Journal of Mathematical Psychology 88 (2019) 15–26

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Characterizing projective geometry of binocular visual space by Möbius transformation

Jun Zhang

Department of Psychology and Department of Mathematics, University of Michigan, 530 Church Street, Ann, Arbor 48109-1109, USA

HIGHLIGHTS

- A new geometric framework on binocular vision is proposed based on Möbius transformation group.
- Four-point cross-ratio as an invariant of Möbius transformation captures the invariant relative disparity as the eyes change fixation.
- Hyperbolic tangent function as the psychophysical function arises naturally from this formulation.
- We conjecture that, along with the relative disparity, the ratio of the distances of the object point to the two eyes (i.e., relative vertical magnification or vertical size ratio), plays an important role in depth perception.

ARTICLE INFO

Article history:
Received 28 September 2012
Received in revised form 4 June 2018

ABSTRACT

Binocular vision involves the projection of objects in the 3-D visual space onto the two retinae and the comparison of spatial layout of objects in these retinal half-images. Here we characterize the unitary representation of the binocular space as a complex half-plane from the perspective of the cyclopean eye. We then investigate its automorphism group, namely the Möbius transformation group, and the associated invariants when the two eye positions are treated as fixed points of the automorphism. A three-point simple ratio from an object point to both eyes is constructed; as a complex number, its angle measures the difference in azimuth of the projected rays from the object point to each eye (i.e., horizontal disparity) while its modulus measures the ratio of the distances of the object point to the two eyes (i.e., relative vertical magnification or vertical size ratio). The four-point cross-ratio of two such simple ratios, as the only four-point invariant under Möbius transforms, reflects the fact that the relative disparity between any two object points remains unchanged when the eyes change fixation. Since the complex half-plane is biholomorphically equivalent to an open unit disk, both Poincaré model and Klein–Beltrami model give rise to a hyperbolic geometry, consistent with the empirically supported Luneburg's (1947; 1950) model of binocular geometry (for the depth plane). Finally, the hyperbolic tangent function $f(z) = \tanh(z)$, with inverse $f^{-1}(z) = \frac{1}{2}\log\frac{1+z}{1-z}$, is shown to act as the psychophysical function relating the physical representation of the binocular space to its cyclopean representation.

1. Introduction

Binocular vision involves the projection of a visual scene onto the retinae of two eyes separately positioned in space. Due to this difference in perspectivity of our imaging devices, the half-images formed on the two retinae record slightly different spatial layouts of objects in the environment. Our unitary, coherent sense of the visual world is derived through the construction of a binocular representation of the 3-D space that respects this projective geometry created by each eye's view. Binocular vision is a field that has been extensively investigated and well understood (see, e.g. Howard & Rogers, 1996). The goal of the present paper is not to present any new empirical findings, but rather to re-visit the projective geometry characterizing the binocular space and study its invariants and their visual relevance using the mathematical tools of complex analysis.

Geometric characterization of the binocular space in modern era is often attributed to the work of Luneburg (1947, 1950) and more contemporarily, to that of T. Indow (see Indow, 1991, 1997 and his recent monograph 2004). Among the central findings is that, to the human visual perceptual system, the geometry of the frontal-parallel plane has zero curvature and hence is Euclidean, whereas that of the depth plane has constant negative Riemannian curvature and hence is hyperbolic (Luneburg, 1950). In the present investigation, we restrict ourselves to the depth plane and ask whether the hyperbolic geometry supporting binocular

E-mail address: junz@umich.edu.

https://doi.org/10.1016/j.jmp.2018.10.007

depth space can be derived from basic considerations of projective geometry — whether it is a necessary consequence of the basic premise upon which binocular vision operates, namely, to establish correspondence between points in the half-images of the eyes and fusing the two half-images into a coherent one. To this end, we resort to the technique of Möbius transform in complex analysis and study its invariants and the automorphism group in the 3-D depth plane modeled as a complex domain, that of the upper half-plane.

Projective geometry, which carries the cross-ratio as its only four-point invariant, is different from affine geometry, which preserves the notion of parallelism. With some rare exceptions (Chan, Pizlo, & Chelberg, 1999; Pizlo, 1994, 2008; Pizlo & Loubier, 2000), the role of projective geometry in binocular vision is not fully appreciated. Cutting (1986), following the Gibsonian viewpoint that perceptual constancy involves the extraction of invariants of the environment, suggested that human visual system may actually detect the cross-ratio of four co-linear points. Having investigated the significance of topology in early visual perception (Chen, 1983, 1982, 1985; Chen & Zhou, 1997; Han, Humphreys, & Chen, 1999; Todd, Chen, & Norman, 1998; Zhuo et al., 2003), L. Chen proposed a processing hierarchy for the perception of visual figures, including the extraction of invariants under (in progressive order) topology, projective geometry, affine geometry, and Euclidean geometry of the visual space (Chen, 2005). In remarkable coincidence with Felix Klein's Erlangen Program (EP), Chen's proposal endows the paramount role of geometrical and spatial reasoning in visual perceptual organization. Our current exploration of the role of projective geometry in binocular vision is motivated in part by this line of theoretical reasoning. More specifically, we set it as our goal to characterize the projective geometry of the binocular space by means of the Möbius transformation group (and its subgroups) over a complex domain, therefore elucidating the origin of negative curvature of the depth plane as revealed by the empirical findings about perceived distance and perceived parallelism in 3-D (Blumenfeld, 1913; Hillebrand, 1902). This is in accord with the Gibsonian thesis that the environment and perception "make an inseparable pair" (Gibson, 1979), in this case, under the projective geometric characterization of the binocular space.

1.1. *Stereopsis, cyclopean eye, and binocular space*

Binocular or stereoscopic vision (or simply "stereopsis") refers to the fusion of the left and the right eye's views of a scene into a single, unitary one. The fact that we perceive a coherent image despite the slightly different perspectives taken by the two eyes suggests that the two overlapping half-images are somehow integrated by a virtual, "cyclopean" eye located midway in-between the two eyes. This cyclopean viewpoint reflects the perceptual reconstruction of the third (i.e., depth) dimension of our visual environment. According to Hering (1879/1964, p. 232) the term "cyclopean eye" was originally coined by Helmholtz, alluding to the one-eyed cyclops from the Greek mythology. Though it is aimed at explaining the unitary binocular direction despite disparate directions originating from the two individual eyes, the logical necessity of such a notion for judging visual direction has been a subject of heated debates (Erkelens & van Ee, 2002; Ono, Mapp, & Howard, 2002). In his book "Foundations of Cyclopean Perception", Julesz (1971) speculated on the existence of the cyclopean eye as a central processing stage inside the brain with neuro-anatomical underpinnings:

> "The mythical cyclops looked out on the world through a single eye in the middle of his forehead. We too, in a sense, perceive the world with a single eye in the middle of the head. But our cyclopean eye sits not in the forehead, but rather some distance behind it in the areas of the brain that are devoted to visual perception. One can even specify a certain site in the visual system as being the location of the cyclopean eye. For instance, we can locate the cyclopean eye at a place where the views of the two external eyes are combined. In this case, normally we think of the information registered by the cyclopean eye as very similar to that presented to our external eyes. Accordingly, we assume that this cyclopean eye receives little more information than each of the external eyes alone obtains. All that is added, it would seem, is a somewhat richer impression of depth".

In the present investigation, the consequence of the assumed existence of a cyclopean eye mediating stereoscopic perception will be further explored from a geometric perspective. Binocular registration and correspondence of object points in stereo vision will be revisited using the tools of complex analysis, specifically, the theory of Möbius transformation (also known in complex analysis as fraction-linear transform, bilinear transform, general linear transform, etc.). The Möbius transform will be shown to provide a natural characterization of the transformation of the binocular space supporting cyclopean perception. It is known that the only projective invariant of a Möbius transform is the four-point cross-ratio. In the current context, two of the four points are taken to be the positions of the two eyes, the third to represent the fixation point, and the fourth to represent an arbitrary object location other than the fixation point. When the eyes change fixation, the invariance of the four-point cross-ratio implies that the relative disparity between the object and the reference/fixation point (or between any two objects for that matter) remains unchanged before and after the eyes change their fixation. Under this interpretation, the significance of projective geometry (as it relates to visual perception) is that it describes how the brain forms a single representation of the visual space where spatial layout of objects are relative stable and consistent despite changes in the point of fixation. The current formulation improves upon Cutting's (1986) original suggestion about our visual system detecting the four-point cross-ratio lying along a straight line — it is, we argue, the invariance of the cross-ratio of *any* four points lying on the (depth) plane that is of true significance to (binocular) vision.

One novel aspect of this mathematical model is the claim that cyclopean representation of the binocular space involves not only the horizontal disparity γ but also the ratio ρ of distances of the point to the two eyes. The latter is essentially equal to vertical size ratio (VSR), namely, the ratio of the vertical magnification factor of the two half-images. The pair of quantities γ, ρ form complex representation of the depth plane. As VSR could be calculated using the knowledge of (the gradient of) vertical disparity (see Howard & Rogers, 1996), our framework implicates VSR as an important quantity in binocular vision.

The remainder of the paper is organized as follows. Section 2.1 will review the geometry involved in binocular vision, and derive the relationship between the positioning of the cyclopean eye and the binocular coordinates of an object point in the *depth plane* (variously called "horizontal plane of eye-level", "epipolar plane"). Section 2.2 further introduces the complex representation, and recasts this relationship using the $\tanh^{-1}(z) \equiv \frac{1}{2}\log\frac{1+z}{1-z}$ function. Section 2.3 studies the three-point simple ratio and establishes that $z \mapsto \tanh^{-1}(z)$ as the psychophysical transformation from the external binocular coordinates to the internal cyclopean representation. Section 2.4 introduces Möbius transforms as an automorphism group on certain complex domains. Section 2.5 studies the (only!) invariant of Möbius transforms, namely the four-point cross-ratio. Section 2.6 concentrates on the subset of those Möbius transforms that keep the two eye positions as fixed points. Section 2.7 restricts the actions of Möbius transforms to the upper half-plane, which is biholomorphically equivalent to the unit disk, and provides a source for (the empirically observed) constant negative curvature as arising out of an invariant metric under the automorphism of the unit disk. Finally, in Section 3, we discuss the implications of our framework with respect to an Erlangen-like program of visual perception articulated by Chen (1985), an axiomatization of binocular coordinates by Heller (1997), and research on vertical size ratio (VSR).

2. Complex characterization of binocular visual space

2.1. Binocular geometry and cyclopean eye

When the optic axes of both eyes converge on a single point in the space (called "fixation point"), these two axes form a plane passing through the optic centers (called "nodal points") of the eyes as well as the fixation point — this plane is variously called "the plane of regard", "epipolar plane", or "horizontal plane of eye-level". If the eyes are torsionally aligned, the vertical axis of the coordinate frame of each eye is orthogonal to the plane of regard and passes through the corresponding optic center. Since eye-movements in the vertical and in the horizontal directions are known to be decoupled and separately controlled, with vertical movements conjugated normally, we may for simplicity only deal with horizontal orbital eye-movement and hence horizontal disparity while ignoring the vertical dimension henceforth. Note that horizontal disparity here is defined as the difference in the azimuth (the "horizontal" component of the 3-D coordinates) of an object point measured in (i.e., projected onto) the plane of regard. In the following, the plane of regard, with depth and horizontal dimensions (the horizontal axis is the inter-ocular line that connects the nodal point of each eye), is called the "upper half-plane" where points are identified by complex numbers $Z = X + iY$ with $Y \geq 0$; we also write $Z = (X, Y)$ under the Cartesian representation of a complex number. Here the head-centered Cartesian coordinate system (X, Y) is set up in such a way that the X-coordinate represents the horizontal position of the 3-D point, where the Y-coordinate represents the distance (range) of the 3-D point to the observer (more accurately, to the inter-ocular line). The origin of the Cartesian coordinate system is placed at the midway along the inter-ocular line; the nodal points of the eyes are separated by a distance of $2k$, with $k > 0$ a real number (Fig. 1).

Denote the angles formed by the optic rays from the object point $Z = (X, Y)$ ("world coordinates") to the left and to the right eye as α, β, respectively, and the distance of the object point to the left and to the right eye as l_L, l_R, respectively. The angles are measured with respect to the Y-axis, $-\frac{\pi}{2} \leq \alpha, \beta \leq \frac{\pi}{2}$, and are positive if the point is to the right of the corresponding eye. Clearly,

$$\begin{cases} l_L &= \sqrt{(X+k)^2 + Y^2} \\ l_R &= \sqrt{(X-k)^2 + Y^2} \end{cases} .$$

Construct a line bisecting the angle γ sustaining the two eyes; such a line will intersect X-axis at a point $(x, 0)$, and the Y-axis at a point $(0, -y)$ (here $y \geq 0$). The angle γ is called the vergence angle

$$\gamma = \alpha - \beta ,$$

whereas the version angle ϕ is given by

$$\phi = \frac{\alpha + \beta}{2} .$$

These two angles, γ giving the binocular parallax encoding egocentric distance and ϕ encoding the direction of the cyclopean eye, are collectively called the "bipolar coordinates" by Luneburg (1947). The trajectories represented by $\gamma = const$ are called the Vieth–Müller circles, whereas the trajectories represented by $\phi = const$ are called the Hillebrand hyperbolae (note that they are not straight lines). Luneburg postulated, based on empirical descriptions of the frontal horopter (von Helmholtz, 1867) and

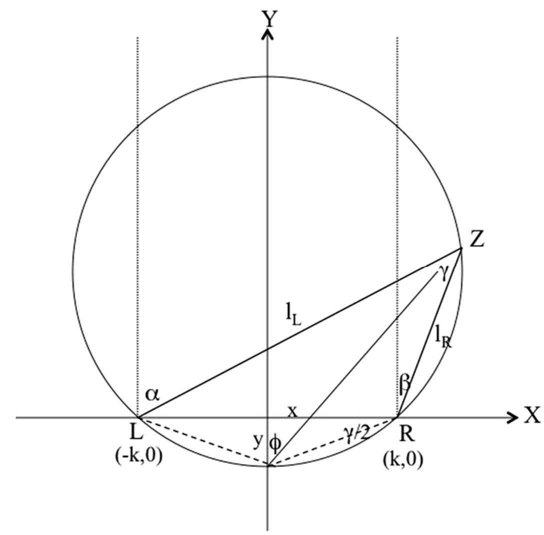

Fig. 1. Binocular viewing geometry. The two eyes, labeled as "L" and "R", fixate on the point Z, with l_L and l_R denoting its distances to the eyes. The "cyclopean eye", on the other hand, is located on the Y-axis behind the inter-ocular line (X-axis). With respect to the cyclopean eye, the distance is determined by y and direction by ϕ. Because $x = y \tan\phi$, it is convenient to use (x, y), called "cyclopean coordinates", to represent an external point in the "fused" perceptual space. There are equivalent representations, though. Luneburg used the "bipolar coordinates" (γ, ϕ), whereas the current paper uses "binocular polar coordinates" (ρ, γ) where $\rho = l_L/l_R$.

of the parallel and distance visual alleys (Blumenfeld, 1913; Hillebrand, 1902), that the visual space is a Riemannian space with constant curvature such that the points lying on a Vieth–Müller circle and on a Hillbrand hyperbola are perceived, respectively, as being equidistant and as being in constant radial direction from the observer. While there is general support for the constant negative curvature aspect of the binocular space (Blank, 1961; Indow, 1997, 2004; Zajaczkowska, 1956a, b), systematic departure from the specific predictions were also reported (Foley, 1966, 1970, 1980; Indow, 1991; Shipley, 1957), though the source of these deviations may be peripheral.[1]

We adopt the convention that positive values of X and x represent a point on the right of the observer, and negative values a point on the left. We refer to the pair of numbers x, y as the cyclopean representation (or "cyclopean coordinates") of the external object point using the world coordinates (X, Y) — there is a 1-to-1 correspondence between the point sets $\{(x, y) \in (-k, k) \times (0, \infty)\} \leftrightarrow \{X, Y \in (-\infty, \infty) \times (0, \infty)\}$ to be explicated below.

The basic relationships of binocular geometry in the physical and cyclopean space, respectively, are

$$\begin{cases} x &= y \tan\phi \\ y &= k \tan\left(\frac{\gamma}{2}\right) \end{cases} .$$

Denote ρ as the ratio of the distances of the object point to the two eyes

$$\rho = \frac{l_L}{l_R} .$$

A well-known result from elementary geometry concerning the line bisecting an angle gives

$$\frac{l_L}{l_R} = \frac{k + x}{k - x} ,$$

[1] Heller (1997); see also Aczél, Boros, Heller, and Ng (1999) advanced an elegantly formulated framework, from axiomatic conjoint measurement theory, to account for this departure, by re-scaling the two monocular angles α, β. Strong empirical support was found for this theory recently (Heller, 2004), see Section 3 for more discussions.

from which we derive

$$\frac{x}{k} = \frac{l_L - l_R}{l_L + l_R} = \frac{\rho - 1}{\rho + 1} .$$

On the other hand, apply the Law of Sine

$$\frac{l_L}{l_R} = \frac{\sin(\frac{\pi}{2} + \beta)}{\sin(\frac{\pi}{2} - \alpha)}$$

which, with some manipulation, equals

$$\frac{\cos\beta}{\cos\alpha} = \frac{\cos(\phi - \frac{\gamma}{2})}{\cos(\phi + \frac{\gamma}{2})} = \frac{\cos\phi\,\cos\frac{\gamma}{2} + \sin\phi\,\sin\frac{\gamma}{2}}{\cos\phi\,\cos\frac{\gamma}{2} - \sin\phi\,\sin\frac{\gamma}{2}} = \frac{1 + \tan\phi\,\tan\frac{\gamma}{2}}{1 - \tan\phi\,\tan\frac{\gamma}{2}} .$$

Therefore

$$\rho = \frac{1 + \tan\phi\,\tan\frac{\gamma}{2}}{1 - \tan\phi\,\tan\frac{\gamma}{2}}$$

or

$$\tan\phi\,\tan\frac{\gamma}{2} = \frac{\rho - 1}{\rho + 1} .$$

This establishes the relationship between the three variables γ, ϕ, and ρ. Luneburg took γ, ϕ as independent variables to specify a point in binocular space; we use γ, ρ instead (see Section 2.2). The cyclopean representation using the pair of numbers x, y sometimes can be convenient. For instance, the family of Hillebrand hyperbolae are mapped to straight lines

$$y = x\,\cot\phi ,$$

where each distinct value of ϕ indexes a particular member of the family. The cyclopean representations for some family of lines are given in Fig. 2 (Fig. 2a for rays emitting from one eye, and Fig. 2b for frontal-parallel lines and for straight-forward lines).

As a comparison, the Hillebrand hyperbolae under the world coordinate (X, Y) are given by

$$(X\cot\phi - Y)(X\tan\phi + Y) = k^2$$

or

$$X^2 - Y^2 + 2XY\,\cot(2\phi) = k^2 ,$$

which is obtained after elimination of intermediary variables in the following geometric relations (D denotes the diameter of the Vieth–Müller circle):

$$\begin{cases} X &= (D\cos\phi)\,\sin\phi \\ Y &= (D\cos\phi)\,\cos\phi - y \\ k^2 &= y(D - y) \end{cases} .$$

In parametric form (t is the parameter of the curve, which can be shown to actually equal y), a particular Hillebrand hyperbola indexed by ϕ is

$$t \mapsto (X(t), Y(t)) = ((k^2/t + t)\,\sin\phi\cos\phi,\ (k^2/t + t)\cos^2\phi - t) .$$

The points with $\gamma = const$ form a family of intersecting circles all passing through the two nodal points — the Vieth–Müller circle, given explicitly by

$$X^2 + (Y - k\,\cot\gamma)^2 = \left(\frac{k}{\sin\gamma}\right)^2 .$$

The points $\rho = const$ form another family of non-intersecting (and non-concentric) circles

$$\left(X - \frac{\rho^2 + 1}{\rho^2 - 1}k\right)^2 + Y^2 = \left(\frac{2\rho}{\rho^2 - 1}k\right)^2 .$$

Table 1
Various coordinate systems for binocular depth space.

Coordinates	Terminology	Term used by
(γ, ϕ)	Bipolar coordinates	Luneburg
$(\gamma, \log\rho)$	Bipolar coordinates	Math community
(γ, ρ)	Binocular polar coordinates	This paper
(x, y)	Cyclopean coordinates	Vision community
(X, Y)	Cartesian (world) coordinates	All
$(\tilde{X}, \tilde{Y})$	Binocular Cartesian coordinates	This paper

All of these circles are orthogonal to the family of Vieth–Müller circles. Two families of circles such that every circle of the first family intersects orthogonally with every circle in the other family are called "Apollonian circles", and $(\gamma, \log\rho)$ are called "bipolar coordinates" in the mathematics community (different from Luneburg's usage of the term).

2.2. Binocular polar coordinates as representation of binocular space

Under complex-number representation, basic viewing geometry (Fig. 1) gives

$$Z = Z_R + l_R\,e^{i(\frac{\pi}{2} - \beta)} = Z_L + l_L\,e^{i(\frac{\pi}{2} - \alpha)}$$

where $Z_L = -k$, $Z_R = k$ are the locations of the left and right eye respectively. Therefore

$$\frac{Z - Z_L}{Z - Z_R} = \frac{l_L\,e^{i(\frac{\pi}{2} - \alpha)}}{l_R\,e^{i(\frac{\pi}{2} - \beta)}} = \rho\,e^{-i\gamma} , \tag{1}$$

or

$$\log\rho - i\gamma = \log\frac{Z - Z_L}{Z - Z_R} = \log\frac{Z + k}{Z - k} .$$

The pair of numbers $(\rho, \gamma) \in (0, \infty) \times (0, \pi)$ will be called "binocular polar coordinates" of the object point (X, Y). They are not to be confused with the bipolar coordinates (ϕ, γ) in Luneburg's sense; nor should they be confused with the polar coordinates (r, θ) of the object point under world coordinate $Z = (X, Y)$, where $X = r\cos\theta$, $Y = r\sin\theta$, $Z = re^{i\theta}$. Using binocular polar coordinates (ρ, γ), we can define "binocular Cartesian coordinates" $(\tilde{X}, \tilde{Y})$ to be

$$\tilde{X} = \rho\cos\gamma , \qquad \tilde{Y} = -\rho\sin\gamma .$$

Table 1 summarizes the various coordinate systems used for binocular vision.

It is convenient at this point to introduce the complex-valued hyperbolic tangent function $f(z) = \tanh(z)$ defined on a complex domain Ω as

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} ,$$

with inverse (when its argument is restricted to a 2π period)

$$\tanh^{-1}(z) = \frac{1}{2}\log\frac{1 + z}{1 - z} .$$

The hyperbolic tangent satisfies

$$\tanh(-z) = -\tanh(z) , \qquad \tanh(z \pm i\frac{\pi}{2}) = (\tanh(z))^{-1} ,$$

and

$$\tanh^{-1}(-z) = -\tanh^{-1}(z) , \qquad \tanh^{-1}(z^{-1}) = (\tanh^{-1}(z)) + i\frac{\pi}{2} .$$

The hyperbolic tangent and the tangent function are related via

$$\tanh(ir) = i\tan(r) , \qquad \tan(ir) = i\tanh(r) .$$

Furthermore,

$$\tanh(z_1 + z_2) = \frac{\tanh(z_1) + \tanh(z_2)}{1 + \tanh(z_1)\tanh(z_2)} .$$

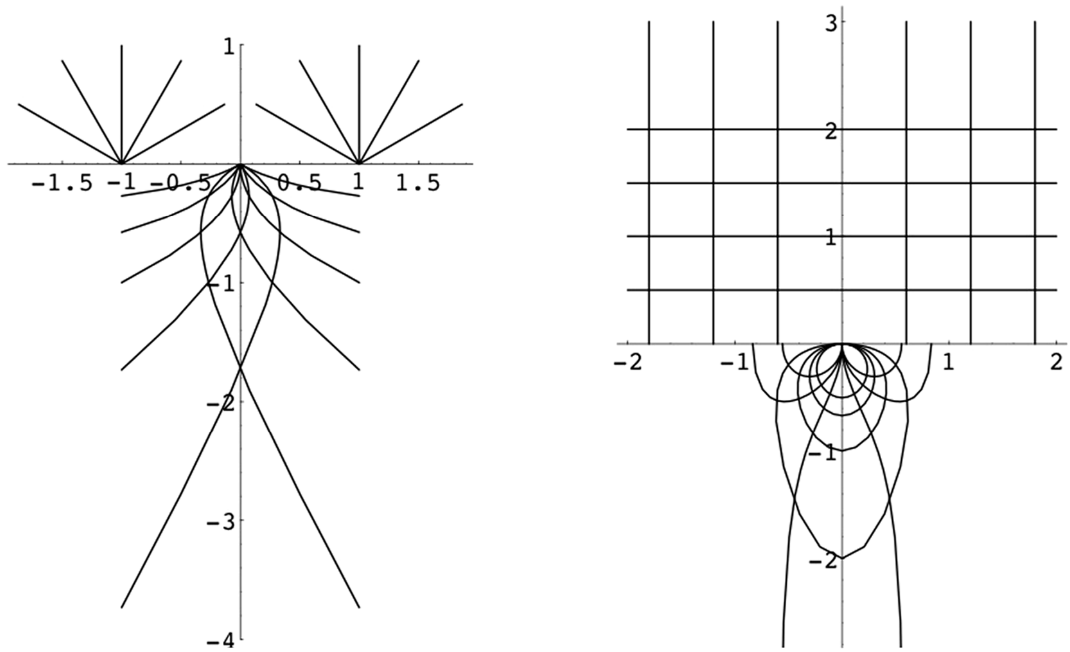

Fig. 2. Cyclopean representation (x, y), shown in the lower half-plane, of certain families of lines, shown in the upper half-plane. The curves above the abscissa are in world or (X, Y)-coordinates. (a) Rays emitting from one eye. (b) Horizontal lines ("frontal-parallel", $Y = constant$) and vertical lines ("straight forward", $X = constant$).

The function $\tanh(\cdot)$, along with its inverse, induces an operation $\oplus$

$$u \oplus v \equiv \frac{u+v}{1+uv}$$

defined on Ω, so that $\tanh^{-1}(\cdot)$ is an "additive" representation

$$\tanh^{-1}(u \oplus v) = \tanh^{-1}(u) + \tanh^{-1}(v) \ .$$

With the hyperbolic tangent function, the cyclopean coordinates $(x, -y)$ of an object point can be cast into

$$\begin{cases} \dfrac{x}{k} &= \dfrac{\rho - 1}{\rho + 1} = \tanh\left(\dfrac{\log \rho}{2}\right) \\ \dfrac{y}{k} &= \tan\dfrac{\gamma}{2} = \dfrac{1}{i}\tanh\left(\dfrac{i\gamma}{2}\right) \end{cases}$$

or

$$\begin{cases} \tanh^{-1}\left(\dfrac{x}{k}\right) &= \dfrac{1}{2}\log\rho \\ \tanh^{-1}\left(i\dfrac{y}{k}\right) &= \dfrac{1}{2}(i\gamma) \end{cases} . \qquad (2)$$

Binocular disparity γ, and hence y, encodes the depth; the Vieth–Müller circles are given by $\gamma = const$, or $y = const$. On the other hand, binocular direction of the cyclopean eye would be encoded by ϕ, since the Hillebrand hyperbolae are given by its intersection with $\phi = const$. The above analysis suggests that the location of the cyclopean eye $(x, -y)$ as it fixates on the object point (X, Y) is such that it "slides" along the Y-axis away from the inter-ocular line, with an amount dependent on the distance from the eyes to the fixation point; its direction, as though looking at the object point, is specified by the x value.

2.3. Simple ratio and its significance under binocular geometry

Binocular polar coordinates have a natural interpretation under projective geometry. Given three points z_1, z_2, z_3 in the completed (also called "extended") complex z-plane $C^* = C \cup \{\infty\}$, their "simple ratio" $(z_1; z_2, z_3)$ is defined as

$$(z_1; z_2, z_3) = \frac{z_1 - z_2}{z_1 - z_3} \ .$$

When $z_1 = z_3$ this ratio is ∞. One can construct other simple ratios $(z_2; z_1, z_3)$, $(z_3; z_1, z_2)$, etc. of the same three points; however they are all related. The simple ratio of a point Z with respect to the two eye positions, $Z_R = k$, $Z_L = -k$, is hereby called "binocular simple ratio"

$$(Z; Z_L, Z_R) = (Z; -k, k) = \frac{Z+k}{Z-k} \ .$$

From (1),

$$(Z; Z_L, Z_R) = \rho\, e^{-i\gamma} \ ,$$

that is, the modulus or the argument of the simple ratio gives, respectively, the ratio of distances to the two eyes or the binocular parallax sustained with respect to them from an object point Z. Stated in another way, binocular simple ratio gives the binocular polar coordinates (ρ, γ) we defined in Section 2.2.

From (1) and (2),

$$\tanh^{-1}\left(\frac{Z}{k}\right) = \tanh^{-1}\left(\frac{x}{k}\right) + \tanh^{-1}\left(i\,\frac{-y}{k}\right) \ .$$

So $\tanh^{-1}$ acts as some kind of scaling (in the complex plane though!) between the physical location (X, Y) and its cyclopean representation $(x, -y)$. The negative sign before y is due to our stipulation $y > 0$. When (X, Y) is being fixated, the cyclopean eye is always "located" behind the inter-ocular axis. Note that $\tanh^{-1}(\cdot)$ is a conformal mapping, so this psychophysical "scaling" is in accord with Drösler's proposal (1988) regarding possible forms of psychophysical laws for the (non-Euclidean) binocular space.

The important insight here for the representation of binocular space is that (ρ, γ), rather than (ϕ, γ), form the pair of coordinates to index any point (X, Y) in the depth plane. The transform is

$$\begin{cases} \dfrac{X}{k} &= \dfrac{(\rho - \rho^{-1})/2}{(\rho + \rho^{-1})/2 - \cos\gamma} \\ \dfrac{Y}{k} &= \dfrac{\sin\gamma}{(\rho + \rho^{-1})/2 - \cos\gamma} \end{cases} ,$$

with inverse transform given by

$$\begin{cases} \rho &= \sqrt{\dfrac{(X+k)^2+Y^2}{(X-k)^2+Y^2}} \\ \gamma &= \cot^{-1}\left(\dfrac{X^2+Y^2-k^2}{2kY}\right) \end{cases} .$$

2.4. Möbius transforms as an automorphism group on complex domains

When the fixation point of the two eyes changes, the spatial layout of object points in the environment would also change, not only in the two retinae but also in its cyclopean representation. However, the configuration of spatial relations of objects should be relatively stable. So it is interesting to investigate the kind of mappings of the binocular space that would preserve some spatial structure (to be explicated below). We say that the change of fixation induces a mapping or transformation f of the binocular space onto itself, so that each point in the pre-transformed coordinate representation (with respect to the old fixation point O) can be identified with a point in the post-transformed coordinate representation (with respect to the new fixation point O'); the point O obviously maps to the point O'. Technically, a transformation that maps a (complex) domain Ω onto itself bijectively is called an automorphism. The set of automorphisms form a group, called the automorphism group.

We require that the mapping f be "sufficiently" smooth — this continuity requirement stems from the fact that our eyes are not only capable of making saccadic eye-movement, where the fixation point jumps abruptly, but also smooth pursuit eye-movement, where the fixation point changes gradually. Here sufficiency means that at minimum the (complex-valued) derivative of f at each point should exist and be well-defined. Technically, f is assumed to be analytic.

It turns out this analyticity assumption is quite constraining for complex-valued functions: their real and imaginary parts satisfy the so-called Cauchy–Riemann condition. As a result, all analytical mappings have the very nice geometric property: they are conformal (at points with non-vanishing first derivative). This is to say, for any two smooth curves that intersect at a point $z_0 \in \Omega$, the direction and magnitude of the angle sustained by these two curves remain invariant under an analytic mapping, while the distance of any two points infinitesimally close to z_0 gets stretched by a constant factor $|f'(z_0)|$ regardless of direction.

When f is analytic on an entire complex domain Ω and $f : \Omega \to \Omega$ is one-to-one, it is called a holomorphic function. In fact, if Ω is either the complex plane $\mathcal{C}$, the completed complex plane $\mathcal{C}^* = \mathcal{C} \cup \{\infty\}$, or the unit disk $\mathcal{D} = \{z : |z| \le 1\}$, the only three types of simply connected Riemann surface (up to holomorphic equivalence), then quite remarkably, f must take a particular form called the Möbius transform.

The Möbius (or Moebius) transform given in (3), also called by various other names such as general linear transform, bilinear transform, fractional linear transform, or simply linear transform, is an important class of analytical mappings $\mathcal{M} : C^* \to C^*$ defined by:

$$\mathcal{M} : Z = \frac{az+b}{cz+d} , \tag{3}$$

where $z, Z \in C^*$, and a, b, c, d are complex constants. It was first studied by the German mathematician Möbius (1790–1868). Basic properties of Möbius transformation are provided in standard textbooks on complex analysis. The monograph by Schwerdtfeger (1962/1979) is a good source for more advanced materials relating Möbius transform to 2-D projective geometry. Below, we recall some basic facts for later use.

Note that the Möbius transform specified by a, b, c, d and by ka, kb, kc, kd (k is any non-zero complex constant) are in fact the same. Writing (3) as

$$Z = \frac{bc-ad}{c} \cdot \frac{1}{cz+d} + \frac{a}{c}$$

we require

$$bc - ad \neq 0$$

to guarantee that $\mathcal{M}$ does not map the entire completed complex plane $\mathcal{C}^*$ to a constant $Z_0 = a/c$. Under this condition, $Z_1 = Z_2$ if and only if $z_1 = z_2$ (and hence $\mathcal{M}$ is bijective), since $Z_1 = (az_1+b)/(cz_1+d)$ and $Z_2 = (az_2+b)(cz_2+d)$ will lead to

$$Z_1 - Z_2 = \frac{(ad-bc)(z_1-z_2)}{(cz_1+d)(cz_2+d)} .$$

Often, one may impose

$$bc - ad = 1$$

to remove the extra degree of freedom in the parameters (and obtain a subgroup of the Möbius group).

The generality of Möbius transform can be appreciated by observing that the following simple transformations are all special types of $\mathcal{M}$: (i) Translation (with complex constant b): $Z = z + b$; (ii) Rotation about origin (with angle α): $Z = e^{i\alpha}z$; (iii) Dilation about origin (with a real positive factor λ representing expansion if $\lambda > 1$ or constriction if $0 < \lambda < 1$): $Z = \lambda z$; and (iv) Reciprocation: $Z = 1/z$, which can be seen as the composition of a reflection about the unit circle $w = 1/\bar{z}$ and a mirror-reflection about the real axis $Z = \bar{w}$ (here $\bar{}$ denotes complex conjugation). Note any similarity (a.k.a. integral) transformation $Z = az + b$ can be decomposed into three successive applications of dilation, rotation, and translation: $Z = e^{i\alpha}(\lambda z) + b = (\lambda e^{i\alpha})z + b$.

Möbius transforms represented by (3), where $bc \neq ad$, have a number of useful properties. Most importantly, they form a group with functional composition as group composition: (i) For any two Möbius transforms $\mathcal{M}_1$ and $\mathcal{M}_2$, their functional composition $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$ is again a Möbius transform; (ii) For any three Möbius transforms $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$, their applications are associative: $\mathcal{M}_1 \circ (\mathcal{M}_2 \circ \mathcal{M}_3) = (\mathcal{M}_1 \circ \mathcal{M}_2) \circ \mathcal{M}_3$; (iii) For every Möbius transform (3), there exists a unique inverse transform $\mathcal{M}^{-1} : z \mapsto (dz-b)/(-cz+a)$ such that $\mathcal{M}^{-1} \circ \mathcal{M} = \mathcal{M} \circ \mathcal{M}^{-1} = \mathcal{I}$, with $\mathcal{I}$ being the identity transform (where $a = d = 1, b = c = 0$).

Hence Möbius transforms form a transformation group acting on the completed complex plane $\mathcal{C}^*$. A convenient isomorphic representation of (the parameters that define) Möbius transforms is through the use of the complex-valued matrix $\mathbf{M}$ with non-zero determinant

$$\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} , \qquad ad - bc \neq 0 .$$

The function/group composition operation $\circ$ is then represented simply as matrix multiplication. Of course, we must keep in mind that $\mathbf{M}$ and $q\mathbf{M}$ represent the same Möbius transform — $\mathbf{M}$ is specified only up to a complex constant q, so without loss of generality, we may require $\det|\mathbf{M}| = ad - bc = 1$. In this case,

$$\mathbf{M}^{-1} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} .$$

The matrix representation $\mathbf{M}$ of $\mathcal{M}$ can be handy in many problems — for instance, it can be immediately shown that a Möbius transform is an involution, namely, it has period-two, $\mathcal{M}(\mathcal{M}(z)) = z$, if and only if $a+d = 0$. We note in passing that when a, b, c, d are integers (satisfying $ad - bc = 1$), a subgroup of Möbius transform is obtained, known as the modular group which is a fundamental object in number theory, geometry, and algebra.

The importance of Möbius transform is that all analytic maps from the completed complex plane $\mathcal{C}^*$ onto itself while preserving the sense of handedness in rotation must take the form of (3); this is a basic result in (single-variable) complex analysis. As a conformal map, Möbius transform preserves angular (and hence orthogonal) relationships of local directions while providing isotropic stretches at each location; it maps circles (including lines, which are equivalent objects in $\mathcal{C}^*$) into circle-lines (i.e., generalized "circle" that also includes a line as a circle with radius ∞). For instance,

$$z = \frac{at+b}{ct+d} , \qquad t \in (-\infty, \infty)$$

represents a circle obtained from mapping the real line.

The fact that Möbius transform is really a mapping with three free (complex) parameters means that $\mathcal{M}$ will be specified completely if three pairs of corresponding points are given: $z_i \leftrightarrow Z_i$, $i = 1, 2, 3$. In this case, the Möbius transform can assume the following form:

$$\frac{Z-Z_1}{Z-Z_2} \div \frac{Z_3-Z_1}{Z_3-Z_2} = \frac{z-z_1}{z-z_2} \div \frac{z_3-z_1}{z_3-z_2} . \tag{4}$$

Lemma 1. *Let z_1, z_2, z_3 be any three distinct points of $\mathcal{C}^*$, and let Z_1, Z_2, Z_3 be any three distinct points in $\mathcal{C}^*$ as well. Then there is a unique Möbius transform $\mathcal{M}$ such that $\mathcal{M}(z_i) = Z_i$ for $i = 1, 2, 3$.*

Proof. The Möbius transform that satisfies $\mathcal{M}(z_i) = Z_i$, $i = 1, 2, 3$ is given by (4). Writing out explicitly in the form of (3), the coefficients can be chosen as

$$a = \frac{Z_3-Z_1}{Z_3-Z_2} Z_2 - \frac{z_3-z_1}{z_3-z_2} z_1 ,$$
$$b = \frac{z_3-z_1}{z_3-z_2} z_2 Z_1 - \frac{Z_3-Z_1}{Z_3-Z_2} Z_2 z_1 ,$$
$$c = \frac{Z_3-Z_1}{Z_3-Z_2} - \frac{z_3-z_1}{z_3-z_2} ,$$
$$d = \frac{z_3-z_1}{z_3-z_2} z_2 - \frac{Z_3-Z_1}{Z_3-Z_2} z_1 .$$

It is easy to verify that for this particular choice of a, b, c, d

$$ad - bc = \frac{(Z_3-Z_1)(Z_1-Z_2)}{Z_3-Z_2} \frac{(z_3-z_1)(z_1-z_2)}{z_3-z_2} \neq 0$$

since z_i, Z_i are all distinct for $i = 1, 2, 3$. ◇

It is easily seen that if a Möbius transform $\mathcal{M}$ has three fixed points, then $\mathcal{M}$ is the identity map. In the language of Narens (1981), the Möbius transform is 3-point homogeneous and 3-point unique, or $(M, N) = (3, 3)$ (Note but: Narens (1981) deals with ordered structure on reals which is very different from the representational structure being dealt with here).

When only two corresponding pairs $\mathcal{M}(z_1) = Z_1, \mathcal{M}(z_2) = Z_2$ are given, the Möbius transform is specified up to the freedom of a (complex) parameter λ:

$$\frac{Z-Z_1}{Z-Z_2} = \lambda \frac{z-z_1}{z-z_2} , \tag{5}$$

or written in the form of simple ratios

$$(Z; Z_1, Z_2) = \lambda(z; z_1, z_2) .$$

Eq. (5) is an alternative representation of any Möbius transform $\mathcal{M}$, in terms of Z_1, Z_2 and λ, if one allows Z_1, Z_2 to coincide and to take the value of ∞.

2.5. Cross-ratio as an invariant of Möbius transform

Given four points z_1, z_2, z_3, z_4 in the completed complex plane C^*, their "cross-ratio" $(z_1, z_2; z_3, z_4)$ is defined as

$$(z_1, z_2; z_3, z_4) = (z_1; z_3, z_4) \div (z_2; z_3, z_4) = \frac{z_1-z_3}{z_1-z_4} \div \frac{z_2-z_3}{z_2-z_4} .$$

Lemma 2. *All cross-ratios constructed from the same four points z_1, z_2, z_3, z_4 are related, i.e., they are functions of one another.*

Proof. Denote

$$\begin{aligned}(z_1, z_2; z_3, z_4) &= (z_2, z_1; z_4, z_3) = (z_3, z_4; z_1, z_2) \\ &= (z_4, z_3; z_2, z_1) = \lambda .\end{aligned}$$

Then it is easy to show that

$$\begin{aligned}
(z_1, z_2; z_4, z_3) &= (z_2, z_1; z_3, z_4) = (z_3, z_4; z_2, z_1) = (z_4, z_3; z_1, z_2) \\ &= \frac{1}{\lambda} , \\
(z_1, z_4; z_2, z_3) &= (z_2, z_3; z_1, z_4) = (z_3, z_2; z_4, z_1) = (z_4, z_1; z_3, z_2) \\ &= \frac{\lambda-1}{\lambda} , \\
(z_1, z_3; z_2, z_4) &= (z_2, z_4; z_1, z_3) = (z_3, z_1; z_4, z_2) = (z_4, z_2; z_3, z_1) \\ &= 1-\lambda , \\
(z_1, z_3; z_4, z_2) &= (z_2, z_4; z_3, z_1) = (z_3, z_1; z_2, z_4) = (z_4, z_2; z_1, z_3) \\ &= \frac{1}{1-\lambda} , \\
(z_1, z_4; z_3, z_2) &= (z_2, z_3; z_4, z_1) = (z_3, z_2; z_1, z_4) = (z_4, z_1; z_2, z_3) \\ &= \frac{\lambda}{1-\lambda} . \quad \diamond
\end{aligned}$$

The notion of cross-ratio has its origin in real projective geometry, whereby it is defined for four points on a straight line. In fact, Cutting (1986), in characterizing environmental invariants associated with a change in the position of the observer, suggested that the visual system might calculate the (real-valued) cross-ratio for any four points aligned (i.e., co-linear) in space. It turns out that when the four points are not necessarily co-linear, but coplanar, a complex-valued cross-ratio may be defined on the completed complex domain $\mathcal{C}^*$ (with the inclusion of the ∞ point) that is a projective invariant — this is our generalization of Cutting's proposal. In particular, it can be shown that λ is real if and only if the four points z_1, z_2, z_3, z_4 are on a circle. The intimate relation of Möbius transform and the four-point cross-ratio is given by the following proposition.

Proposition 3. *Let f be a function: $f : (\mathcal{C}^*)^4 \to \Omega \subseteq \mathcal{C}^*$ such that fixing any three of its four variables, the mapping (restriction of f) is one-to-one. Let $h : \mathcal{C}^* \to \Omega$ and $\mathcal{M} : \mathcal{C}^* \to \mathcal{C}^*$ both be one-to-one. Then any two of the following statements imply the third (here $z_1, z_2, z_3, z_4 \in \mathcal{C}^*$ are distinct points)*
(i) $f(z_1, z_2, z_3, z_4) = h((z_1, z_2; z_3, z_4))$;
(ii) $f(z_1, z_2, z_3, z_4)$ is invariant under $\mathcal{M}$;
(iii) $\mathcal{M}$ is a Möbius transform.

Proof. From (ii) and (iii) to (i): Since

$$f(\mathcal{M}(z_1), \mathcal{M}(z_2), \mathcal{M}(z_3), \mathcal{M}(z_4)) = f(z_1, z_2, z_3, z_4)$$

holds for any arbitrary Möbius transform $\mathcal{M}$, we can take a particular $\mathcal{M}$

$$\mathcal{M}(z) = \frac{(z-z_3)(z_2-z_4)}{(z-z_4)(z_2-z_3)} ,$$

so that

$$f(z_1,z_2,z_3,z_4)=f\left(\frac{(z_1-z_3)(z_2-z_4)}{(z_1-z_4)(z_2-z_3)},1,0,\infty\right)$$
$$=h\left(\frac{(z_1-z_3)(z_2-z_4)}{(z_1-z_4)(z_2-z_3)}\right)$$

where the function h is some single-variable function mapping $\mathcal{C}^* \to \Omega$ which is, according to the assumption on f, one-to-one. This proves (i).

From (i) and (ii) to (iii): Because of (i), we only need to prove that if

$$(z_1,z_2;z_3,z_4)=(\mathcal{M}(z_1),\mathcal{M}(z_2);\mathcal{M}(z_3),\mathcal{M}(z_4))$$

for any four distinct points z_1, z_2, z_3, z_4, then $\mathcal{M}$ is a Möbius transform. To show this, fix z_1, z_2, z_3 and denote $Z_i = \mathcal{M}(z_i)$, $i = 1, 2, 3$. Then the mapping $\mathcal{M}(z)$ given by

$$(z_1,z_2;z_3,z)=(Z_1,Z_2;Z_3,\mathcal{M}(z))\ ,$$

or equivalently (4), is in the form of a Möbius transform.

From (i) and (iii) to (ii): Let the Möbius transform of the four points z_1, z_2, z_3, z_4 be

$$Z_i=\frac{az_i+b}{cz_i+d}\ ,\qquad i=1,2,3,4\ .$$

From

$$Z_i-Z_j=\frac{(ad-bc)(z_i-z_j)}{(cz_i+d)(cz_j+d)}\ ,$$

straight-forward substitution leads to

$$(Z_1,Z_2;Z_3,Z_4)=\frac{Z_1-Z_3}{Z_1-Z_4}\div\frac{Z_2-Z_3}{Z_2-Z_4}=\frac{z_1-z_3}{z_1-z_4}\div\frac{z_2-z_3}{z_2-z_4}$$
$$=(z_1,z_2;z_3,z_4)\ .$$

Because of (i), it immediately follows that

$$f(z_1,z_2,z_3,z_4)=f(\mathcal{M}(z_1),\mathcal{M}(z_2),\mathcal{M}(z_3),\mathcal{M}(z_4)).\quad\diamond$$

The implication of Proposition 3 for representation of binocular space will be discussed in the next subsection.

Finally, we recall the notion of "Schwarzian derivative" $S(h)$ of a function h of single complex variable, defined as

$$(Sh)(z)=S(h)(z)=\left(\frac{f''(z)}{h'(z)}\right)'-\frac{1}{2}\left(\frac{h''(z)}{h'(z)}\right)^2$$
$$=\frac{h'''(z)}{h'(z)}-\frac{3}{2}\left(\frac{h''(z)}{h'(z)}\right)^2.$$

The Schwarzian derivative of a linear fractional transform is zero. Therefore, Schwarzian derivative of h describes how h "deviates" from a fractional linear (that is, Möbius) transform), just as an ordinary derivative describes how a function deviates from linearity.

For any two differentiable functions h, g, the Schwarzian derivative satisfies

$$S(h(g))(z)=(Sh)(g(z))\cdot(g')^2+(Sg)(z)$$

In particular, if g is a Möbius transform, then

$$S(h(g))(z)=(Sh)(g(z))\cdot(g')^2;$$

if h is a Möbius transform, then

$$S(h(g))(z)=(Sg)(z).$$

2.6. Möbius transform of upper half-plane and with two eyes as fixed points

In the above investigations (Sections 2.4 and 2.5), the complex domain we deal with is the unrestricted complex plane $\mathcal{C}^*$, where the Möbius transform is an automorphism group. In binocular vision, only the upper half plane, denoted $\mathcal{C}_+ = \{(X, Y) : Y \geq 0\}$, is involved. It is known from complex analysis that the automorphism group of the upper half plane $Aut(\mathcal{C}_+)$ is a subgroup of the full Möbius group, constrained by the stipulation that a, b, c, d in (3) are all real numbers, with $ad - bc > 0$ (when $ad - bc < 0$, the upper half-plane is mapped to the lower half-plane). In the form of (5), when λ is real, the Möbius transform $z \mapsto Z$ maps the upper half-plane $\mathcal{C}_+$ onto $\mathcal{C}_+$.

We now investigate a special subgroup of Möbius transforms which keep the two eye positions (strictly speaking, the nodal points of the eyes) as the fixed points. (Not fixing the two eye-positions allowing them to map to two points separated by $2k$ distance apart correspond to a change of perspective-viewing. This will be investigated in a future project.)

Lemma 4. *The Möbius transform $m : \mathcal{C}^* \to \mathcal{C}^*,\ z \mapsto m(z)$ that keeps $k, -k$ ($k > 0$) as fixed points, i.e., with $m(\pm k) = \pm k$, obeys*

$$\frac{m(kz)}{k}\,\frac{m(kz^{-1})}{k}=1\ . \tag{6}$$

Proof. We first prove an identity $(z, z^{-1}; 1, -1) \equiv -1$ for all z, by observing

$$(z,z^{-1};1,-1)=\frac{(z;1,-1)}{(z^{-1};1,-1)}=\frac{z-1}{z+1}\div\frac{z^{-1}-1}{z^{-1}+1}\equiv-1\ .$$

Since simple ratios are scale invariant, we have

$$(kz,kz^{-1};k,-k)=-1\ .$$

Now apply Möbius transform m to the four points $kz, kz^{-1}, k, -k$ and use the above identity, we have

$$(m(kz),m(kz^{-1});m(k),m(-k))=-1\ .$$

Writing out explicitly while noting $m(\pm k) = \pm k$:

$$\frac{m(kz)-k}{m(kz)+k}\div\frac{m(kz^{-1})-k}{m(kz^{-1})+k}=-1\ .$$

Rearranging terms yields

$$m(kz)\,m(kz^{-1})=k^2\ .\quad\diamond$$

Introducing

$$\tilde{m}(z)=\frac{m(kz)}{k}\ ,$$

we have

$$(\tilde{m}(z))^{-1}=\tilde{m}(z^{-1})\ .$$

Proposition 5. *The Möbius transform from $\mathcal{C}_+ \to \mathcal{C}_+$ with two eye positions $k, -k$ as fixed points takes the following form*

$$\frac{Z+k}{Z-k}=\frac{1+\lambda}{1-\lambda}\,\frac{z+k}{z-k} \tag{7}$$

where λ is some real number.

Proof. The formula (7) is expected from (5) except for the challenge to prove the proportional constant is real. Substituting $z_1 = Z_1 = k$ and $z_2 = Z_2 = -k$ into (3), we obtain the following two equations

$$\pm k(d-a)=b-ck^2\ ,$$

where a, b, c, d are all real. This gives $a = d$ and $b = ck^2$. Therefore the Möbius transform in this case is

$$Z = \frac{az + ck^2}{cz + a} = \frac{z + \lambda k}{1 + \lambda k^{-1} z}$$

where $\lambda = kc/a$. Rearranging the above gives (7). $\diamond$

It is easily seen that the values of the λ-parameter form a group on R, with group composition $\circ$ given by $\oplus$

$$\lambda_1 \oplus \lambda_2 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 \lambda_2} .$$

That is,

$$\frac{1 + (\lambda_1 \oplus \lambda_2)}{1 - (\lambda_1 \oplus \lambda_2)} = \frac{1 + \lambda_1}{1 - \lambda_1} \frac{1 + \lambda_2}{1 - \lambda_2} \quad (8)$$

or, from Section 2.2,

$$\tanh^{-1}(\lambda_1 \oplus \lambda_2) = \tanh^{-1}(\lambda_1) + \tanh^{-1}(\lambda_2) .$$

Note that, unlike Proposition 3 and (5) which describes any Möbius transform mapping $\mathcal{C}^* \to \mathcal{C}^*$ while preserving two fixed points (i.e., an arbitrary change of fixation), Proposition 5 considers the additional requirement of mapping $\mathcal{C}_+ \to \mathcal{C}_+$. In the latter case, the change of fixation points must be along the same horopter (i.e., Vieth–Müller circle) and quantified by a real number λ, resulting in the binocular polar coordinates (1) to satisfy

$$\gamma_Z = \gamma_z .$$

The relationship expressed by (8) simply reflects a rotation affecting ϕ (and ρ); these Ḿobius transforms form a subset of the ones given by (5).

2.7. Hyperbolic geometry of binocular space

Recall from the theory of holomorphic mappings that the upper half-plane $\mathcal{C}_+$ is one of the three types of complex domains where "type" here refers to an equivalent class (under bijective, holomorphic mapping) of all simply connected Riemann surfaces with complete metric.

The well-known *Riemann Mapping Theorem* states that any simply connected open subset $\Omega \subset \mathcal{C}$ of the complex plane can be mapped biholomorphically (i.e., holomorphic in both directions) onto the *open* unit disc. Biholomorphicity implies that the mapping is conformal, namely, preserving angles and shapes of sufficiently small figures with possible rotation and stretching.

The binocular space $\mathcal{C}_+$ itself can be biholomorphically mapped onto the unit disk $\mathcal{D} = \{(X, Y) : \sqrt{X^2 + Y^2} \leq 1\}$, given by

$$w = e^{i\theta} \frac{z - a}{z - \bar{a}}$$

where a is a complex number with positive real part $\mathrm{Re}(a) \geq 0$ and θ is a real number. An alternative expression is

$$w = \frac{\gamma z - \delta}{\bar{\gamma} z - \bar{\delta}}$$

where $\frac{\gamma}{\bar{\gamma}} = e^{i\theta}$ and $\delta = \gamma a$.

As a consistency check, we look at the Möbius transform $w : \mathcal{C}_+ \to \mathcal{D}$ mapping the two eye positions $\pm k$ to ± 1: $w(k) = 1$, $w(-k) = -1$. It is not difficult to show (similar to the proof of Proposition 5) that w is characterized by

$$\frac{w + 1}{w - 1} = \frac{\nu + \bar{\nu}}{\nu - \bar{\nu}} \frac{z + k}{z - k}$$

where ν is an arbitrary complex number. Any automorphic mapping $h : \mathcal{D} \to \mathcal{D}$ preserving the points ± 1 as fixed points (i.e., $h(\pm 1) = \pm 1$) takes the form

$$h = \frac{z + \lambda}{1 + z\lambda} = z \oplus \lambda ,$$

where λ is a real constant. Rewriting the above as

$$\frac{1 + h}{1 - h} = \frac{1 + \lambda}{1 - \lambda} \frac{1 + z}{1 - z} ,$$

it is clearly seen that the $\oplus$ operation given by (8) characterizes the group composition operation for both $Aut(\mathcal{D})$ and for $Aut(\mathcal{C}_+)$ when restricted under two fixed points (± 1 for the former case, $\pm k$ for the latter case).

The geometry of binocular space, modeled as that of the upper half-plane $\mathcal{C}_+$ or the biholomorphically equivalent open unit disk $\mathcal{D}$, can also be endowed with a metric and with affine connections. It is well-known that there can be two models (and therefore two Riemannian metrics) of the unit disk: the Poincaré model and the Klein–Beltrami model. Both models lead to a hyperbolic space, consistent with the well-established empirical finding (Blank, 1961; Luneburg, 1950; see also Indow, 2004) that the binocular depth space is of constant negative curvature. Compared with spherical geometry, which provides one-point compactification of the complex plane (i.e., identifying the infinity as *one* single point ∞), hyperbolic geometry is a better model for binocular visual space because there is one "infinity" point for each visual direction. Though binocular space has been axiomatized (Blank, 1958; Heller, 1997), a thorough analysis of which of the two hyperbolic models (Poincaré versus Klein–Beltrami model) provides better insight for binocular perception is an interesting question for future research.

3. Discussion

A geometric framework for binocular visual space is advanced — more specifically, a projective geometric description of the fusion of two half-spaces by the cyclopean eye. Möbius transforms are used to characterize the automorphic mappings of the binocular space. In particular, we identified the class of Möbius transforms that preserve the two eye positions, and used that to model a change of fixation point. We derived the hyperbolic tangent function $f(z) = \tanh(z)$, with inverse $f^{-1}(z) = \frac{1}{2} \log \frac{1+z}{1-z}$, as the psychophysical function relating the physical representation of the binocular space to its cyclopean representation.

To discuss the implications of our framework for binocular representation, let us recall the notion "saccadic remapping", i.e., a remapping of the binocular space, whether modeled as (i) the completed complex plane $\mathcal{C}^*$ or at least (ii) the upper half-plane $\mathcal{C}_+$, as a result of a change of the fixation point either abruptly as in saccade or continuously as in smooth pursuit. Our mathematical analysis shows that any such remapping of the space onto itself (assuming bijectivity and analyticity) would have to be a Möbius transform in the form of (5) for case (i) or the Möbius transform in the form of (7) for case (ii). The binocular polar coordinates (γ, ρ) used by our framework, as opposed to the bipolar coordinates of Luneburg, are uniquely suited in expressing such Möbius transforms, and hence saccadic remapping. This is because, under such coordinates, a change of fixation is simply a change in one or both of the values of γ, ρ which parameterize the Möbius transforms as the automorphism group of $\mathcal{C}_+$ or C^*.

In terms of invariants under saccadic remapping (whether along Vieth–Müller circle or not), our investigation shows (Proposition 3) that the four-point ratio is the only quantity that remains constant. Taking two of these points to be the two eye positions as mandated by saccadic remapping, this result means that despite fixational change, the spatial relation between *any* two points A and B in the binocular space is preserved, in terms of their relative disparity $\gamma_A - \gamma_B$ and relative lateralization ρ_A/ρ_B. That is to say, when the fixation point O is changed to fixation point O', their relative disparity as determined by the projective geometry of A and B with O as fixation and their relative disparity under the new geometry with O' as fixation remains unchanged — a constant amount of vergence is subtracted/added when fixation is changed from O to O'. In this regard, our theory connects two inter-related statements about the representation of binocular space, one concerning arbitrary fixational change affecting the calculation of binocular parallax (as emphasized by Luneburg's theory) and the other concerning relative disparity of any two points as an invariant (as emphasized by Blank's axiomatization). The subtle difference of these two viewpoints is that, though binocular parallax of a point is defined with respect to the two eye positions, without reference to the fixation point, in terms of its calculation by the visual system, the fixation point is the reference zero of visual incoming rays for each eye, and hence binocular parallax is calibrated by fixation.

3.1. Erlangen program in visual perception

Geometry is intrinsic to vision because visual perception is, after all, the perception of relational structure in the spatial (as opposed to verbal or symbolic) domain. Different types of geometric spaces, however, have been invoked in the past, including Euclidean, affine, projective, and even topological spaces, to explain different visual phenomena (Chen, 2005). In discussions below, we have excluded any considerations of how the presence of visual objects will alter the geometric property of the perceptual space and consider the latter only as a "container" or ambient space that supports perception.

Euclidean geometry defines the rigidity of an object under translatory or rotational movements and under mirror-reflection through the well-defined notion of length. The work of Carlton and Shepard (1990a, 1990b) laid the foundation for mental rotation and other rigid operations on visual object in a Euclidean geometry. Specifically, Carlton and Shepard applied the principles of kinematic geometry (in particular, Charles's theorem stating that for any two configurations of an asymmetric object there exists a unique axis in space that carries the object from one configuration to another by a helical combination of a translation along and rotation about that axis, two movements collectively called the "screw motion") and the Law of Least Action to derive the geodesics in a six-dimensional manifold (three degrees of freedom in translation and three in rotation). These shortest transformational paths map directly onto the mental paths of objects undergoing apparent motion (see also Foster, 1973a, 1978). The theory of Carlton and Shepard, when applied to symmetric objects, allow the computation of the relative "strength" of various apparent motion paths, predictions that can be evaluated by experimental data using the paradigm of path competition in apparent motion (Chen, 1985; Foster, 1973b, 1975).

Affine geometry defines the parallelism of contours (or segments of contour) of an object viewed under different perspective conditions. Leyton (1987b), based on his earlier work on local symmetric axis (the symmetry-curvature duality theorem, Leyton, 1987a), proposed a description by Lie group action on the object's symmetric axis as a general way of achieving different shapes corresponding to the same object. He demonstrated how a combination of simple "stretch", "shear", and "rotation" operations (which themselves form appropriate subgroups) will result in the positive general linear transformation group acting on the object plane, with the parameters of the subgroups completely characterizing the shape of the object under affine transformations (but not involving mirror reflection).

Topology studies those geometrical properties that are preserved through arbitrary twisting, stretching, or other deformation of the space. Though its abstract axiomatization is founded on the set-theoretic notions like open sets and their unions and intersections, topological characterizations do convey in a precise way our intuitions about continuity, connectedness, and boundaries of visual objects in a scene. L. Chen believed that the processing of topological information about objects is an important function of pre-attentive early vision (Chen, 1983, 1982, 2001, 2005; Chen & Zhou, 1997; Han et al., 1999; Todd et al., 1998; Zhuo et al., 2003). To paraphrase this theory of topological visual perception, Zhang (2005) proposed an interpretation of topological computation as one establishing figure-ground relations, which includes the determination of (i) continuity, connectivity and closure of the region occupied by an object; (ii) its boundaries and their ownership in determining the relative status of background/foreground; (iii) the nature of apparent "holes" and the occlusion relationship it may indicate. Viewed in this context, the differential geometric framework proposed by the author for motion-based figure-ground segregation (Zhang, 1995; Zhang & Wu, 1990; see also Zhang, 1994) is aimed at establishing the large-scale topology of objects and their surroundings by linking ("binding") adjacent locations via intrinsic comparison of motion vectors through parallel transport.

Viewed in this way, our current work can be seen as taking another step towards fulfilling the Erlangen Program in visual perception as envisioned in Chen (2005), namely a hierarchy stream of visual information processing from topological to projective-geometric, to affine geometric, and finally to Euclidean geometric properties. Neurophysiologically, the establishment of binocular correspondence happens at a stage prior to object recognition and its mental manipulation (such as translation and rotation). We thus claim that the extraction of projective-geometric invariants occurs prior to the extraction of affine and Euclidean invariants characteristic of, respectively, object shape recognition and mental rotation.

3.2. Relation to Blank's and Heller's axiomatization

Blank (1958) provided an axiomatic foundation of binocular space from orderings of perceived length (of line segments) and perceived alignment or colinearity (of points), and showed how a metric necessarily arises from these empirically defined relations. Compared to Luneburg's (1947) conceptualization emphasizing binocular parallax with respect to the fixation point, Blank's (1958) work treated the above two order relations of arbitrary points in the binocular space as being primitive relations. In this regard, the statement about automorphism of the upper-half plane upon change of fixation along the horopter (Proposition 5) seems to be in the spirit of the former, while the statement about the invariance of four-point cross-ratio under arbitrary Möbius transformation (Proposition 3) seems to be in the spirit of the latter.

Heller (1997) offered a new axiomatization of the use of binocular angles (α, β) (and hence Luneburg's bipolar coordinates) in stereopsis. His theory also started with two kinds of orderings between any two points in the binocular space. Quite different from Blank's (1958) postulates, Heller (1997) considered the ordering of perceived egocentric distance of points ("nearer/farther") and the ordering of egocentric direction ("lefter/righter"). Starting with these two binary relations, Heller showed that under a few rather innocent conditions plus one technical but empirically supported condition (the so-called "Reidemeister" condition, based on for equidistance relation among pairs of points), that the bipolar coordinates admit a conjoint representation, i.e., there exist strictly increasing functions f, g such that the transformed variables $(\alpha, \beta) \mapsto (f(\alpha), g(\beta))$ participates in the distance and direction judgments. So the generalized Luneburg transforms are

$$\Gamma = f(\alpha) - g(\beta),$$
$$\Phi = \frac{f(\alpha) + g(\beta)}{2}$$

with $\Gamma = const$ as loci of perceptually equidistant to the observer, and $\Phi = const$ as loci of radial line with perceptually constant direction. The observed discrepancies between the empirically measured locus of perceived equidistance and the Vieth–Müller circle, and between the empirically measured locus of constant visual direction, are attributed to the above input transformations which operate on the monocular information. The effects captured by those monotonic transformations may be caused by the optical properties of the eye, such as different magnification factors of the lens or aspherical retina shape, dissociation of optical node and rotation center of eye; the transformations may also capture effects of the neural processing that occurs prior to binocular combination. These observer-specific transformations may provide an account for individual differences in the performance of visual direction and distance judgments. Aczél et al. (1999) further illuminated iseikonic transformation as automorphism groups of this conjoint structure.

According to Heller's (1997) conjoint measurement framework, monocular direction of the eyes is a fundamental measure for stereo vision — equidistant judgment of two points in the epipolar plane need not involve binocular combination, but rather merely across-eye comparison, of monocular information about these two points. Such emphasize of fundamental independence of the two eyes for extraction of stereo information is a novel perspective that stresses the two independent ordered structures in stereo vision, namely, the relative distance judgement ("far/near") between a point and the viewing self, and relative direction ("left/right") judgement between any two points with respect to the viewing self. However, how these two structures interact to give rise to a sense of absolute distance between any two points in the depth plane and a sense of parallelism between any four points in the depth plane has not been addressed. In other words, starting with the conjoint measurement structure of the two eyes, Heller showed with rigor and ingenuity how the sense of (ego-centric) direction (Φ) and the sense of (ego-centric) distance (Γ) can be constructed. How to turn Φ and Γ into a geometric structure in the depth plane that supports the notion of metric and parallelism is an entirely orthogonal research question. What we have shown here is that Luneburg's original intuition about this negatively-curved visual space is essentially correct — perhaps after taking (the transformed values) Φ and Γ as the input to the analysis. Future research will illuminate deeper insights as to how the freedom in the transformation (f and g) can be characterized in the automorphisms of the upper half-plane.

3.3. *Implication for vertical size ratio (VSR)*

In terms of substantive claims, our framework suggests that the complex representation of a point in the depth plane (by the cyclopean eye) involves not only the horizontal disparity γ but also the ratio ρ of distances to the two eyes. For a pair of closely spaced points (roughly, < 5deg when $\tan\alpha = \alpha$ in radians holds), ρ is essentially equal to vertical size ratio (VSR), namely, the ratio of the vertical magnification factor of the two half-images, a quantity related to vertical disparity. So our framework implicates VSR as an important quantity in binocular vision. This prediction is elaborated below.

That the eyes are separated horizontally means that the principal disparities are horizontal, i.e., parallel to the inter-ocular axis. However, similarly to the definition of azimuth angle from the two eyes and the horizontal disparity defined by their difference, one can define the absolute vertical disparity as the difference of the elevation (the "vertical" component of the 3D coordinates) of a point. That vertical disparity is perceptually involved in coding depth was demonstrated by the so-called "induced effect" of Ogle (1938), that is, a surface lying in the frontal plane appears slanted about (or away from) a horizontal axis when the image in one eye is magnified (or shrunk), in the direction of vertical meridian, relative to the image in the other eye, just as a surface would appear slanted about a vertical axis when the magnification/shrinkage is made along the horizontal meridian.

All points lying in the plane of regard have zero (relative) vertical disparity, since they have a zero angle of elevation by definition. All points lying in the median (sagital) plane of the head also have zero (relative) vertical disparity because they are the loci of intersection between binocular cones of elevation. The only points that have both zero (relative) horizontal disparity and zero (relative) vertical disparity are: the Vieth–Müller circle in the plane of regard, plus the (ocular-centrically defined) vertical line through the fixation point. For all other points, vertical disparity is non-zero and its measurement will provide useful information about the object's 3-D position. For instance, Mayhew and Longuet-Higgins (1982) suggested a way how measurement of absolute vertical disparity can be used to estimate both viewing distance (from the cyclopean point) and the direction of gaze.

An alternative use of vertical disparity is to calculate the vertical size ratio (VSR). The basic viewing geometry dictates that the vertical component of the vertical disparity gradient equals the ratio (minus 1) of vertical sizes of the same object as imaged by the two eyes (VSR), see Howard and Rogers (1996). Gillam and Lawergren (1983) demonstrated that the gradient of the VSR versus eccentricity can be used to compute absolute distance to a surface. Contrary to Mayhew and Longuet-Higgins, Gilliam and Lawergren's estimate of the distance to a surface is purely based on local computations (see also Rogers & Bradshaw, 1993). Subsequent psychophysical experiments found that VSR indeed contributes to the perception of the slant of a stereoscopically presented surface (Backus, Banks, van Ee, & Crowell, 1999), its curvature (Rogers & Bradshaw, 1995), or its size (Brenner, Smeets, & Landy, 2001), but not its azimuth (Banks, Backus, & Banks, 2002).

Our investigation of the binocular geometry suggests that VSR (ρ) is an important quantity to compute, second perhaps only to horizontal disparity (γ) — this is because, they together form the (complex-valued) cyclopean representation of the binocular coordinates of a point in the 3-D space. So far, the author is uninformed of any computational theory of binocular vision that *exclusively* exploits these two quantities (ρ, γ), and their spatial gradients, to recover viewing parameters and/or scene structure. More attention in future research, both psychophysical and computational, will hopefully be drawn towards their usage in stereopsis.

Acknowledgments

The research reported in this paper was initiated at the State Key Laboratory of Cognitive Sciences, Institute of Biophysics, Chinese Academy of Sciences, Beijing China, directed by Lin Chen. The author acknowledges support from an international collaborative grant from K.C. Huang Foundation (1995, 2002) and an Overseas Team Collaboration Grant from the Chinese Academy of Sciences (2004–2008) awarded to Lin Chen, and a block travel grant from National Natural Science Foundation (NNSF) of China managed through the China Bridge Program (1998–2000). The mathematical framework grew out of initial discussions with Lin Chen on geometry and vision, with a first draft completed and circulated by the author in 2005. The final draft resulted from many rounds of editorial comments leading to improved presentation (thanks to J. Heller).

References

Aczél, J., Boros, A., Heller, J., & Ng, C. (1999). Functional equations in binocular space perception. *Journal of Mathematical Psychology*, *43*, 71–101.

Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, *39*, 1143–1170.

Banks, M. S., Backus, B. T., & Banks, R. (2002). Is vertical disparity used to determine azimuth?. *Vision Research*, *42*, 801–807.

Blank, A. A. (1958). Axiomatics of binocular vision. The foundations of metric geometry in relation to space perception. *Journal of the Optical Society of America*, *48*, 328–333.

Blank, A. A. (1961). Curvature of binocular visual space: An experiment. *Journal of the Optical Society of America*, *51*, 335–339.

Blumenfeld, W. (1913). Untersuchungen über die scheinbare Größe im Sehraume. *Zeitschrift Für Psychologie*, *65*, 241–404.

Brenner, E., Smeets, J. B. L., & Landy, M. (2001). How vertical disparities assist judgements of distance?. *Vision Research*, *41*, 3455–3465.

Carlton, E. H., & Shepard, R. N. (1990a). Psychologically simple motions as geodesic paths. I. Asymmetrical objects. *Journal of Mathematical Psychology*, *34*, 127–188.

Carlton, E. H., & Shepard, R. N. (1990b). Psychologically simple motions as geodesic paths. II. Symmetrical objects. *Journal of Mathematical Psychology*, *34*, 189–228.

Chan, M. W., Pizlo, Z., & Chelberg, D. M. (1999). Binocular shape reconstruction: Psychological plausibility of the 8 point algorithm. *Computer Vision and Image UnderstandIng*, *74*, 121–137.

Chen, L. (1983). What are the units of figure perceptual representation. Studies of Cognitive Science, No. 22 (Sponsored by school of social sciences, University of California, Irvine).

Chen, L. (1982). Topological structure in visual perception. *Science*, *218*, 699–700.

Chen, L. (1985). Topological structure in the perception of apparent motion. *Perception*, *14*, 197–208.

Chen, L. (2001). Perceptual organization: To reverse back the inverted (upside down) down question of feature binding. *Visual Cognition*, *8*, 287–303.

Chen, L. (2005). Topological approach to perceptual organization. *Visual Cognition*, *12*, 553–637.

Chen, L., & Zhou, W. (1997). Holes in illusory conjunctions. *Psychonomic Bulletin and Review*, *4*, 507–511.

Cutting, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.

Drösler, J. (1988). The psychophysical function of binocular space perception. *Journal of Mathematical Psychology*, *32*, 285–297.

Erkelens, C. J., & van Ee, R. (2002). The role of the cyclopean eye in vision: sometimes inappropriate, always irrelevant. *Vision Research*, *42*, 1157–1163.

Foley, J. M. (1966). Locus of perceived equidistance as a function of viewing distance. *Journal of the Optical Society of America*, *56*, 822–827.

Foley, J. M. (1970). Loci of perceived, equi-, half- and double-distance in stereoscopic vision. *Vision Research*, *10*, 1201–1209.

Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, *87*, 411–434.

Foster, D. H. (1973a). A hypothesis connecting visual pattern recognition and apparent motion. *Kybernetik*, *13*, 151–154.

Foster, D. H. (1973b). An experimental examination of a hypothesis connecting visual pattern recognition and apparent motion. *Kybernetik*, *14*, 63–70.

Foster, D. H. (1975). Visual apparent motion and some preferred paths in the rotation group *SO*(3). *Biol. Cybernetics*, *18*, 81–89.

Foster, D. H. (1978). Visual apparent motion and the calculus of variations. In E. L. J. Leeuwenberg, & H. F. J. M. Buffart (Eds.), *Formal theories of visual perception* (pp. 67–82). New York: Wiley.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton-Mifflin.

Gillam, B., & Lawergren, B. (1983). The induced effect, vertical disparity, and stereoscopic theory. *Perception and Psychophysics*, *24*, 121–130.

Han, S., Humphreys, G. W., & Chen, L. (1999). Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of closure. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1411–1432.

Heller, J. (1997). On the psychophysics of binocular space perception. *Journal of Mathematical Psychology*, *41*, 29–43.

Heller, J. (2004). The locus of perceived equidistance in binocular vision. *Perception and Psychophysics*, *66*, 1162–1170.

Hillebrand, F. (1902). Theorie der scheinbaren Größe beim binokularen Sehen. *Denkschrift Der Kaiserlichen Akademie Der Wissenschaften Wien, Mathematisch-Naturwissenschaftliche Classe*, *72*, 255–307.

Howard, I. P., & Rogers, B. J. (1996). *Binocular vision and stereopsis*. Oxford University Press.

Indow, T. (1991). A critical review of Luneburg's model with regard to global structure of visual space. *Psychological Review*, *98*, 430–453.

Indow, T. (1997). Hyperbolic representation of global structure of visual space. *Journal of Mathematical Psychology*, *41*, 89–98.

Indow, T. (2004). *The global structure of visual space*. Singapore: World Scientific Publishing Co..

Julesz, B. (1971). *Foundations of cyclopean perception*. University of Chicago Press.

Leyton, M. (1987a). Symmetry-curvature duality. *Compter Vision Graphics, and Image Processing*, *38*, 327–341.

Leyton, M. (1987b). A limitation theorem for the differentiable protypification of shape. *Journal of Mathematical Psychology*, *31*, 307–320.

Luneburg, R. K. (1947). *Mathematical analysis of binocular vision*. Princeton, NJ: Princeton Univ. Press.

Luneburg, R. K. (1950). The metric of binocular visual space. *Journal of the Optical Society of America*, *40*, 627–642.

Mayhew, J. E., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature*, *297*, 376–378.

Narens, L. (1981). On the scale of measurement. *Journal of Mathematical Psychology*, *24*, 249–275.

Ogle, K. N. (1938). Induced size effect. I. A new phenomenon in binocular space perception associated with the relative size of the images of the two eyes. *Archives of Ophthalmology*, *20*, 604–623.

Ono, H., Mapp, A. P., & Howard, I. P. (2002). The cyclopean eye in vision: the new and old data continue to hit you right between your eyes. *Visual Research*, *42*, 1307–1324 (Erratum in **43**: 2331) .

Pizlo, Z. (1994). A theory of shape constancy based on perspective invariants. *Vision Research*, *34*, 330–350.

Pizlo, Z. (2008). *3D Shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.

Pizlo, Z., & Loubier, K. (2000). Recognition of a solid shape from its single perspective image obtained by a calibrated camera. *Pattern Recognition*, *33*, 1675–1681.

Rogers, B. J., & Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature*, *361*, 243–255.

Rogers, B. J., & Bradshaw, M. F. (1995). Disparity scaling and the perception of frontoparallel surfaces. *Perception*, *24*, 155–179.

Schwerdtfeger, H. (1962/1979). *Geometry of complex numbers: circle geometry, Möbius transformation, non-euclidean geometry*. New York: Dover.

Shipley, T. (1957). Convergence function in binocular visual space. II. Experimental report. *Journal of the Optical Society of America*, *47*, 804–821.

Todd, J., Chen, L., & Norman, J. (1998). On the relative salience of Euclidean, affine and topological structure for 3D discrimination. *Perception*, *27*, 273–282.

von Helmholtz, H. (1867). *Handbook of Physiological Optics (Vol. 9)*. Voss.

Zajaczkowska, A (1956a). Experimental test of Luneburg's theory. Horopter and alley experiments. *Journal of the Optical Society of America*, *46*, 514–527.

Zajaczkowska, A. (1956b). Experimental determination of luneburg's constants σ and κ. *Quarterly Journal of Experimental Psychology*, *8*, 66–78.

Zhang, J. (1994). Image representation using affine covariant coordinates. In Y.-L. O, A. Toet, D. Foster, H. J. A. M. Heijmans, & P. Meer (Eds.), *Shape in Picture: Mathematical Description of Shape in Grey-Level Images* (pp. 353–362). Berlin: Springer-Verlag.

Zhang, J. (1995). Motion detectors and motion segregation. *Spatial Vision*, *9*, 261–273.

Zhang, J. (2005). Object oneness: the essence of the topological approach to perception. *Visual Cognition*, *12*, 683–690.

Zhang, J., & Wu, S. (1990). Structure of visual perception. *Proceedings of the National Academy of Sciences of the United States of America*, *87*, 7819–7823.

Zhuo, Y., Zhou, T. G., Rao, H. Y., Wang, J. J., Meng, M., Chen, M., et al. (2003). Contributions of the visual ventral pathway to long-range apparent motion. *Science*, *299*, 417–420.