NEUROCOMPUTING

ELSEVIER

Neurocomputing 44–46 (2002) 907–914

www.elsevier.com/locate/neucom

Why is binocular visual space distorted compared to physical space ? ☆

Taro Maeda[a,b,*], Eimei Oyama[c], Susumu Tachi[a]

[a] *School of Engineering, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, 113-8656 Tokyo, Japan*

[b] *PREST, JST, Honmachi 4-1-8, Kawaguchi-shi, Saitama 332-0012, Japan*

[c] *National Institute of Advanced Industrial Science and Technology, Namiki 1-2, Tsukuba-shi, Ibaraki 305–8564, Japan*

Abstract

We propose the independent scalar learning elements summation (ISLES) model, a neural network model for the developmental learning of binocular visual space. When applied to the phenomena of Hering's horopter and the locus of perceived egocentric equidistance, this model provides a better qualitative explanation than the Luneburg theory. This model provides physical explanations for the differences among individual subject data, by accounting for the spatial distribution of the visual experiences for learning the perception. When the perceptions are classified via psychophysical scaling, their class logically determines the learning signal in this model. As a result, the same ISLES model can predict each phenomenon when a subject learns the physical loci under various conditions. © 2002 Published by Elsevier Science B.V.

Keywords: Binocular visual space; Horopter; ISLES model; Psychophysical scaling

1. Introduction

From a psychophysical standpoint, human binocular visual space is distorted and does not completely coincide with physical space. Even in darkness, humans can perceive the locations of points of light and their separation distance with binocular vision. In such a situation, the subjective straight line to the objective point becomes the reference. However, the subjective straight line is observed to have a certain physical curve which is generally not straight in the physical sense. This phenomenon is well known

☆ This paper was presented at CNS*2001.

* Corresponding author. Tel.: +81-3-5841-6916; fax: +81-3-5841-8606.
E-mail address: maeda@star.t.u-tokyo.ac.jp (T. Maeda).

0925-2312/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0925-2312(02)00490-3

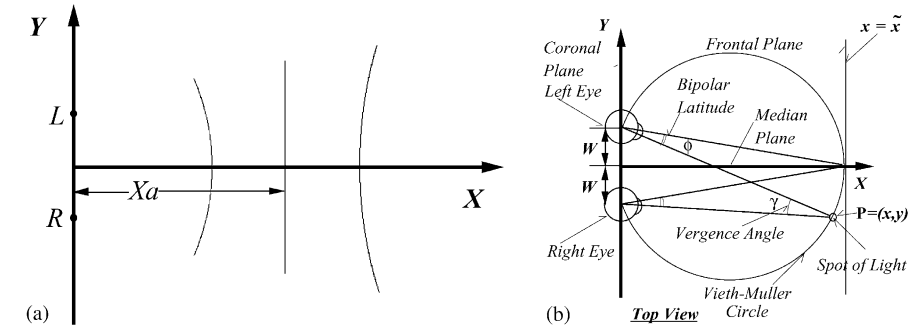


Fig. 1. Hering's horopter and the human oculomotor system. (a) Top view of typical horopter. (b) Definition of parameters.

as Hering's front parallel horopter [3]. A horopter curve is a subjective frontal plane. Fig. 1(a) shows a top-view of some typical front parallel horopter curves. In the figure, L and R are the left and right eyeball positions of the observer, respectively. The shape of the horopter curve depends on the distance from the observer. At a certain distance (X_a in this figure), it is practically straight. At closer distances, the horopter curves are concave to the observer, while at greater distances they are convex [9]. Discussions of binocular disparity are not necessary for this phenomenon. It can also be observed by alternately gazing at each point of light in darkness. In this case, the cues for the perception of the vertical are only the vergence angle γ, and the bipolar latitude ϕ. With these cues, binocular visual space can be described as a perceptual space endowed with a rich geometrical structure, which according to Luneburg [7] is a non-Euclidean Riemannian geometry of constant curvature. This hypothesis, together with certain psychophysical assumptions, provides a qualitative explanation of classical empirical phenomena. Luneburg's metric of binocular visual space provides a good description for the phenomena in the horizontal plane, but it does not give the reason as to why perceptual space is distorted. A new hypothesis is proposed here, wherein perceptual space is distorted as a result of some physiological constraint that renders the learning for the perception incomplete, since the phenomenon is observed independent of the subject's mathematical knowledge of the geometry. In order to confirm this hypothesis, the independent scalar learning elements summation (ISLES) model is proposed, which is a neural network model for the developmental learning of perceptual space.

2. Background—geometry on binocular visual space

When a human subject gazes at a point of light, the location and orientation of the eyeballs are identified by the sensory signals of the vergence angle γ and the bipolar latitude ϕ (Fig. 1(b)). In the process of human visual space perception, a transformation is required to map γ and ϕ to physical world orthogonal coordinates x and y that describe planes and lines. Under Euclidean geometry, this transformation is as follows:

$$x = \frac{W}{\sin\gamma}(\cos\gamma + \cos 2\phi) \equiv X(\gamma,\phi), \quad y = \frac{W}{\sin\gamma}\sin 2\phi \equiv Y(\gamma,\phi). \tag{1}$$

On the other hand, in Luneburg's metric, which is based on Riemannian geometry η and ξ, the horopter is defined as a geodesic line. The line meets the ξ-axis at a right angle, and is asymptotical to one direction at infinity. The geodesic line is described as Eq. (2). K is the constant curvature of the space. σ is the constant for scaling the distance. These are defined as the individual constants of subjects in Luneburg's metric. These constants are used to explain the difference among the individual data of the horopters. Here, C is the invariant to define a curve of horopter. When the curve meets the ξ-axis at ξ_0, C is defined as follows:

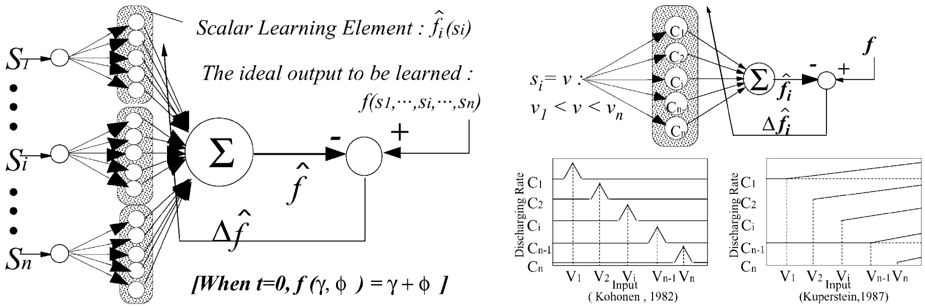
$$\frac{K}{4}(\xi^2 + \eta^2) - 1 = C\xi, \quad \begin{cases} \xi = 2\mathrm{e}^{-\sigma\gamma}\cos\phi \\ \eta = 2\mathrm{e}^{-\sigma\gamma}\sin\phi \end{cases} \Rightarrow C = \frac{K}{4}\xi_0 - \frac{1}{\xi_0}. \tag{2}$$

From the standpoint of psychological scaling, some invariant is necessary in order to define a subjective line. This is the necessary condition for the nominal scale, which is the first class of the psychological scaling. The human perception process also needs to have some function from the sensory cues to such an invariant, since the human perception process constantly performs psychophysical measurement.

3. Independent scalar learning elements summations model

A plane and a straight line are abstract concepts acquired developmentally. Therefore, the function for the invariant should be obtained from some learning mechanism involving a visual space perception process. If the human learning mechanism could completely learn the function from Eq. (1), then human sensory space would coincide with physical space. However, the subjective straight line of a human operator differs from a physical one, so the human learning mechanism is apparently incomplete in learning such a function. So what kind of mechanism generates the characteristics for human visual space? The incompleteness should mostly comprise physiological factors of neural networks of human brains, since everyone has the same tendency. We propose here an assumption of the physiological learning rule that the physiological learning mechanism cannot propagate error signals backward to any layer except the final (output) layer. It is such a basic and natural constraint from a physiological viewpoint that such learning mechanisms are actually found in a human brain [8]. In this case, the training signal for learning is not a vector but a scalar signal, because only one scalar signal is made from a scalar evaluating function. We call this learning rule the scalar learning rule. We provide a neural network model for the developmental learning of perceptual space, which we call the ISLES model. (Fig. 2)

The ISLES model has similar constraints for learning as a biological neural network from a scalar learning rule. This model does not apply the back-propagation method,



(a) ISLES Model (b) Scalar Learning Element

Fig. 2. ISLES model. (a) ISLES model. (b) Scalar learning element.

but utilizes the Hebbian learning rule as a three-layer Perceptron. Further, the cells of the hidden layer are simple cells with only one input signal, so the ability of this model to learn arbitrary functions is mathematically incomplete. When the model has n independent input signals, it has n independent groups of simple cells. Each group is called a scalar learning element. This model has n scalar learning elements $\hat{f}_1(s_1)\hat{f}_n(s_n)$ and only one summation unit for an output $\hat{f}$:

$$\hat{f}(s_1, s_2, \ldots, s_i, \ldots, s_n) = \sum_{i=1}^{n} \hat{f}_i(s_i) + C, \tag{3}$$

where each $\hat{f}_i(x)$ is a nonlinear continuous scalar function acquired through training. Each scalar function is made to learn its output with the error signal: $\Delta\hat{f}_i = \Delta\hat{f} \equiv f - \hat{f}$. Here, $f(s_1, s_2, \ldots, s_i, \ldots, s_n)$ is a training function to be learned. These scalar learning functions can be implemented by a neural network model with the constraint mentioned above [5,6]. If the neural network model's learning method is like the Perceptron or a method of steepest decent, then after sufficient training, each function $\hat{f}_i(x)$ converges to each expectation as follows:

$$\begin{aligned}\lim_{t\to\infty} E[\Delta\hat{f}^2] = C_{\min} \;\Rightarrow\; & \lim_{t\to\infty} E[\hat{f} - f] = C'_{\min},\\ & \Rightarrow \lim_{t\to\infty} E\left[\sum_{j=1}^{n} \hat{f}_j - f + C\right] = C'_{\min}\\ & \Rightarrow \lim_{t\to\infty} \sum_{j=1}^{n} E[\hat{f}_j] = E[f] + C'.\end{aligned} \tag{4}$$

Here, S is defined as the whole of the learning domain, and S_i is defined as a subdomain where $s_i = x$. The output of $\hat{f}_i(x)$ is determined with the subdomain S_i, as follows:

$$\lim_{t\to\infty}\sum_{j=1}^{n} E_{S_i}[\hat{f}_j] = E_{S_i}[f] + C' \;\Rightarrow\; \lim_{t\to\infty}\hat{f}_i(x) + \sum_{j=1}^{n} C_j = E_{S_i}[f] + C',$$
$$\Rightarrow \lim_{t\to\infty}\hat{f}_i(x) = E_{S_i}[f] + C''$$
$$\Leftrightarrow E_{S_i}[f(s)] = \frac{\int_{S_i}\rho(s)f(s)\,\mathrm{d}s}{\int_{S_i}\rho(s)\,\mathrm{d}s}. \quad (5)$$

Every other function except $\hat{f}_i(x)$ becomes constant because each function is dependent on a single input signal that is independent of every other input. The expected value $E_{S_i}[f(s)]$ can be calculated directly by integrating over the region S_i. Therefore, simulations calculated by this model do not need successive training when the distribution function $\rho(s)$ of learning points is known a priori. In the ISLES model, input signals are "isled" with each other until they are summed up in the output layer. Therefore, while this model cannot completely learn all the mathematical functions, it can make some differences from ideal ones. When the differences are similar to those reported in psychophysical experiments, the neural network model can be considered a good approximation to the physiological process of spatial perception.

4. Simulation for the horopter curves calculated by ISLES model

ISLES model is incomplete in learning arbitrary functions. However, this incompleteness and the class of psychological scaling for learning made the curves of the frontal parallel horopter. In order to learn a frontal parallel plane, the training function is not necessarily the $X(\gamma,\phi)$ of Eq. (1), but it is necessarily invariant on the frontal plane. In addition, the function must be monotonic to the depth x in order to conserve the order. This is the necessary condition for the ordinal scale, which is the second class of the psychological scaling, and is necessary and sufficient to measure parallelity. Here, the value of the invariant is given as f_0, the output at the point P_0 crossing the median plane on the same frontal parallel plane. As the definition of the invariant, f_0 and P_0 correspond to C and ξ_0 of Eq. (2) in Luneburg's metric. In this case, the error signal for learning is defined as $\Delta\hat{f} = f_0 - \hat{f}$. The domain for learning in the ISLES model is defined as the domain where the cues γ and ϕ are effective spatial cues in human binocular perception. In this study, the domain was defined as $-\pi/4 < \phi < \pi/4$, $0.003 < \gamma < 0.3$. This domain is equivalent to the domain of ordinal human binocular visual experience. Luneburg's individual constants K, σ are explained as the distribution of the learning points in ISLES model. In this result, the distribution was homogeneous along γ and ϕ. After training, the experiments for measurement of horopter curves were simulated by the ISLES model. The results are shown in Fig. 3(a).

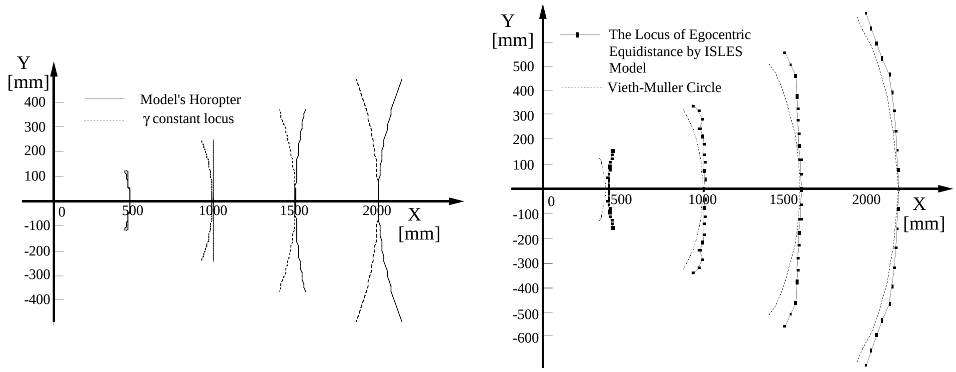


Fig. 3. Results of simulation by ISLES model. (a) Hering's horopter. (b) Egocentric equidistance.

5. Discussion and additional simulation: egocentric equidistance

The shapes of the horopter curves are quite similar to those of humans that the ISLES model provides a qualitative explanation for the phenomena of Hering's horopter as well as Luneburg's model. Further, this model provides a physical explanation for individual data as the distribution of the learning points. In this result, the σ of Luneburg's individual constants was slightly greater than the empirical human average, which implies that this distribution of learning points in the near domain is slightly greater than average. This distribution on the ISLES model is concerned with the visual experiences of the human subjects in physical space, while Luneburg's individual constants K, σ can be defined only in Riemannian geometry. From this result, it seems that ISLES model can provide a better explanation in the physical sense than the Luneburg theory. Although the Luneburg theory provides a qualitative explanation of classical empirical phenomena, the theory has been less successful in quantitatively predicting individual data [4], a fact that has sometimes been taken as evidence against the presumed geometrical structure. Some of these shortcomings, however, may also be attributed to psychophysical assumptions that do not depend on the geometry of visual space. The locus of perceived egocentric equidistance, for example, has been found to deviate systematically from the Vieth–Muller circle postulated by Luneburg [2]. The locus is always observed at the outside of the Vieth–Muller circle. However, the locus of perceived egocentric equidistance is easily defined under the ISLES model. An interval scale is necessary for the perception of distances, which is the third class of psychological scaling. In order to learn egocentric distances, the training function f has to conserve the interval. Here, the value of the invariant is given as the interval Δf_0, the difference of the outputs at the points whose interval on the physical egocentric distance is invariant ΔD_0. In this case, the error signal for learning is defined as $\Delta\hat{f} = \Delta f_0 - \Delta f'_0$, the difference between the Δf_0 at a set of points and $\Delta f'_0$ at another set of points. The training function f must conserve the Δf_0 at any ΔD_0. Therefore, this function is necessarily equivalent to $aD + b$, where a and b are constants and D is $\sqrt{x^2 + y^2}$, the Euclidean norm of the learning point as the egocentric distance. After learning the front parallel horopter as well, the loci of perceived egocentric equidistance were simulated by the ISLES model. The results are shown in Fig. 3(b). The loci by the ISLES model were always observed to be outside the Vieth–Muller circles. This result means that the same ISLES model can generate both Hering's frontal parallel horopter and the locus of egocentric equidistance when it learns the physical loci under each psychological scaling. This confirms that the ISLES model can provide a better qualitative explanation than the Luneburg theory. Lately, a computational model of depth perception based on headcentric disparity was proposed by Erkens and Ee [1]. This model is based on three types of retinal disparity, and integrates the contradictions among those disparities into the error of the oculomotor signals, in order to compensate each other's disparity. The approach of that model may ultimately be closely related to our study of the ISLES model, although the main depth cues for the two models differ.

6. Conclusions

In this paper, we provided the ISLES model, a neural network model for the developmental learning of perceptual space. This model was applied to the phenomena of Hering's front parallel horopter and the locus of perceived egocentric equidistance, and provided a better qualitative explanation than the Luneburg theory, especially for the perception of egocentric equidistance. In addition, this model provides the physical explanation for the differences among individual data as the spatial distribution of the visual experiences for learning the perception. Furthermore, the perceptions are classified by psychological scaling, and the class of scale determines the learning signal logically. As a result, the same ISLES model can predict each phenomenon when the model learns the physical loci under each psychological scaling. This result suggests that the class of psychological scaling in perception is significant not only for psychophysical measurement but also for the human learning process, since the perception system of a human brain is in itself a psychophysical measurement system.

References

[1] C.J. Erkelens, R. van Ee, A computational model of depth perception based on headcentric disparity, Vision Res. 38 (1998) 2999–3018.

[2] J.M. Foley, Locus of perceived equidistance as a function of viewing distance, J. Opt. Soc. Amer. 56 (1966) 822–827.

[3] H. Hering, Die Raumsinn und die Bewegungen des Auges, in: L. Hermann (Ed.), Handbuch der Physiologie, Vol. 3, (Vogel, Leipzig) 1879.

[4] T. Indow, An approach to geometry of visual space with no a priori mapping functions: multidimensional mapping according to Riemannian metrics, J. Math. Psychol. 26 (1982) 202–236.

[5] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybernet. 43 (1982) 59–69.

[6] M.A. Kuperstein, An adaptive neural model for mapping invariant target position, Behavioral Neurosci. 102 (1) (1988) 148–162.

[7] R.K. Luneburg, The metric of binocular visual space, J. Opt. Soc. Amer. 40 (10) (1950) 627–642.
[8] P.K. Stanton, T.J. Sejnowski, Associative long-term depression in hippocampus induced by hebbian covariance, Nature 339 (1989) 215–218.
[9] A. Zajaczkowska, Experimental test of Luneburg's theory, Horopter and Alley experiments, J. Opt. Soc. Amer. 46 (7) (1956) 514–527.