Global structure of visual space as a united entity

Tarow Indow*

Department of Cognitive Sciences and Institute for Mathematical Behavioral Sciences,
University of California, Irvine, CA 92697, USA

Abstract

Visual space (VS) is the final product of the series of processes, physical, physiological, and perceptual. Two planes in VS are discussed. Though each can be regarded as Riemannian space of constant Gaussian curvature $K$, a plane passing through the eyes that extends in the depth direction is hyperbolic $(K < 0)$ whereas a plane appearing fronto-parallel is Euclidean $(K = 0)$. Empirical and mathematical bases for this structure of VS are presented. How VS is related to the stimulus condition is complex and dynamic, but the intrinsic structure of VS per se is homogeneous in the sense that we can see global congruence and/or similarity between figures, provided the corresponding physical objects are appropriately adjusted. All percepts appear at finite distances and VS is closed. Perceptual structure of natural scene, e.g. appearance of the horizon, is discussed. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Visual space; Geometry; Luneburg’s model

1. Global structure of visual space

The visual space is the final product of the series of processes, physical, physiological, and perceptual (Fig. 1). The following abbreviations will be used throughout in this article.

- $E$: a Euclidean space.
- $R$: a Riemannian space in which the Gaussian total curvature $K$ is a constant. $R$ is elliptic if $K > 0$, Euclidean $E$ if $K = 0$, and hyperbolic if $K < 0$.
- $VS$: visual space regarded as an $R$.
- $\delta$: a perceptual distance in VS (a latent variable).
- $d$: a scaled value representing $\delta$ (a manifest variable).

*Tel.: +1-714-824-6218.
E-mail address: tindow@uci.edu (T. Indow)

0165-4896/99/$ – see front matter © 1999 Elsevier Science B.V. All rights reserved.
PII: S0165-4896(99)00026-8
physical space as an E from which light stimulus comes.
e a physical distance in X.

EM Euclidean map of VS. If VS is not an E \((K \neq 0)\), we cannot depict VS in a sheet of paper (E) without distortion. In the present article, VS as an R will be represented in an E in the way of Poincaré, and this representation will be denoted as EM (Section 2).

\(\rho\) a Euclidean distance in EM.

\(\{Q_j\}\) a configuration of stimulus points \(Q_j\) in X, \(j = 1, \ldots, n\).

\(\{P_j\}\) a configuration of perceived points \(P_j\) in VS, \(j = 1, \ldots, n\). The same symbol will be used for its representation in EM.

Whenever necessary, the dimensionality of a space, \(m\), will be designated by the superscript of the symbol of space such as \(X^3\), \(VS^2\).

Features of VS were discussed in previous articles (Indow, 1991, 1997). Fig. 1 depicts perception of an outdoor scene. The perceived scene is an organization due to multiple glances, but it is stable (Howard, 1991) and we take it for granted that, when one stands

![Fig. 1. Visual space (VS) and physical space (X). A natural scene. Reproduced (with a slight modification) with permission from Indow (1991).](Fig. 1. Visual space (VS) and physical space (X). A natural scene. Reproduced (with a slight modification) with permission from Indow (1991).)
at the same place, everyone will have VS of the same macroscopic structure. In the present article, a mathematical model on this structure will be presented. The model is purely phenomenological and no discussion will be made about the physical and/or physiological processes underlying this perception. Features of VS relevant to the subsequent discussion are as follows.

1.1. VS 1

VS as a total entity extends around the self in three directions, namely VS³. The self is a percept as the final product of the long series of processes mainly stemming from proprioceptive stimulations in the body in X. When we say ‘P appears far to the right direction’, for example, the distance and direction are meant from the self as the origin of VS. Similarly, the body plays the role of origin of X. Respective origins will be written as O, and radial distance from O will be denoted by δ₀ in VS (its scaled value by d₀) and by e₀ in X. Its representation in EM will be written as r₀. Non-radial distances between two points, P₁ and P₂, in VS as well as in EM will be, respectively, denoted by δjk (its scaled value by djk) and by ejk, or simply by δk (d) and e. Corresponding distances between two stimulus points, Qj and Qk, in X are written as ejk or e.

1.2. VS 2

VS is closed. At the end of our line of sight, there is always a percept appearing at a finite distance. That is to say, we cannot see ‘void’ and ‘infinity’. In Fig. 1, we see a ship on the ocean, an island at the horizon, and the sky above. The sky and ocean meet at the horizon, and the horizon always appears in VS at the height of eye-level of the self, irrespective of whether the eyes are directed upward or downward. The perceptual distance to the boundary of VS at a given direction ω will be denoted by max δ₀(ω).

1.3. VS 3

VS is dynamic. If the eyes are exposed to homogenous light of a sufficiently low intensity, one feels as if being surrounded by fog of light (Metzger, 1930; Avant, 1965). In order to have a structured VS, the image of patterns or the texture of a surface in X must be formed on the retina. Under ordinary conditions, VS consists of individual percepts, some are close and some are far. In the vicinity of the self, the patterns we see in VS are veridical so that we can guide our body to reach, manipulate, or avoid objects in X. This isomorphic correspondence in pattern between X and VS does not hold for the entire range. In the outdoor scene as shown in Fig. 1, the distance to the perceived sky is max δ₀ and the corresponding distance in X will be called the max effective e₀. If an airplane flies away in Fig. 1, the perceived plane in VS behaves in that way when e₀ to the plane is not too large. When e₀ reaches the max effective e₀, δ₀ to the airplane in VS cannot increase anymore and the plane if visible or its vapor trail will go down along the perceived sky. Most people say that the sky appears as a vault flattened in the zenith direction, which means max δ₀(ω) changes according to direction ω. When we are in a room, VS in that direction of wall ends at the perceived wall. However, the physical
distance $e_0$ to the wall is smaller than the max effective $e_0$ and the perceived distance $\delta_0$ to the wall is not what is called max $\delta_0$. Though invisible, max $\delta_0$ may be somewhere beyond the perceived wall. If a tree is visible in a window, it appears far than the window. How max $\delta_0$ changes according to the stimulus configuration in indoor frameless spaces was discussed before (Indow, 1991, 1997).

2. VS as a Riemannian space $\mathbb{R}$

As to the structure of VS, most people have either one of the following intuitive opinions. Some regard it obvious that VS is Euclidean whereas others think that VS is too elusive to be formulated in terms of any mathematically well-established geometry. On the other hand, we take it for granted that we can see the course of a moving ball and catch or fit it according to this perception and that everyone sees the same constellations in the curved surface of night sky. Most studies of visual perception are concerned with local phenomena in VS without referring to the geometry of VS. Though under a restricted condition, Luneburg (1947, 1950), a geometer, was the first to propose the idea that a part of VS is an $\mathbb{R}$ having a constant curvature $K$. He tried to account for the results of Ames demonstration and alley experiments. The hypothesis was reiterated by Hardy et al. (1953) and by Blank (1953), a theoretical physicist etc. More recent articles are Chapter 13 in Suppes et al. (1989), Indow (1991) and Lukas (1996). The basic conditions that lead to this hypothesis will be explained in Section 4.

Mathematically, there are many geometries and spaces of different properties. Intuitive tree structures of these are given, for instance, in Shepard (1974) and Laugwitz (1965). Riemannian space is a metric space that is locally Euclidean. To say that VS is a metric space means that the length $d$ a subject sees between any two points in VS, satisfies the Fréchet’s conditions

$$\delta_{ik} \sim \delta_{ij} > 0, \quad \delta_{ji} > 0$$ (symmetry and non-negative)

$$\delta_{ij} \oplus \delta_{jk} > \delta_{ki}$$ when three points are not collinear (triangular inequality)

$$\delta_{ij} \oplus \delta_{jk} \sim \delta_{ki}$$ if three are collinear in the order, $i, j, k$ (additivity)

where ‘$\sim$’, ‘$>$’, and ‘$\oplus$’ respectively means that ‘to appear equal to, longer than, and to be concatenated’. Because $\delta$s are latent variables that cannot be observed by anybody else than the subject, it is impossible to present the direct proof that $\delta$s satisfy these conditions. Many will agree that the conditions are not contradictory to our perceptual experiences. We can ask the subject to make judgments on $\delta$s to have scale values for $\delta$s. Designate any three points, $Q_i$, $Q_j$, and $Q_k$ in $X$ and ask the subject to assign such a numeral $r_{i,j,k}$, representing the ratios of perceptual distances from $P_i$ to $P_j$ and from $P_j$ to $P_k$, $r_{i,j,k} = \delta_{ij}/\delta_{ik}$. Because these are lengths of two segments starting from a common point $P_i$ in VS, the subject, especially naive one, does not hesitate to say such that ‘$r_{i,j,k} = 0.4$ or 1.5, etc.’, provided the ratio is within a certain range, say 0.2 to 5.0. That is true even when the self is counted as a point $P_i$ and radial distance $\delta_0$ is included. We can select such triplet, $Q_i$, $Q_j$, $Q_k$ that $r_{i,j,k}$ does not exceed the range in which the
subject feels comfortable in assessing the ratios. If these ratio assessments are systematically made for a configuration of \( n \)-points \( \{Q_j\} \), we can have scale values \( d_{jk} \) with an arbitrary common unit that reproduce data \( r_{ij} \) by the ratio \( d_{ij}/d_{jk} \) with sufficient accuracy (Indow and Ida, 1975; Indow, 1982). Furthermore, these \( d_s \) were shown to satisfy the Fréchet’s conditions with respect to ‘=’, ‘>’, and ‘+’. Plots showing this fact are given in Figs. 1 and 2 in Indow (1968), Fig. 4 in Indow and Watanabe (1988), and Fig. 4 in Indow (1990). Namely, we can regard \( d \) as a proportional representation of latent variable \( \delta \). Fig. 13 in Indow (1991) provides a more direct evidence that \( d_i + d_j = d_k \) for three perceptually collinear points in VS. Based on these findings, let us regard VS to be a metric space. The condition that VS is locally Euclidean is not against our intuition, though we have no way to directly test it. Hence, let us assume that VS is a Riemannian space.

The geometrical properties of any region in an R are characterized by the Gaussian total curvature \( K \) therein. In a general Riemannian space, the value of \( K \) can vary from region to region. A physical object can move from one position to another in the physical space without changing its size and it is called the Helmholtz-Lie problem to conclude that the free mobility is not possible unless \( K \) is constant in all regions of this space. The sign of \( K \) cannot be specified by the free mobility condition. Namely, the physical space is an R. As to X in which visual stimuli are, we can safely regard it as an E (\( K = 0 \)). As to VS, the free-mobility condition is more subtle, because we have to deal with the invariance of a perceived figure. In order to keep the size of a figure moving away from the self in VS, the size of its physical counterpart in X must be appropriately adjusted according to \( e_a \). The free mobility condition can be interpreted in various ways. This problem and the implication of the fact that congruence and similarity is possible in subspaces of VS will be discussed in Section 4. Luneburg did not make explicit in which part of VS he assumed it to be an R. Let us proceed with the assumption that VS consists of such subspaces that can be regarded as an R.

3. Theoretical equations in EM and mapping functions

Riemannian space of constant curvature R is of three kinds: hyperbolic (\( K > 0 \)), Euclidean E (\( K = 0 \)), and elliptic (\( K > 0 \)). It is well known how to represent the structure of \( \mathbb{R}^2 \) of \( K \neq 0 \) in a 3-D Euclidean space \( \mathbb{E}^3 \): a sphere for \( K > 0 \) and a saddleback for \( K < 0 \). We can map \( \mathbb{R}^m \) in \( \mathbb{E}^m \), \( m = 2, 3 \), without adding an extra-dimension. For our purpose, the most convenient way of mapping is one given by Poincaré (e.g. 19.6 in Berger, 1977). This representation of VS\(^m\) will be called the Euclidean map and abbreviated as EM\(^m\). The logic underlying this representation was explained in Indow (1979). The following symbols will be used in addition to \( \{P_j\} \), \( \rho, d \) defined in Section 1 (see the right side in Fig. 1, and Fig. 2 II to be shown in Section 4).

\( (x, y, z) \) Cartesian coordinates of a point \( Q \) in \( X^3 \). The origin O is the midpoint between the two eyes.

\( (\gamma, \phi, \theta) \) bipolar coordinates of \( Q \) in \( X^3 \) in terms of the reference system having the
right and left eyes of subject as origins. The variables respectively represent the bipolar parallax, the bipolar latitude, and the angle of elevation.

\((\xi, \eta, \zeta)\) Cartesian coordinates of \(P\) in \(\text{EM}^3\). The origin \(O\) represents the self in \(\text{VS}\). The variables respectively represent directions, from near to far (depth), from left to right (latitude), and from above to below (elevation).

\((\rho_0, \varphi, h)\) polar coordinates of \(P\) in \(\text{EM}^3\). The variables respectively represent the Euclidean radial distance from \(O\) and the latitude and elevation angles.

D-plane a plane \(X^2\) that is slanted with an angle \(\theta\) (spanned by the \(y\)- and \(x'_y\)-axes), and a plane \(\text{VS}^2\) that extends from the self in the depth direction with an elevation angle \(\vartheta\).

H-curve a curve or plane appearing fronto-parallel in \(\text{VS}\).

or plane \(\text{BC}\) Basic circle in \(\text{EM}^2\) with an angle \(\vartheta\) that represents, when \(K<0\), perceptual radial infinity \((\delta_0=\infty)\). This has no counterpart in \(\text{VS}\) \((\text{max } \delta_0 < \infty)\), but plays an important role in defining theoretical curves.

When \(K<0\), the entire \(\text{VS}^3\) is represented within the \(\text{BC}\), because \(\text{max } \delta_0\) is finite. The relationship between \(\text{VS}\) and \(\text{EM}\) is not isometric but conformal. A non-radial straight line in \(\text{VS}\) is represented in \(\text{EM}\) by a circle orthogonal to the \(\text{BC}\) when extended and a radial straight line in \(\text{VS}\) by a radial line that is also orthogonal to the \(\text{BC}\) when extended. Two distances, \(\delta\) in \(\text{VS}\) and the Euclidean length \(\rho\) of the corresponding curve or radial line in \(\text{EM}\), are not proportional (non-isometric). On the other hand, angles in \(\text{VS}\) can be preserved in \(\text{EM}\) without any distortion (conformal). All theoretical equations for experimental results are formulated in \(\text{EM}\). In order to fit these equations to experimental data, \(\{Q_j\}\) in \(\text{X}\), the theoretical curves must be mapped from \(\text{EM}\) to \(\text{X}\).

Luneburg assumed a simple functional relationship between variables in \(\text{X}^3\) and \(\text{EM}^3\).

\[
\rho_0 = g(\gamma; \sigma) = 2 \exp\{-\sigma \gamma\}, \quad \varphi = \phi, \quad \vartheta = \theta
\]

These will be called the Luneburg’s mapping functions, where \(\sigma\) is an individual constant as \(K\) is. These forms are a priori assumptions. The characteristic of this mapping may be called completely ego-centric and rigid. The position of \(P\) in \(\text{EM}\) and hence how far and in what direction the point is perceived with regard to the self in \(\text{VS}\) are solely determined by its relation to the body, \((\gamma, \phi, \theta)\) of \(Q\), regardless of the conditions of remaining points in \(\text{X}\). Should this context-free mapping hold, it must be only under a frameless situation. This is the reason that all the relevant experiments were performed with small light points in the dark or small objects on a table at the eye-level \((\theta=0)\) where the edge of the table and the wall of the room were kept invisible. Even under this condition, however, the following predictions do not hold exactly (Foley, 1980; Lukas, 1996; Heller, 1997). According to the first equation of (1), the loci of \(Q\) giving equal \(\delta_0\) should be given by a Vieth-Miller circle and according to the second equation, the loci of \(Q\) appearing in the same direction should be given by a Hillebrand hyperbola.

An approach was proposed to construct \(\{P_j\}\) in \(\text{EM}\) without using the Luneburg’s
mapping functions (Indow, 1982, 1991, 1995, 1997). Its relationship with \( \{Q_j\} \) in X can be made explicit after \( \{P_j\} \) has been constructed in EM. The relationship may change according to the context in X and the position of \( P_j \) may not be formulated in terms of \( (\gamma_j, \phi_j, \theta_j) \) alone. It is important to keep two problems separate; whether a subspace in VS under a given condition can be regarded as an R and how it is related to X. No doubt, VS is complex and dynamic (VS 3) and it may seem impossible to think about a mathematical model for its structure. However, it is because the relationship between VS and X is taken into account. Once the VS is formed under a given condition in X, its intrinsic structure seems to show such a regularity that is amenable to mathematical formulation. A triangle we see is a simple perceptual entity. However, to discuss its relation to the physical object and the retinal image is a complicated problem, as exemplified by size constancy and shape constancy.

4. Geometrical structures of D-plane and H-plane

Fig. 2 I is the result of an alley-experiment in the frameless horizontal D-plane in VS. All stimulus points \( Q_j \) are presented in the D-plane of \( \theta = 0 \) in X. The furthest pair \( \{Q_{jr}, Q_{jl}\} \) are fixed. Two series of filled symbols, \( Q_{jr} \) and \( Q_{jl} \), are the results when the subject adjusted their positions in the direction of y-axis so that the two series appear perceptually straight and parallel to each other in VS. The configuration is called a parallel alley, P-alley \( \{Q_{jr}, Q_{jl}\}_{D}, j = 1, \ldots, 6 \), in this experiment. Two series of unfilled symbols are the results when \( Q_j \) are adjusted so that each pair \( \{Q_{jr}, Q_{jl}\} \) appear to have the same lateral distance in VS. The configuration is called a distance alley, D-alley \( \{Q_{jr}, Q_{jl}\}_{D}, j = 1, \ldots, 6 \). The fact that the two alleys do not coincide and D-alley lies outside of P-alley was first noticed by Hillebrand (1902) and systematically studied by Blumenfeld (1913). The same discrepancy between the two alleys are shown on a slanted D-plane passing the eyes (\( \theta > 0 \)). Luneburg showed that the observed discrepancy between P- and D-alleys in X is accounted for if D-plane in VS is assumed to be \( R^2 \) of \( K < 0 \) (hyperbolic). If \( K > 0 \) (elliptic), then P-alley is expected to lie outside of D-alley. If \( K = 0 \) (Euclidean), the two alleys should be the same, of course. That the discrepancy between P- and D-alleys is not an artifact was shown in Indow and Watanabe (1984a). In the experiment in Fig. 2, the subject also adjusted positions of \( Q_{jr} \) and \( Q_{jl} \) in the direction of the \( x \)-axis so that each set of five \( Q_j \), \( \{Q_{jr}, Q_{jl}\}_a \) including a fixed \( Q_a \) in the center (a symbol with a dot), appear to be fronto-parallel (a H-curve). The theoretical curves in EM for P- and D-alleys and H-curves are described in Indow (1979, 1988, 1991), Luneburg (1950), Hardy et al. (1953). Fig. 2 II is a schematic illustration of the theoretical curves in EM with an individual constant \( K \). The curves in Fig. 2 I are those curves that are mapped into \( X^2 \) through the Luneburg’s mapping functions (1) in which the other individual constant \( \sigma \) is involved. Values of the two individual constants were optimized so that the curves in X fit all the sets of \( \{Q_{jr}, Q_{jl}\}_a \)s. The stimuli \( Q_j \)s in this experiment were small light points in the dark and the subject was a naive student who participated in experiments of this kind for the first time. It is true that mapping through Eq. (1) gives satisfactory results for \( \{Q_j\} \) in a frameless D-plane, in spite of the fact that Eq. (1) is not well supported in their direct experimental tests.
In Section 2, a method was explained to obtain such a scale value \( d \) with an arbitrary unit that is regarded as a proportional representation of the latent perceptual distance \( \delta \), i.e., \( d = u\delta \), where \( u \) is a constant depending upon the unit of \( d \). Since \( \delta \) is assumed to be a Riemannian metric, \( d \) is also a Riemannian metric. If a matrix \( (d_{ij}) \) of scaled distances is obtained for \( \{Q_j\} \) constructed by the subject, then it is possible, by the use of an
extended form of multidimensional scaling (metric MDS in this case), to optimize the value of $K$ and to construct such $\{\hat{P}_j\}$ in EM that meets the following requirements. First, Riemannian distances $d_{jk}$ are converted through a constant, $c = q/u$ where $q = \sqrt{-K/2}$, to Euclidean distances $\tilde{d}_{jk}$ in EM. Second, from the matrix $\{\rho_{jk}\}$ such $\{\hat{P}_j\}$ is constructed in EM that is best fitted by the theoretical curves in EM that are solely determined by $K$. Denote by $\hat{d}_{jk}$ interpoint distances of $\{\hat{P}_j\}$. Third, using the same constant $c$, we can transform Euclidean distances $\tilde{d}_{jk}$ back to the Riemannian distances $d_{jk}$. These are theoretical values for $d_{jk}$. What is to be optimized is only the one parameter $c$. The mapping relationship between $\{Q_j\}$ in X and $\{\hat{P}_j\}$ in EM automatically becomes explicit. In order to separate $K$ and $u$ in $c$ and to determine the numerical value of $K$, we have to introduce in EM an appropriate unit (Indow, 1982). In the present context, however, it suffices to know that we can obtain, without using the mapping functions (1), $\{\hat{P}_j\}$ that describes the results of alley experiment $\{Q_j\}$ and $\hat{d}_{jk}$ that are proportional to distance assessment data $d_{jk}$.

As an example, Fig. 2 III and IV show the results of distance assessment by the same subject. The theoretical curves in X shown in I, that are based on the mapping function (1), are irrelevant to this analysis. The configuration $\{\hat{P}_j\}$, the representation of $\{Q_j\}$ in I in EM, and the theoretical curves in EM are shown in IV. The proportionality between assessment data $d$ and $\hat{d}$ predicted from $\{\hat{P}_j\}$ is given in III. The results of alley experiments, $\{Q_j\}$ in I, is symmetric. Although $Q_{jk}$ adjusted by the subject were not necessarily symmetric, the two were averaged to define $\pm Q_j$ and the theoretical curves in X were fitted to this symmetrized configuration $\{Q_j\}$. It was this $\{Q_j\}$ that was presented to the subject in the distance assessment experiment. The configuration $\{\hat{P}_j\}$ in III was not symmetrized. If $\{P_j\}$ is symmetrized, its fit to the symmetric theoretical curves becomes slightly better (Fig. 11 in Indow (1982)). Compared with the fit between $\{Q_j\}$ and theoretical curves in X, the fit of $\{\hat{P}_j\}$ to the theoretical curves in EM is in general not completely satisfactory. One reason may be ascribed to the fact that, when the subject pays attention to a designated triples, $P_i$, $P_j$, $P_k$, it tends to introduce a perturbation into the total perceived configuration. The same trouble was noticed in a distance assessment experiment with stars in the night sky (Indow, 1968, 1991). This is inevitable because of the dynamic property of VS (VS 3 in Section 1), and a more appropriate scaling method must be invented. In constructing $\{P_j\}$ from $(d_{jk})$, $c$ was not constrained to be negative. However, the best fit was always obtained with $c < 0$, i.e. $K < 0$. In an experiment in which no scaling procedure is involved, it was shown with naive subjects that we must assume $K < 0$ for all D-planes with various slant ($\theta > 0$) (Watanabe, 1996). More will be discussed in the next section.

Parallel alleys we see in our daily life are extended not in the form of P-alley in D-plane. Most cases, they are running horizontally or vertically on a fronto-parallel plane (H-plane). Experiments to construct horizontal P- and D-alleys on a H-plane and also to ask the subject ratios between $d$s with $\{Q_j\}$ thus constructed were performed. These two approaches, one using and the other not using the Luneburg’s mapping functions (1), showed unequivocally that it is not necessary to consider any other geometry than Euclid to describe the geometry of H-plane (Indow, 1982, 1988, 1991; Indow and Watanabe, 1984b, 1988). Then, it is necessary to distinguish in VS$^3$ two qualitatively different subspaces VS$^2$, though each can be regarded as an $R^2$; D-plane.
following hyperbolic geometry \((K<0)\) and H-plane following Euclidean geometry \((K=0)\).

5. Congruence and similarity in VS

Busemann (1942, 1955, 1959) discussed the Helmholtz-Lie problem (Section 2) without touching upon differentiability of space. This formulation is useful for us because we have no way to prove the differentiability of VS. First, he defined \(G\)-space. The symbol ‘G’ was chosen to suggest that geodesics have all the properties apart from differentiability. He called Euclidean, hyperbolic and spherical spaces to be elementary. These are spaces denoted as R in the present article. The axioms for a \(G\)-space are as follows.

1. The space is a metric space. Denote in this section the distance between two points \(x, y\) by \(d(x, y)\).
2. The space is finitely compact, i.e. a bounded infinite set has an accumulation point.
3. Any two distinct points \(x \neq y\) can be connected by a segment.
4. Every point \(p\) has a neighborhood \(S(p, \Sigma_p), \Sigma_p > 0\), in which a geodesic is uniquely defined.

For any five points \(a, a', b, c, x\) in \(S(p, \Delta), 0 < \Delta < \Sigma_p\), the relations ‘\(d(a, b) = d(a', b)\), \(d(a, c) = d(a'c)\), and \(x\) is on the geodesic between \(b\) and \(c'\) entail \(d(a, x) = d(a', x)\). Then, \(S(p, \Delta)\) is isometric to an open sphere of radius \(\Delta\) in \(R\), and the Helmholtz-Lie problem is formulated in two ways.

Global: if for any two given isometric triples, \(a_1, a_2, a_3\) and \(a'_1, a'_2, a'_3\), i.e. \(d(a_i, a_j) = d(a'_i, a'_j)\), of a \(G\)-space, a motion (isometric mapping that preserves distance) exists taking \(a_i\) to \(a'_i, i = 1, 2, 3\), then this \(G\)-space is an R.

Local: if every point of a \(G\)-space has a \(S(p, \Delta)\) such that for any four points \(a_1, a_2, a'_1, a'_2\) in \(S(p, \Delta)\) with \(d(p, a_1) = d(p, a'_1), d(p, a_2) = d(p, a'_2)\) and \(d(a_1, a_2) = d(a'_1, a'_2)\), a motion of \(S(p, \Delta)\) exists which takes \(a_i\) into \(a'_i, i = 1, 2\), then the universal covering space of this \(G\)-space is an R.

Wang (1951a,b) gave the following theorem. If a \(G\)-space is two point homogeneous (for any four points \(a, a', b, b'\), with \(d(a, a') = d(b, b')\), there exists an isometric mapping carrying \(a, a'\) to \(b, b'\), then the space is congruent to an R, provided its dimension is two or odd. Mathematically speaking, if the dimension is even and greater than two, there can be a two point homogeneous space that is not an R. Notice that we are concerned with only VS\(^2\) or VS\(^3\).

In alley experiments described before, \(\{Q_j\}\) is a configuration of stationery stimulus points. The same discrepancy between P- and D-alleys occurs when the trajectories of moving points, e.g. \(Q_{1R}, Q_{1L}\) in Fig. 2 I, are adjusted. It was shown in a dark or illuminated D-plane (Indow and Watanabe, 1984a) and in a dark H-plane (Indow and Watanabe, 1984b; 1988). The subject adjusted the trajectories in X so that two points appeared to move in VS (apparent movement), keeping the paths straight and parallel (P-alley) or the lateral distance invariant (D-alley), from or toward the self in a D-plane,
and left to right or right to left in a H-plane. Constructing a D-alley, including stationary and moving situations, corresponds to the homogeneity condition in the Wang’s theorem, and hence we can conclude that D-plane as well as H-plane are respectively an R, and, from the relationship between P- and D-alleys, we can say that $K<0$ for D-plane and $K=0$ for H-plane.

In the above discussion, motion of a pattern is not explicitly referred to. Congruence between two patterns includes invariance of the angles in addition to invariance of the side lengths. The structure of VS can be approached from the possibility of congruence between patterns, e.g. Foster (1975), Lapin and Wason (1991) etc. To our perception of a figure, the angles are distinctive features as the lengths of sides are. On a H-plane I in Fig. 3, we can see a figure to move keeping its shape and size invariant, provided its counterpart in X is appropriately adjusted. Our daily life is based upon the possibility of perceptual congruence in this plane. Moreover, we take it for granted that we recognize similarity between figures, not only in the same H-plane, but also between different H-planes. We have no difficulty to recognize objects in H-planes, II and II’, in their pictures placed on a H-plane I. This fact is in agreement with the idea that $K=0$ in all these H-planes. When we see congruence or similarity between figures, however, we do not pay much attention to exact equalities of corresponding angles or corresponding sides. If these equalities are directly tested in an experiment, it is not a surprise to find some violations. Nevertheless, we cannot deny the possibility of seeing overall similarity

Fig. 3. Congruence and similarity on fronto-parallel planes (H) and on the horizontal D-plane in VS.
in addition to overall congruence in H-planes, and the result of the alley experiment in the H-plane is in accordance with this possibility. The alley experiment on D-planes with different slant suggests that $K < 0$ in each of these subplanes of VS. Mathematically speaking, this is a plane in which congruence holds but similarity does not. Hence, to perform the following two sets of experiments will be interesting. Suppose a standard figure $S_0$ is presented at a place on a D-plane. Then ask the subject to adjust another figure presented at a different place in one set of experiments with two instructions; (A) the figure can be superimposed over $S_0$ (congruence) and (B) the figure consists of same angles and sides with $S_0$. Denote the adjusted figure by $S$. Of course, $S$ and $S_0$ will be physically different. The question is whether $S$ under (A) and $S$ under (B) are physically the same or not. In the other set of experiments, two instructions are (A') the figure appears to be the same but with different size as $S_0$ (similarity) and (B') the figure consists of same angles and proportional sides with $S_0$. Denote the adjusted figure by $S'$. The question is to compare two $S'$ under (A') and (B') and also to see the relationship between $S'$ and $S$.

It is not a problem that H- and D-planes have different values of $K$, because we cannot see a motion across the two planes. It is inconceivable in VS that a figure vertically moves on a H-plane to the joint with a D-plane and then moves on the D-plane keeping its form constant throughout. It has been the problem of perspective how to represent the ground surface in front of the object in II (the hatched area in Fig. 3) in the picture in a H-plane I. It is a problem because it is similarity of figures in two planes having different $K$. In a previous article (Indow, 1997), it was pointed out that the idea of VS being Euclidean is prevailing in the discussion of linear perspective and also in the discussion of shape of the perceived sky. Some suggestions were made to reconsider these problems in the framework that VS is an $R^3$.

6. Perception of natural scene

In Section 1, it was emphasized that the perceived scene under natural conditions is stable despite that it is the result based on multiple glances. Cutting and Vishton (1995) divide X into three zones according to what cues therein are effective to produce difference in radial distance $d_0$ from the self in VS. Ego-centric cues such as convergence $\gamma$ in (1) is effective only in the personal space, $e_0$ of which is less than, say, 2 m. Binocular disparities and motion perspective are effective until the next zone called the action space ($e_0 < 30$ m or so). In the vista space beyond this level, except occlusion and relative size of familiar objects, only aerial perspective and relative density of surface may be factors to affect $d_0$. Heelan (1983) distinguishes near and distant zones. Stimulus length $e$ perpendicular to the line of sight is overestimated in the near zone whereas underestimated in the distant zone. Both of these classifications are based on the change of the correspondence between X and VS and have nothing to do with the internal structure of VS. We do not see any discontinuity in VS even if the underlying cues are changed.

The question raised in this article is whether a H-plane can be interpreted as an $R^2$. 
irrespective of its location in $VS^3$ and a D-plane as an $R^2$ irrespective of its direction and extension. The natural scene in Fig. 1 is a phenomenon in the vista space. Mathematically it may be a problem to assert that $VS^3$ is an $R^3$ up to its boundary, max $\delta_\nu$. An experiment was carried out to assess perceptual distances $\delta_{ik}$ with a set of stars $\{Q_j\}$, $j = 1, 2, \ldots, 11$, where the subject was included as $Q_0$. The subject faced toward the dark ocean and assessed ratios $r_{ijk} = \delta_j/\delta_{ik}$ with triplets designated (Section 3). Then, configuration $\{P_i\}$ was obtained according to a Riemannian metric that forms a curved surface in $R^2$ according to the procedure described in Section 5. Three metrics with $K > 0$, $K = 0$, and $K < 0$ gave almost the same $P_i$ and almost the same goodness of fit (Indow, 1968, 1990, 1991, 1995). Three geometries cannot be differentiated because there are no percepts in between the self and stars. In so far as distances are concerned, difference among the three geometries is very small. The difference comes to fore only when a pattern such as alleys are dealt with. It is not well understood how the length of max $\delta_i(\omega)$ in a direction $\omega(\varphi, \theta)$ is determined. Clearly max $\delta_i(\omega)$ depends upon what is visible in the direction $\omega(\varphi, \theta)$ (Indow, 1991, 1997). The perceptual distance $\delta_0$ to an object in $X$ is determined by the cues related to it, and max $\delta_0$ in that direction cannot be smaller than that. In this representation of the perceived night sky, max $\delta_i(0, 0) < \max \delta_i(0, 90^\circ) < \max \delta_i(\pm 90^\circ, 0)$. In the horizontal front direction $\omega(0, 0)$ no object was visible, and in the zenith direction $\omega(0, 90^\circ)$ there were stars, whereas in the peripheral directions $\omega(\pm 90^\circ, 0)$ lights of houses were in the sight. The daytime sky is perceived to extend least in the zenith direction (Section 1). In this case, we see objects on the ground in front and in the peripheries but nothing above.

When one stands on the top of a hill going down to the cliff as shown in Fig. 4 I, if the texture of the surface of the ocean is not visible, the ocean appears as if it is a blue wall standing vertically and the upper edge is straight from left to right. Namely max $\delta_0$ to the horizon is determined by $\delta_0$ to the hills on both sides and the ocean only fills the gap. There is such a spot in the vicinity of Irvine, and someone simply cannot believe the blue standing wall is the ocean. When one stands on the tip of cliff (II), the ocean appears as a slanted surface with the horizon horizontally curved; max $\delta_i(\varphi)$ is shortest in the center ($\varphi = 0$) and becomes larger to the left and right. As shown in the left lower corner of Fig. 1, the ocean and sky seems to meet at the distance where texture gradient of the water becomes ineffective to make $\delta_i$ larger. The horizon under this condition appears vertically straight (II). When one stands at the seashore, then the ocean is like a D-plane. If the ocean is occluded by objects on both sides, the horizon appears slightly curved vertically (III). Examples that the artists depict the scene in this way were presented in the meeting. The perceptual distances to these objects are determined by a different factor other than the limiting texture density of the water in the center. When the ocean is observed from a boat at night where no other objects are visible and the ocean is evenly illuminated by the moon at the zenith, the horizon appears, horizontally and vertically, as shown by the solid curve and line in IV. When a part of the ocean is illuminated by the moon in the front direction and the texture of water is more visible in that direction, the ocean extends horizontally further and appears slightly higher in that direction (dotted curves in IV). It will be an interesting project to systematically collect observational data and scenes depicted in pictures. Another interesting project will be to construct configurations $\{Q_j\}$ for P- and D-alleys and H-curves etc. with boats on the
Fig. 4. Appearances of the ocean under various conditions.
ocean for the subject at the seashore. It will be readily carried out if boats with communication system and a device to determine their positions in $X^2$ are available.

References

Indow, T., 1982. An approach to geometry of visual space with no a priori mapping functions: multidimensional mapping according to Riemannian metrics. Journal of Mathematical Psychology 26, 204–236.