SIAM J. IMAGING SCIENCES
Vol. 5, No. 4, pp. 1213–1233

Gauge Fields in Pictorial Space*

Jan Koenderink† and Andrea van Doorn‡

Abstract. "Pictorial space" is the mental structure that appears to be the scaffold for the visual awareness when looking "into" (as opposed to "at") a picture. Its structure differs from the "visual space" that is the scaffold for the visual awareness when looking into the scene in front of the observer. The structure of pictorial space has been probed empirically and explored theoretically. Here we propose a framework that allows one to handle cases that have been encountered empirically, but thus far have not been explored in a formal, geometrical setting. The framework allows one to handle many idiosyncrasies of human visual observers, as well as to characterize the (frequent) individual differences in a principled manner. This opens the door to a principled formalism of the structure (e.g., quality) of the pictorial spaces evoked by various methods of presentation, as required for applications.

Key words. visual perception, image understanding, pictorial space, geometry of perception, image quality

AMS subject classifications. 53Z99, 68T45, 92F99, 94A08

DOI. 10.1137/120861151

1. Pictorial space and visual space.

Although pictorial space $\mathbb{P}^3$ (notice that $\mathbb{P}$ is not used in its conventional sense as projective space throughout this paper) and visual space $\mathbb{V}^3$ are both three-dimensional spaces related to visual perception, they are categorically different [20, 21]. We open this introduction with a concise account of the differences.

Visual space is experienced when one opens one's eye in broad daylight [21]. In generic circumstances one experiences the scene in front of the eye. The physical scene in front of the eye has (for all practical purposes) the structure of three-dimensional Euclidean space $\mathbb{E}^3$. If vision were veridical, then visual space would be a faithful copy of Euclidean space. In practice this is only approximately the case. Although visual space shares its dimensionality (three, or maybe two plus one) with Euclidean space, it is definitely not a homogeneous space.

Visual space is *perspectival*, in that it is organized about one's vantage point. This may be a single vantage point, as when one is at rest in an otherwise static scene, with one eye open, or it may be based on multiple vantage points. This happens in binocular viewing, or when the observer and the environment are in relative motion. In various cases the structure of visual space is likely to vary, whereas Euclidean space remains just what it is.

Pictorial space is experienced when one "looks into" a picture [29, 20]. This is different from "looking at" a picture. When one looks *at* a picture, one experiences an object in visual

*Received by the editors January 3, 2012; accepted for publication (in revised form) July 30, 2012; published electronically October 31, 2012. This work was supported by the Methusalem program by the Flemish Government (METH/08/02), awarded to Johan Wagemans.
http://www.siam.org/journals/siims/5-4/86115.html
†Laboratory of Experimental Psychology, University of Leuven (K.U. Leuven), BE-3000 Leuven, Belgium, and Delft University of Technology, EEMCS, MMI, NL-2628 CD Delft, The Netherlands (j.j.koenderink@tudelft.nl).
‡Industrial Design, Delft University of Technology, NL-2628 CE Delft, The Netherlands (a.j.vandoorn@tudelft.nl).

space, a photograph, a painting, or whatever the case may be. When one looks *into* a picture, one becomes visually aware of the existence of another space that is evidently not part of visual space and does not "represent" aspects of the Euclidean structure of the scene in front of the eye. Like visual space, pictorial space is a mental entity. It differs from visual space in that there is no immediate relation to the scene in front of the eye. The observer's eye is not at the origin of pictorial space, for pictorial space has no natural origin. In visual space the eye is the perspective center; in pictorial space there is no such thing [45, Proposition 5.633]. The *picture plane* is an object in physical space and also does not belong to pictorial space.

Euclidean space $\mathbb{E}^3$ is, of course, well understood [6, 8] and will be taken for granted here. Visual space, understood as some representation of Euclidean space seen from one or more particular points, can be approached as a subfield of Euclidean geometry. There exists an extensive literature on this, starting with Euclid and significantly expanded in recent times, for instance, in fields such as machine vision [7]. There has been remarkable progress in this area.

Visual space $\mathbb{V}^3$ from the experiential point of view (the ontologically correct viewpoint) has been studied in experimental psychology and psychophysics. There exist both an extensive literature on empirical approaches and a rather smaller one on theoretical, formal approaches [28]. The progress in this area has not been remarkable, to put it mildly.

Pictorial space $\mathbb{P}^3$ is categorically different from both $\mathbb{E}^3$ and $\mathbb{V}^3$ because it is neither the space in front of the observer nor necessarily some representation of it. The eye, most importantly, is not in the pictorial space. It is *elsewhere* with respect to that space. Neither is the picture plane in pictorial space.

Pictorial space is evidently based on some *picture*. A picture is a physical object such as a painting, a photograph, or a computer screen with some simultaneous arrangement of colored pixels on it. For convenience we will speak of a "picture" as a physical surface, whose points are instantiated through a pigment (as in a painting), a silver deposit (as in an old-fashioned photograph), a radiance (as on an old-fashioned cathode ray tube), and so forth.

We will treat the (physical) "picture plane" as a Euclidean plane $\mathbb{E}^2$. Pictorial space is clearly based on the picture plane, the picture plane evoking the "visual field," which is an object of visual awareness. The visual field appears when one assumes a "painter's view," for instance. In daily life visual awareness is dominated by visual space. For the purposes of this paper we will simply equate the visual field with the picture plane, even if these are ontologically disparate entities. Pictorial space is different from the picture plane in that each point of the picture plane somehow carries another spatial dimension, the "depth domain."

"Depth" is a basic feeling of "remoteness," that is, a feeling of separateness from the observer. It appears as a one-dimensional, ordered entity in awareness, although it is capable of degrees of definiteness and vagueness. In its most articulate form it is like a one-dimensional line with perhaps the structure of the affine line $\mathbb{A}^1$. There certainly is no natural origin, and there is hardly a natural unit of depth. The affine structure is often evident, though. For instance, it often makes sense to consider the midpoint of two fiducial points.

Thus, pictorial space may be described as a fiber bundle $\mathbb{P}^3 = \mathbb{E}^2 \times \mathbb{A}^1$ [11]. Its coherence suggests that nearby fibers are somehow coordinated. We will refer to the fibers as "depth threads." One should not conceive of the depth threads as somehow standing in a geometrical relation to the picture plane. They are in a different universe. The relation is purely formal.

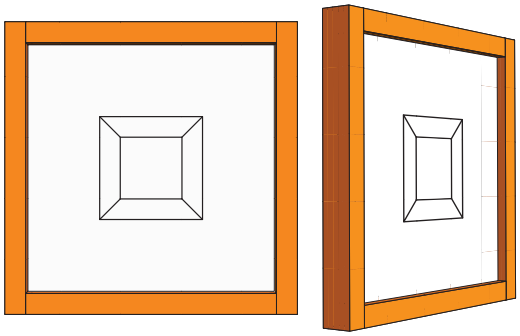

Figure 1. *At left a framed copy of a wireframe cube, at right the same picture viewed obliquely. Notice that the cube (when seen as a solid object, not a flat glyph) still points right at the viewer; it merely "grew somewhat thinner." What is happening here? The cube (or rectangular block maybe) is seen frontally, both in the straight on and in the obliquely viewed picture! This is possible only because the viewer's eye is not in the pictorial space. (It often takes human observers a few seconds to become fully aware of the depth; if the pictorial mode is not dominant, the wireframe glyph at right will look flat and oblique.)*

For instance, if one views a painting obliquely (Figure 1), the idea that the depth threads would be inclined with respect to the picture plane is simply nonsense. The depth threads have to be understood in a purely abstract sense. They model an aspect of visual awareness and are purely mental entities that do not correlate with some obvious structure of the environment.

The fiber bundle $\mathbb{P}^3 = \mathbb{E}^2 \times \mathbb{A}^1$ is the formal starting point of the present investigation (Figure 3 (left)).

2. Formal structure and coordinatization of pictorial space. The picture plane is conveniently modeled as the Euclidean plane $\mathbb{E}^2$. It makes practical sense to introduce a Cartesian coordinate system in it. Pictures are usually viewed right-side-up; thus one discriminates between the horizontal (or left-right) and the vertical (or up-down) dimensions. These dimensions are mutually orthogonal, and one uses the same yardstick for distances in any direction. Given the typically rectangular shape of pictures, there is a natural origin, the lower-left corner. Then we reckon the X-dimension from left to right, and the Y-dimension from bottom to top. The unit of length may be defined as the pixel size, a degree of angular extent in the visual field, or some other measure. We leave this open.

We affix a standard frame of unit vectors $\{\mathbf{e_x}, \mathbf{e_y}\}$ to every point of the picture plane. This is a natural and extremely simple Cartan moving frame that allows differential geometry in the most elementary sense [17]. For reasons of conciseness we will ignore the fact that pictures are typically of limited extent throughout this paper.

The depth threads are modeled as the affine line $\mathbb{A}^1$. For ease of reference we assume that these are fitted out with scales such that their origins are mutually coordinated ("in the picture plane" if we draw an illustration) and their unit points are equally coordinated. This will be referred to as the "standard picture" of pictorial space. The unit vectors $\mathbf{e_z}$ attached to the depth threads complement the Cartesian frame field to a frame field of three-dimensional pictorial space. Of course the units of length in the picture plane and along the depth threads are categorically different (the former being of a physical, the latter of a mental origin), so no comparison is possible.

This categorical difference becomes perhaps more intuitive if one considers how one might directly compare a stretch in the picture plane with a stretch along a depth fiber. Here "direct comparison" implies aligning the two stretches and judging the coincidences of their end points. No operational method has ever been suggested. Intuition tells one that this is not going to change either. Depth threads (of potentially "infinite extent") are mere points in the picture plane. There are no physical methods for changing this. The extension in depth exists only in the visual awareness of an observer. Many (maybe most) animals do not experience pictorial depth at all.[1] Yet pictorial depth is of great importance in numerous applications involving human observers. Engineering fields such as computer graphics should acquire the means to deal with pictorial depth in a formal sense.

A point $\{x, y, z\}$ has trace $\{x, y\}$ in the picture plane. The bundle projection simply "forgets the depth." The observer that lacks pictorial space (one's cat or dog perhaps) is never aware of the threads, only of their projections.

Two points $\{x, y, z_1\}$ and $\{x, y, z_2\}$ (with $z_1 \neq z_2$) will be distinct for you, and identical for your cat or dog. Such points we propose calling "parallel" (see section 2.1 for a formal account). Such cases abound for the human observer, for instance consider a figure like this:  . There are two line crossings that intersect in the picture plane, but *do not intersect* in pictorial space. At least when one "sees" the figure as a three-dimensional cube (as most human observers would) and not as a planar glyph (as one's cat or dog probably would). If one sees the figure as a cube, then the cross-over points are not one, but two points. These are at zero distance from each other (in the picture plane), yet distinct! Some continued viewing is likely to reveal that the depth order may change suddenly [27]. Apparently pictorial space is a volatile entity, constructed by the mind anew from one moment to the next.

Useful as this coordinatization is, it does not imply that observers are able to compare depth threads (for instance, a human observer and his or her cat or dog will experience difficulties communicating, even nonverbally), neither with respect to the origins, nor with respect to the unit. This is purely a matter of empirical research. Experiments reveal huge differences between generic human observers [22, 19, 23, 24, 20, 25, 18].

The standard frame field, when drawn as a sampling of pictorial space at unit intervals $\{i, j, k\}$, $i, j, k \in \mathbb{Z}$, displays an array of unit cells (but notice that the Z-units are different from the XY-units!) as in a simple tetragonal (drawn as simple cubic) crystalline lattice (illustrated in Figure 8 (right)). We use this as the default structure.

2.1. Pictorial space as a homogeneous space. Since the picture plane is Euclidean, all entities of interest are invariant with respect to Euclidean congruences and—in many cases—similarities of the picture plane. Since these are well known and of only minor interest here, we will ignore them in our discussion. More interesting transformations (for the present discussion at least) involve transformations in the depth threads.

Transformations that respect the structure of pictorial space should evidently respect the affine structure of the depth threads. Although this is a strong constraint, it actually does not have a very deep meaning. The affine structure of the threads is essentially forced by the conceivable methods of operationalization of "depth." In visual space depth can be correlated with physical range (distance reckoned from the eye), and this correlation may turn out to

[1] See http://www.pigeon.psy.tufts.edu/avc/toc.htm.

be nonlinear. (It usually turns out to be.) In pictorial space there is no such convenient comparison available (since the eye is not in pictorial space, there is no notion of "physical range"), so one has only the depth. One uses linear scales on the depth threads as a matter of definition. It is a mere convenience and superficiality. There is no a priori relation between the scales on different threads. Their origins and unit distances may be different.

When one considers pictures as renderings of physical scenes, it makes sense to apply additional constraints. For instance, the relief of surfaces is usually noticed because of the gradients of shading according to Lambert's law [1, 6, 10, 32, 47]. For the case of Lambertian surfaces, which are perfect scatterers, scattering equal radiance in all directions depending only on the irradiance of the surface, it is the case that planar surfaces will appear in a uniform shading when illuminated with a homogeneous light field. Homogeneous light fields are due to sources at infinite distances. Thus uniform regions in the picture can be understood as due to planar surfaces in the scene, although the spatial attitude of such planar surfaces remains fully indeterminate. Analogous discussions may be based on other "depth cues" [2, 28]. The upshot is that it makes sense to consider transformations that conserve planarity in pictorial space as congruences, or similarities.

Combining these ideas, one obtains the following group of similarities of pictorial space:

$$\begin{aligned} x' &= h(x\,\cos\varphi - y\,\sin\varphi) + t_x, \\ y' &= h(x\,\sin\varphi + y\,\cos\varphi) + t_y, \\ z' &= g_x x + g_y y + kz + t_z, \end{aligned} \tag{1}$$

where the parameters h, φ, and $\mathbf{t} = \{t_x, t_y\}$ have their conventional Euclidean significance as scaling, rotation, and translation in the picture plane. The parameters $\mathbf{g} = \{g_x, g_y\}$ specify an isotropic rotation, whereas the parameter k specifies a "similarity of the second kind," which is a scaling of isotropic angles. (The Euclidean similarity parameterized by h is a "similarity of the first kind" in that it scales distances.) This is an eight-parameter group S^8, having one parameter more than the corresponding Euclidean group. The reason is that there exist two kinds of scaling in pictorial space because both the distance and angle metrics are parabolic, whereas the Euclidean angle metric is elliptic (periodic).

According to Klein, a group of motions and similarities defines a geometry [13, 14, 15]. In this case it is one of the Cayley–Klein geometries [5, 46]. It is a well-researched geometry [35, 36, 37, 38, 39, 40]. English textbooks are available [31, 30, 46]. In the work of Pottmann and Liu [30] on architectural design, the isotropic direction is set by the direction of gravity, similar to the depth dimension in our application.

The following subgroup of S^8 leaves the picture plane invariant (Figure 2):

$$\begin{aligned} x' &= x, \\ y' &= y, \\ z' &= g_x x + g_y y + kz + t_z. \end{aligned} \tag{2}$$

The parameters $\mathbf{g} = \{g_x, g_y\}$, k, and t_z denote a (non-Euclidean) rotation, an (equally non-Euclidean) scaling of angles, and a translation in depth (see below). This subgroup is the four-parameter group D^4 of pure depth similarities and motions.

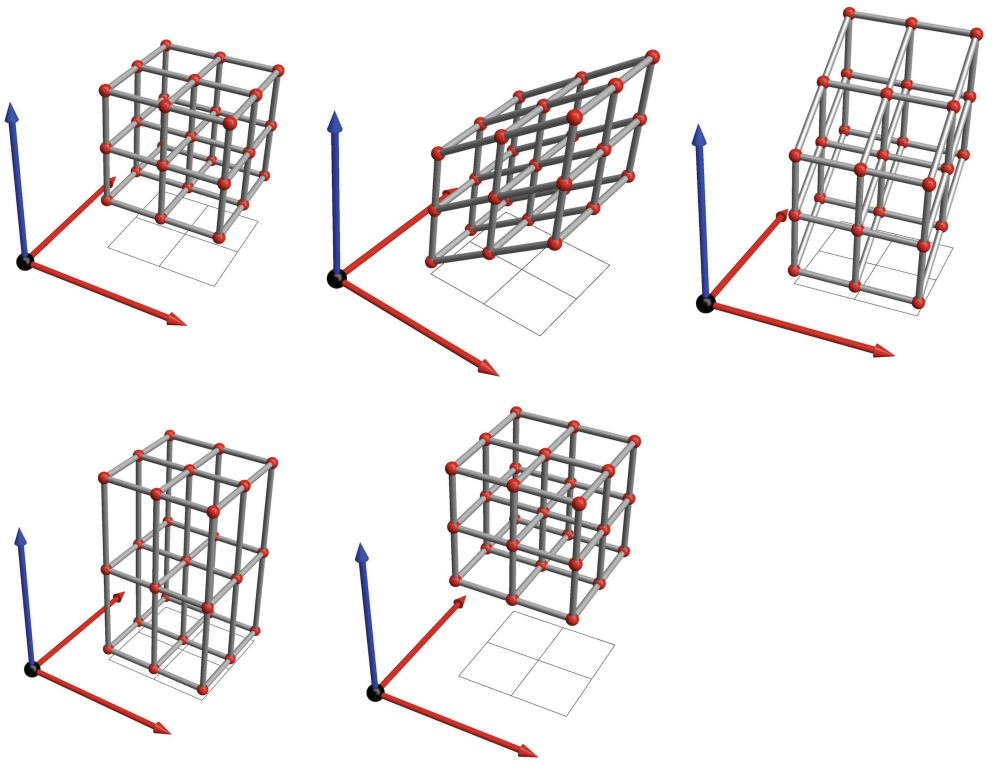

Figure 2. *Illustration of the basic actions of the subgroup* D^4. *In these figures the basis vector* $\mathbf{e_x}$ *is the red arrow pointing to the right,* $\mathbf{e_y}$ *is the red arrow pointing into the paper, and* $\mathbf{e_z}$ *is the blue arrow pointing upwards. The subfigure at top left shows a default configuration, that is,* $\mathbf{g} = \{0, 0\}$, $k = 1$, $t_z = 0$. *The subfigure at top center shows the influence of* g_x, *and that at top right the influence of* g_y. *These are non-Euclidean rotations. The subfigure at bottom left shows the influence of* k, *a non-Euclidean similarity ("of the second kind"). Finally, the subfigure at bottom right shows the influence of* t_z, *a translation in depth.*

The metric is that of the picture plane; that is, the distance between points $\{x_1, y_1, z_1\}$ and $\{x_2, y_2, z_2\}$ is $d_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$. It is easily seen to be conserved by the group actions. That is, the group S^8 scales the distance by h. For $h = 1$ one has congruencies, for $h \neq 1$, $h > 0$, similarities ("of the first kind"; see below). Notice that this distance fully ignores the depth structure. It captures only the physical structure, that is, the Euclidean geometry of the picture plane, and ignores any "mental additions."

When $d_{12} = 0$ the points might still be distinct, at least to an observer who assigns a depth structure. In that case—and only in that case—we denote the distance $\delta_{12} = z_2 - z_1$. The group actions of S^8 scale the distances δ_{12}, for point pairs with $d_{12} = 0$, by the parameter k. For $k = 1$ one has congruencies and for $k \neq 1$, $k > 0$ similarities ("of the second kind"; see below). The distance δ_{12} is a depth difference, that is, a purely mental entity. It is likely to differ between observers and might well change over time for any specific observer.

"The" distance D_{12} is defined as either d_{12} ("proper distance") or δ_{12} ("special distance"), whichever applies (Figure 3 (right)). In the latter case the points are said to be "parallel." One easily checks that the distance is indeed a movement invariant for $h = k = 1$. The similarities

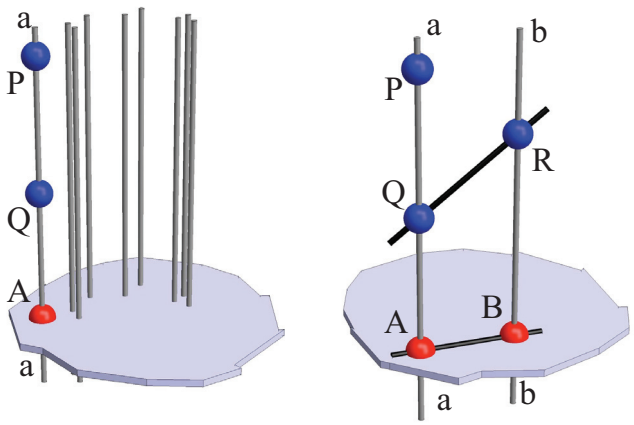

Figure 3. *At left an illustration of the fiber bundle notion. The fiber a is attached to the point A of the base space (the picture plane). The points P and Q have the same trace A; they coincide in the picture, but are at distinct depths. At right an illustration of the notion of distance. Points P and Q on the fiber a have the same distance to point R, the distance AB in the picture plane. Points P and Q have zero proper distance, yet are different ("parallel points"). They have a special distance PQ measured along the fiber.*

have two moduli, scalings of the first kind ($k = 1$) and those of the second kind ($h = 1$). This is an essential difference from Euclidean space $\mathbb{E}^3$, which knows only a single type of similarity.

Two generic planes in pictorial space, p_1, p_2 say, meet in a line ℓ_{12}, say, with trace $\overline{\ell}_{12}$. Let Q be a point in the picture plane not on $\overline{\ell}_{12}$, and let q_1, q_2 be points on the depth thread at Q and on p_1 and p_2, respectively. Let $d_{Q\ell}$ be the Euclidean distance from Q to the line $\overline{\ell}_{12}$. Then the (proper) angle mutually subtended by the planes p_1, p_2 is defined as $\delta_{q_1q_2}/d_{Q\ell}$. When the planes are parallel their (special) angle is defined as $\delta_{q_1q_2}$ for any pair of parallel points q_1, q_2 on the planes p_1 and p_2, respectively. One easily checks that the angle defined in this way is indeed a movement invariant (for $h = k = 1$), and that it scales with k. (See Figure 4.) Scalings of the second kind (with $h = 1$, $k \neq 1$) scale the angles between planes, whereas scalings of the first kind (with $h \neq 1$, $k = 1$) scale proper distances of points.

This angle between generic planes is a mental entity. It cannot be defined in the projection. Only two nongeneric planes, both of whose traces are lines, subtend a well-defined Euclidean angle in the picture plane. However, such planes are evidently singular, because they are completely made up of depth threads, and are seen "edge-on."

In this paper we will ignore these singular cases[2] as not belonging to pictorial space. We consider all planes to be of the generic variety. That is to say, the trace of a plane is the full picture plane.

The angles are in the range $\{-\infty, +\infty\}$; thus they are not periodic. One says that the angle measure in pictorial space is "parabolic."

[2]In the differential geometry of surfaces in pictorial space one has occasion to consider such singular planes. They correspond to the case where surfaces "turn over in depth," and thus the points where the tangent planes of the surface are seen "edge-on." Such points correspond to the "occluding edge," that is, the boundary of the "front-" and the "back-side" of the surface. The trace of this curve is the "contour."

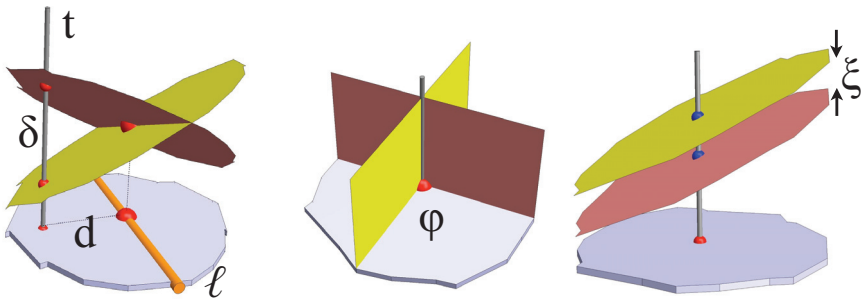

Figure 4. *The notion of "angle." At left two generic planes intersect in a line, which has trace ℓ. The arbitrary fiber t intersects the planes in two parallel points at special distance δ. The proper distance of the trace of t to ℓ is d. Then the proper angle between the planes is defined as the ratio δ/d. At the center two singular planes have Euclidean angle φ. At right two parallel planes have special angle ξ.*

Thus the essential difference with Euclidean space $\mathbb{E}^3$ is the parabolic angle measure for the proper angles between planes. In $\mathbb{P}^3$ both the distance and angle measures are parabolic, whereas $\mathbb{E}^3$ has a parabolic distance but an elliptic angle measure.

It is easy to check that one has a perfect metrical duality between points and planes. Thus one has both "parallel points" and "parallel planes" in pictorial space. The reason is that both angle metric and distance metric are parabolic. Because Euclidean angles are periodic, scalings of the second kind do not apply in $\mathbb{E}^3$, where angles cannot be scaled at all, although distances can. This is the reason why S^8 is an eight-parameter group, whereas the equivalent group of Euclidean similarities is only a seven-parameter group. It is also the reason that the geometry of pictorial space is simpler[3]— because it is more symmetric—than that of Euclidean $\mathbb{E}^3$.

Every generic plane of pictorial space has the full picture plane as its trace. The singular planes with a line as trace subtend infinite angles with every generic plane. The depth threads are isotropic lines of pictorial space. They are singular lines that subtend an infinite angle with any generic plane. Generic lines have well-defined slopes with regard to generic planes (defined in a way similar to the way the angle between generic planes is defined).

A rotation by $\mathbf{g}$ ((1) and (2)) is about a singular plane whose trace is $-g_y\mathbf{e_x} + g_x\mathbf{e_y}$. The slope angles of all lines in singular planes perpendicular to this trace change by the same amount, $\|\mathbf{g}\| = \sqrt{g_x^2 + g_y^2}$. This is the reason to refer to this subgroup ($k = 1$, $t_z = 0$ in (2)) as the group of (non-Euclidean) rotations of pictorial space.

The group of rotations of pictorial space is different from the group of (Euclidean) rotations in the picture plane (parameter φ in (1)). The pictorial rotations cannot "turn things over": a certain surface element shows only its frontal side, the back side being forever out of view. This is of course as it should be: in the *en face* painting of a face you are never going to see the back of the head, for the simple reason that it was never painted. This illustrates the necessarily non-Euclidean nature of pictorial space.

[3]The notion that the geometry of pictorial space is "simpler" than that of Euclidean space may take some getting used to. This is due to one's familiarity with Euclidean concepts. Most of the numerous "exceptions" mentioned in Euclidean theorems do not occur in pictorial space. Hence the claim of simplicity.

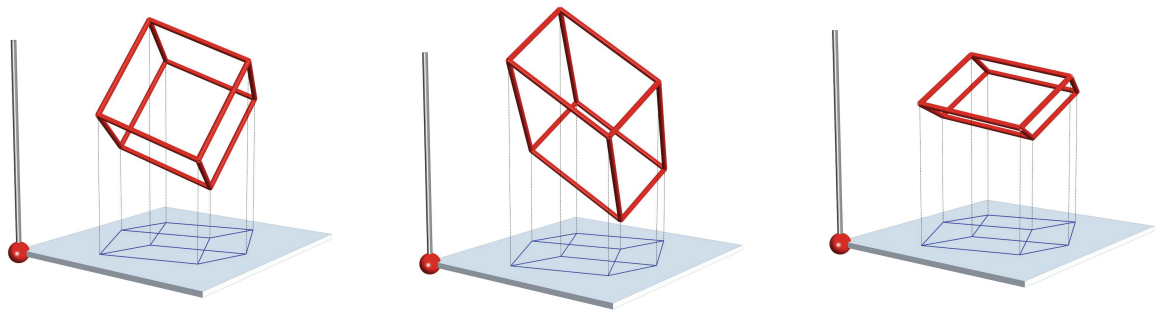

Figure 5. *Some of the actions of* D^4. *At left a line drawing in the picture plane has evoked a three-dimensional image in pictorial space. (The vertical line indicates the direction of the depth dimension.) At center we show another image, which is congruent to that at left, being connected to it by a rotation. At right we show yet another image, this one being similar to the image at left. It is connected to it by a similarity of the second kind. These images are representative of the variations one finds between the perceptions of different human observers.*

3. Global gauges. The application of a transformation from the group D^4 will be designated as a "global gauge transformation." Consider its effect: It defines a certain oblique plane as the "subjective frontal plane" and introduces a certain unit of depth difference. Thus, it is simply a copy of the default framework that has been rotated and scaled (by a similarity of the second kind). For $k = 1$ the copy is isometric (congruent) to the default framework.

Thus, one obtains a copy of the default structure that is related to it in the simplest manner. Because of the group structure of D^4 two copies obtained through different gauges will be mutually related by a transformation of the same kind. (See Figure 5.)

3.1. Measurements in pictorial space. In order to profit from a knowledge of the geometrical structure of pictorial space, one requires a toolbox of various geometrical measurements. Since pictorial space is a mental entity, it is evidently out of the question to apply the familiar tools from physical space such as yardsticks, dividers, protractors, straightedges, templates, and sieves.

A variety of tools has nevertheless been developed, and they are detailed in the literature. Here we merely mention the major methods. These can be divided into two major classes: The first class addresses the shape of smooth submanifolds, such as surfaces or "pictorial reliefs," and the second class addresses the configuration of sets of mutually dispersed points.

Surfaces are measured through sets of discrete samples, typically a few hundred. One method samples the spatial attitudes of tangent planes at a set of points. From a dense field of samples one constructs the best fitting surface and checks for consistency. The method yields detailed measurements of the shapes of pictorial surfaces [20, 22, 19, 23, 25, 24].

Configurations of points can be addressed in various ways. A simple way is to sample depth order for point pairs [42, 41]. For N points[4] one samples $N(N-1)/2$ pairs and explains

[4]A typical experiment might involve a few to several dozens of points. The possible (and useful) values of N range from about 5 to 50, depending upon the particular method. Here "possible" refers to the duration of an experimental session, an hour being the limit.

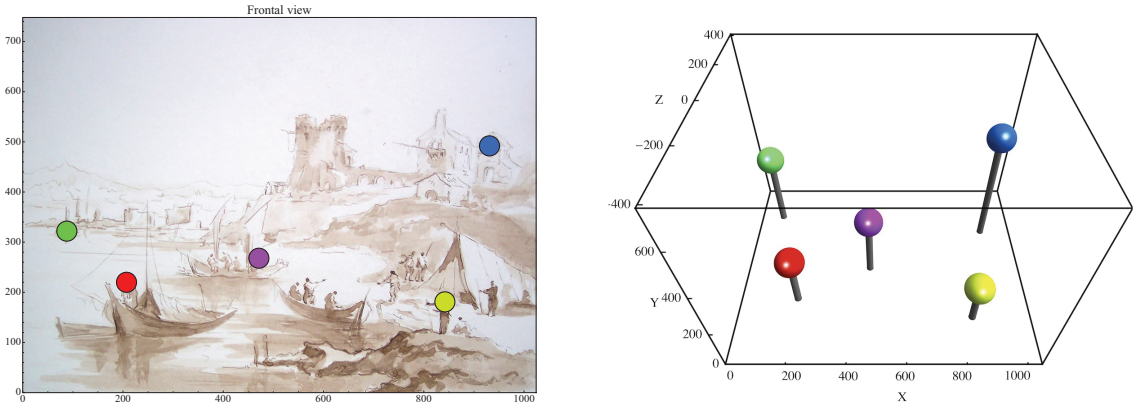

Figure 6. *At left a picture with five fiducial points. Observer JK (one of the authors) used a pointing method to find the depth configuration as shown at right for frontal viewing.*

the result in terms of $N-1$ parameters (the depths up to a common constant). This yields an ordinal scale and a consistency check. A different method involves superimposing the pictures of a pointer and a target on the picture plane [42, 43]. One attaches the pointer and target to certain locations in pictorial space. The observer controls the apparent spatial attitude of the pointer so as to point it to the target *in pictorial space*. Repeated for many point pairs, this yields a metrical depth scale and a consistency check. Yet another method exploits the relative size cue [44].

Of course wielding these tools requires a basic familiarity with psychophysical methods. Diverse details can be found in the above-mentioned literature.

3.1.1. Empirical validation of the model. Although a subgroup of D^4 was discovered by the German sculptor Hildebrand [9] at the close of the 19th century, the full group was discovered (empirically, that is) only about a century later [24].

The effects of these transformations are very significant. Both the scaling and the rotation may have huge effects. They are nearly always detectable (but most often dominantly present) when two observers are compared.

In many experiments observers are compared by correlating depths at identical locations. Here "identical location" is well defined by the traces in the picture plane. Depths are defined only up to arbitrary additive constants; one conveniently sets the average depth to zero for each observer. It is common to find moderate (e.g., 0.5–0.9 range) coefficients of variations, but often values that are even lower and not significantly different from zero (no correlation) are found. Perhaps surprisingly, in the majority of cases a correlation modulo some element of D^4 will raise the coefficient of variation in the 0.9–1.0 range (values like 0.99 being no exception), even if the straight correlation proved insignificant. (Here a "correlation modulo some element of D^4" simply means a multivariate correlation involving the picture plane coordinates.) A typical example is shown in Figures 6 and 7.

These results are striking and yield a strong confirmation of the applicability of the model of $\mathbb{P}^3$ introduced above.

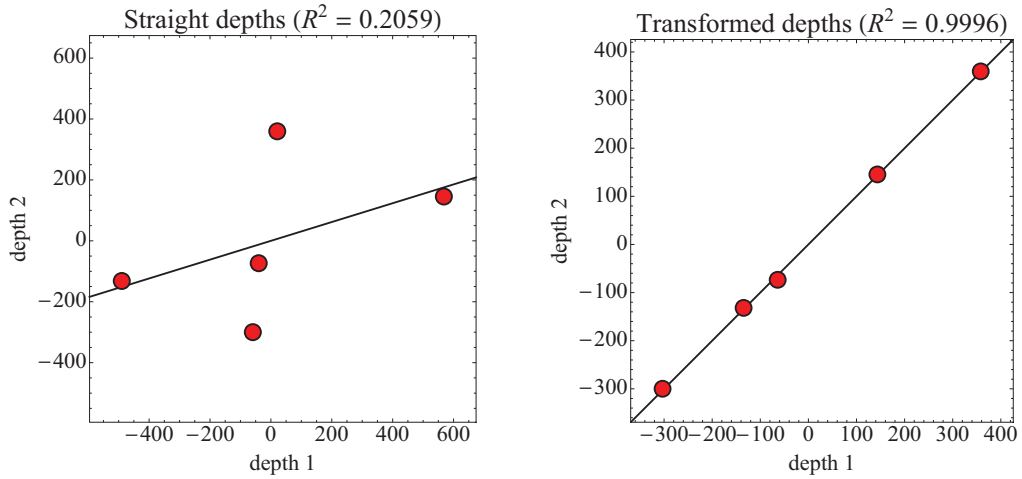

Figure 7. *Observer JK repeated the task discussed in the caption of Figure 6, this time looking obliquely at the picture (at 45° obliquity), once from the left, once from the right. The resulting depths correlate only weakly, as shown in the left scatter plot. A transformation from* D^4 *immediately restores excellent correlation, as shown at right. In this case the coefficient of variation changes from a mere* 0.21 *to* 1.00. *Without the gauge transformation one might easily have drawn erroneous conclusions on the basis of the raw results (left). Such cases are not at all rare in the literature.*

4. Space-variant gauges. The overall transformations (elements of D^4) are commonly encountered when the pictures are "obvious," such as photographs or realistic paintings. When the pictures are more "difficult" (see below), one sometimes finds results that call for more complicated deformations. In this section we consider general space-variant, smooth transformations.

The most general, smooth, space-variant moving frame, in terms of the standard coordinatization, is evidently

$$\begin{aligned} \mathbf{f_1} &= \mathbf{e_x} + f(x,y,z)\mathbf{e_z}, \\ \mathbf{f_2} &= \mathbf{e_y} + g(x,y,z)\mathbf{e_z}, \\ \mathbf{f_3} &= \mathbf{e_z} + h(x,y,z)\mathbf{e_z}, \end{aligned} \tag{3}$$

where f, g, and h are smooth real functions of $\{x,y,z\}$. Remember that $\{\mathbf{e_x},\mathbf{e_y}\}$ span the picture plane, whereas $\mathbf{e_z}$ spans the depth threads. The traces of $\{\mathbf{f_1},\mathbf{f_2},\mathbf{f_3}\}$ are obviously $\{\mathbf{e_x},\mathbf{e_y},\mathbf{0}\}$. (See Figure 8.)

We consider the differential geometry of pictorial space. The basis vectors $\{\mathbf{f_1},\mathbf{f_2},\mathbf{f_3}\}$ form a so-called Cartan moving frame [3]. The differential geometry analysis proceeds by considering the changes in the moving frame in terms of the moving frame itself. These changes are conveniently described in terms of the components of the "affine connection." The Christoffel symbols of the second kind [26, 17, 34] are defined such that $\Gamma_{ij}^{k}\mathbf{f_k}$ denotes the $\mathbf{f_k}$ component of the change in the basis vector $\mathbf{f_j}$ as $\mathbf{f_j}$ is moved over a unit distance along the $\mathbf{f_i}$ basis vector.[5]

[5]This defines the so-called covariant derivative of one vector by another [16], from which the Christoffel symbols derive much of their practical utility.

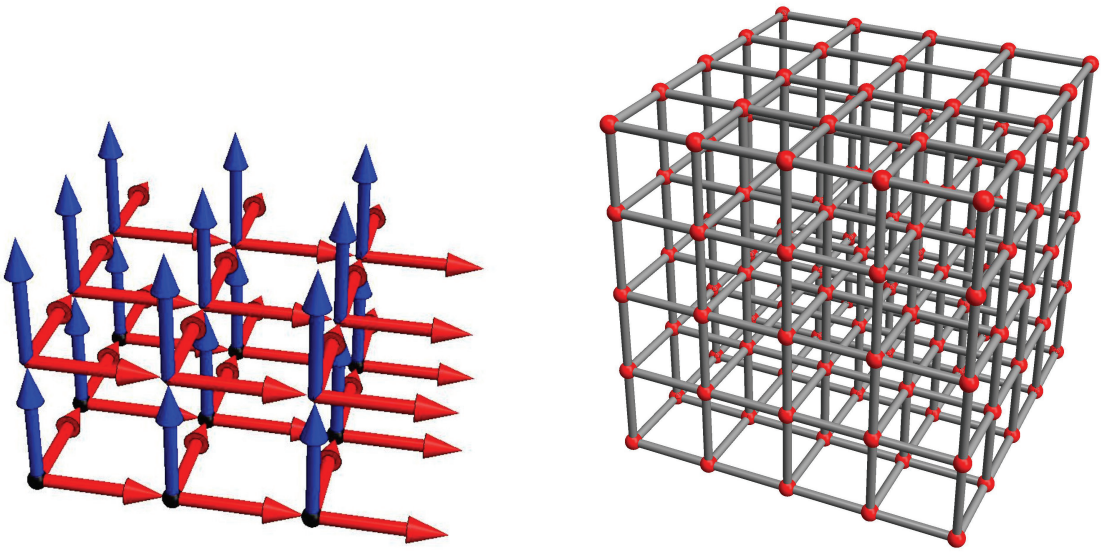

Figure 8. *At left the default frame field. The red vectors are a conventional Cartesian frame of the picture plane. The blue vectors are in the depth direction. In the default frame field the origins and unit points of the fibers are neatly coordinated. At right the tetragonal lattice defined by the frame vectors.*

Notice that all the changes are in the $\mathbf{f_3}$ direction; thus the components Γ^1_{ij}, Γ^2_{ij} $(i,j = 1,2,3)$ are identically zero. In order to describe the connection one needs only the components Γ^3_{ij} for $i,j = 1,2,3$. They can be found through differentiation of the moving frame equations (3). The nine nonvanishing components of the affine connection are given in Table 1.

Notice that we do not use a metric for this construction, as in the theory of Riemann spaces.[6] We describe the geometry purely in terms of the moving frame (equations (3)).

Table 1

The Christoffel symbols of the second kind Γ^k_{ij}. Many of the coefficients (those for $k = 1,2$) vanish identically.

$k=3$	$i=1$	$i=2$	$i=3$
$j=1$	$\frac{f_x}{1+h}$	$\frac{g_x}{1+h}$	$\frac{h_x}{1+h}$
$j=2$	$\frac{f_y}{1+h}$	$\frac{g_y}{1+h}$	$\frac{h_y}{1+h}$
$j=3$	$\frac{f_z}{1+h}$	$\frac{g_z}{1+h}$	$\frac{h_z}{1+h}$

Table 1 is clearly not symmetric. The components of the torsion tensor [16] are

$$T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji} - \gamma^k_{ij}, \quad i,j,k = 1,2,3, \tag{4}$$

where the $\gamma^k_{ij} = [\mathbf{f_i}, \mathbf{f_j}]$ are the commutator coefficients of the basis [16]. These commutator coefficients vanish in a holonomic (or coordinate) basis [16].

[6] The Christoffel symbols are familiar mainly from the theory of Riemannian spaces. There they are defined in terms of the Riemann metric. Here we do not use a metric, and we define the Christoffel symbols in terms of the moving frame. The difference is important; for instance, Riemannian spaces are torsion-free (see below).

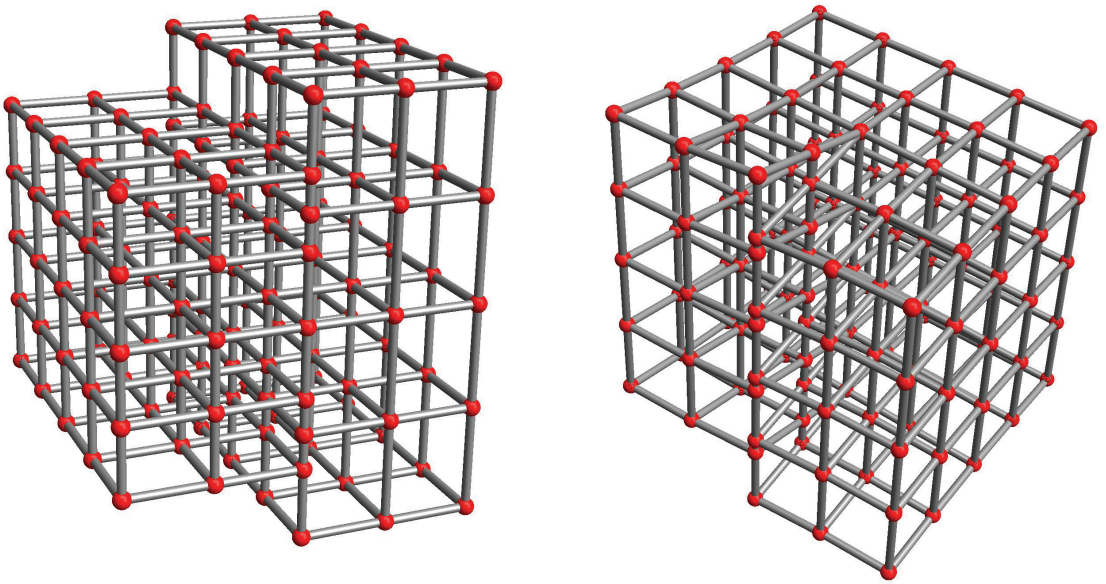

Figure 9. *The two types of dislocations defined by the torsion. At left a "horizontal" Burgers vector characterizes an "edge dislocation." At right a "vertical" Burgers vector characterizes a "screw dislocation." Here the dislocations are localized singularities. In the general case one has to reckon with a continuous distribution of edge and screw dislocations.*

Because so many of the Christoffel symbols vanish identically, all coefficients of the torsion tensor T^k_{ij} for $k = 1, 2$ also vanish identically. Moreover, due to its antisymmetry, the torsion tensor has only three independent coefficients. It is conveniently summarized through the "torsion vector"

$$\mathbf{t} = \frac{1}{1+h}\left((h_y - g_z)\mathbf{f_1} + (f_z - h_x)\mathbf{f_2} + (g_x - f_y)\mathbf{f_3}\right). \tag{5}$$

The torsion vector is similar to the Burgers vector in the theory of crystal lattice deformations [12, p. 599]. (See Figure 9.)

The component of the torsion vector in the depth thread direction describes so-called screw dislocations. It is $t_3 = g_x - f_y$; thus it vanishes if the functions f, g satisfy a Cauchy–Riemann condition [4, 33]. Then the vector fields $\mathbf{f_1}$, $\mathbf{f_2}$ are tangent to an integral surface.

We have extensive empirical data on smooth surfaces (or "reliefs") in pictorial space. These reliefs tend to be well-defined surfaces, even if measured through mutually independent, pointwise samples. Empirically, we have never encountered any evidence of nonintegrability of such samples. It seems likely that screw dislocations do not exist.

Since there is no empirical evidence for anything like "screw dislocations," it makes sense to require the Cauchy–Riemann condition of f, g; thus one wants to impose the constraint $g_x(x, y, z) = f_y(x, y, z)$. Then the $\mathbf{f_3}$ component of the torsion vector vanishes identically. The remaining dislocations are of the edge dislocation type. (See section 4.1.)

The space implied by the general model has both torsion and curvature. The Riemann

curvature tensor [16][7] is defined as

$$R^a_{bcd} = \left(\frac{\partial \Gamma^a_{bd}}{\partial x^c} - \frac{\partial \Gamma^a_{bc}}{\partial x^d} \right) + \sum_{e=1}^{3} \Gamma^a_{ec} \Gamma^e_{bd} - \sum_{e=1}^{3} \Gamma^a_{ed} \Gamma^e_{bc}, \tag{6}$$

where $a, b, c, d = 1, 2, 3$ and x^i denotes the ith coordinate. The Riemann curvature tensor has twelve (out of eighty-one) components that do not vanish identically. These depend critically on the z-dependence. The curvature vanishes if one imposes our basic constraint that the allowable transformations should respect the affine structure of the depth threads. Thus we may reckon with zero curvature in all cases of interest.

4.1. The "Lasagna model" of gauge fields in pictorial space. The most general model that conforms to the empirical constraints—the existence of reliefs and the affine structure of the depth threads—is

$$\begin{aligned} \mathbf{f_1} &= \mathbf{e_x} + F_x(x,y)\mathbf{e_z}, \\ \mathbf{f_2} &= \mathbf{e_y} + F_y(x,y)\mathbf{e_z}, \\ \mathbf{f_3} &= \mathbf{e_z} + h(x,y)\,\mathbf{e_z}, \end{aligned} \tag{7}$$

where F and h are smooth real functions of $\{x, y\}$. The first constraint is satisfied because the Cauchy–Riemann condition is enforced ($(F_x)_y = (F_y)_x$). The second constraint implies that h does not depend upon the z-coordinate. This again implements the basic constraint that the allowable transformations respect the affine structure of the depth threads. For $F(x, y) = \text{CONSTANT}$ and $h(x, y) = 0$, one regains the default space model.

The Christoffel symbols are collected in Table 2. Notice that there are only six nonvanishing coefficients, which reduce to four in the case when h is a constant function.

Table 2

The Christoffel symbols of the second kind Γ^k_{ij} in the case defined by equations (7). The coefficients for $k = 1, 2$, as well as those for $j = 3$, vanish identically.

$k=3$	$i=1$	$i=2$	$i=3$
$j=1$	$\frac{F_{xx}(x,y)}{1+h(x,y)}$	$\frac{F_{xy}(x,y)}{1+h(x,y)}$	$\frac{h_x(x,y)}{1+h(x,y)}$
$j=2$	$\frac{F_{xy}(x,y)}{1+h(x,y)}$	$\frac{F_{yy}(x,y)}{1+h(x,y)}$	$\frac{h_y(x,y)}{1+h(x,y)}$

The torsion vector is

$$\mathbf{t}(x,y) = \frac{1}{1+h(x,y)} \{h_y(x,y), -h_x(x,y), 0\}. \tag{8}$$

Notice that this torsion vector is always "horizontal." It describes a pure edge dislocation, there being no screw dislocation by design (the Cauchy–Riemann condition is forced).

So far we have found no empirical evidence for the existence of torsion. Thus the simplest structure that would suffice to explain the data has the additional constraint that $h(x, y) = \text{CONSTANT}$, thus forcing zero torsion.

[7] Due to the nonvanishing torsion, not all of the conventional symmetries (from the theory of Riemannian spaces) pertain.

The curvature tensor vanishes identically. This implies that the space defined by the frame field specified by equations (7) is *flat* (has a vanishing curvature tensor).

4.1.1. The frame field as a coordinate basis. From here on we will set $h(x,y) = H$, a constant. Notice that $\|\mathbf{f_3}\| = 1 + H = G$, say.

The frame field defined by equations (7) has a very simple geometrical structure; it is a (curvilinear) *coordinate basis*. We discuss it here in some detail because of its practical importance.

Consider the three functions

$$U(x,y,z) = x, \quad V(x,y,z) = y, \quad \text{and} \quad W(x,y,z) = z + F(x,y) + Hz. \tag{9}$$

Define curvilinear coordinates $\{u = U(x,y,z), v = V(x,y,z), w = W(x,y,z)\}$. One has

$$\begin{aligned} \mathrm{d}u &= \frac{\partial U}{\partial x}\mathrm{d}x + \frac{\partial U}{\partial y}\mathrm{d}y + \frac{\partial U}{\partial z}\mathrm{d}z = \mathrm{d}x, \\ \mathrm{d}v &= \frac{\partial V}{\partial x}\mathrm{d}x + \frac{\partial V}{\partial y}\mathrm{d}y + \frac{\partial V}{\partial z}\mathrm{d}z = \mathrm{d}y, \\ \mathrm{d}w &= \frac{\partial W}{\partial x}\mathrm{d}x + \frac{\partial W}{\partial y}\mathrm{d}y + \frac{\partial W}{\partial z}\mathrm{d}z = F_x\mathrm{d}x + F_y\mathrm{d}y + G\mathrm{d}z. \end{aligned} \tag{10}$$

Thus (using (7) and (10)),

$$\mathrm{d}u\,\mathbf{e_x} + \mathrm{d}v\,\mathbf{e_y} + \mathrm{d}w\,\mathbf{e_z} = \mathrm{d}x\,\mathbf{f_1} + \mathrm{d}y\,\mathbf{f_2} + \mathrm{d}z\,\mathbf{f_3}. \tag{11}$$

Apparently the coordinates $\{u,v,w\}$ induce the frame $\{\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3}\}$ as a coordinate basis.

The coordinate functions $U(x,y,z)$ and $V(x,y,z)$ trivially induce the Cartesian basis of the picture plane, so we do not have to consider these in detail. The coordinate function $W(x,y,z) = z + F(x,y) + Hz = F(x,y) + Gz$ encodes all the action. In case the function $F(x,y)$ is linear we regain the frame field of the homogeneous space considered earlier in this paper. The curvilinear coordinates generalize the structure of the homogeneous space in a natural manner.

The surface $\mathcal{O}$ defined by $\{x, y, F(x,y)\}$ and the parallel surface $\mathcal{I}$ defined by $\{x, y, F(x,y) + G\}$ define a space-variant gauge on pictorial space. Each thread $\{x,y\}$ meets these surfaces $\mathcal{O}, \mathcal{I}$ in two points, say $\mathbf{o}(x,y)$ on $\mathcal{O}$ and $\mathbf{i}(x,y)$ on $\mathcal{I}$. One takes the point $\mathbf{o}(x,y)$ as the origin and $\mathbf{i}(x,y)$ as the unit point on the thread $\{x,y\}$. The origin $\mathbf{o}$ evidently depends upon the location in the picture plane, whereas the unit (special) distance $\delta_{\mathbf{io}} = G$ is the same throughout. In this model the depth calibration varies smoothly from one depth fiber to the next.

In defining the surfaces $\mathcal{O}$ and $\mathcal{I}$ we have used the fact that absolute depth remains empirically undefined. This freedom allows one to consider the integral surface $F(x,y)$ as a section that defines the origin for each fiber, and the surface $F(x,y) + G$ as a section that defines the unit points.

More generally, the surfaces

$$\{U(x,y,k), V(x,y,k), W(x,y,k)\} = \{x, y, F(x,y) + kG\}, \tag{12}$$

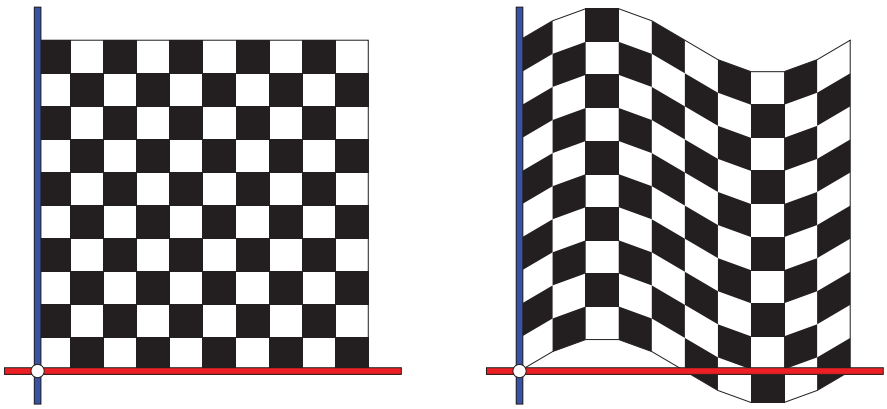

Figure 10. *These figures show a section of the XZ-plane. At left the tetragonal lattice of the default geometry. At right this has been deformed by a conformal transformation. The verticals are just the depth threads, and the (curvilinear) "horizontals" are the leaves of the lasagna model. Note that the cells are variously rotated, and all are scaled by the same factor (here, one). Thus all the cells in the deformed version are mutually congruent and similar to the cells of the default structure.*

for $k = -\infty, \ldots, -1, 0, +1, \ldots, +\infty$, foliate pictorial space, much like the surfaces of pasta in the conventional Italian dish lasagna.[8] They represent sections for the integral depth values $\ldots, -2, -1, 0, +1, +2, \ldots$. Hence we refer to this model of the geometry of pictorial space as the "lasagna model."

The surface $\mathcal{O}$, that is, $\{x, y, F(x, y)\}$, determines the field of local rotations by $\{F_x, F_y\}$ of the default frontal planes $z(x, y) = C$ to $z(x, y) = C + xF_x + yF_y$ for some arbitrary constant C. Thus, one way to understand the surface $\mathcal{O}$ geometrically is to regard it as the *apparent frontoparallel plane* in its attitude changes from place to place in the picture plane. The surfaces $\{x, y, C+F(x, y)\}$ replace the frontoparallel "planes." In the simple homogeneous geometry the function F is linear; thus the surface $\mathcal{O}$ is planar, and the apparent frontoparallel plane is simply an oblique plane.

The curves on the surfaces $\{x, y, C + F(x, y)\}$ that have straight traces in the picture plane may be considered frontoparallel pregeodesics (geodesics without parameterization). The depth threads themselves may also be considered pregeodesics, since they are autoparallels.[9] This yields a simple geometrical overview of the pregeodesics, or perceptual "straight lines." They are the images of the straight lines in the default geometry under the conformal transformations imposed by the lasagna model (see below).

The spacing of the surfaces $\mathcal{O}$ and $\mathcal{I}$, that is, $G\mathbf{e_z}$, determines an overall dilation of the second kind. For the default homogeneous geometry case one simply has $G = 1$; thus $H = 0$.

The unit cells defined by the default structure are identical cubes in identical attitudes. The gauge field deforms them into rotated and scaled cubes. Thus the gauge field imposes a *conformal transformation* on pictorial space. This characterizes the lasagna model. It considers conformal deformations of the default geometry (Figure 10). The surfaces in the lasagna

[8] Also "lasagne." The word possibly derives from an ancient Greek dish $\lambda\acute{\alpha}\gamma\alpha\nu o\nu$.

[9] Autoparallels are curves such that the tangents at different points of the curve are related by parallel transport along the curve. Indeed, one easily checks that $\nabla_{\mathbf{f_3}}\mathbf{f_3} = \mathbf{0}$.

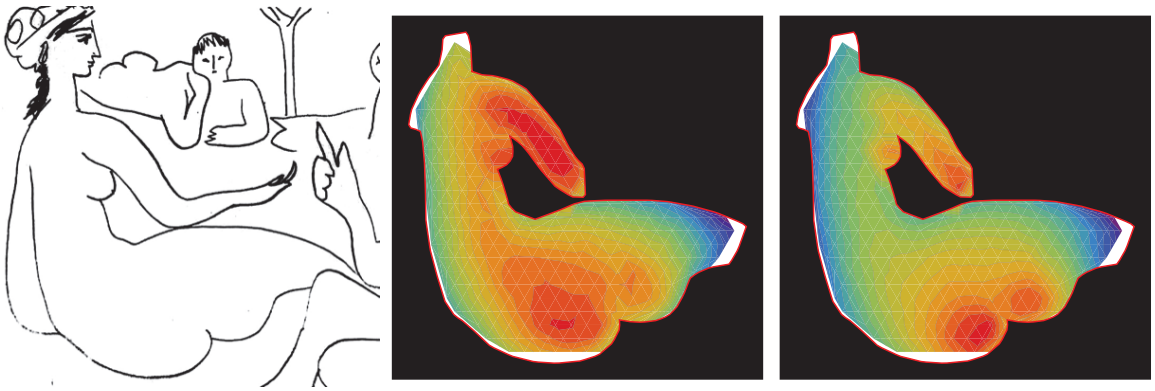

Figure 11. *At left a cutout from a Picasso drawing. At center and right two depth maps for observer JK obtained on different days. Note that these depth maps are different. On comparison one notices local mutual rotations.*

model do have an immediate intuitive content. They visually specify how neighboring parts of space are mutually related. Each (small enough) part is like the homogeneous geometry of the default model, but the parts are mutually related through rotations and dilations of the second kind. The most useful visualization of the structure is through the pair of surfaces $\mathcal{O}$, $\mathcal{I}$, a "gauge field."

4.1.2. Empirical validation of the lasagna model. In the pointing method one selects a point pair $\{\mathcal{P}, \mathcal{Q}\}$ and points once from $\mathcal{P}$ to $\mathcal{Q}$, then from $\mathcal{Q}$ to $\mathcal{P}$. One finds that (in typical cases) the two pointings fail to be collinear. In the default model they determine a unique parabolic arc with the observed directions as limiting tangents. In the lasagna model such an arc may be regarded as the autoparallel that connects the points in pictorial space. The arc specifies a depth difference, which is the basis for the measurement tool mentioned in section 3.1.

We find that in the large majority of cases the empirical pregeodesics (as described in terms of the default model) are significantly curved. The degree of curvature varies quite a bit over observers, but is consistent for a given observer in a given session. We conclude that deviations often occur from the simple homogeneous model.

A strong test of the lasagna model is to use pictures considered "difficult" by typical human observers. We will show some results for a drawing by Picasso that shows a human body in a physically impossible pose (Figure 11 (left)). In these drawings Picasso tries to escape from central projection in that he combines views from different vantage points. Picasso simultaneously aims at an interesting composition in the picture plane. The drawing is evidently (and no doubt intentionally) very different from a generic view of a model. The drawing breaks with the "picture as a window on the world" convention.

It is hardly surprising that one encounters large differences between observers. In Figures 11 (center) and 11 (right) we show two depth maps obtained for a single observer JK (one of the authors) on different days. A scatter plot of the depths for two sessions at corresponding points (same trace in the picture plane) reveals a low though significant correlation (Figure 12 (left); $R^2 = 0.73$). In this case the differences cannot be fully accounted for by an

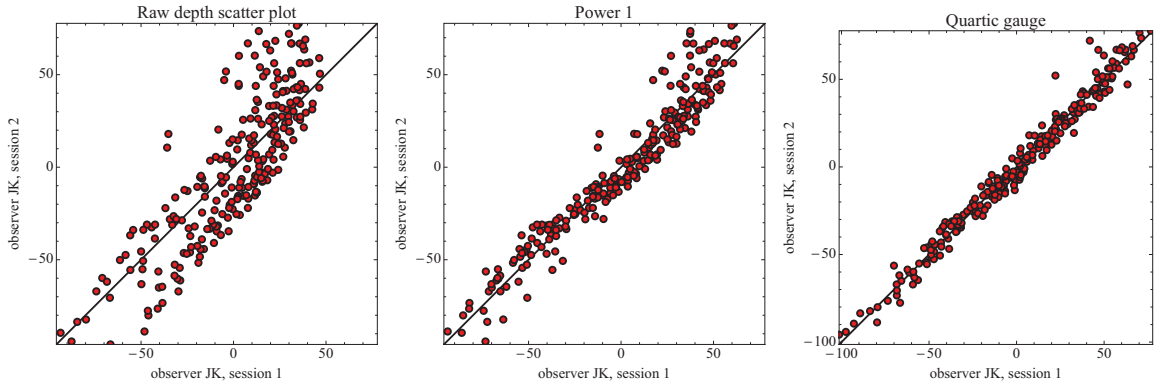

Figure 12. *Scatter plots of depths at corresponding points of the reliefs shown in Figures 11 (center) and 11 (right). At left a straight scatter plot. Notice the low correlation ($R^2 = 0.73$). At center a global transformation from D^4 reveals some correlation, although such a global transformation cannot remove all scatter. The coefficient of variation is raised to 0.93. At right a quartic (see text) conformal transformation according to the lasagna model reveals excellent agreement between the two reliefs ($R^2 = 0.98$).*

element of D^4, although this certainly removes much of the difference (Figure 12 (center)). The coefficient of variation is raised to 0.93.

In an attempt to account for the persisting differences by a suitable lasagna structure, we applied the nonlinear transformation

$$z_B(x, y) = h\, z_A(x, y) + B_k(x, y), \tag{13}$$

where h is a constant and $B_k(x, y)$ is a polynomial in $\{x, y\}$ of maximum degree k. In this case we used a quartic polynomial. The correlation now becomes very high (Figure 12 (right); $R^2 = 0.98$), albeit at the cost of a lasagna-type transformation between the two pictorial spaces. (See Figure 13.) A chi-square test reveals that the lasagna model does indeed do an excellent job of explaining the individual differences. Apparently the observer JK has somewhat similar pictorial spaces in the two sessions, but these are related not by a global similarity but by a certain conformal transformation that changes from place to place. The observer apparently perceives the relative spatial attitude of the major rigid body parts (rib cage, pelvic girdle, upper and lower extremities, head) somewhat differently in the two sessions.

5. Conclusions. Pictorial space is the intended end result of any visualization process. The end results of computer graphics, for instance, are not the pictures per se, but the images these evoke in the minds of their observers. Ideally one would design for this end result. Unfortunately, this ideal is still far off. One reason is that the set of prospective observers is far from homogeneous. Another reason is that pictorial space is outside the realm of entities for which the computer graphics engineer possesses the necessary formal tools.

The former problem is common enough in applications involving human users. It is usually solved "for all practical purposes" by

- defining a "standard" generic user,
- formally describing the nature of the variation typically encountered in the intended audience, and
- ignoring singular cases and outliers.

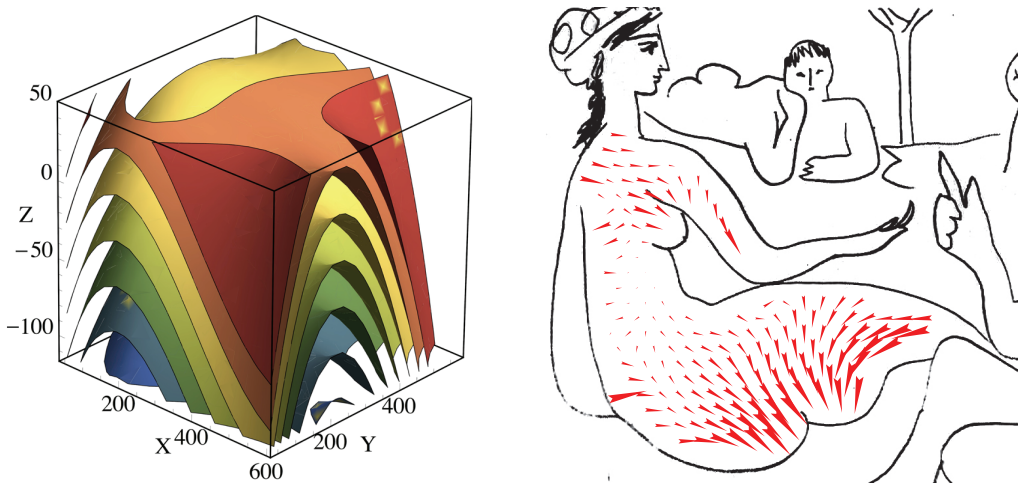

Figure 13. *At left the leaves of the lasagna foliation of pictorial space for the best quartic transformation. At right the local rotations (the vector* $\mathbf{g} = \{g_x, g_y\}$ *field (2) plotted in the picture plane) as calculated from the depth variations in small neighborhoods. These can be correlated with the local slopes of the lasagna leaves.*

This is how one currently deals with audio quality, image quality, and so forth. There is no reason why pictorial space should not be handled in a similar way. Methods that allow one to probe the structure of visual space in human observers have been developed and are easy to use. What is still lacking is an extensive investigation of a group of "typical" observers. It is perfectly possible to run such an investigation; i.e., all necessary tools are in place. It implies a major investment, though.

The latter problem has been the topic of this paper. We have proposed a toolbox of geometrical methods that allows one to handle the structure of pictorial spaces in a formal way. Most of these tools are simple. They can easily be applied to the empirical observations. In most cases one will want to reformulate them as statistical methods. This is straightforward enough in virtually all applications. For instance, fitting a lasagna-type gauge transformation amounts to a multiple correlation using a nonlinear model (although often a linear model will suffice).

As we said in the introduction, the computer graphics engineer (advertisement designer, photographer, lecturer, or various other professionals) ultimately designs for the visual awareness of the observer. The "end result" is neither the picture nor the designer's visual experience of it, but rather what will be in the minds of its *potential observers*.

What does it take to "design for the end result?" Among other requirements, one must

- be able to predict the realm of structures of the pictorial spaces that will be evoked in the visual awareness of the intended audience;
- be able to measure the structure of pictorial spaces in specific cases;
- be able to find the necessary changes in the picture that will induce intended changes in the viewer's visual awareness;
- and be able to understand the range of likely visual experiences for the population at large.

Perhaps one might add a few more topics, but all are likely to require similar tools.

As compared to such (equally important!) issues as (basic) "image quality," the possibility of engineering for the structure of pictorial spaces is still in a very preliminary state. Visual artists are far ahead of the engineers, but their knowledge is partly implicit and certainly hard to mine. So far we have designed effective methods for the quantitative measurement of various geometrical properties. In this paper we have proposed the required formal tools. A major component that is still lacking is the collection of serious databases of baseline data over the generic population. The tools that allow one to do so are now in place.

Acknowledgment. We thank Johan Wagemans for his advice, inspiring discussions, and critical reading of the text.

REFERENCES

[1] P. N. Belhumeur, D. Kriegman, and A. Yuille, *The bas-relief ambiguity*, Int. J. Comput. Vision, 35 (1999), pp. 33–44.
[2] G. Berkeley, *An Essay Towards a New Theory of Vision*, printed by Aaron Rhames at the back of Dick's Coffee-House for Jeremy Pepyat, Bookseller in Skinner-Rows, Dublin, Ireland, 1709.
[3] E. Cartan, *La théorie des groupes finis et continus et la géométrie différentielle traitées par la méthode du repère mobile*, Gauthier-Villars, Paris, 1937.
[4] A. L. Cauchy, *Mémoire sur les intégrales définies*, Oeuvres complètes Ser. 1, Vol. 1, Gauthier-Villars, Paris, 1882, pp. 319–506.
[5] A. Cayley, *A sixth memoir upon the quantics*, Philos. Trans. Royal Soc. London, 149 (1859), pp. 61–90.
[6] H. S. M. Coxeter, *Introduction to Geometry*, Wiley Classics Lib., John Wiley & Sons, New York, 1989.
[7] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA, 1993.
[8] D. Hilbert and S. Cohn-Vossen, *Geometry and the Imagination*, 2nd ed., Chelsea, New York, 1952.
[9] A. von Hildebrand, *Das Problem der Form in der bildenden Kunst (The Problem of Form in Painting and Sculpture)*, translation of 1893 original by M. Meyer and R. M. Ogden, G. E. Stechert, New York, 1945.
[10] B. K. P. Horn, *Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View*, Ph.D. thesis, Department of Electrical Engineering, MIT, Cambridge, MA, 1970.
[11] D. Husemoller, *Fiber Bundles*, Grad. Texts in Math. 20, Springer, London, 1993.
[12] C. Kittel, *Introduction to Solid State Physics*, 3rd ed., John Wiley & Sons, New York, 1967.
[13] F. Klein, *Über die sogenannte nicht-Euklidische Geometrie*, Math. Ann., 6 (1871), pp. 112–145.
[14] F. Klein, *Vergleichende Betrachtungen über neuere geometrische Forschungen*, Math. Ann., 43 (1893), pp. 63–100.
[15] F. Klein, *Vorlesungen über nicht-Euklidische Geometrie*, Springer, Berlin, 1928.
[16] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Wiley-Interscience, New York, 1963.
[17] J. J. Koenderink, *Solid Shape*, MIT Press, Cambridge, MA, 1990.
[18] J. J. Koenderink, L. Albertazzi, A. J. van Doorn, W. A. van de Grind, A. M. L. Kappers, J. L. Lappin, J. F. Norman, A. H. J. Oomes, S. P. te Pas, F. Phillips, S. C. Pont, W. A. Richards, J. T. Todd, R. van Ee, F. A. J. Verstraten, and S. de Vries, *Does monocular visual space have planes?*, Acta Psychologica, 134 (2010), pp. 40–47.
[19] J. J. Koenderink, and A. J. van Doorn, *Relief: Pictorial and otherwise*, Image Vision Comput., 13 (1995), pp. 321–334.
[20] J. J. Koenderink and A. J. van Doorn, *Pictorial space*, in Looking into Pictures: An Interdisciplinary Approach to Pictorial Space, H. Hecht, R. Schwartz, and M. Atherton, eds., MIT Press, Cambridge, MA, 2003, pp. 239–299.
[21] J. J. Koenderink and A. J. van Doorn, *The structure of visual spaces*, J. Math. Imaging Vision, 31 (2008), pp. 171–187.

[22] J. J. Koenderink, A. J. van Doorn, and A. M. L. Kappers, *Surface perception in pictures*, Perception & Psychophysics, 52 (1992), pp. 487–496.
[23] J. J. Koenderink, A. J. van Doorn, and A. M. L. Kappers, *Pictorial surface attitude and local depth comparisons*, Perception & Psychophysics, 58 (1996), pp. 163–173.
[24] J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, and J. T. Todd, *Ambiguity and the 'mental eye' in pictorial relief*, Perception, 30 (2001), pp. 431–448.
[25] J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, and J. T. Todd, *Pointing out of the picture*, Perception, 33 (2004), pp. 513–530.
[26] C. W. Misner, K. Thorne, and J. A. Wheeler, *Gravitation*, Freeman, New York, 1973.
[27] L. A. Necker, *Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solids*, Philosophical Magazine, 1 (1832), pp. 329–337.
[28] S. E. Palmer, *Vision Science: Photons to Phenomenology*, MIT Press, Cambridge, MA, 1999.
[29] M. H. Pirenne, *Optics, Painting, and Photography*, Cambridge University Press, London, 1970.
[30] H. Pottmann and Y. Liu, *Discrete surfaces in isotropic geometry*, in Mathematics of Surfaces 2007, Lecture Notes in Comput. Sci. 4647, R. Martin, J. Sabin, and J. Winkler, eds., Springer-Verlag, Berlin, Heidelberg, 2007, pp. 341–363.
[31] H. Pottmann and K. Optitz, *Curvature analysis and visualization for functions defined on Euclidean spaces or surfaces*, Comput. Aided Geom. Design, 11 (1994), pp. 655–674.
[32] V. S. Ramachandran, *Perception of shape from shading*, Nature, 331 (1988), pp. 163–166.
[33] B. Riemann, *Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen komplexen Grösse*, in Riemann's gesammelte mathematische Werke, H. Weber, ed., Dover, New York, 1953, pp. 3–48.
[34] G. F. B. Riemann, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, in Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Vol. 13, 1867; *On the hypotheses which lie at the bases of geometry*, Nature, 8 (1854), pp. 14–17, 36, 37; translated by W. K. Clifford.
[35] H. Sachs, *Ebene Isotrope Geometrie*, Friedrich Vieweg & Sohn, Braunschweig, Germany, 1987.
[36] H. Sachs, *Isotrope Geometrie des Raumes*, Friedrich Vieweg & Sohn, Braunschweig, Germany, 1990.
[37] K. Strubecker, *Differentialgeometrie des isotropen Raumes.* I. *Theorie der Raumkurven*, Akad. Wiss. Wien, S.-B. IIa, 150 (1941), pp. 1–53.
[38] K. Strubecker, *Differentialgeometrie des isotropen Raumes.* II. *Die Flächen konstanter Relativkrümmung* $K = rt - s^2$, Math. Z., 47 (1942), pp. 743–777.
[39] K. Strubecker, *Differentialgeometrie des isotropen Raumes.* III. *Flächentheorie*, Math. Z., 48 (1942), pp. 369–427.
[40] K. Strubecker, *Differentialgeometrie des isotropen Raumes.* IV. *Theorie der flächentreuen Abbildungen der Ebene*, Math. Z., 50 (1944), pp. 1–92.
[41] A. J. van Doorn, J. J. Koenderink, and J. Wagemans, *Rank order scaling of pictorial depth*, *i*-Perception, 2 (2011), pp. 724–744.
[42] A. J. van Doorn, J. Wagemans, H. de Ridder, and J. J. Koenderink, *Space perception in pictures*, in Proceedings of the SPIE, Vol. 7865, 2011, 786519.
[43] J. Wagemans, A. J. van Doorn, and J. J. Koenderink, *Measuring* 3*D point configurations in pictorial space*, *i*-Perception, 2 (2011), pp. 77–111.
[44] J. Wagemans, A. J. van Doorn, and J. J. Koenderink, *Pictorial depth probed through relative sizes*, *i*-Perception, 2 (2011), pp. 992–1013.
[45] L. J. J. Wittgenstein, *Tractatus Logico-Philosophicus*, Kegan-Paul, London, 1922.
[46] I. M. Yaglom, *A Simple Non-Euclidean Geometry and Its Physical Basis*, Springer, New York, 1979.
[47] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, *Shape from shading: A survey*, IEEE Trans. Pattern Anal. Mach. Intell., 21 (1999), pp. 690–706.