A Failure of the Proximity Principle in the Perception of Motion*

Sergei Gepshtein[†]
sergei@salk.edu

Ivan Tyukin[‡]
I.Tyukin@le.ac.uk

Michael Kubovy[§]
kubovy@virginia.edu

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, however, 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 are not more likely to be perceived as parts of the same object than elements separated by longer spatiotemporal distances.

The Proximity Principle

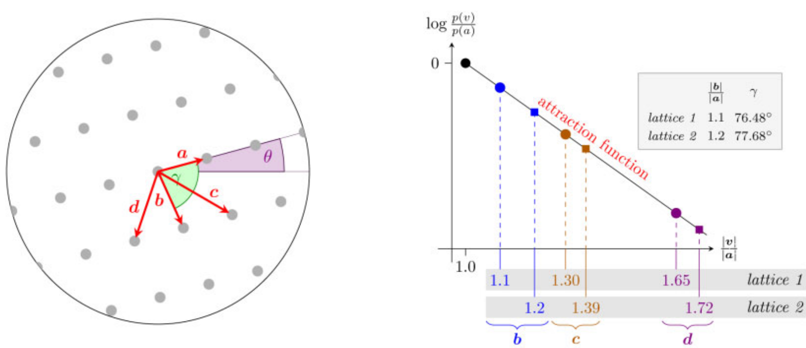
The proximity principle, advanced by the Gestalt psychologists as one of a few foundational perceptual facts, has been a staple of the study of perceptual or-

* This research was supported by US NSF, grants to Salk (SG, PI) and UVa (MK, PI).

[†] The Salk Institute for Biological Studies, USA.
[‡] University of Leicester, UK.
[§] University of Virginia, USA.

ganization. It is an empirical law that holds in the perception of static scenes (Hochberg & Silverstein, 1956; Kubovy, Holcombe, & Wagemans, 1998; Kubovy & van den Berg, 2008; Oyama, 1961; Wertheimer, 1923): the closer elements of a scene to one another, the more likely it is that they will appear to belong to the same object. Studies of the statistics of natural images have revealed its ecological utility: image regions (or elements) that correspond to the same object are likely to be closer to each other than elements that correspond to different objects (Brunswik & Kamiya, 1953; Elder & Goldberg, 2002; Geisler, Perry, Super, & Gallogly, 2001; Martin, Fowlkes, Tal, & Malik, 2001).

We illustrate the proximity principle in Fig. 1a using a regular array of dots called a *dot lattice* (for the nomenclature of dot lattices, see Kubovy, 1994). Any dot of the lattice is surrounded by eight neighbors at four different distances from it, shown by the four red arrows in the figure, and labeled by lower–case bold letters, $\boldsymbol{a}, \ldots, \boldsymbol{d}$ (which we simplify by introducing generic vector $\boldsymbol{v}$ for vectors other than $\boldsymbol{a}$: $\boldsymbol{v} \in \{\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}\}$. Lengths of these vectors are $|\boldsymbol{a}| \leq |\boldsymbol{b}| \leq |\boldsymbol{c}| \leq |\boldsymbol{d}|$.



(a) A dot Lattice

(b) An attraction function

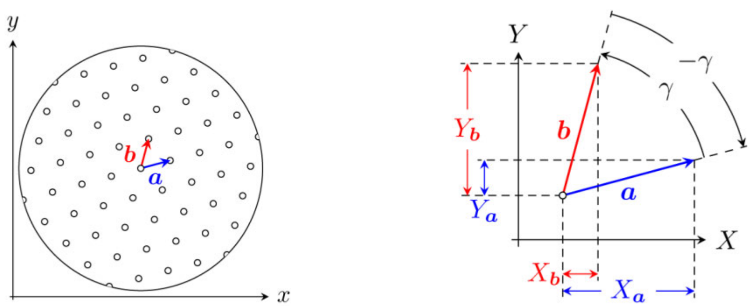
Figure 1. Perceptual grouping in spatial dot lattices.

Dot lattices can be seen organized into strips along $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, or $\boldsymbol{d}$. If $|\boldsymbol{b}|/|\boldsymbol{a}| \lesssim 1.5$, the lattice is multistable; the perceived organizations are in competition and they can spontaneously change (or can be voluntarily changed) even though the stimulus does not. If we wish to preclude such changes during a single viewing of a dot lattice, we can show it for 300 ms or less, too short for a reorganization to occur. Nevertheless, it is still multistable: the same stimulus is seen differently on different presentations.

Kubovy and Wagemans (1995) and Kubovy et al. (1998) manipulated $|\boldsymbol{b}|/|\boldsymbol{a}|$ and γ in briefly–exposed dot lattices, and asked observers to report their organization. Fig. 1b shows schematic data for such an experiment. We denote the four possible responses by lower–case italic letters, $a, \ldots, d$ and a generic response by v (where $v \in \{b, c, d\}$). The x-axis of this figure is $|\boldsymbol{v}|/|\boldsymbol{a}|$, and the y-axis is $\log[p(v)/p(a)]$ (i.e., the log–odds of responding v rather than a).

The figure shows the results for two dot lattices, denoted *lattice 1* and *lattice 2* (whose $|\boldsymbol{b}|/|\boldsymbol{a}|$ and γ values are shown in the inset). We first consider the *b* responses. Recalling that in *lattice 1*, $|\boldsymbol{b}|/|\boldsymbol{a}|=1.1$ and in *lattice 2*, $|\boldsymbol{b}|/|\boldsymbol{a}|=1.2$, we mark their locations on the $|\boldsymbol{v}|/|\boldsymbol{a}|$ axis. The frequency of *b* responses relative to the frequency of a responses for each lattice is represented by blue data points, which show the corresponding values of $\log[p(v)/p(a)]$.

We then consider the c responses. In *lattice 1*, $|\boldsymbol{c}|/|\boldsymbol{a}| = 1.3$; in *lattice 2*, $|\boldsymbol{c}|/|\boldsymbol{a}|=1.39$. The brown data points show the corresponding values of $\log[p(c)/p(a)]$. Turning to the d responses, the purple data points show the corresponding values of $\log[p(d)/p(a)]$. Finally, there is one point for which we don't need data: when $|\boldsymbol{b}|/|\boldsymbol{a}| = 1$, is it inevitable that $\log[p(b)/p(a)]=0$ (the black data point), because $p(b)= p(a)$ when $|\boldsymbol{b}|/|\boldsymbol{a}|$.



(a) Competing organizations are seen equally often if the distances that define them, $|\boldsymbol{a}|$ and $|\boldsymbol{b}|$ are equal.

(b) When $\boldsymbol{a}$ is rotated to obtain $\boldsymbol{b}$, and vice-versa, their projections on the axes tradeoff their lengths.

Figure 2. Tradeoff of spatial distance components.

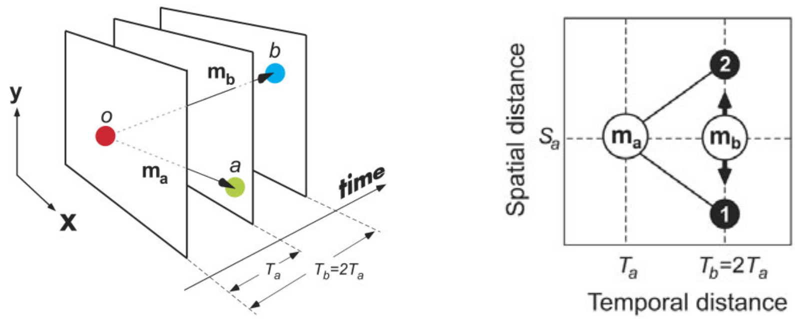
It is striking that all these data points are aligned on a single straight line, known as the *attraction function*, which shows that grouping by proximity follows a *pure–distance law*. This means that grouping by proximity is determined by the three distance ratios $|\boldsymbol{b}|/|\boldsymbol{a}|$, $|\boldsymbol{c}|/|\boldsymbol{a}|$, and $|\boldsymbol{d}|/|\boldsymbol{a}|$, and that it is unaffected by the symmetries of the lattice (as described by Kubovy, 1994). As Kubovy et al. (1998) show, this means that the organization of dot lattice can be modeled as if it were a collection of *unconfigured dots in an isotropic Cartesian space*.

Tradeoff of Distance Components

The proximity principle implies the *tradeoff of distance components*. To explain this concept we consider a dot lattice in which $|\boldsymbol{a}| = |\boldsymbol{b}|$ (Fig. 2a); as we have seen, this means that $p(a) = p(b)$. In Fig. 2b, we show the vectors in a Cartesian plane with coordinates x and y (Fig. 2a). The projections of $\boldsymbol{a}$ and $\boldsymbol{b}$ onto the X–axis are X_a and X_b and onto the Y–axis are Y_a and Y_b. There are two ways to visualize a transformation that will turn $\boldsymbol{a}$ into $\boldsymbol{b}$ and $\boldsymbol{b}$ in to $\boldsymbol{a}$: (a) A clockwise rotation of $\boldsymbol{b}$ by γ and a concurrent counter–clockwise rotation of $\boldsymbol{a}$, also by γ. (b) A tradeoff between the lengths $|X_a|$ and $|X_b|$, and a concurrent tradeoff between $|Y_a|$ and $|Y_b|$. The latter is called the *tradeoff of distance components* (Appendix A).

We now ask the same question about space–time. Suppose one of the two dimensions in Fig. 2b is time. To preserve an equality of distances in space–time, the spatial and temporal components of spatiotemporal distance must trade off, just as they did in space: an increment or decrement in the spatial distance between elements must be accompanied by a decrement or an increment in temporal distance.

Such a tradeoff was found by Burt and Sperling (1981): the longer they made the spatial gap between dots, the more they had to shorten the temporal interval between dots for apparent motion to be seen. In contrast, however, according to Korte's Third Law of Motion (Korte, 1915; Koffka, 1935/1963), the larger the spatial gap between alternating lights, the slower the rate at which they need to be flashed in alternation for apparent motion to be seen. Koffka (1935/1963, p. 293) himself found this result counterintuitive:



(a) A stimulus for ambiguous apparent motion. Element *O* has two potential matches, *a* and *b*, giving rise to potential motion paths m_a and m_b.

(b) Procedure to find the equilibrium between competing motion paths m_a and m_b. Each motion path is represented by a coordinate in time-space, $(T_i;\ S_i)$; where $i \in \{a, b\}$. The double-headed arrow represents the manipulation of path O→ b.

Figure 3. Tradeoff and coupling of spatiotemporal distance components.

> [...] when Korte and I discovered it, I was surprised [...]: if one separates the two successively exposed objects more and more, either spatially or temporally, one makes their unification more and more difficult. Therefore increase of distance should be compensated by decrease of time interval, and *vice versa*.

In an attempt to resolve these inconsistent results, Gepshtein and Kubovy (2007) devised the following procedure. Three short-lived dots, O, a, and b, appear and disappear sequentially at three locations in space (Fig. 3a). Nothing prevents us from seeing apparent motion O→a or O→b. (The distance between a and b is too long for a→b.) We call the O→a motion $\boldsymbol{m}_{\mathrm{a}}$, and the O→b motion $\boldsymbol{m}_{\mathrm{b}}$. Each of these has a temporal and a spatial component: (T_{a}, S_{a}) and ($T_{\mathrm{b}}, S_{\mathrm{b}}$). This allows us to represent each motion as a point in a plot of distances (Fig. 3b).

Why did Korte's law puzzle psychologists while the result of Burt and Sperling does not appear surprising? It is probably because space-time coupling contradicts the widespread intuition of distance: the fact that to preserve distance its components must tradeoff (Appendix B).

If we allow S_{b} to vary (represented by the interval between ❶ and ❷ connected by the double-headed arrow in Fig. 3b), while holding S_{a}, T_{a} and T_{b} constant, with the condition that $T_{\mathrm{b}} = 2T_a$. We can vary S_{b} until we find a value $S_{\mathrm{b}} = S_{\mathrm{b}}^\star$ for which $p(\boldsymbol{m}_b) = p(\boldsymbol{m}_a)$. In light of the previous literature, we pit two hypotheses against each other (an intermediate hypothesis is discussed in Gepshtein & Kubovy, 2007):

- *Space–time tradeoff* ($S_{\mathrm{b}}^\star < S_{\mathrm{a}}$), which supports the proximity principle in space-time (because $T_{\mathrm{b}} > T_{\mathrm{a}}$). In Fig. 3b this result is represented by outcome ❶ where the line connecting the conditions of equilibrium has a negative slope.
- *Space–time coupling* ($S_{\mathrm{b}}^\star < S_{\mathrm{a}}$), where the proximity principle is not applicable. In Fig. 3b this result is represented by outcome ❷ where the line connecting the conditions of equilibrium has a positive slope.

Using the manipulation represented by the double-headed arrow in Fig. 3b, Gepshtein and Kubovy (2007) varied T_{a} and S_{b}, as shown in the lower half of Fig. 4. The graphs are plotted as a function of motion speed (S_{b}/T_a) in panel A and as a function of the reciprocal of motion speed, i.e., slowness (T_{a}/S_b) in panel B. The response variable is the ratio $r_{13}^\star = S_{\mathrm{b}}^\star / S_{\mathrm{a}}$. When $r_{13}^\star < 1$, we have space–time tradeoff, whereas when $r_{13}^\star > 1$, we have space–time coupling. Since the functions in panels A and B cross the boundary $r_{13}^\star = 1$, both tradeoff and coupling occur, depending on the speed (or slowness) of the motion. Tradeoff occurs at low speeds (i.e., at small spatial and large temporal distances), but as the speed is increased (i.e., toward large spatial and small temporal distances), eventually we observe coupling.

In Fig. 5 we transfer the data of Fig. 4 to a representation similar to figure Fig. 3b. The thin lines on the background are the empirical *equivalence contours of apparent motion* we derived 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). That is, the results are consistent with the proximity principle at some conditions, where tradeoff is observed. But the results are inconsistent with the proximity principle at other conditions, where coupling is observed.

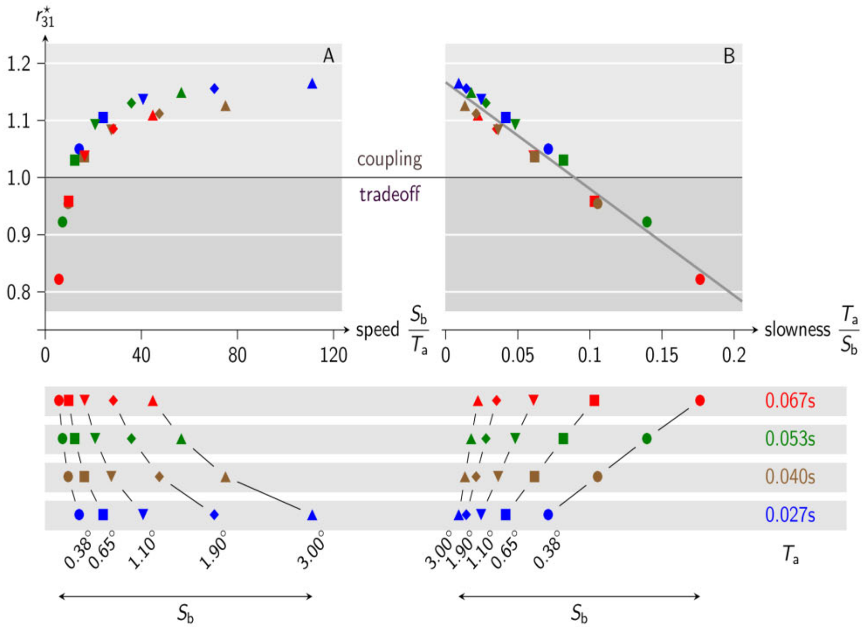


Figure 4. An idealized representation of the data of Gepshtein and Kubovy (2007). We jittered the "data" points vertically to improve the legibility of the figure.

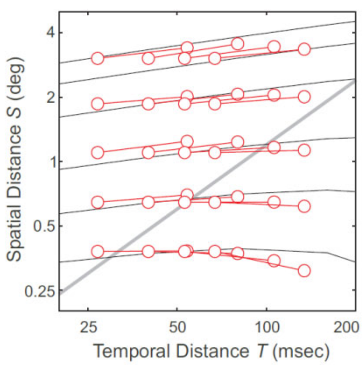


Figure 5. Equivalence classes of motion perception. The pairs of red connected circles represent the pairs of conditions of apparent motion that were seen equally often in the displays of Gepshtein and Kubovy (2007). The slopes of the lines connecting the equilibria are positive and negative in different parts of the parameter space.

Summary and Resolution

The failure of the proximity principle in dynamic displays undermines its generality as a law of perceptual organization. The principle holds only when components of distance between visual elements trade off to preserve strength of perceptual grouping. This requirement is not met in the perception of motion. When stimulus elements are separated by spatiotemporal distances, strength of grouping is preserved sometimes when the spatial and temporal distance components tradeoff, and sometimes when distance components are coupled: both increase or both decrease.

This insight led us to ask what general characteristic of visual systems may supplant the proximity principle (Gepshtein, Tyukin, & Kubovy, 2007). We showed that the results summarized in Figs. 4–5 are predicted by two properties of visual systems: (a) intrinsic limitations of visual measurements (Gabor, 1946; Daugman, 1985; Jones & Palmer, 1987) and (b) constraints on the number of measurements visual systems can perform concurrently. To account for the failure of the proximity principle, this point of view appeals to facts more basic and general than perceptual organization or perception of motion. Thus the tensions created by the apparent inconsistency of experimental findings (Korte, 1915; Burt & Sperling, 1981), and the contradiction between experimental findings and one's intuitive concept of distance (Korte, 1915; Koffka, 1935/1963) find a simple resolution.

Appendix A. Decomposability of Distance Components

We demonstrate that tradeoff of distance components is a necessary property of a proximity metric. As we illustrated in Figure 2, distances δ of $\boldsymbol{a}$ and $\boldsymbol{b}$ can be mapped onto each other by rotation while preserving distance equality. This property is called *rotation invariance* (a case of *metric equivalence*; Mendelson, 1974). It holds in the familiar Euclidean metric.

The Euclidean metric is a special case of the power metric. Although rotation invariance does not generally 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

$$\delta(MN) = [\psi_s(M_s, N_s)^r + \psi_t(M_t, N_t)^r]^{1/r}, \qquad (1)$$

where

- ψ_s and ψ_t are the spatial and temporal differences between locations M and N in space-time, $\psi_i = |\phi(\mathrm{M}_i) - \phi(\mathrm{N}_i)|$ satisfying, $\psi_i(\mathrm{M}_i, \mathrm{N}_i) > \psi_i(\mathrm{M}_i, M_i)$ whenever $\mathrm{M}_i \neq \mathrm{N}_i$,
- 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 define function F:

$$\delta(MN) = F[\psi_s(M_s, N_s), \psi_t(M_t, N_t)], \qquad (2)$$

which must increase whenever $\psi_s(\mathrm{M}_s, \mathrm{N}_s)$ or $\psi_t(\mathrm{M}_t; \mathrm{N}_t)$ increases. According to decomposability, if one of the arguments of distance function (e.g., the X-projection in Figure 2b) increases, then distance is preserved only if the other argument (the Y-projection in Figure 2b) decreases. If the second argument had not decreased, then the distance would necessarily have increased.

We now apply this argument to the case of multistability in motion perception (Figure 3), where the spatiotemporal distances of competing motion paths are $\delta(\mathbf{m}_a)$and $\delta(\mathbf{m}_b)$. Let the spatial and temporal coordinates of points *o*, a, and *b* be $(\mathrm{N}_{s,o}, \mathrm{N}_{t,o})$, $(\mathrm{M}_{s,a}, \mathrm{N}_{t,a})$, and $(\mathrm{M}_{s,b}, \mathrm{N}_{t,b})$. Suppose that:

1. $\psi_s(\mathrm{M}_{s,a}, \mathrm{M}_{s,o}) = S_a$, $\psi_s(\mathrm{M}_{s,b}, \mathrm{M}_{s,o}) = S_b, S_b = S_a + \Delta S$ and
2. $\psi_t(\mathrm{N}_{t,a}, \mathrm{N}_{t,o}) = T_a$, $\psi_t(\mathrm{N}_{t,b}, \mathrm{N}_{t,o}) = T_b$ where $T_b = T_a + \Delta T$

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

$$F[S_a, T_a] = F[S_a + \Delta S, T_a + \Delta T]. \qquad (3)$$

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 figure 2b we interpret axis X as space, and axis Y as time, then the lengths of spatial and temporal projections of perceptually equivalent spatiotemporal segments $\boldsymbol{a}$ and $\boldsymbol{b}$ will trade off. Applied to the apparent-motion display in Figure 3b, Equation 3 becomes:

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

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

Appendix B. Tradeoff of Distance Components

As mentioned in Appendix A, the intuitive Euclidean metric is a special case of the power metric. Although rotation invariance illustrated in Fig. 2b does not generally hold in the power metric, the tradeoff of distance components does. The generality of this tradeoff follows from the decomposability property of the power metric (Suppes et al., 1989), according to which a distance function must be a strictly monotonically increasing function in each of its arguments.

To see why tradeoff of distance components is a general property of the power metric, let $d(|x_1 - x_2|, |y_1 - y_2|)$, be the function of distance between points (x_1, y_1) and (x_2, y_2). We shall only assume that $d(\cdot,\cdot)$ satisfies the requirement of decomposability. Let us fix the initial distance between (x_1, y_1) and (x_2, y_2), and let

$$d\,(\,|\,x_1 - x_2\,|\,,\,|\,y_1 - y_2\,|\,) = d_0.$$

Now consider another point, (x_3, y_3), such that $d(|x_1 - x_3|, |y_1 - y_3|) = d_0$ and $|x_1 - x_3| > |x_1 - x_2|$. This means that one of the distance components is increased but the distance between two points did not. Let $|y_1 - y_2| \leq |y_1 - y_3|$, which represents the hypothesis that the other distance component did not increase. From this we have, because of the decomposability of $d\,(\cdot,\cdot)$:

$$\begin{aligned} d_0 = d\,(|x_1 - x_2|, |y_1 - y_2|) &< d\,(|x_1 - x_3|, |y_1 - y_2|) \\ &\leq d(|x_1 - x_3|, |y_1 - y_3|), \end{aligned}$$

which leads to a contradiction:

$$d_0 < d(|x_1 - x_3|, |y_1 - y_3|) = d_0.$$

Hence, our hypothesis that $|y_1 - y_2| \leq |y_1 - y_3|$ is false. That is, $|y_1 - y_2|$, must be strictly larger than $|y_1 - y_3|$. Hence the tradeoff of distance components in the power metric.

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.

Daugman, J.G. (1985). Uncertainty relation for the resolution in space spatial frequency, and orientation optimized by two-dimensional visual cortex filters. *Journal of the Optical Society of America A*, *2* (7), 1160–1169.

Elder, J.H., & 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. (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.

Hochberg, J., & Silverstein, A. (1956). A quantitative index of stimulus-similarity: Proximity vs. differences in brightness. *American Journal of Psychology*, *69* (3), 456–458.

Jones, A., & Palmer, L. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*, 1233–1258.

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

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

Kubovy, M. (1994). The perceptual organization of dot lattices. *Psychonomic Bulletin & Review*, *1* (2), 182–190.

Kubovy, M., Holcombe, A.O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35* (1), 71–98.

Kubovy, M., & van den Berg, M. (2008). The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychological Review*, *115* (1), 131–154.

Kubovy, M., & Wagemans, J. (1995). Grouping by proximity and multistability in dot lattices: A quantitative gestalt theory. *Psychological Science*, *6* (4), 225–234.

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Iccv'01* (Vol. 2, p. 416). Eighth International Conference on Computer Vision.

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

Oyama, T. (1961). Perceptual grouping as a function of proximity. *Perceptual & 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. (Translated extract in Ellis, 1938, pp. 71–88).