WHY DOES THE PROXIMITY PRINCIPLE FAIL IN PERCEPTION OF MOTION?

Sergei Gepshtein[1,2], Ivan Tyukin[3], and Michael Kubovy[4]

[1]Brain Science Institute, RIKEN, Japan; [2]The Salk Institute, USA

[3]Department of Mathematics, University of Leicester, UK

[4]Department of Psychology, University of Virginia, USA

Abstract

The proximity principle is a fundamental fact of spatial vision. It has been a cornerstone of the Gestalt approach to perception, it is supported by overwhelming empirical evidence, and its utility has been proven in studies of the "ecological statistics" of optical stimulation. We show that the principle fails in the perception of motion, which means that the standard (Minkowski) notion of proximity does not apply to the perceptual combination of space and time the way it applies to the combination of spatial dimensions in perceptual organization of static scenes. We demonstrate that in perception of motion the proximity principle should be supplanted by a more general notion – the equilibrium principle – which is related to the minimum principle championed by the Gestaltists.

Introduction

The proximity principle is an empirical law that holds in the perception of static scenes (Wertheimer, 1923; Hochberg & Silverstein, 1956; Oyama, 1961; Kubovy, Holcombe, & Wagemans, 1998): the closer elements of a scene to one another, the more likely it is that they will appear parts of the same object. Studies of the statistics of natural images have revealed its ecological utility: Image regions (or "elements") from one object are likely to be closer to each other than elements from different objects (Brunswik & Kamiya, 1953; Geisler, Perry, Super, & Gallogly, 2001; Elder & Goldberg, 2002).

Here we show that the principle does not generalize to dynamic scenes, i.e., no *spatiotemporal* proximity principle governs the perception of motion. In other words, elements of a dynamic display separated by short spatiotemporal distances do not appear as parts of the same object more readily than elements separated by longer spatiotemporal distances. We demonstrate this in three steps. First, we examine how spatial dimensions combine in grouping by spatial proximity. Second, we ask what consequences it would have for the perception of motion if space and time had combined similarly, i.e., if a spatiotemporal proximity principle held. Then, we show that the predictions from proximity hold only for some conditions of motion perception, which means that a proximity principle is not a general rule for the perceptual space-time. This result is consistent with other results on motion perception and with a normative theory of motion perception.

The structure of spatial proximity

In Fig. 1A we show a spatial dot lattice, which is a multistable visual stimulus. Human observers see the dots spontaneously group into strips, and the perceived groupings change while the stimulus does not change. The multistable perception of dot lattices is lawful (Kubovy et al., 1998). Thus, when spatial distances between elements along vectors **a** and **b** are equal, the probabilities of seeing the dots organize into strips parallel to **a** and **b** are approximately equal.[1] In other words, the two organizations are in *perceptual equilibrium.*

To illustrate the metric structure of grouping by spatial proximity, let us embed the lattice in a pair of orthogonal axes K and L in the plane, so the projections of vectors **a** and **b** on these axes are a_K, a_L, and b_K, b_L (Fig. 1A). Suppose we can independently manipulate the projections until we

[1]For simplicity, we disregard orientation biases in perceptual grouping (Gepshtein & Kubovy, 2005).

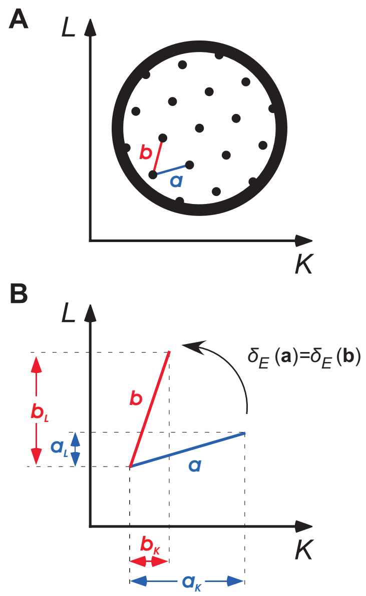

Figure 1. Tradeoff of distance components. **A**. In a dot lattice the competing organizations are seen equally often if the vectors that define them, **a** and **b**, are equally long. **B**. Vector **a** can be rotated to obtain segment **b**: $\delta_E(\mathbf{a}) = \delta_E(\mathbf{b})$, where the Euclidean distances are $\delta_E(\mathbf{a}) = \sqrt{a_K^2 + a_L^2}$ and $\delta_E(b) = \sqrt{b_K^2 + b_L^2}$, a_K and a_L are the projections of segment **a**, and b_K and b_L are projections of segment **b**. To maintain the distance, increasing the projection of a vector on axis **K** (from a_K to b_K) must be accompanied by decreasing its projection on axis **L** (from a_L to b_L), i.e., the two projections must trade off. This invariance of distance under rotation does not generalize to other power metrics (Appendix), whereas the tradeoff of distance components does.

obtain the equality of distances. We will be able to equate the distances only when the projections trade off their lengths, i.e., when one of the two conditions holds:

$$a_K > b_K \quad \text{and} \quad a_L < b_L$$

or

$$a_K < b_K \quad \text{and} \quad a_L > b_L.$$

In words, to achieve the same distance (the same spatial proximity) between elements along different orientations in space, we must increase the projection of this distance on one spatial dimension and decrease the projection on the other dimension (Fig. 1B). We call this property *tradeoff of distance components*. We formulate it explicitly in the Appendix. Now we apply this argument to space-time.

Regimes of motion perception

Suppose one of the two dimensions in Fig. 1 is time. To sustain the equality of distances in the two-dimensional space-time, the spatial and temporal components of spatiotemporal distance will have to trade off, just as they did in space. Thus, to maintain the same spatiotemporal distance, increasing the spatial distance between the elements will have to be accompanied by decreasing the temporal distance, and decreasing the spatial distance will have to be accompanied by increasing the temporal distance.

It turns out that only some of the results on motion perception are consistent with this prediction. We illustrate this using a simple case of perceptual multistability in an apparent motion display (Fig. 2). Three short-lived dots sequentially appear and disappear at three loci: O, A, and B, so that O has two potential matches: at A and B. We denote the potential motion paths from O to A and from O to B by $\mathbf{m_a}$ and $\mathbf{m_b}$. (For simplicity, suppose the distance between A and B is long, and motion from A to B is unlikely.) Each path has a temporal component (T_a, T_b) and a spatial component (S_a, S_b), so in a plot of distances in Fig. 2B each motion path is represented by a point. Suppose we manipulate S_b (double-headed arrow in Fig. 2B), while holding constant all other parameters (S_a, T_a and T_b), such that $T_b = 2T_a$. We determine the value of S_b for which the probabilities of competing motions are equal. If the equilibrium obtains at $S_b < S_a$ (space-time tradeoff) then we found evidence supporting the proximity principle in space-time, but if we find equilibrium at $S_b > S_a$ (space-time coupling) then we find support to the claim that the proximity principle does not generalize to space-time.[2] Both possibilities are supported by empirical evidence.

[2]Generally, $S_b \geq S_a$ is evidence against a proximity principle in space-time. See Gepshtein and Kubovy (2007) for a discussion of the regime of *time independence*, $S_b = S_a$.

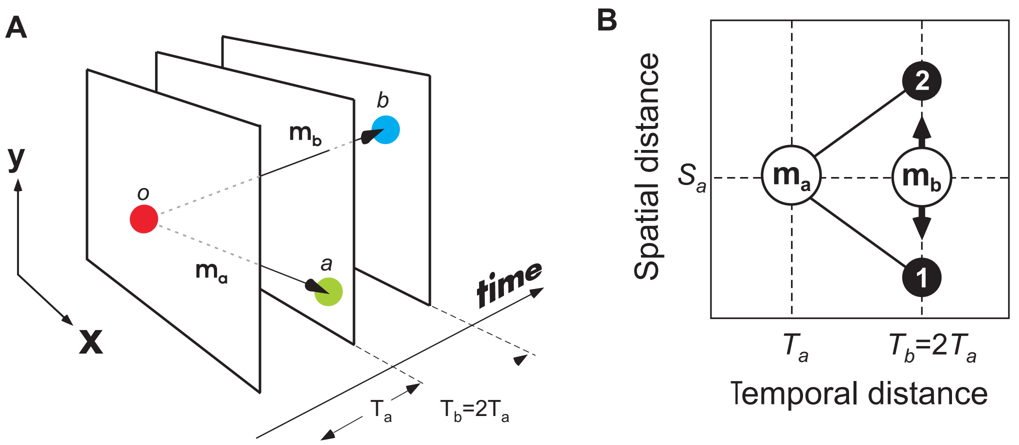

Figure 2. Regimes of apparent motion. **A**. A stimulus for ambiguous apparent motion. Element O has two potential matches, A and B, giving rise to potential motion paths $\mathbf{m}_a$ and $\mathbf{m}_b$. **B**. Plot of spatial and temporal distances ("distance plot") where each motion path is represented by a point $\{T_i, S_i\}$. The double-headed arrow represents the manipulation used by Gepshtein & Kubovy (2007) who varied S_b to find the conditions of perceptual equilibrium between the competing motion paths. The qualitatively distinct outcomes of this manipulation are space-time tradeoff ① and coupling ②. Tradeoff is consistent with a proximity metric but coupling is not. Note that the regimes of tradeoff and coupling correspond to, respectively, negative and positive slopes of the lines connecting the conditions of equilibrium.

The classical Korte's Third Law of Motion (Korte, 1915; Koffka, 1935/1963) is an instance of space-time coupling, and the results of Burt and Sperling (1981) is an instance of space-time tradeoff.

Using the manipulation represented by the double-headed arrow in Fig. 2B, Gepshtein and Kubovy (2007) measured the conditions of perceptual equilibrium under multiple spatial and temporal scales of multistable apparent motion. The authors found evidence of both regimes of tradeoff and coupling in the same stimulus. Tradeoff was found at low speeds of motion, but it gradually changed to the regime of coupling as the speed grew.

Gepshtein and Kubovy (2007) also showed that the gradual transition was consistent with other results on motion perception (Nakayama, 1985), including a comprehensive characteristic of visual sensitivity to *continuous motion* (Kelly, 1979). We illustrate this in Fig. 3B, where we plot human isosensitivity contours along which the visual system is equally sensitive to spatiotemporal modulations of luminance. Note that the slopes of contours change systematically across the distance plot. Thus, if the ability to see apparent motion related monotonically to the ability to detect continuous motion (we dispense with this assumption below), one could predict the different regimes of apparent motion under different conditions of stimulation from the shapes of isosensitivity contours.

Equilibrium theory of motion perception

Gepshtein, Tyukin, & Kubovy (2007) proposed a normative theory of motion perception that explains why the isosensitivity contours have the shapes they do and why the different regimes of apparent motion hold under different conditions of stimulation. By this theory, the reliability of motion measurement depends on the balance of two kinds of uncertainty: (i) *measurement uncertainty*, which is a consequence of the uncertainty principle of measurement (Gabor, 1946), and which affects any measurement, not only visual or biological, and (ii) *stimulus uncertainty*, which depends on the statistical properties of stimulation. Specifically, Gepshtein et al. demonstrated the following.

1. The hyperbolic shape of the maximal sensitivity set (the gray curve in Fig. 3B) is a consequence of the uncertainty principle of measurement. For every point of this set, the spatial and temporal measurement uncertainties are balanced exactly.

2. Outside of the maximal-sensitivity set, there exist equivalence classes (or equivalence contours) of uncertainty where the measurement uncertainty and stimulus uncertainty are *imbalanced* to the

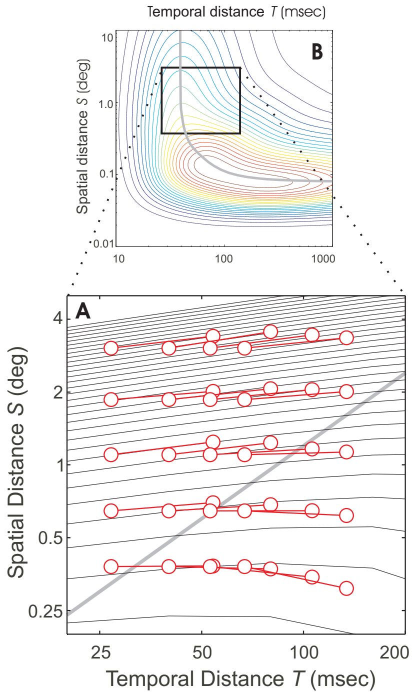

Figure 3. Equivalence classes of motion perception. **A**. The pairs of red connected circles represent the pairs of conditions of apparent motion that were seen equally often in the multistable displays of Gepshtein and Kubovy (2007). The thin lines on the background are the empirical *equivalence contours of apparent motion* derived by Gepshtein & Kubovy from the pairwise equilibria. The slopes of these contours gradually change across the plot, indicating a gradual change from the regime of tradeoff (negative slope) to the regime of coupling (positive slope), in qualitative agreement with predictions of the equilibrium theory (Gepshtein, Tyukin, & Kubovy, 2007). **B**. Human isosensitivity contours (Kelly, 1979) converted into the distance plot. The grey hyperbolic curve is the set of maximal spatiotemporal sensitivity. The rectangle marks the region of conditions in which Gepshtein & Kubovy could find the points of equilibrium of apparent motion. The slopes of isosensitivity contours gradually change across conditions as do the slopes of equivalence contours of apparent motion in panel A.

same degree, which is why the conditions that make up the equivalence classes are equally suboptimal (equally suitable) for motion measurement. The shapes of these equivalence classes in the space of parameters are determined by the degree of balance of uncertainties.

In other words, from the normative considerations it follows that whether tradeoff or coupling are expected at a particular condition of stimulation (i.e., whether the proximity principle stands of falls at that condition) depends on the degree to which the different uncertainties balance each other.

The equilibrium principle

The proximity principle was proposed by the Gestaltists a part of an explanatory framework in which a small number of principles would capture a great variety of perceptual phenomena. The proximity principe is often cited as a staple of this framework because of its simplicity and alleged universality supported by the argument from ecological statistics (Brunswik & Kamiya, 1953).

Now, having demonstrated that the principle does not generalize to perception of motion, we should look for another explanatory framework. A strong candidate is the normative-economic framework, because the theory of motion measurement derived in this framework (Gepshtein et al., 2007) explains why spatial and temporal distances combine differently under different conditions of stimulation. On this view, the proximity principle holds under some conditions of stimulation as an accidental outcome of the optimization process whose goal is to minimize errors associated with motion measurement.

Interestingly, the normative-economic framework is related to another basic idea of the Gestaltists: the *minimum principle* (Hatfiled & Epstein, 1985). Empirical regularities, such as the spatial proximity principle, have been viewed as instantiations of the minimum principle. Gepshtein et al. (2007) demonstrated that the minimum of uncertainties — the "local optimal set" in their theory — is also a balance of uncertainties. Since both the minimal set and the equivalence sets (whose shapes imply the different regimes of motion perception) constitute equilibria of uncertainties, we propose that the *equilibrium principle*, rather than the minimum principle, should serve as a foundation of the new explanatory framework.

The two notions — equilibrium conditions and minimal (extremal) conditions — are closely related also in the optimization theory. Consider, for instance, the case in which the minimal value of a smooth function $U(S,T) : \mathbb{R}\times\mathbb{R} \rightarrow \mathbb{R}$ is searched for minima over a bounded domain. A minimal (extremal) value of $U(S,T)$ is obtained at a condition (S^*, T^*) where the following equilibrium holds:

$$dU(S^*, T^*) = \left.\frac{\partial U}{\partial S}\right|_{S=S^*,\ T=T^*} dS + \left.\frac{\partial U}{\partial T}\right|_{S=S^*,\ T=T^*} dT = 0.$$

The notion of equilibrium is more suitable for analysis of stable systems than the notion of minimum also from the perspective of optimization theory. Systems whose domains are closed sets and whose global minima are reached on domain boundaries (where the minima are not equilibria) are vulnerable to arbitrarily small perturbations, which makes the systems unstable. Thus, the equilibrium principle affords a more suitable framework for analysis of stable biological systems than the minimum principle.

Appendix

We presently demonstrate that tradeoff of distance components is a necessary property of a proximity metric. As we illustrated in Fig. 1, distances δ of **a** and **b** can be mapped onto each other by rotation while preserving distance equality. This property is called *rotation invariance*. The only distance metric for which rotation invariance holds is the Euclidean metric (Mendelson, 1974). However, suppose we relaxed the requirement of rotational invariance. The Euclidean metric is a special case of the *power metric*. Although rotation invariance does not hold in power metrics, the tradeoff of distance components does. The tradeoff follows from the *decomposability* property of power metrics, according to which a distance function must be a strictly monotonically increasing function in each of its arguments (Suppes, Krantz, Luce, & Tversky, 1989). To formalize this idea we write the distance between some space-time locations M and N as

$$\delta(\mathsf{MN}) = [\psi_s(\mathsf{M}_s, \mathsf{N}_s)^r + \psi_t(\mathsf{M}_t, \mathsf{N}_t)^r]^{1/r}, \tag{1}$$

where

- ψ_s and ψ_t are the spatial and temporal differences between locations M and N in space-time, $\psi_i = |\phi(\mathsf{M}_i) - \phi(\mathsf{N}_i)|$, satisfying $\psi_i(\mathsf{M}_i, \mathsf{N}_i) > \psi_i(\mathsf{M}_i, \mathsf{M}_i)$ whenever $\mathsf{M}_i \neq \mathsf{N}_i$,
- ϕ is a real-valued function (the "scale") that represents a mapping between a physical location and its perceptual counterpart, and
- $r \geq 1$ is an integer.

We can introduce function F:

$$\delta(\mathsf{MN}) = F[\psi_s(\mathsf{M}_s, \mathsf{N}_s), \psi_t(\mathsf{M}_t, \mathsf{N}_t)], \tag{2}$$

which must increase whenever $\psi_s(\mathsf{M}_s, \mathsf{N}_s)$ or $\psi_t(\mathsf{M}_t, \mathsf{N}_t)$ increases. According to decomposability, if one of the arguments of distance function (e.g., the L-projection in Fig. 1B) increases, then distance is preserved only if the other argument (K-projection in Fig. 1B) decreases. If the second argument had not decreased, then the distance would necessarily have increased.

We now apply this argument to the elementary case of multistability in motion perception in Fig. 2, where the spatiotemporal distances of competing motion paths are $\delta(\mathbf{m}_a)$ and $\delta(\mathbf{m}_b)$. Let spatial and temporal coordinates of points o, a, and b be $(\mathsf{N}_{s,o}, \mathsf{N}_{t,o})$, $(\mathsf{M}_{s,a}, \mathsf{N}_{t,a})$, and $(\mathsf{M}_{s,b}, \mathsf{N}_{t,b})$. Suppose that:

$\psi_s(\mathsf{M}_{s,a}, \mathsf{M}_{s,o}) = S_a$, $\psi_s(\mathsf{M}_{s,b}, \mathsf{M}_{s,o}) = S_b$, $S_b = S_a + \Delta S$, and

$\psi_t(\mathsf{N}_{t,a}, \mathsf{N}_{t,o}) = T_a$, $\psi_t(\mathsf{N}_{t,b}, \mathsf{N}_{t,o}) = T_b$, where $T_b = T_a + \Delta T$.

If paths $\mathbf{m}_a$, $\mathbf{m}_b$ are in equilibrium, then we can apply Equation 2:

$$F[S_a, T_a] = F[S_a + \Delta S, T_a + \Delta T]. \tag{3}$$

From decomposability it follows that whenever $\Delta T > 0$, the equilibrium of the two paths is possible only when $\Delta S < 0$.

Thus, if the spatial proximity principle generalizes to space-time, under power metric (1) or its generalization (2), then a tradeoff of distance components between the dimensions of space and time must follow. If in Fig. 1B we interpret axis K as space, and axis L as time, then the lengths of spatial and temporal projections of perceptually equivalent spatiotemporal segments **a** and **b** will trade off. Applied to the apparent-motion display in Fig. 2, Equation 3 becomes:

$$F[S_a, T_a] = F[S_b, 2T_a]. \tag{4}$$

The equality of distances can be achieved only when $S_b < S_a$.

References

Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of "proximity" and other Gestalt factors. *American Journal of Psychology*, *66*, 20–32.

Burt, P., & Sperling, G. (1981). Time, distance, and feature tradeoffs in visual apparent motion. *Psychological Review*, *88*, 171–195.

Elder, J., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*, 324–353.

Gabor, D. (1946). Theory of communication. *Institution of Electrical Engineers*, *93 (Part III)*, 429–457.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*, 711–724.

Gepshtein, S., & Kubovy, M. (2005). Stability and change in perception: Spatial organization in temporal context. *Experimental Brain Research*, *160*(4), 487–495.

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision*, *7*(8), 1–15.

Gepshtein, S., Tyukin, I., & Kubovy, M. (2007). The economics of motion perception and invariants of visual sensitivity. *Journal of Vision*, *7*(8), 1–18.

Hatfiled, G., & Epstein, W. (1985). The status of the minimum principle in the theoretical analysis of vision. *Psychological Bulletin*, *97*, 155-186.

Hochberg, J., & Silverstein, A. (1956). A quantitative index of stimulus-similarity: Proximity versus differences in brightness. *American Journal of Psychology*, *69*, 456–458.

Kelly, D. H. (1979). Motion and vision II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, *69*(10), 1340–1349.

Koffka, K. (1935/1963). *Principles of Gestalt psychology.* New York, NY, USA: A Harbinger Book, Harcourt, Brace & World, Inc.

Korte, A. (1915). Kinematoskopische Untersuchungen [Kinematoscopic investigations]. *Zeitschrift für Psychologie*, *72*, 194–296.

Kubovy, M., Holcombe, A. O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35*, 71–98.

Mendelson, B. (1974). *Introduction to topology.* Boston, MA, USA: Allyn & Bacon.

Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research*, *25*(5), 625–660.

Oyama, T. (1961). Perceptual grouping as a function of proximity. *Perceptual and Motor Skills*, *13*, 305–306.

Suppes, P., Krantz, D. H., Luce, R. D., & Tversky, A. (1989). *Foundations of measurement* (Vols. II: Geometrical, threshold, and probabilistic representations). New York: Academic Press.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II [Investigations of the principles of Gestalt, II]. *Psychologische Forschung*, *4*, 301–350.