-independent mechanism indicated in Fig. 2 relates to the fixed vantage point viewing of static monoscopic images in which s-bd is experienced without recourse to the spatial disparities associated with binocular parallax. This modality is usually associated with monocular or synoptic binocular viewing—but, when an observer has familiarity in experiencing s-bd, it may also apply to direct binocular viewing.

The structure indicated in Fig. 2 provides a useful basis for discussion. However, in practice more than one category of retinal disparity may contribute to the overall perceptual experience and consequently they cannot necessarily be considered in isolation. Furthermore, this structure focuses exclusively on retinal image characteristics—the final stage in the visual pipeline that operates on the basis of straightforward physical optics. Eaton [19] remarks: ‘*Our knowledge of retinal images tends to make us exaggerate the importance of a particular sub-station on what is merely a system of communication. The fact that the*

Fig. 2 Summary of generalized retinal disparity mechanisms that have the potential to deliver s-bd. In practice, more than one form of disparity mechanism may contribute to the perceptual experience. Here we focus on approaches that operate with both color and grayscale images. Consequently, the ‘color rivalry’ mechanism is excluded



[retinal] images are this shape or that at the place where light radiation transfers the message to nerve conduction, gives no indication ipso facto of our ultimate perception, either as to its superficial form or solidity... Ordinarily we are not conscious of retinal images however like or unlike they be. We are conscious of nothing within us, only of an object without...’ In this sense, retinal images are considered to strongly impact on, but not define, our experience of vps.

In relation to our visual impression of 3D physical space, conventional stereoscopic 3D perception can occur in the region in which the visual fields of the eyes overlap [the binocularly active region (BAR)]. This is surrounded by a binocularly inactive region (BIR). In practise when viewing objects in near field physical space, situations are regularly encountered in which all or part of an object is visible to only a single eye. For example when a finger is held in front of one eye so that a more distant object is visible only to the other eye, a sub-BIR (termed ‘occlusion constraint zone’ by Nakayama and Shimojo [42]) is formed in the BAR. Consequently, the latter cannot necessarily be regarded as a continuum but rather as a complex arrangement of BAR and BIRs. However, in most situations the visual system operates to maintain uniform 3D perception (Nakayama and Shimojo [42] use the term ‘da Vinici stereopsis’ when referring to this perceptual experience—also see Makino and Yano [39]). Introspective scrutiny by the author suggests that in some situations ‘occlusion constraint zones’ may give rise to false geometries and are regions in which negative space may perhaps be less evident. Despite this there is a seamless transition

between BAR and BIR regions such that monoscopic and stereoscopic content are integrated so as to support unified 3D in vps.

Covering one eye eliminates the BAR from the visual field. However, temporal dynamics associated with observer/object motion facilitate retention of our spatial perception of 3D physical space. In the absence of such motion, near-field spatial perception is significantly impaired but is often not eliminated. Familiar objects demonstrating clear-cut perspective tend to best retain spatial characteristics whereas the spatial form of static, near field foliage, tree branches and the like tends to be considerably flattened/eliminated. Overall near/mid field space appears compressed (for example, looking immediately downwards reveals that both the ground and feet appear closer).

The visual system’s ability to resolve depth ambiguities associated with occlusion constraint zones through to many monocular viewing scenarios in which—although the BAR is completely eliminated—a degree of spatial perception is retained, reinforces the notion that support for 3D perception represents the visual system’s most natural (preferred) mode of operation. With this in mind, here we employ the term ‘stereopsis by default’—although over the years other terminology has been adopted. For example, ‘paradoxical monocular stereopsis’—although the approach can only be deemed paradoxical if it is (incorrectly) assumed that binocular parallax is a fundamental requirement for 3D spatial perception. Furthermore, not all situations necessitate monocular viewing. Münsterberg [41] refers to the ‘plastic effect’ (also see, for example, Vishwanath [51]). In this vein,

references are also made in literature to the perception of a ‘plastic space’.

3 Viewing Techniques

Well before the advent of stereoscopic imaging, it was recognised that the spatial realism of perspective paintings could be enhanced through the adoption of certain viewing strategies. An early example is Filippo Brunelleschi’s (1377–1446) demonstration of accurate, mathematically-based perspective painting using a ‘peepshow’ technique: *‘The necessary conditions for viewing were that the spectator should peep from the back of the panel [painting] through a small hole at a mirror, in such a way that the painted surface was visible in reflection’* (quoted in Kemp [33]). Edgerton [21] writes: *‘The shrewd master may have realised something which has received attention from perceptual psychologists in recent times: that perspective illusion is strong only when the observer’s awareness of the painted picture surface is dispelled. When the viewer loses his “subsidiary awareness”..., he tends to believe the picture surface does not exist and that the illusionary space depicted is actually three-dimensional.’*

In addition, Brunelleschi’s use of a viewing hole suggests an appreciation that monocular viewing from a predefined vantage point coupled with the imposition of a restricted visual field (thereby increasing the degree of immersion) could enhance the spatial perception of images crafted in a 2D pictorial space.

In the seventeenth century, anamorphic art was used to support the formation of monoscopic images exhibiting strong spatial effects. For example, in connection with a peepshow created by van Hoogstraten (1627–1678), Pirenne [47] observes: *‘[It depicts] a Dutch interior consisting a hall with a black and white tiled pavement, opening on two furnished rooms with a view of a street and a canal. All this appears in three dimensions when viewed through the peephole. This peepshow looks very much like a real interior, extending far beyond the dimensions of the cabinet’*. This was achieved through the creation of a geometrically precise anamorphic rendition of an image scene on the inner surfaces of an enclosure—the monocular viewing position being defined by the positioning of the viewing hole.

In the late seventeenth century Fra Andrea Pozzo (1642–1709) achieved considerable success in the painting of the hemicylindrical ceiling (up to ~30 m in height) of St. Ignazio Church, Rome. Pirenne [46] describes viewing the work from the optimal vantage point (corresponding to the centre of projection (COP) which is indicated by a disc): *‘From the position marked by the yellow marble disc, the arches supported by columns at both ends of the ceiling are seen to stand upright into space. They are seen in three dimensions, with a strength of illusion similar to that given by the stereoscope.’*

The sale of optical aids intended to enhance the viewing of single monoscopic images by providing greater realism (spatial characteristics), and viewer immersion can be traced back to ~1745. Below exemplar approaches are outlined (also see Schlosberg [49], Ames Jr [2, 3], and Coe [16]).

The Zograscope (also referred to as the ‘Optique’ ([14, 17], ‘Optical Diagonal Machine’ and ‘Optical Pillar Machine’ [32]) was commercialised in ~1745 and the spatial VR experience offered by this device rapidly gained popularity. A Zograscope comprises a suitably mounted plane mirror and converging lens (Fig. 3a). A monoscopic image is reflected in the mirror and is viewed via the lens whose diameter is sufficient to enable binocular viewing. From the perspective of the eye’s accommodation and vergence, the lens functions so as to place the image at ‘infinity’.

The mirror enables images to be viewed from a physically horizontal orientation. Further, Ponzo [48] and Schlosberg [49] suggest that it contributes to the perceptual experience (although this is not in agreement with [32]). For example: *‘Le miroir et les lentilles réunis de la manière décrite donnent un effet plastique qui rappelle celui des miroirs concave, mais en l’augmentant et en l’améliorant de beaucoup, comme j’ai pu l’établir par une série d’observations comparatives’*⁴ [48].

Ames Jr [3] also outlines the role of image reflection in supporting s-bd and attributes this to the mirrors’ impact on an observer’s ability to intuitively

⁴ ‘The mirror and lenses used in the manner described above give a plastic effect which brings to mind that of the concave mirror, while greatly augmenting and improving it, as I have been able to establish by a series of comparative observations’ (translation courtesy P. J. Walker).

gauge screen/scene distance. Ponzo [48] refers to an enhanced perception of depth that can occur in the cinema environment when content cast onto the screen is subsequently viewed via its reflection in a mirror. This is supported by Schlosberg [49] and also by Higashiyama and Shimono [27] who report a controlled trial involving 22 participants who observed pictures directly and by mirror reflection. They indicate that across all perceptual judgements $\sim 73\%$ associated the latter scenario with exhibiting greater perceived depth. The authors suggest: ‘...we assumed that a mirror that reflects a picture localises it in two places...where the mirror is placed and where the virtual image is formed. We also assumed that the distance from the observer to the reflected picture is unsettled by the double locations of the picture, so that instead the effects of pictorial depth cues in the reflected picture are emphasised... If the visual system faces such a cue conflict, it may tend to take the cues for processing, that provide reliable or powerful information.’ Ames Jr [3] outlines experimental work which also suggests that ‘...it is our uncertainty as to the exact position of the picture when looking at its reflection that is responsible for the increase in the depth illusion.’

The commercial success of the Zogrscope was supported by the creation of suitable content. Topographical images were popular and employed strong perspective with buildings and streets/roads looming large in the foreground and, in depth, rapidly

diminishing in size (see, for example Fig. 3b—also see Blake [7, 8], Kaldenbach [32], Letkiewicz [38]). The Graphoscope was patented by Charles Rowse in 1864 and can usually be operated in two modes. In one scenario it is equivalent to the Zogrscope but does not include a mirror (Fig. 4).

S-bd may be stimulated by viewing a monoscopic image with a single eye from a vantage point which provides an undistorted perspective view (i.e., a perspective that conforms to natural expectations): ‘...if a photograph is viewed, at the proper distance, through a lens of about the same focal length as that employed on the camera used for obtaining the photograph, the objects will be observed undistorted and with a marked natural effect... so that a natural plastic impression [s-bd] is obtained.’ [31].

Thus monocular viewing from the COP can facilitate perception of s-bd (although initially, some adjustment time may be required). In ~ 1903 , Carl Zeiss devised the Verant viewing aid and in one form this enables single photographs ($\sim 9 \times 12$ cm) to be viewed via a small lens placed in front of one eye—the other eye being shaded. The device is adjusted in order to effectively position the observer’s eye at the original camera lens location, and was primarily intended for viewing images captured using an objective lens of short focal length (i.e. less than the near point distance). Guarini [25] indicates: ‘The result is an impression of reality that leads the observer to a correct appreciation of distances. He

Fig. 3 In **a**, a Zogrscope—comprising converging lens and plane mirror. In **b** ‘The Monument’ by Sutton Nicholls (circa 1753). Zogrscope images were often coloured, and laterally inverted text was used to accommodate image reversal when viewed via the Zogrscope

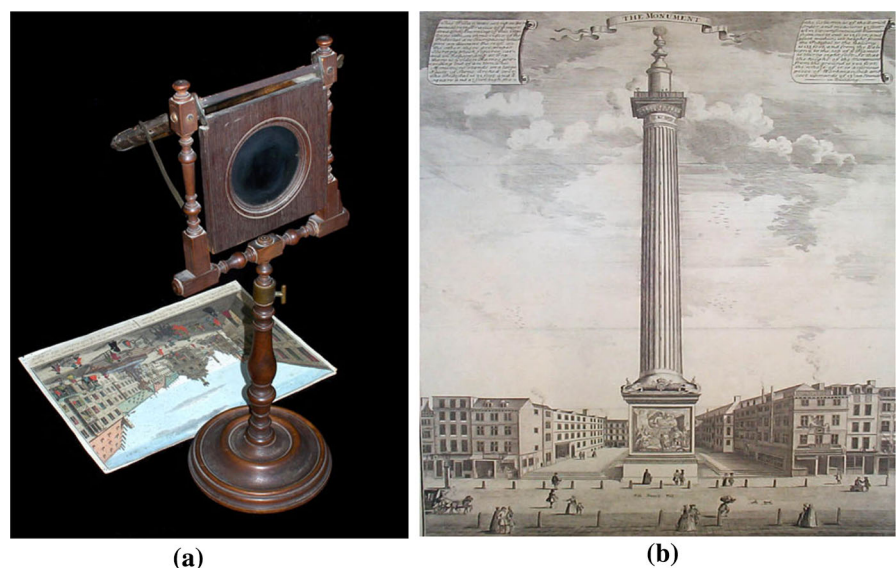




Fig. 4 A nineteenth century Graphoscope employing a bi-convex lens with a diameter of 14 cm. Two smaller viewing lenses are located immediately below the main lens (photograph courtesy of Rose Young at Auckland War Memorial Museum, New Zealand)

sees the photograph with its natural relief.’ For related discussion also see Münsterberg [41].

Although the name Perspectoscope has been applied to several unrelated devices, its original use related to a pinhole camera operated in conjunction with an inclined photographic plate. Once developed, the anamorphic image is viewed through the pinhole in such a way that the original geometry is restored [31]. A second form of Perspectoscope employs two mirrors and lenses to deliver identical views of a monoscopic image to the two eyes [30]. The device supports immersive viewing—see Fig. 5.

The theme of presenting ‘identical’ images to the two eyes underpins the operation of the synopter (see, for example Black [6], Koenderink et al. [36]). This is achieved using a beam splitter and appropriately positioned reflective elements (Fig. 6). Several exemplar embodiments are described in Zeiss [56] Also see the Myopter and Cyclopter (see www.myopter.org and www.phantascope.co.uk, respectively). The Brewster form of stereoscope [13] provides a simple means of synoptically viewing identical monoscopic image pairs.

Over the years various other viewing strategies have been recommended. These include the

minimization of frame visibility,⁵ the use of a curved mirror for image reflection which underpinned the operation of the Reflectoscope as marketed in ~1935 [31], binocular viewing with a cylindrical lens held before one eye [2]⁶ and the counter rotation of the retinal images. In the case of this latter scenario, Ames Jr [3] describes the use of two sets of mirrors to counter rotate the retinal images in the two eyes. The degree of rotation is limited to maintain image fusion. In the context of monoscopic viewing, he writes: ‘If an image is viewed through this instrument and the images coming to the two eyes are rotated in opposite directions to almost the limit of fusion the depth illusion is enhanced... if natural objects are looked at through the mirrors set at the rotations which give the best depth with a picture the perspective is decreased and the objects appear flattened.’ The synopter demonstrates similar characteristics—enhanced 3D perception of 2D images and flattened 3D perception when used to observe physical 3D space.

4 Stereopsis by Default: The Perceptual Experience

Binocular observation of the animated monoscopic content indicated in example of the Pulfrich effect (see footnote 3) (using a filter from a pair of sunglasses held before the dominant eye) provides a vivid demonstration of s-bd based on luminance disparity (the Pulfrich effect). Kinetic objects appear to move within three spatial dimensions and a strong sense of emergence/regression is evident.

The ability of both the temporal and disparity independent mechanisms to yield s-bd can be readily experienced using a ‘tunnel’ viewing approach or magnifying glass. Tunnel viewing is particularly effective for kinetic scenes (e.g. animations, television and cinema) but can also work with many forms of static image. In this latter respect high quality colour

⁵ Informal observation by the author suggests that framing elements do not *necessarily* impair perception of negative space in the monoscopic 3D experience, as demonstrated by mullion visibility in Video Wall displays.

⁶ For viewing distances of less than a metre, a 2.5D cylindrical convex lens is recommended by Ames [2]. Informal trials by the author have failed to confirm Ames’ observations in respect of the efficacy of this technique.



He Sees No Photograph
but the scene exactly as it appears
to the one who pressed the button.

The Perspectoscope

quadruples the pleasure of photography and changes any collection of camera pictures into an assortment of living scenes. Your best pictures will seem, without the Perspectoscope, flat and uninteresting after once seen through this instrument.

Accommodates any size of single pictures.

Price, morocco finish, \$2.00 prepaid. For circular and name of nearest dealer, address

THE PARLIAMENT PUBLISHING CO.

Fig. 5 The Perspectoscope appeared in a number of forms. The embodiment illustrated here closely resembles a device patented by Hill [28] (UK Patent 15,299)—although in the patent it is confusingly referred to as a Graphoscope

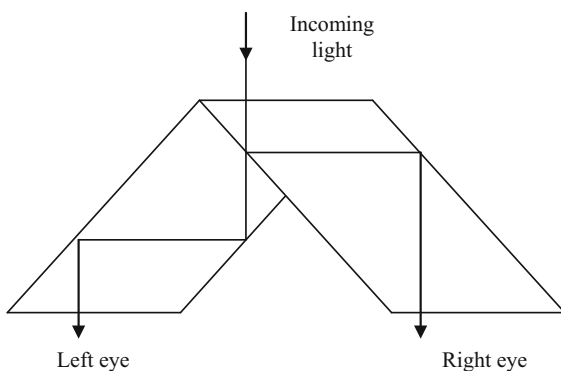


Fig. 6 Exemplar implementation of a synopter constructed using prisms. Adapted from Zeiss [56]

images such as those reproduced in National Geographic form a good starting point—although as Enright [22] remarks ‘...one of the most interesting aspects of paradoxical monocular stereopsis [s-bd] is how difficult it is to predict whether a given illustration

will be effective in evoking the response.’ The simplest viewing method is to close one eye and view through a ‘hole in the fist’. Alternatively viewing through either a tapering tube (the taper, coupled with the use of dark matt paper, reduces internally reflected light) or circular/oval aperture ($\sim 2\text{--}3$ mm in diameter punched in a piece of card) can give satisfactory results. In each case, the ‘tunnel’ restricts the visual field such that it is possible to ensure that vps essentially comprises only the image scene.

Vishwanath and Hibbard [54] and Vishwanath [52] report a trial carried out with 23 stereo-normal participants who viewed monoscopic near field images both binocularly and via a monocular aperture. Of these, 20 reported that the latter approach provided a better qualitative impression of stereopsis: ‘...a clearer impression of depth separation, a sense of realness, a sense of protrusion towards the viewer, the feeling that things appeared touchable...’. A further trial solicited responses from two groups to obtain feedback comparing qualitative binocular viewing of monoscopic images with either stereo viewing (via filtered glasses) of anaglyph images or monocular aperture viewing of monoscopic images. In both cases similar results are reported.

It has long been recognized that monocular viewing of both static and kinetic monoscopic images via a converging lens provides a simple and often effective means of experiencing s-bd (see, for example, Fig. 7): ‘It is a well known fact that a large magnifying glass is itself an important aid to the perception of the third dimension in a photograph’ [30]. This viewing technique is frequently used in pavement (anamorphic) art.

When viewing computer screen output, a Fresnel lens (of the type found in discarded OHP machines) is generally effective. The lens acts as a window onto the image scene and confounds judgement of image distance via the oculomotor cues. Additionally, it provides a degree of immersion—see, for example, the early portion of the animated monoscopic content indicated in example of s-bd/immersion⁷.

Techniques underpinned by spatial disparity provide the most widely usable, vivid and flexible method of supporting perception of 3D space—with, for

⁷ See example of s-bd/immersion (the early part of this content is recommended): <https://www.youtube.com/watch?v=9SphgMxwD5M>. Accessed 20 March 2015.

example, stereoscopic imaging supporting the controlled incorporation of negative space, emergence and recession. When viewed in an appropriate way, some forms of kinetic monoscopic image can give rise to a qualitative perceptual experience exhibiting holistic spatial effects generally associated with stereoscopic techniques—although there are key differences. While negative space is usually clearly evident, the degree of emergence and recession cannot be readily controlled. Further, quantitative egocentric judgement of distance appears to be somewhat impaired.

In general terms, disparity independent techniques provide the ‘gentlest’ support for spatial perception. For example, as indicated previously, switching from binocular to monocular viewing (and in the absence of motion and/or a visible ground plane (or equivalent)) results in an apparent compression of near/mid field physical 3D space. Conversely, under appropriate conditions monocular viewing of 2D content can result in a perceived expansion of depth, although in comparison to stereo imaging, depth effects are usually mellower. Thus in both cases binocular vision provides the stronger perceptual experience of the third dimension—although the profundity of s-bd appears to increase with familiarity. In comparing the effectiveness of the spatial, temporal and disparity independent approaches, Enright [22] observes: ‘...there can be little doubt that the most effective representations of the third dimension are those which involve [conventional] stereopsis; and that the second most effective way to convey a feeling of depth is through the use of image motion: optical flow patterns, image shear, motion parallax and the like. When both [conventional] stereopsis and image motion are excluded, [disparity independent] one is dealing with no more than third best.’

5 Underlying Mechanisms and Controlled Studies

S-bd is most frequently explained on the basis of enhanced coherence between depth cues through the reduction/elimination of cues (designated ‘counter-cues’ by Adams [1]) which suggest to the visual system that the image under observation actually exists in a 2D pictorial space (the ‘cue-coherence’ theory). Additionally the need to minimise physical attributes that enhance image surface awareness

(‘surface awareness’ theory) is considered to be a key ingredient and embraces the elimination of frame visibility, avoidance of specular reflections, minimisation of visually obtrusive surface textures (in, for example, painting) and maximisation of image quality. Schlosberg [49] writes: ‘*In normal binocular inspection of a picture the ‘flatness’ cues are strong enough to force the observer to see a flat picture; but if ‘flatness’ cues can be eliminated or weakened, or if the depth cues that are present can be sufficiently exaggerated, the perception takes on depth.*’

Consider the simple scenario in which a well-focused static planar image is viewed via one eye through a converging lens. Monocular viewing ensures that binocular parallax and *natural* vergence cues are suppressed. The lens serves to reduce any contribution that accommodation may make to judgement of depth, may amplify the perspective framework present in the 2D image scene, and will generally make extraneous physical space objects less noticeable. In the case that a Fresnel lens is employed, and putting to one side evident chromatic aberrations, the lens tends to convey an impression equivalent to the presence of a physical window through which a 3D scene is being observed. This serves to reduce the observer’s subsidiary awareness of the image surface.

As with the tunnel viewing technique, the ensuing visual impression is then based on pictorial cues associated with a single retinal image, image dynamics, secondary vergence effects, prior experience, familiarity and expectations. Under these conditions (and assuming that head movement is minimized), given appropriate image content and composition, the visual system will strive to provide a spatial interpretation of the monoscopic image scene.

Koenderink et al. [37] report on a controlled study intended to verify Zograscopic viewing. The optical arrangement comprised a 12 cm diameter lens used for iPad viewing. The mirror associated with the Zograscope was omitted and so the configuration more closely resembled that of the Graphoscope. Comparisons were made of three sample images viewed with the lens in situ and absent (unfortunately the use of original forms of Zograscope image is not reported). The study employed the use of a gauge figure method [35] to enable four researchers to judge surface gradient at numerous points on entities comprising the image scene. This enabled the formulation of a perceived depth map and allowed estimated pictorial



Fig. 7 Monocular viewing of this grayscale image via, for example, a magnifying glass or the ‘hole in the fist’ provides a simple example of s-bd as derived from a static monoscopic image. Here, the use of a grayscale image ensures that spatial effects are not influenced by color rivalry. The strong perspective framework inherent in this image parallels that used in many eighteenth century Zograscopes images. In the case

relief to be compared between the two viewing modes. Reported results suggest that viewing via the lens significantly enhanced the perceived depth of pictorial relief—although effectiveness varied between participants and with image content.

In extensive publications, Vishwanath and Hibbard [54] and Vishwanath [52] question the conventional explanations of s-bd and describe an alternative hypothesis. In essence, this is based on depth scaling—specifically that the impression of tangible separation in depth is linked to the precision with which absolute (egocentric) depth can be determined: ‘...I propose that 3D structure can be perceived solely on the basis of relative depth estimates but that the impression of stereopsis is induced only when absolute depth values can be estimated’ [52]. Thus in the case that a static monoscopic image is observed using one eye—via a small viewing hole (so as to eliminate frame visibility) Vishwanath suggests: ‘In the absence of a visible picture surface distance information deriving from residual cues such as accommodation...or default perceptual tendencies is postulated to

of static images, introspective observation by the author suggests that s-bd is usually associated with specific image entities (see, for example the overhang of the thatched roof) and less frequently with the image as a whole. This is confirmed by Schlosberg [49] ‘...we may get the plastic effect [s-bd], but find the depth more adequate in certain parts of the view than others’

be assigned to the pictorial object allowing a derivation of size and absolute depth values.’

In the context of the depth scaling hypothesis, Vishwanath reports a number of controlled studies. One of these focuses on the qualitative perception of depth during the observation of near field monoscopic images under the four conditions: direct monocular viewing, monocular viewing via an aperture, direct binocular viewing and binocular viewing via two apertures. The two monocular viewing scenarios correspond to maximum cue coherence while those involving binocular viewing exhibit minimum cue coherence. Thus on the basis of the cue coherence theory, comparisons of monocular and binocular viewing should yield the greatest change in 3D perception while the monocular versus monocular aperture and binocular versus binocular aperture should yield the smallest changes (since, in principle, there is no change in cue coherence). However as summarized in Fig. 8, the largest reported changes in 3D perception corresponded to the monocular and monocular aperture scenarios and the smallest

changes to monocular versus binocular viewing. Although these results suggest that further verification of the cue coherence theory is required, the elimination of frame visibility in moving from monocular to monocular aperture viewing must also be considered. This may impact on surface awareness and so influence an observer's perception of s-bd.

Eby and Braunstein [20] discuss controlled studies undertaken in connection with the impact of frame visibility on depth and shape perception—although this work relates to viewing physical objects located in 3D space. In summary they write: ‘*Collectively these observations suggest that the presence of a frame, even one surrounding a 3D scene, can reduce perceived depth, while the elimination of a frame, even one surrounding a 2D scene, can increase perceived depth. Both effects are consistent with a hypothesis that a frame provides flatness information to an observer.*’ This leads to the notion of the frame acting a ‘flatness cue’—such that it reduces overall cue-coherence. Frame impact can be significant—for example, in one study Eby and Braunstein report the presence of a visible frame resulting in a 20 % reduction in perceived depth.

Informal observation by the author indicates that the perception of s-bd appears to be strongly facilitated by experience. A sound starting point in building such experience is the observation of appropriate kinetic monoscopic image content using, for example, the ‘hole in the fist’ to ensure that the visual field is fully occupied by the image scene. Over time it becomes easier to identify the type of content that will yield a strong spatial experience and frequently once s-bd is

induced it is possible to relax viewing constraints by, for example, returning to direct monocular, or on occasion even binocular, observation. For example, in relation to the observation of monoscopic content via its reflection in a plane mirror Higashiyama and Shimono [27] observe: ‘*...the plastic effect [s-bd] that had been formed by observing reflected pictures transferred to directly observed pictures and persisted after the reflected pictures had been removed.*’

Additionally, a memory effect may apply. For example, monocular viewing of an oil painting using the ‘hole in the fist’ approach over a period of several minutes provided the author with a strong experience of s-bd in which, for example, the branches of a tree (a major component in the work) appeared to emerge from the canvas in a natural manner. Subsequently this effect could be effortlessly evoked by direct monocular viewing and eventually by natural binocular viewing.

This confirms that as experience builds, image frame visibility is often no longer problematic and in this general context Vishwanath [52] suggests: ‘*The induction of stereopsis [s-bd] appears to be linked to a reduction in surface visibility. However, stereopsis [s-bd] appears to be induced even when the picture surface is visible.*’

Increasing familiarity in experiencing s-bd appears to facilitate its initiation when exposed to new content. In some cases this can enable immediate perception—but often an adaptation time is required. For example, in a controlled study requiring 22 participants to compare directly viewed kinetic content with equivalent content reflected in a mirror, Higashiyama and

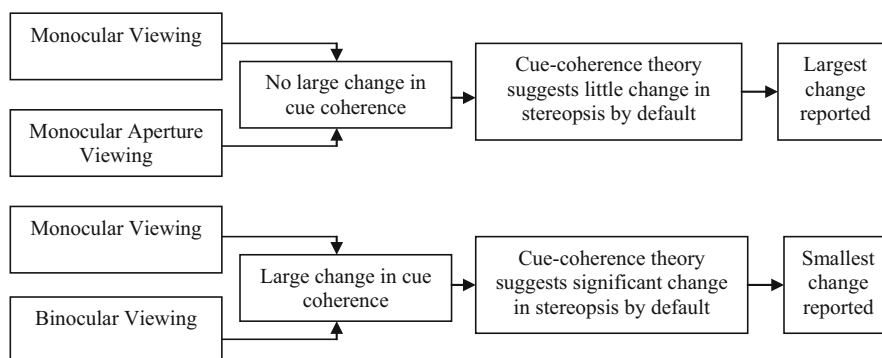


Fig. 8 Vishwanath [52] reports a number of controlled studies including one in which participants reported on their perception of depth when viewing single monoscopic images in four ways:

direct monocular viewing, monocular viewing via an aperture, direct binocular viewing and binocular viewing via apertures

Shimono [27] report that at the outset 55 % judged the latter to exhibit greater depth. After 5 minutes this increased to 86 %. However even for those who are familiar with experiencing s-bd, content that works well for one observer does not necessarily work well for others. For example, Ames Jr [2] discusses the monocular viewing of a number of paintings using a cylindrical lens. In relation to the work ‘Rain, Steam and Speed’ (which does not employ a strong perspective framework) he writes: ‘Turner...has produced a depth of atmosphere which, in my limited search, is unapproached. If you look with the lens at the atmosphere over the train it seems to have actual thickness. I can get this same effect after I have ceased to use the lens.’ However, for the author, neither this viewing method nor image stimulates s-bd.

In general terms, when presented with a monoscopic image scene (and given an appropriate period of adaptation), s-bd tends to be clearly evident or otherwise—some images remaining firmly locked within the confines of 2D space while others vividly span three spatial dimensions. This is confirmed by several researchers. For example: ‘We do not have a simple addition or subtraction of factors, with more or less depth resulting. The perception seems to exist in two modes. In one it is still a picture. In the other mode we find objects in depth... It is true that there is some variation around the modes... The plastic effect [s-bd], however, seems usually to be either clearly present or absent’ [49].

Although monocular viewing from the COP can facilitate the induction of s-bd, this is not a requirement and is a condition that can be relaxed with experience. In fact, the benefit to be derived from this viewing location may link to image surface awareness. Pirenne [46] discusses the notion of psychological compensation (also referred to as ‘La Gournerie’s paradox’ by Cutting [18]) in relation to viewing geometric entities in a monoscopic image from a non-COP location: ‘When the shape and position of the picture surface can be seen, an unconscious intuitive process of psychological compensation takes place, which restores the correct view when the picture is looked at from the wrong position.’

Conversely, when characteristics of the image surface are not perceived, Pirenne considers the absence of psychological compensation. For example, in the case of the painting by Pozzo in the church of St. Ignazio, Pirenne observes: ‘...if the spectator walks

away from the yellow disc [denoting the COP]... the illusion of depth does remain, but the scene represented, still seen in 3D becomes deformed... This deformation continually varies as one walks about... the light flux now received by his eyes becomes very nearly the same for each new position as the light flux which his eyes would receive if the deformed scene he sees from this position were actually there in three dimensions.’

However, in this example the disc (denoting the intended viewpoint) represents a coign of vantage in as much as both the physical architecture and rendered image are in geometric harmony. This is an important consideration since the image does not fully occupy the visual field and so geometric disharmony is likely to exacerbate perceived image distortion. More recent research relating to the notion of psychological compensation confirms that image surface visibility appears to play a key role in supporting perceptual geometric invariance when images are viewed from non-COP locations—see, for example Vishwanath et al. [53].

Even when one eye is shaded, appropriate image content can evoke vergence movements and this may be a further mechanism that underpins s-bd. This was loosely suggested by Claparède [15], and Enright [22] appears to have been the first to undertake a controlled study in connection with this notion. Tidbury et al. [50] suggest: ‘In a flat image, a cloud will elicit a different vergence response compared with a near object. A cloud is known to be more distant, so a compelling presentation will induce a ‘vergence memory’ response, to diverge gaze to the distance. ... When the surface of an image is not discernible by resolution or context, the visual system appears to process the information as more real than a flat 2D image.’ In this scenario, vergence movements are fundamentally driven by prior experience/familiarity.

When observing physical 3D space, the visual system endeavors to bring the subject of fixation into sharp focus. The depth of focus provides a measure of the range of distances over which other entities within the visual scene are also perceived as being satisfactorily focused and increases with fixation distance. Vishwanath [52] discusses the incorporation of depth of focus blur in static monoscopic images and describes a controlled trial in which participants gauged the impression of depth derived from pairs of images which differed only in terms of sharpness about the region of interest. A significant number of

participants reported a better impression of depth for images containing depth of focus blur—although the blur was deliberately created in such a way as to be inconsistent with the depth gradient. As Vishwanath remarks: ‘*Based on standard cue-combination regimes [cue coherence], the inconsistent blur should have resulted in a decrease, not an increase, in perceived depth.*’

6 Discussion

Research to date clearly demonstrates that it is an oversimplification to consider conventional stereopsis as mediating the de facto qualitative experience of spatial perception. Equally, the fact that s-bd does not conform to the standard model of stereoscopic vision dating back to the first half of the nineteenth century does not necessarily form a basis for its dismissal as a curious ‘illusionary’ experience. In fact as discussed, stereopsis in general fails to provide an externalized perceptual experience that truly reflects the nature of 3D physical space, and visual illusions abound. Certainly in the case of s-bd, spatial perception is inferred from monoscopic content but, due to the presence of occlusion constraint zones, such inference (albeit to a much more limited extent) is also associated with conventional stereopsis.

Stereopsis is not a single well-defined perceptual experience. Even in the case of monocular observation, different image presentation scenarios couple with experience and familiarity to elicit different perceptual responses. Consequently in the case of 3D display technologies, there is a need to more closely assess image presentation methodologies with particular reference to their ability to effectively capitalize on the visual sense and to the attributes of the perceptual experience they are intended to evoke.

As noted previously, investigations into spatial perception based on s-bd have largely been overshadowed by more widespread interest in binocular parallax based approaches and with the advent of the stereoscope, simple geometrical constructs have generally provided a reassuring means of explaining stereopsis and are reproduced in most relevant texts. Unfortunately, these are superficial as they only account for the relatively simple matter of retinal image formation. The real challenges begin when the precise nature of retinal disparities and corresponding

points are questioned and crucially when consideration is given to the complex mechanisms that operate beyond the retina. In this general context, Vishwanath [52] writes: ‘*...it is clear that current models are inadequate for attaining a full understanding of the complex nature of the human experience of the third dimension.*’ Koenderink [34] observes: ‘*It [s-bd] is evidently due to complicated (largely automatic and subconscious) operations on the image structure and cannot be explained away as a simple application of elementary geometry as with binocular [conventional] stereopsis. In the latter case the actual image structure is immaterial as long as there is any as is evident from the success of random image structure.*’

Although controlled trials continue to play a key role in developing understanding of s-bd, it is important to recognize that they do not generally reflect our everyday perceptual experiences. In this context it is useful to consider the sense of sight from the perspectives of subconsciously and consciously driven experiences. The former represents holistic visual perception which, as we navigate through our physical surroundings, is ‘effortlessly’ derived from the seamless and generally synergistic mediation of multiple stimuli. This invariably yields a unified experience in which physical 3D space is perceived as a spatial and temporal continuum such that familiar in-context objects appear to be experienced holistically: ‘*...it is with strict accuracy that we say we see objects and not the light of objects...*’ Eaton [19].

In contrast the *conscious* analysis of the visual impression is experienced as we pause and critically appraise a view or seek to understand its characteristics. In these circumstances natural holistic flow is often disrupted. Controlled studies are clearly more likely to represent a consciously rather than subconsciously driven experience. Additionally, our current level of understanding of s-bd is such that results from controlled studies may be confounded by a potentially complex interplay of variables relating to image content and composition, method of presentation, viewing technique and human factors (including enhanced perceptual malleability as personal viewing experience increases).

Since differences exist in the mechanisms that underpin conventional stereopsis and s-bd, it is evident that they are unlikely to evoke identical experiences. For example, the former is able to support both qualitative and quantitative perception whereas the

latter predominantly links to the qualitative domain. Further, the stereographer is supported by the geometrical constructs which underpin binocular vision and so is able to control the perceived distance of image components from the stereo plane. In contrast, when dealing with both static and kinetic monoscopic images, s-bd does not lend itself to known rules defining perceived depth.

The cue coherence theory provides a straightforward model by which some, but not all, aspects of s-bd can be explained. For example, consider a person visually impaired from birth such that fused binocular vision is not possible. According to the cue coherence theory, this should reduce the degree of cue conflict and so facilitate s-bd based perception when viewing monoscopic images. However this appears not to be the case. For example, Barry [4, 5] was born with strabismus and vividly describes personal experiences of binocular spatial perception when, in adulthood, the condition was corrected: *'I got the powerful sense of being immersed in a three-dimensional world that I never experienced before. ... Before, if I looked at a snowfall, the snowflakes appeared a little distance away from me in one flat plane. When I gained stereovision, I felt myself within the snowfall: the snow was falling all around me in many levels of depth'* [5]. Such a marked contrast does not suggest that 3D perception based on s-bd was experienced prior to the correction of strabismus. One explanation (albeit of a speculative nature), is that binocular parallax based vision provides crucial experience in facilitating the formulation of our everyday spatial impression of physical 3D space. Without such familiarity the visual system may well be less adept at mediating the various data sources from which the visual impression is derived. In this scenario, stereopsis based on binocular vision would play a vital role in the development of the visual skills needed to first experience s-bd.

Poorly crafted and/or presented stereoscopic content has the potential to cause visual strain and also may give rise to the *coulisse* effect. In contrast, observations in relevant research literature confirm those of the author in suggesting that s-bd does not appear exhibit either of these effects. The visual 'solutions' associated with s-bd are invariably plausible in both form and geometry.

Continued rigorous research into s-bd has the potential to greatly enhance understanding of the mechanisms supporting our perception of 3D. Current

immersive virtual reality technologies present numerous opportunities for the study of the approaches summarized in Fig. 2. Ultimately, in the case of kinetic image content this could lead to new image depiction opportunities and may facilitate the seamless presentation of 2D and 3D content. Research to date suggests that display attributes can strongly influence the initiation/perception of s-bd and in this context it is possible that large, ultra high definition, screen formats (including curved screen profiles) may play an important role. In relation to display resolution, Masaoka et al. [40] describe a controlled study in which participants reported on the sensation of 'realness' when presented with electronic images (exhibiting angular pixel resolutions in the range 19–155 cycles per degree) as compared to viewing corresponding physical objects. Monoscopic viewing of both forms of content was achieved via a synopter. The study indicates a steady rise in perceived realness up to ~60 cycles per degree after which the increase becomes more gradual. Some participants reported on differences in perceived three-dimensionality (although this may perhaps have been influenced by the use of the synopter).

Display design features should reduce an observer's awareness of the presence of the screen surface and minimize cues suggesting that it coexists in the same plane as the frame/surround. This latter consideration was adopted in the implementation of Russian second generation glasses-free cinema which opened in 1947 [11] in as much as the screen was set back by a significant distance from surrounding surfaces.

Despite the passage of ~600 years since Filippo Brunelleschi employed image presentation techniques which facilitate s-bd, relatively little progress has been made in understanding the fundamental principles which enable some forms of monoscopic image to exhibit qualitative spatial characteristics akin to, although usually mellower than, those routinely ascribed to binocular-based stereopsis. Without doubt many interesting questions await further research.

References

1. Adams, K. R. (1972). Perspective and the viewpoint. *Leonardo*, 5(3), 209–217.
2. Ames Jr., A. (1925). The illusion of depth from single pictures. *J Opt Soc Am*, 10(2), 137–147.

3. Ames Jr., A. (1925a). Depth in pictorial art. *The Art Bulletin*, 8(1), 4–24.
4. Barry, S. R. (2009). *Fixing my gaze*. New York: Basic Books.
5. Barry, S. R. (2009b). An interview with 'stereo Sue'. *Review of Optometry*, 48–49.
6. Black, R. (2010). An optical device. UK Patent Application GB 2478358.
7. Blake, E. C. (2002). Topographical prints through the Zogrscope. *Imago Mundi*, 54, 120–124.
8. Blake, E. C. (2003). Zogrscopes, virtual reality, and the mapping of polite society in eighteenth-century England. In L. Gitelman & G. B. Pingree (Eds.), *New media 1740–1915*. Cambridge, MA: Massachusetts Institute of Technology.
9. Blundell, B. G. (2008). *An introduction to computer graphics and creative 3-D environments*. London: Springer.
10. Blundell, B. G. (2011). *3D displays and spatial interaction. Exploring the science, art, evolution and use of 3D technologies*. Walker and Wood Ltd. Retrieved January 12, 2015, from <https://www.dropbox.com/sh/r96nhin3od4z81t/hepxE7-jul>.
11. Blundell, B. G. (2014). On aspects of glasses-free 3D cinema ~70 years ago. Retrieved January 12, 2015, from <https://www.dropbox.com/sh/uw3yjnwn3qo9z1e/SyPYAkHSqc>.
12. Blundell, B. G., & Schwarz, A. J. (2006). *Creative 3-D display and interaction interfaces: A trans-disciplinary approach*. Hoboken, NJ: Wiley.
13. Brewster, D. (1856). *The stereoscope its history, theory and construction*. London: John Murray.
14. Chaldecott, J. A. (1953). The Zogrscope or optical diagonal machine. *Annals of Science*, 9(4), 315–322.
15. Claparède, E. (1904). Steréoscopie monoculaire paradoxale. *Annales d'Oculistique*, 132, 465–466.
16. Coe, B. (1981). *The history of movie photography*. London: Ash and Grant.
17. Court, T. H., & von Rohr, M. (1935). On old instruments both for the accurate drawing and the correct viewing of perspectives. *The Photographic Journal*, 75, 54–66.
18. Cutting, J. E. (1987). Rigidity in cinema seen from the front row and side aisle. *Journal of Experimental Psychology*, 13(3), 323–334.
19. Eaton, E. M. (1919). The visual perception of solid form. *The British Journal of Ophthalmology*, 3, 349–408.
20. Eby, D. W., & Braunstein, M. L. (1995). The perceptual flattening of three-dimensional scenes enclosed by a frame. *Perception*, 24, 981–993.
21. Edgerton, S. Y. (1976). *The Renaissance rediscovery of linear perspective*. New York: Harper and Row.
22. Enright, J. T. (1993). Paradoxical monocular stereopsis and perspective vergence. In S. Ellis (Ed.), *Pictorial communication in virtual and real environments* (2nd ed., pp. 567–576). Bristol: Taylor and Francis, Inc.
23. Fichte, J. G. (1997). *Die Bestimmung des Menschen*. Stuttgart: Reclam, Philipp Jun. GmbH.
24. Gabor, D. (1960). Three-dimensional cinema. *New Scientist*, 8(191), 141–145.
25. Guarini, E. (1904). The Verant: A new device for viewing photographs. *Scientific American*, 91(19), 312.
26. Hibbard, P. (2008). Can appearance be so deceptive? Representationalism and binocular vision. *Spatial Vision*, 21(6), 549–559.
27. Higashiyama, A., & Shimono, K. (2012). Apparent depth of pictures reflected by a mirror: The plastic effect. *Attention, Perception, and Psychophysics*, 74, 1522–1532.
28. Hill, H. H. (1898). Improved Graphoscope for obtaining stereoscopic effects. UK Patent Number 15,299.
29. Howard, I. P. (2002). *Seeing in depth, Volume I basic mechanisms*. Thornhill, ON: I Porteous.
30. Jastrow, J. (1898). Some aids to the study of stereoscopic vision. *Science, New Series*, 7(175), 615–622.
31. Judge, A. W. (1935). *Stereoscopic photography, its application to science, industry and education*. London: Chapman and Hall Ltd.
32. Kaldenbach, C. J. (1985). Perspective views. *Print Quarterly*, 2(2), 87–104.
33. Kemp, M. (1978). Science, non-science and nonsense: The interpretation of Brunelleschi's perspective. *Art History*, 1(2), 134–161.
34. Koenderink, J. J. (1998). Pictorial relief. *Philosophical Transactions of the Royal Society of London*, 356(1740), 1071–1086.
35. Koenderink, J. J., & van Doorn, A. J. (1995). Relief: Pictorial and otherwise. *Image and Vision Computing*, 13, 321–334.
36. Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1994). On so-called paradoxical monocular stereopsis. *Perception*, 23, 583–594.
37. Koenderink, J. J., Wijntjes, M., & Van Doorn, A. (2013). Zogrscopic viewing. *i-Perception*, 4, 192–206.
38. Letkiewicz, M. (2013). Cudowna maszyna optyczna zogrscope. *Annales Universitatis Mariae Curie-Sklodowska Lublin-Polonia*, 11(1), 27–43.
39. Makino, Y., & Yano, M. (2006). Pictorial cues constrain depth in Da Vinci stereopsis. *Vision Research*, 46, 91–105.
40. Masaoka, K., Nishida, Y., Sugawara, M., Nakasu, E., & Nojiri, Y. (2013). Sensation of realness from high-resolution images of real objects. *IEEE Transactions on Broadcasting*, 59(1), 72–83.
41. Münsterberg, H. (1904). Perception of distance. *Journal of Philosophy, Psychology and Scientific Methods*, 1(23), 617–623.
42. Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, 30, 1811–1825.
43. Ohmori, K., Sakamoto, K., Nomura, S., Hiroto, T., Shiwaku, K., & Hirakawa, S. (2010). Monocular 3D vision with correct depth using Polypyrrole film actuator for lightweight display unit. In *Proceedings of IMECS* (Vol. II), Hong Kong, 17–19 March 2010.
44. Patterson, R. E. (2011). Binocular vision and depth perception. In W. Cranton, M. Fihn, & J. Chen (Eds.), *Handbook of visual display technology* (pp. 121–127). Berlin: Canopus Academic Publishing and Springer GmbH.
45. Petic, P. (Ed.). (2009). *Hermann Weyl, mind and nature: Selected writings on philosophy, mathematics and physics*. Princeton, NJ: Princeton University Press.
46. Pirenne, M. H. (1970). *Optics, painting and photography*. Cambridge, MA: Cambridge University Press.
47. Pirenne, M. H. (1975). Vision and art. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception, Vol 5, seeing* (pp. 433–490). New York: Academic.

48. Ponzo, M. (1911). Un appareil pour la vision plastique de photographies. *Archives Italiennes de Biologie*, 56, 125–126.
49. Schlosberg, H. (1941). Stereoscopic depth from single pictures. *The American Journal of Psychology*, 54(4), 601–605.
50. Tidbury, L. P., Black, R. H., & O'Connor, A. R. (2014). Perceiving 3D in the absence of measurable stereo-acuity. *The British and Irish Orthoptic Journal*, 11, 34–38.
51. Vishwanath, D. (2011). Visual information in surface and depth perception: Reconciling pictures and reality. In L. Albertazzi, G. van Tonder, & D. Vishwanath (Eds.), *Perception beyond inference: The information content of visual processes* (pp. 201–240). Cambridge, MA: MIT Press.
52. Vishwanath, D. (2014). Towards a new theory of stereopsis. *Psychological Review*, 121(2), 151–178.
53. Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, 8(10), 1401–1410.
54. Vishwanath, D., & Hibbard, P. B. (2013). Seeing in 3D with just one eye: Stereopsis without binocular vision. *Psychological Science*, 24, 1673–1685.
55. Wheatstone, C. (1838). Contributions to the theory of vision—Part the first, on some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society (London)*, 128, 371–394.
56. Zeiss, C. (1903). Linsensystem zum einaugigen Betrachten in der Brennebene befindlichen Photographie. German Patent Number 151312.