

Gestalt neurons and emergent properties in visual perception: A novel concept for the transformation from local to global processing

Lothar Spillmann

Department of Neurology, University of Freiburg,
Freiburg, Germany



Li-Chuan Hsu

China Medical University, Taichung, Taiwan



Wei Wang

Institute of Neuroscience, Chinese Academy of Sciences,
Shanghai, China



Chien-Chung Chen

Department of Psychology, National Taiwan University,
Taipei, Taiwan



Chung-I Yeh

Department of Psychology, National Taiwan University,
Taipei, Taiwan



Chia-Huei Tseng

Research Institute of Electrical Communication,
Interdisciplinary ICT Research Center for Cyber and
Real Spaces, Tohoku University, Sendai, Japan



Dedicated to the memory of Richard L. Gregory, who would have been 100 years on July 24, 2023.

Gestalten in visual perception are defined by *emergent properties* of the *whole*, which cannot be predicted from the *sum of its parts*; rather, they arise by virtue of inherent principles, the *Laws of Seeing*. This review attempts to assign neurophysiological correlates to select *emergent properties* in motion and contour perception and proposes parallels to the processing of *local versus global* attributes by *classical versus contextual* receptive fields. The aim is to identify *Gestalt neurons* in the visual system to account for the *Laws of Seeing* in causal terms and to explain “Why do things look as they do” (Koffka, 1935, p. 76).

(Metzger, 1936; Metzger, 2006; Todorović, 2008). These laws are inherent principles responsible for the way in which (a) scenes are segmented, (b) figures are segregated, and (c) elements are grouped, thereby becoming foreground and background. They are effective both in humans as well as species of animals (Lorenz, 1951). This is best documented by camouflage, where the same Gestalt factors that are responsible for making an animal stand out in its environment are being used by nature to render it invisible (Metzger, 2006). These factors are assumed to be largely innate. Metzger called them *Bedingungen der Möglichkeit für Erfahrung* (prerequisites for enabling experience) in the Kantian sense, not just in vision but also in touch and other sense modalities (Gallace & Spence, 2011).

From observations made in simple pencil drawings, Max Wertheimer (1923), founder and protagonist of the Gestalt movement, identified the factors responsible for intrafigural and figure-ground perception, later expanded by Metzger (1936). These factors are symmetry, good continuation, and closure for *segmentation* (i.e., for organizing contiguous parts into a structured whole or *Gestalt*) and proximity, similarity, and common fate for *grouping* (i.e., for linking spatially separated parts into a coherently grouped perceptual

Introduction

Starting at the beginning of the 20th century, Gestalt psychologists famously proposed that perceptual mechanisms organize, or disambiguate, the physical stimulus through segmentation, segregation, and grouping. The principles according to which the visual input is so constrained are called the *Laws of Seeing*

Citation: Spillmann, L., Hsu, L.-C., Wang, W., Chen, C.-C., Yeh, C.-I, & Tseng, C.-H. (2023). Gestalt neurons and emergent properties in visual perception: A novel concept for the transformation from local to global processing. *Journal of Vision*, 23(12):8, 1–15, <https://doi.org/10.1167/jov.23.12.8>.



pattern). Common to these principles is the claim that *the whole is different from*, or superordinate to, *the sum of its parts* and that the whole follows the law of good Gestalt or *Prägnanz* (i.e., exhibiting simplicity, regularity, and inner balance). Strong long-range correlations in the distribution of oriented edges or line segments have been demonstrated in natural scenes (Sigman, Cecchi, Gilbert, & Magnasco, 2001), suggesting a preponderance for *smooth continuation* and *collinearity*. Local image features such as lines, edges, and corners are represented by the response properties of simple and complex cells in visual areas V1 and V2 of the cat and monkey (Hubel & Wiesel, 1959; Hubel & Wiesel, 1968), before they are integrated into object properties in the inferior temporal cortex (Desimone, Albright, Gross, & Bruce, 1984).

There have been many studies demonstrating that visual perception can be described by Gestalt principles, but a concerted effort to find neurophysiological correlates of such principles to explain perception in causal terms is lacking. This is astonishing as Wertheimer (1912, p. 247; 2012, p. 57), in his pioneering article on apparent motion, already suggested that this percept could be interpreted “by physiological transverse processes [*Querfunktionen*] of a special kind that serve as the physiological correlate of the ϕ -phenomena.” His interpretation of apparent motion as an *emergent property* that cannot be reduced to the sequential presentation of two static stimuli started Gestalt psychology. Wertheimer (1923) called this “*Ganzbestimmtheit der Teile*” (determination of the parts by the whole).

In a noteworthy article, Westheimer (1999) writes, “Is there a better example of Wertheimer’s vision of a whole determining the behavior of its constituent parts than an experimental verification of the fact that what a visual cortical neuron responds to best depends more on the properties of the overall configuration in the visual field than on the parameters of the stimulus in its receptive field” (p. 11).

Surveys of the literature (Spillmann, 1997; Spillmann, 1999; Spillmann, 2006; Spillmann, 2009; Spillmann, 2012) show that there are neuronal correlates for many of the striking phenomena described by the Gestaltists, often characterized by *emergent properties* and a corresponding neuronal response pattern. They also show that today’s leading visual neuroscientists and neurocomputational theorists are increasingly using Gestalt terms to assign visual phenomena to neuronal mechanisms (e.g., von der Heydt, 1984; Allman, Miezin, & McGuinness, 1985; Grossberg & Mingolla, 1985; Singer, 1989; Lamme, 1995; Li, 1998; Roelfsema & Singer, 1998; Albright & Stoner, 2002; Kourtzi, Tolia, Altmann, Augath, & Logothetis, 2003; Li, Piëch, & Gilbert, 2006; Li, Piëch, & Gilbert, 2008; Pan et al., 2016) This article uses select examples from the fields of motion and contour perception to raise awareness for

neurophysiological correlations of phenomenological observations (Spillmann, 2009).

Apparent motion

Wertheimer’s (1912) *apparent motion* arises from the presentation of two spatially separated bars (or dots) presented in alternation (Figure 1). Under appropriate conditions, continuous motion from location *a* to location *b* is perceived on a straight path traversing the intervening space, although physically there is none. This percept is illusory and is an *emergent property*, which is different from the sum of its parts. The motion path may be filled in either with object motion (called *beta*), where the stimulus appears to traverse the nonstimulated interspace, or with pure motion (called *phi*), where only back-and-forth motion is perceived without a stimulus. Object motion is utilized by the zoetrope, where a band of individual static pictures is presented on the inside of a rotating cylinder (Metzger, 1953, p. 411; Veras, Quang-Cuong, & Maus, 2017). When the figures are viewed sequentially through a set of vertical slits passing by in front of the eye, the observer sees a single figure in continuous motion, the forerunner of today’s movie theater. Speed is critical; Charlie Chaplin’s jerky gait is an example of what happens when the number of frames/s is too low. Apparent motion need not be straight; it also occurs along a curvilinear path if supported by a series of stimulus exposures in appropriate places and at appropriate times.

To explain apparent motion, Wertheimer (1912, p. 62) writes, “If place *a* is stimulated and a certain short time thereafter the nearby place *b*, then a kind of physiological *short circuit* would occur from *a* to *b*. . . . The closer the two places *a* and *b* are to each other, the

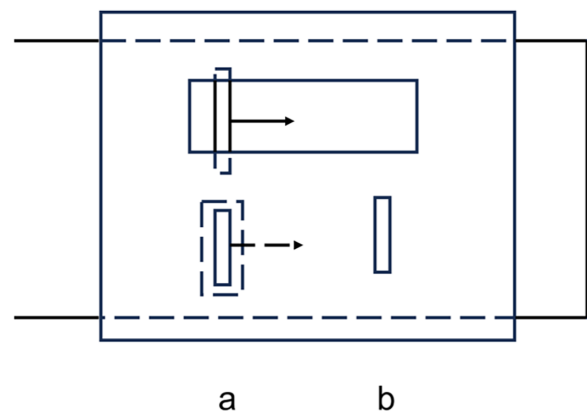


Figure 1. Sliding frame for displaying real motion (top) and apparent motion (bottom). The two vertical bars, *a* and *b*, on the bottom are static and are presented in succession. (From Metzger, 1953, p. 413; after Wertheimer, 1912.)

more favorable the conditions for the emergence of the ϕ -process” (Gestalt factor of *proximity*). Wertheimer does not go into the seeming paradox that “the direction of [apparent motion] which is given by the fact that *a* and its spreading activation were there first” is seen before the second stimulus is exposed.

In search of a physiological explanation, Jung and Spillmann (1970) argued that apparent motion is perceived as long as the two stimuli interact within the same receptive field; when one stimulus falls outside the receptive field, apparent motion breaks down. Using this assumption, they defined the spatial range or receptive field size of a motion detector in human vision by determining the maximum distance between stimuli *a* and *b* for which apparent motion prevailed. Receptive fields for apparent motion measured in this manner were up to 20 times larger than for contrast stimuli and increased in size with increasing retinal eccentricity (Jung & Spillmann, 1970; Spillmann, 2006, Figure 5). This finding anticipated neurophysiological measurements in the rhesus monkey, which show receptive fields of area MT neurons much larger than those of retinal ganglion cells or V1 neurons (Britten, 2004).

There is now evidence in support of this interpretation from functional magnetic resonance imaging (fMRI) experiments on long-range apparent motion (Liu, Slotnick, & Yantis, 2004). Specifically, an increase of the blood oxygenation–level dependent (BOLD) response was found along the V1 representation of the apparent motion path and was attributed to top-down activity from area MT/V5, where neurons have receptive fields large enough to span the distance between the two stimuli inducing the apparent motion percept (Muckli, Kohler, Kriegeskorte, & Singer, 2006; Sterzer, Haynes, & Rees, 2006).

This prompts the question whether real and apparent motion are equivalent, not only in appearance but also in the way in which the different motion stimuli are processed in the brain. Early on, Grüsser and Grüsser-Cornehls (1973) recorded the response of motion-sensitive neurons in the frog and found that the neuronal response for real and apparent motion was the same, regardless of whether the motion stimulus was continuous or successive (i.e., separated in space and time). The observation that the two kinds of motion elicited in this way were indistinguishable is consistent with the motion detector explanation proposed by Hassenstein and Reichardt (1956) for the compound eye of the beetle *Chlorophanus*.

Aperture motion

Real motion differs from apparent motion by continuous stimulation of the eye, yet its perceived

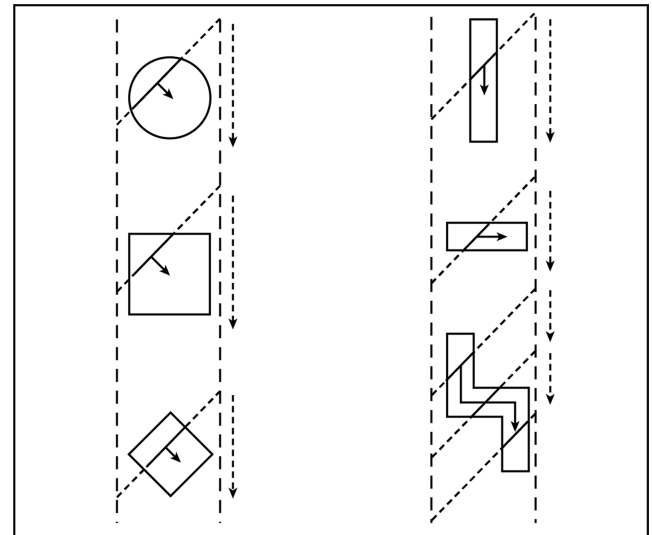


Figure 2. Aperture motion. The real motion direction of the line in each case is vertical (dashed arrows), whereas the perceived motion direction (solid arrows) occurs at right angles to the moving line (left) or conforms to the shape of the aperture (right). (From Metzger, 1953, p. 407; after Wallach, 1935.)

direction is not uniquely determined. Indeed, whereas under normal circumstances, the perceived direction of a moving line depends on the movement of its *extrinsic* endpoints, the perceived direction in an aperture is governed by the shape of the window (Wallach, 1935). If the aperture is round or square shaped (Figure 2, left), the line will always appear to move at right angles to its orientation, regardless of its true direction. This is because in this case, the only cues for a motion-sensitive neuron are the two *intrinsic* endpoints, where the line intersects the frame (Adelson & Movshon, 1982).

However, if the aperture consists of three interlinked rectangles such as in Figure 2 (right), the perceived direction of the line will be constrained by the overall shape of the aperture, moving first downward, then horizontally and thereafter downward again, although, in reality, it always moves in the same vertical direction. Perceived motion direction in an aperture therefore is an *emergent property* that depends on the interaction between the local movement of the line and the enclosing aperture. Again, the whole is different from the sum of its parts. Notwithstanding the change in local movement direction at each bend, the global percept of the moving line refers to one and the same stimulus (object identity).

Meanwhile, MT cells have been reported to be capable of solving the aperture problem. Unlike area V1 neurons, which have small receptive fields and thus permit only a limited *view* of a moving stimulus, MT neurons, which have large receptive fields, initially

respond primarily to a motion direction, which is perpendicular to the orientation of the stimulus, but over a period of approximately 60 ms, the response gradually shifts to encode the true stimulus direction, regardless of orientation (Pack & Born, 2001).

Coherent motion

The strongest factor for grouping is coherent motion such as illustrated by a dotted pattern on a static random dot background (Figure 3). When the dots shown in black move in unison, the two populations will instantaneously segregate according to the Gestalt factor of *common fate* (Singer, 1989). This even works on a dynamic noise field, so that a few select dots (the parts) moving together in one direction and at the same speed will *pop out* as a group (the whole). Stürzel and Spillmann (2004) have shown that only four such dots moving on a dynamic noise field of 100 dots will be seen as a group, even if the exposure time is as short as 430 ms. Global motion can be seen when as few as 35% of the dots move coherently (Chen, Ashida, Yang, & Chen, 2020). Grouping of local events by *coherent motion* is an *emergent property*. In real life, a few people walking together as a group would be expected to similarly stand out in a crowd of a hundred people if seen from above.

Is there any neurophysiological evidence for this kind of grouping? In a carefully designed experiment, Britten, Shadlen, Newsome, and Movshon (1992) showed that neurons in primate area MT respond strongly to coherent movement. They even demonstrated that the neuronal threshold and

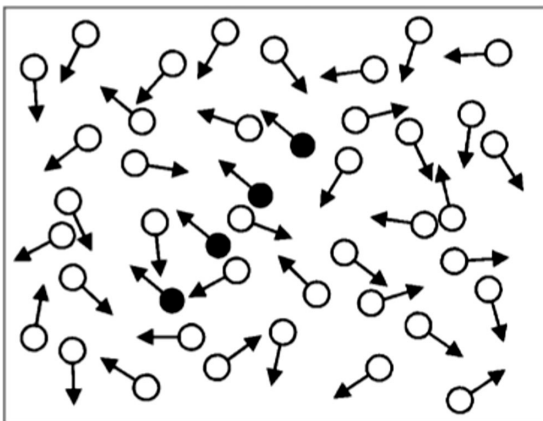


Figure 3. Coherent motion, static display. Four dots (marked black for better visualization) moving in the same direction on a background of 100 randomly moving dots will pop out as a group. The speed and contrast to the ground (white on black) were the same for all dots. (From Stürzel & Spillmann, 2004, Copyright © 2004 Elsevier Ltd.)

the behavioral threshold measured simultaneously were comparable in the same animal. These results demonstrate that the Gestalt factor of *common fate* is a basic mechanism of our neuronal inventory. Neuroimaging in the human has confirmed area hMT(V5) as the brain locus responsible for mediating perception of motion coherence (Braddick et al., 2001).

Biological motion

Another movement phenomenon, for which the whole is different from the sum of its parts, is *biological motion* (Johansson, 1950; Johansson, 1973). Here, a point walker, represented by small light bulbs affixed to the major body joints (Figure 4A) when moving in the dark, serves as a surrogate for a stick walker in daylight where all dots are connected by short sticks (Figure 4B). The percept of a dancing, climbing, and jumping person can be generated in this manner from local motions of the head, arms, and legs (the parts) and will easily and instantaneously be perceived as a global figure in coherent motion (the whole). This kind of motion owes to the fixed relationship between pairs of neighboring dots, which serve as subunits from which global motion of the entire figure derives. This is an *emergent property*. Evidently, the shape of each figure cannot be predicted from the sum of the individual trajectories. Biological motion here occurs due to grouping by the Gestalt factors of *proximity*, *good continuation*, and *common fate*, although the individual dots of the point walker go in different directions and have different motion vectors.

Common fate applies in conjunction with the factor of exhaustiveness (*Gestalt Faktor des Aufgehens ohne Rest*), which states that in a spatiotemporal scene, the components of the previous one are included in the following one in the interest of preserving a continued percept. The perception of biological motion is not tied to the perception of a human being but occurs equally well with animals and presumably also with dynamically changing objects.

Is there a neurophysiological correlate to support these perceptual observations? Neurons have been reported in area STPa (the anterior superior temporal polysensory area) of the macaque, which respond to biological motion patterns (Oram & Perret, 1994). A study by Peterhans, Heider, and Baumann (2005) further shows strong responses to rows of coherently moving dots already in macaque areas V2 and V3. Here we may ask whether learning and memory provide templates, which prime the response to sparse stimulus patterns such as the point walker in Figure 4.

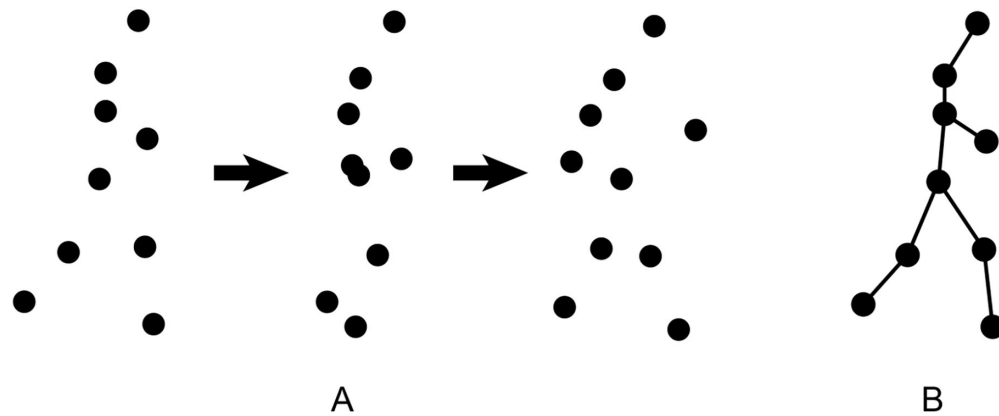


Figure 4. Biological motion. Sequential shots of a point walker. (A) Despite different motion vectors of the dots, one perceives a coherently moving figure not unlike a person walking in real life. (B) A stick figure is shown for comparison to illustrate how local motion between pairs of adjacent dots may lead to a unified percept of global motion. (After Johansson, 1973.)

Three-dimensional structure from motion

The examples discussed so far were all confined to a flat surface. Yet, coherent motion also works for three-dimensional (3D) surfaces. Bradley, Chang, and Andersen (1998) used a transparent vertical cylinder covered with black dots and rotating around its vertical axis to demonstrate scission in 3D and structure-from-motion (Figure 5). This kind of stimulus creates the impression of two superimposed populations of dots moving coherently in opposite

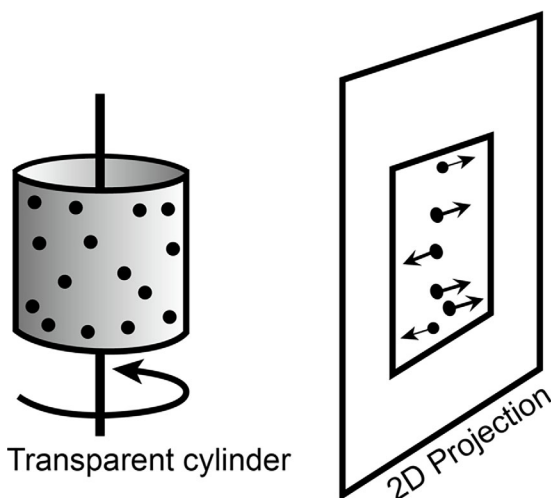


Figure 5. Three-dimensional structure from motion. Random dots presented on a transparent cylinder revolving around its axis segregate in depth and direction depending on whether they are seen in front or in back. The two-dimensional projection of the moving dots is shown on the right. Note that the linear speed of each dot is not constant across. (from Bradley, Chang, & Andersen, 1998.)

directions and at different depths according to the Gestalt factor of *common fate*. While the dots in front (the parts) appear to move in one direction, the dots in the back seem to move in the other, generating a unified percept of a rotating cylinder (the whole). The sense of rotation is ambiguous but subjects generally perceive one or the other. Human observers as well as monkeys perceive the depth almost immediately even with one eye. This is called motion parallax or motion transparency and is an *emergent property*.

In a neurophysiological experiment combined with behavioral testing, monkeys signaled the direction in which the front of the cylinder appeared to be moving for them. At the same time, neuronal activity in brain areas V1 and MT was recorded from single cells. In area V1, 20% of the neurons reacted to the stimulus as compared to more than 60% in brain area MT. Thus, MT neurons are more involved in the coding of structure-from-motion than V1 neurons, which may receive feedback from MT (Grunewald, Bradley, & Andersen, 2002).

Illusory rotation (Pinna illusion)

Illusory motion arises in a pattern consisting of two concentric rings of small static rhombi (lozenges), delineated by narrow black and white L-shaped edges of opposite polarity (Figure 6) (Pinna & Brelstaff, 2000). When one approaches or recedes from this pattern, while fixating in the center, one perceives an apparent rotation of the two rings in a counterclockwise and clockwise direction, respectively. Note that these rhombi move radially outward or inward when one changes the observation distance, not circularly. The illusory counterrotation is also observed with small inclined squares instead of rhombi and is

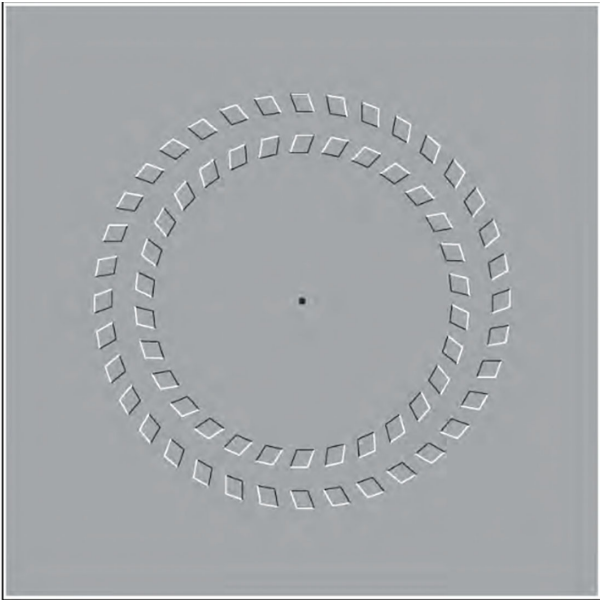


Figure 6. Pinna's illusion of apparent rotation. With the observer bobbing forward and backward while fixating in the center, the two rings composed of small static rhombi appear to rotate in opposite directions. With forward motion of the head, the inner ring appears to rotate counterclockwise, the outer ring clockwise. With backward motion of the head, all directions are reversed. (Reprinted from Pinna & Brelstaff (2000). A new visual illusion of relative motion. *Vision Research*, 40, 2091–2096, Copyright © 2000 Elsevier Science Ltd., with permission from Elsevier.)

likely due to the opposite juxtaposition of the black and white edges delineating the gray stimulus elements and their oblique orientation relative to the radial expansion and constriction. The precise origin of the illusory movement, however, in terms of processing by an orientation- and direction-sensitive neuron is still unclear (Pinna, 2009). Although the transformation from radial stimulus movement to apparent rotation has not yet been fully understood, the perceptual grouping of the local elements (the parts) into a global ring-shaped percept (the whole) may be attributed to the Gestalt factors of *proximity*, *similarity*, and common fate and constitutes an *emergent property*.

There is neural evidence that the illusory rotation in the Pinna–Brelstaff figure is initiated in human area MST (Pan et al., 2016; Luo et al., 2019), which normally encodes physical rotary motion (Tanaka & Saito, 1989). Both percepts presumably rely on a cascade of integrative processes from earlier visual areas.

Contour integration

We here change the topic to the discussion of contour phenomena. Grouping of collinear elements also

occurs when a string of static aligned Gabor patches is embedded on a randomly oriented static Gabor background. Field, Hayes, and Hess (1993) studied how strings connect to each other in visual perception. To find out, they varied the overall orientation of the path, the orientation of the individual elements, and the distance between the elements. These variables imply a combined effect of the Gestalt factors of *good continuation*, *similarity*, and *proximity*. Deviation from collinearity more than ± 30 degrees affected detectability the most, suggesting that the Gestalt factor of *good continuation* was critical for contour integration. Field and Hayes (2004) proposed that the detection of elements forming a path involves long-range neuronal interactions within a local *association field* in area V1, which integrates information across neighboring receptive fields tuned to similar orientations.

Figure 7A shows a string of aligned Gabor patches resembling a snake embedded in a field of randomly oriented Gabor patches, which becomes visible by virtue of the Gestalt factor of *good continuation* (Field et al., 1993). Kovács and Julesz (1993) independently demonstrated a similar effect for a semicircular arc (Figure 7B) and a ring popping out due to the additional Gestalt factors of *closure* and *Prägnanz* (Figure 7C). This change from local patches to a global shape is an *emergent property*, showing that the whole is different from the sum of its parts. This structure has been shown to guide attentional distribution in visual search tasks (e.g., Jingling, Tseng, & Zhaoping, 2013; Tseng & Jingling, 2015), where the underlying mechanism is unlikely to be subserved by early vision exclusively (Chow, Jingling, & Tseng, 2013; Chow, Jingling, & Tseng, 2016; Tseng, Chow, Liang, Shioiri, & Chen, 2021).

Long-range horizontal connections in the primary visual cortex of the cat and monkey (Nelson & Frost, 1985; Gilbert & Wiesel, 1989) suggest that contour integration is based on contextual neurons with similar orientation preferences in striate cortex (Li, Piëch, & Gilbert, 2006). However, fMRI findings by Kourtzi et al., (2003) suggest that the unified perception of collinear patterns also involves visual areas V2 and lateral occipital cortex (LOC). Importantly, Li, Piëch, and Gilbert (2008) used embedded contours such as described in Figure 7 for training monkeys and found contour integration in area V1 to depend strongly on *perceptual learning* and top-down influences specific to contour detection. Furthermore, by using simultaneous recordings in both areas V1 and V4 in the awake monkey, Chen et al. (2014) reported that embedded contours emerged initially in V4 and only 40 ms later in V1. Therefore, they propose that contour integration owes to both bottom-up and top-down (reentrant) processes.

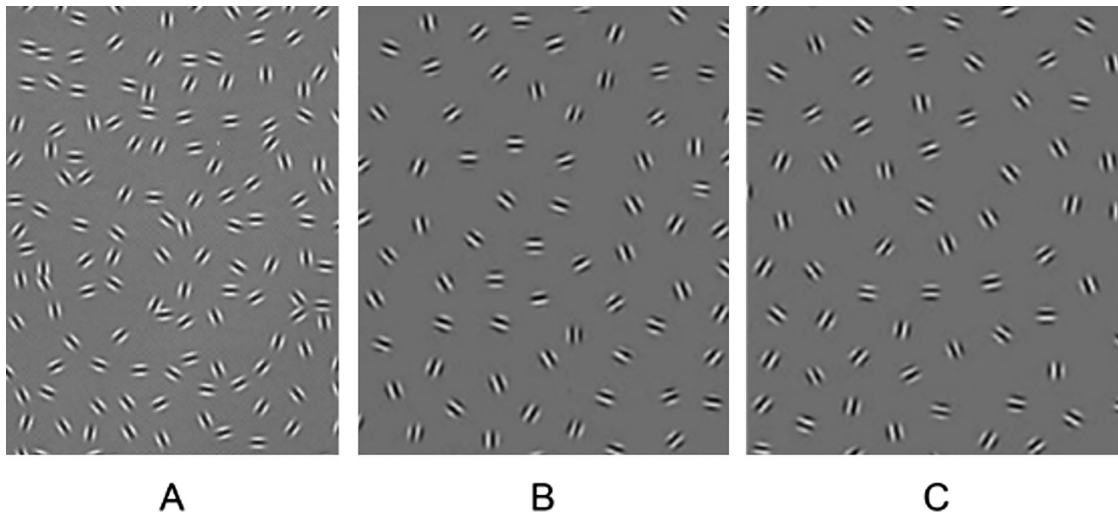


Figure 7. Contour integration. The hidden “snake” in (A) composed of approximately collinear Gabor elements connects due to the Gestalt factor of *good continuation* (from Field, Hayes, & Hess, 1993, available under a CC BY-NC-ND license). Compared to the semicircular arc in (B), the complete circle in (C) is more conspicuous due to the additional Gestalt factor of *closure* and *Prägnanz* (Kovács & Julesz, 1993).

Orientation contrast

Collinearity of textural elements is the prerequisite for orientation contrast, which is a strong factor for segregating a figure from the ground. Lamme (1995) demonstrated that neurons in the primary visual cortex respond to an oriented texture in a “figure” region defined by cross-orientation more strongly than to the same texture being iso-orientated with the ground (Figure 8). This response indicates that the neurons must be sensitive to orientation contrast at the border (Knierim & van Essen, 1992; Kastner, Nothdurft, & Pigarev, 1999). The Gestalt factor of *similarity* may be at work here, involving long-range interaction beyond the classical receptive field (Spillmann & Werner, 1996; Spillmann, Dresch-Langley, & Tseng, 2015). Consistent with this finding, short stimulus elements (textels) adjacent to the contrast border were found to elicit a stronger response than textels placed further away (Nothdurft, Gallant, & van Essen, 2000). Also, the perceived contrast of a Gabor probe placed on the figure versus the background texture was higher (Self, Mookhoek, Tjalma, & Roelfsema, 2015). This pop-out or figure enhancement effect (Lamme, 1995) is an *emergent property*, which derives from the interaction of the central patch (the part) and its cross-oriented surround (the whole).

The question is, how does a neuron in area V1 “know” what is figure and what is ground and whether the edge stimulating its receptive field belongs to the figure?

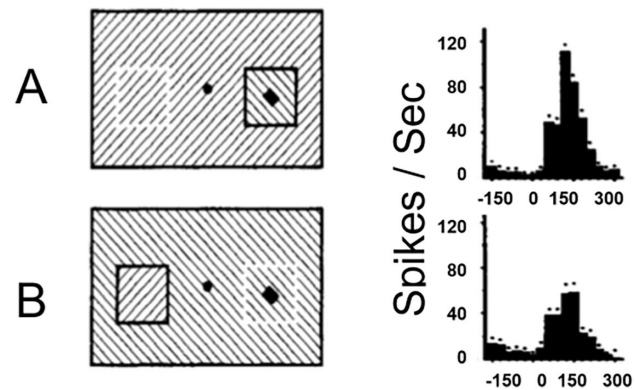


Figure 8. Orientation contrast (top). A neuron whose receptive field (black rectangle) is located (A) inside the cross-oriented patch responds more strongly than (B) the same neuron when the patch is iso-oriented with the surround. Neuron responses on the right. (From Lamme, 1995, Copyright © 1995 by Society for Neuroscience.)

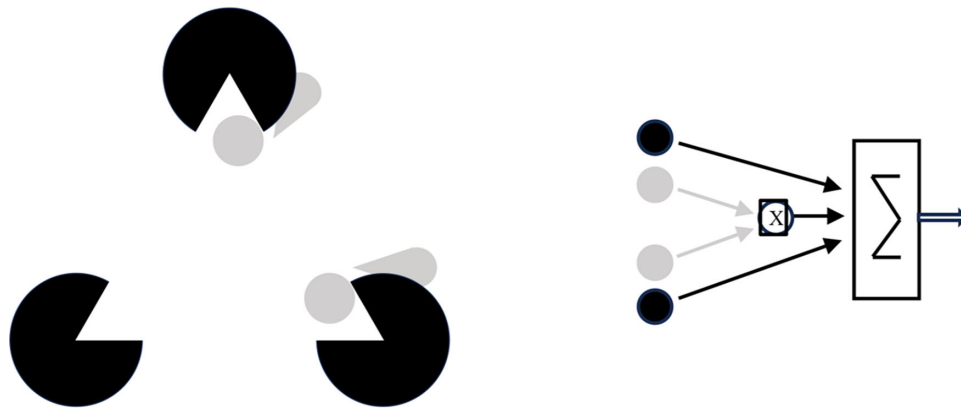


Figure 9. Contour completion in the Kanizsa triangle. Left: A bright illusory triangle delineated by illusory contours is seen superimposed onto three black disks (pacmen). Shaded patches illustrate orthogonal receptive fields of end-stopped neurons in area V2 located at right angles to the aligned edges (gray oval: excitatory region, gray disk: inhibitory end zone). Right: Model explanation of illusory contours; see text. (Modified from Peterhans & von der Heydt, 1989.)

Contour completion

Most contours in the visual field are incomplete due to occlusion of one object by another. Kanizsa's (1955) and Kanizsa's (1979) illusory triangle illustrates how such gaps may be perceptually filled in by illusory contours in the interest of completing the foreground object according to the Gestalt factor of *good continuation*. In this way, the three local corners (the parts) of the Kanizsa figure are seen not in isolation but as the limiting endpoints of a global triangle (the whole). This is shown in Figure 9 (left) and is an *emergent property*. Furthermore, the triangular surface delineated by the contours not only looks uniformly brighter but also appears nearer. In this way, brightness enhancement known to move surfaces forward (aerial perspective) affords 3D figure-ground stratification and modal Gestalt completion. This again is an *emergent property*. The Kanizsa triangle led Baumgartner (1990) to ask, "Where do visual signals become a perception?"

Filling in gaps with illusory contours is not a privilege of human observers. There is evidence that illusory contours can also be seen by nonhuman animals (Nieder, 2002). End-stopped neurons have been found in area V2 of the macaque, whose response pattern suggests that they may be capable of filling in a gap with a collinear illusory contour, provided the edges are aligned (Baumgartner, von der Heydt, & Peterhans, 1984; von der Heydt, Peterhans, & Baumgartner, 1984). Figure 9 (right) illustrates how signals from the aligned edges (black arrows) and occlusion signals from the end-stopped neurons (gray arrows) are combined in a higher-order neuron (box), where they are summed (Σ). In this way, an illusory contour for which there is no physical equivalent is thought to emerge at right angles to the inducing cues. Peterhans and von der

Heydt (1991) called subjective contours a phenomenon bridging the gap between psychophysics and physiology. Only illusory contours, not brightness enhancement, are explained by their model.

Border ownership

Figure-ground segregation typically occurs at a unilateral border that belongs to the figure, not the ground (Rubin, 1915). Zhou, Friedman, and von der Heydt (2000) reported that a neuron in area V2 responds differently, depending on whether the border inside the receptive field belongs to a figure on one side of the receptive field or the other (Figure 10). For example, they found that the neuron responds strongly to the luminance step demarcating the light square on the lower left (A) but much less so to the dark square on the upper right (B), although both edge profiles are

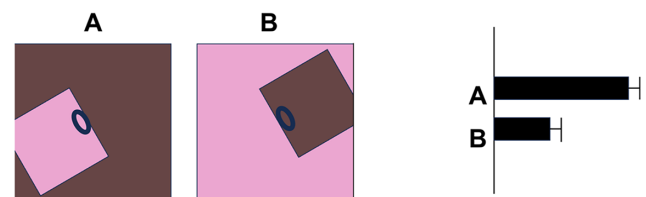


Figure 10. Border ownership. A neuron in area V2 responds more vigorously (horizontal bars on the right) to pattern A than to pattern B, although the local contrast at the edge is identical. The difference is that in A, the border belongs perceptually to the light square, whereas in B, it belongs to the dark square. This result is suggestive of border ownership or belongingness (Zhou, Friedman, & von der Heydt, 2000, Copyright © 2000 by Society for Neuroscience.)

locally identical. This neuronal response thus reflects *border ownership* or *belongingness*, which again is an *emergent property*.

Border ownership implies that the neuronal response depends not only on local stimulus analysis within the classical receptive field but rather on global feature integration and that such contextual influences extend over relatively large regions (4 degrees and more). Spillmann and Werner (1996) called this long-range interaction. These results suggest a global mechanism for segregating figures on the ground that may involve “grouping cells,” which sum responses of feature neurons (such as simple and complex cells) and enhance the responses of the same feature neurons via back projection (von der Heydt & Zhang, 2018).

Gestalt neurons

A central question in Gestalt vision is how our brain creates *wholes* from *parts* or how the visual system integrates *local* cues into a *global* representation (Allman et al., 1985; Kourtzi et al., 2003; Pan et al., 2012; Luo et al., 2019). The distinction between local and global information becomes apparent if one looks, for example, at a tire track (Figure 11) where the local profile points to the side, whereas the global profile points straight ahead (Lu et al., 2018). Spatial scale and orientation distinguish the two features. Clearly, the whole or Gestalt is different from the sum of its parts. Accordingly, the global direction is seen first, before we notice the local direction of the tire tracks.

We here suggest, for the first time, that global phenomena from local information are based on stimulus processing by *Gestalt neurons*. Such neurons would enable the correlation between visual brain and perception (Spillmann & Ehrenstein, 1996; Ehrenstein, Spillmann, & Sarris, 2003; Spillmann & Ehrenstein, 2004). More than a hundred years after (Wertheimer, 1912; Wertheimer, 1923), these neurons would change the understanding of a given perceptual phenomenon from merely descriptive Gestalt principles to putative explanations based on neuronal mechanisms.

Since Gestalt vision depends on the organizational principles governing the visual brain, there must be neural units, which are responsible for the resulting *emergent properties* characteristic of the interplay between *parts* and *wholes*. Neuronal networks of Gestalt neurons, rather than single neurons, might be responsible for the oversummative properties of the whole. Such neurons are creative as they produce something new that does not derive by simple summation from the parts. Their input would be local but their output global.

The distinction between local and global is reminiscent of low-level *classical* receptive fields and

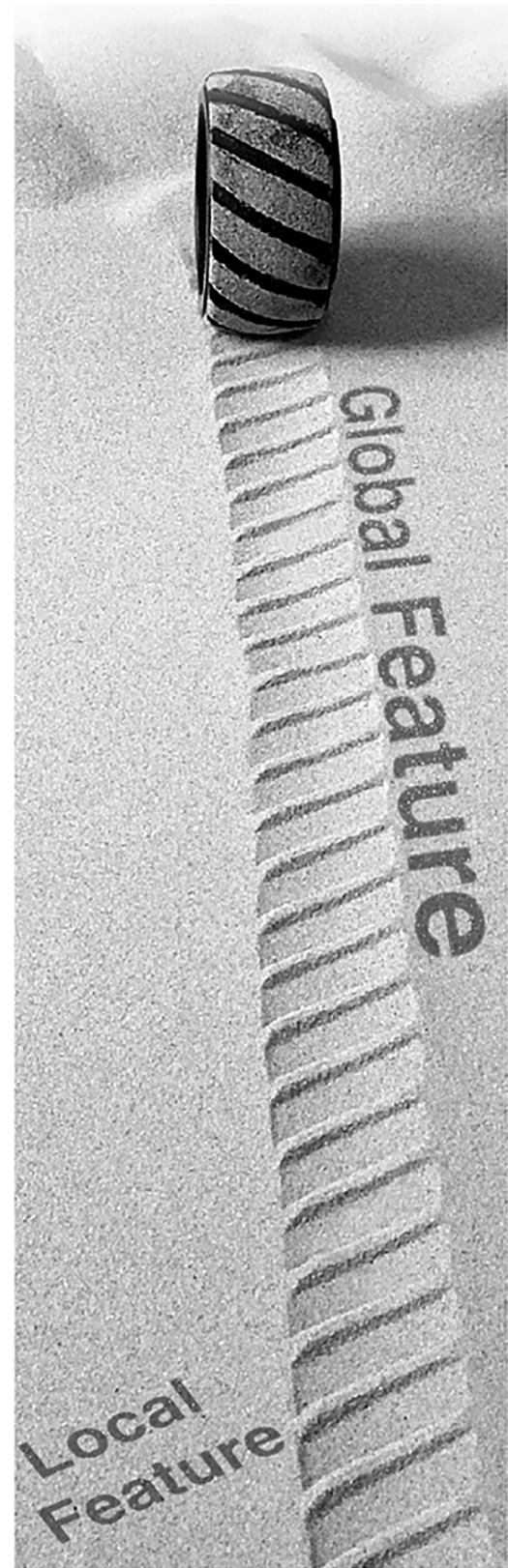


Figure 11. Global versus local direction. The global orientation of a tire track perceptually overrides the local orientation of the tire profile. Whereas the local tire prints point rightward, the track points straight ahead. This is an emergent property. (From Lu et al., 2018, Copyright © 2018 Elsevier Inc.)

higher-level *contextual* receptive fields (Allman et al., 1985; Albright & Stoner, 2002). Classical receptive fields are small and process dots, lines, and edges, whereas contextual receptive fields are large and take the surround into account (Spillmann, 2014; Spillmann et al., 2015). The transformation from the classical to the contextual response pattern is a possible waystation for the resulting *emergent properties*. Unlike simple and complex cells, the Gestalt neuron is a concept not defined by parameters such as orientation specificity, stereo, and motion sensitivity but by its ability to transform a stimulus percept from local to global, thereby giving it Gestalt properties.

Such properties come about by the integration of low-level input via hierarchical processing to higher-level areas, thereby converting local features to global features. Feed-forward, horizontal, and feed-backward mechanisms may participate in a recurrent network responsible for the conversion from stimulus to percept (Lamme, Supèr, & Spekreijse, 1998). Faces are examples of local features such as eyes, nose, and mouth (the parts) processed in areas V1 and V2 and integrated into a whole in area IT. Gestalt neurons thus represent “signature” responses for the complex stimuli in our world (Perrett, Rolls, & Caan, 1982; Desimone et al., 1984). The inability to see a face in prosopagnosia is an example of Gestalt blindness (Spillmann, Laskowski, Lange, Kasper, & Schmidt, 2000).

Koffka’s (1935) question “Why do things look as they do?” might thus be answered that things look as they do, owing to Gestalt factors, which have their origin in Gestalt neurons.

Conclusions

Several well-known phenomena have been presented in this article to demonstrate *emergent properties*, which are tentatively ascribed to Gestalt neurons. Candidates for such neurons have been identified as follows.

Table 1 lists candidates of Gestalt neurons for the phenomena described here within the context of known neurophysiology. Other examples, such as the perception of boundaries from accretion/deletion (Shipley & Kellman, 1994), surface transparency (Metelli, 1974), and the perception of depth planes in random dot textures (Julesz, 1971), could readily be added.

Johansson (1950), who created the point walker (Figure 4A), best captured the reality of emergent properties in perception when he wrote about “the radical changes that the [perceived] motion of an element undergoes when the next element arises. The direction of motion is reversed: A rectilinear path

Brain area	Phenomenon (discussed in subchapter)	Technique
V1/V4	Orientation contrast	Single-cell recording
V2	Contour completion	Single-cell recording
V2	Border ownership	Single-cell recording
V1/V2/V4	Contour integration	Single-cell recording
LOC		fMRI (human)
V1/MT	3D structure from motion	Single-cell recording
MT/V5	Apparent motion	fMRI (human)
MT/V5 (MST)	Coherent motion	fMRI (human)
MT/MST	Aperture motion	Single-cell recording
MST	Apparent rotation	Single-cell recording
STPa	Biological motion	Single-cell recording

Table 1. Potential Gestalt neurons for visual phenomena and their respective brain locations. *Notes:* LOC = lateral occipital cortex. STPa = anterior superior temporal polysensory area.

becomes circular, a circular path becomes rectilinear, etc. . . . We have here an exemplification of the [Gestalt] thesis that the whole is something qualitatively different from the summation of its component parts” (p. 78).

Following Mach (1865), who already had provided a prescient account of the illusory bands named after him in terms of lateral inhibitory interactions in the retina, (Jung, 1961a; Jung, 1961b) and Baumgartner (1961) came up with empirical evidence of center-surround antagonism of ON- and OFF-center neurons as the physiological basis for the bright and dark bands, respectively. They termed this kind of explanation a *neurophysiological correlate*. The time has come to extend this approach to high-level visual phenomena such as *Gestalten*. Undoubtedly, the aforementioned examples must have neurophysiological correlates in the visual brain that go beyond the mere phenomena. To reveal, understand, and explain them will be the task of neuro-Gestalt theory.

Keywords: Gestalt, laws of seeing, segmentation, segregation, grouping, emergent property, parts versus wholes, local versus global, receptive fields, Gestalt neurons

Acknowledgments

This article is based on a graduate seminar entitled *Foundations of Vision Sciences* (PSYC0036) held at the University of Hong Kong, Second Semester (2012–2013); graduate seminars on W. Metzger’s *Laws of Seeing*, R. Gregory’s *Eye and Brain*, and J. J. Gibson’s *The Perception of the Visual World* at China Medical University and National Taiwan University;

and in-depth discussions on *Neuro-Gestalt Theory* at the Chinese Academy of Sciences, Shanghai.

LS thanks Li-Chuan Hsu for initiating the contact and CMU President Jong-Tsun Huang, Su-Ling Yeh, and Keng-Chen Liang for inviting him to Taiwan and for offering unparalleled inspiration, generosity, and hospitality over a period of 8 years. He likes to acknowledge the generous support by an anonymous NTU-donor. He also thanks the Serena Yang Foundation for granting a visitorship to the University of Hong Kong and the Chinese Academy of Sciences, Shanghai, for ongoing laboratory collaborations. He wishes to express his gratitude to the students in his classes for their enthusiasm, cooperation, hard work, and loyalty. Thanks go to Dejan Todorović for commenting on the manuscript and for providing the quote by Gunnar Johansson and to John S. Werner for reading the manuscript. Yingfan Liu very kindly improved the figures, while the librarians of the Medical School of the University of Freiburg procured much needed literature. We thank them all.

Supported by Ensemble Grant and by the Cooperative Research Project Program from the Research Institute of Electrical Communication at Tohoku University to CT.

Commercial relationships: none.

Corresponding author: Chia-Huei Tseng.

Email: ch_tseng@alumni.uci.edu.

Address: Research Institute of Electrical Communication, Interdisciplinary ICT Research Center for Cyber and Real Spaces, Tohoku University, Sendai, Japan.

References

- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525, <https://doi.org/10.1038/300523a0>.
- Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience*, *25*, 339–379, <https://doi.org/10.1146/annurev.neuro.25.112701.142900>.
- Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, *8*, 407–430, <https://doi.org/10.1146/annurev.ne.08.030185.002203>.
- Baumgartner, G. (1961). Neuronale Grundlagen der visuellen Kontrastverschärfung und die Signalübertragung vom Auge zur Hirnrinde [Neuronal foundations of the visual enhancement of contrast and the signal transduction from the eye to the visual cortex.] In VDE Nachrichtentechnische Gesellschaft (Ed.), *Aufnahme und Verarbeitung von Nachrichten durch Organismen* (pp. 100–108). Stuttgart, Germany: Hirzel.
- Baumgartner, G. (1990). Where do visual signals become a perception? In J. C. Eccles, & O. Creutzfeldt (Eds.), *The principle of design and operation of the brain* (pp. 99–114). Berlin, Germany: Springer.
- Baumgartner, G., von der Heydt, R., & Peterhans, E. (1984). Anomalous contours: A tool in studying the neurophysiology of vision. *Experimental Brain Research (Suppl.)*, *9*, 413–419.
- Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception*, *30*, 61–72, <https://doi.org/10.1068/p3048>.
- Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature*, *392*, 714–717, <https://doi.org/10.1038/33688>.
- Britten, K. H. (2004). The middle temporal area: Motion processing and the link to perception. In L. M. Chalupa, & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1203–1216). Cambridge, MA: MIT Press, <https://doi.org/10.7551/mitpress/7131.003.0081>.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*, 4745–4765, <https://doi.org/10.1523/jneurosci.12-12-04745.1992>.
- Chen, C.-C., Ashida, H., Yang, X., & Chen, P.-Y. (2020). Seeing global motion in a random dot image sequence. *i-Perception*, *11*(5), 1–10, <https://doi.org/10.1177/2041669520961104>.
- Chen, M., Yan, Y., Gong, X., Gilbert, C. D., Liang, H., & Li, W. (2014). Incremental integration of global contours through interplay between visual cortical areas. *Neuron*, *82*, 682–694, <https://doi.org/10.1016/j.neuron.2014.03.023>.
- Chow, H. M., Jingling, L., & Tseng, C. H. (2013). Collinear integration affects visual search at V1. *Journal of Vision*, *13*(10), 24, <https://doi.org/10.1167/13.10.24>.
- Chow, H. M., Jingling, L., & Tseng, C. H. (2016). Eye of origin guides attention away: An ocular singleton column impairs visual search like a collinear column. *Journal of Vision*, *16*(1), 12, <https://doi.org/10.1167/16.1.12>.
- Desimone, R., Albright, T. D., Gross, C., & Bruce, C. (1984). Stimulus-selective properties of

- inferior temporal neurons in the macaque. *Journal of Neuroscience*, 4, 2051–262, <https://doi.org/10.1523/jneurosci.04-08-02051>.
- Ehrenstein, W. H., Spillmann, L., & Sarris, V. (2003). Gestalt issues in modern neuroscience. *Axiomathes*, 43, 433–458.
- Field, D. J., & Hayes, A. (2004). Contour integration and the lateral connections of V1 neurons. In L. M. Chalupa, & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1069–1079). Cambridge, MA: MIT Press, <https://doi.org/10.7551/mitpress/7131.003.0081>.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research*, 33, 173–193, [https://doi.org/10.1016/0042-6989\(93\)90156-Q](https://doi.org/10.1016/0042-6989(93)90156-Q).
- Gallace, A., & Spence, C. (2011). To what extent do Gestalt grouping principles influence tactile perception? *Psychological Bulletin*, 137, 538–561, <https://doi.org/10.1037/a0022335>.
- Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *Journal of Neuroscience*, 9, 2432–2442, <https://doi.org/10.1523/jneurosci.09-07-02432.1989>.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, 92, 173–211, <https://doi.org/10.1037/0033-295X.92.2.173>.
- Grunewald, A., Bradley, D. C., & Andersen, R. A. (2002). Neural correlates of structure-from-motion perception in macaque V1 and MT. *Journal of Neuroscience*, 15, 6195–6207, <https://doi.org/10.1523/jneurosci.22-14-06195.2002>.
- Grüsser, O.-J., & Grüsser-Cornehls, U. (1973). Neuronal mechanisms of visual movement perception and some psychophysical and behavioral correlations. In R. Jung, (Ed.), *Handbook of sensory physiology. Central processing of visual information A: Integrative functions and comparative data* (Vol. 7/3/3 A, pp. 333–429). Berlin, Germany: Springer, https://doi.org/10.1007/978-3-642-65352-0_6.
- Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung B*, 11, 513–524, <https://doi.org/10.1515/znB-1956-9-1004>.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *The Journal of Physiology*, 148, 574–591, <https://doi.org/10.1113/jphysiol.1959.sp006308>.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243, <https://doi.org/10.1113/jphysiol.1968.sp008455>.
- Jingling, L., Tang, D. L., & Tseng, C. H. (2013). Salient collinear grouping diminishes local salience in visual search: An eye movement study. *Journal of Vision*, 13(12), 6, <https://doi.org/10.1167/13.12.6>.
- Jingling, L., Tseng, C. H., & Zhaoping, L. (2013). Orientation is different: Interaction between contour integration and feature contrasts in visual search. *Journal of Vision*, 13(3), 26, <https://doi.org/10.1167/13.3.26>.
- Johansson, G. (1950). *Configurations in event perception*. Uppsala: Almqvist & Wiksell.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Attention, Perception, & Psychophysics*, 14, 201–211, <https://doi.org/10.3758/BF03212378>.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.
- Jung, R. (1961a). Neuronal integration in the visual cortex and its significance for visual information. In W. A. Rosenblith, (Ed.), *Sensory communication* (pp. 627–674). Cambridge, MA: MIT Press, <https://doi.org/10.7551/mitpress/9780262518420.003.0032>.
- Jung, R. (1961b). Korrelationen von Neuronentätigkeit und Sehen. In R. Jung, & H. Kornhuber (Eds.), *Neurophysiologie und Psychophysik des visuellen Systems* (pp. 410–434). Berlin, Germany: Springer, https://doi.org/10.1007/978-3-662-22221-8_46.
- Jung, R., & Spillmann, L. (1970). Receptive-field estimation and perceptual integration in human vision. In F. A. Young, & D. B. Lindsley (Eds.), *Early experience and visual information processing in perceptual and reading disorders* (pp. 181–197). Washington, DC: National Academy of Sciences.
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, 49, 7–30.
- Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York, NY: Praeger.
- Kastner, S., Nothdurft, H. C., & Pigarev, I. N. (1999). Neuronal responses to orientation and motion contrast in cat striate cortex. *Visual Neuroscience*, 15, 560–587, <https://doi.org/10.1017/s095252389916317x>.
- Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology*, 67, 961–980, <https://doi.org/10.1152/jn.1992.67.4.961>.

- Koffka, K. (1935). *Principles of Gestalt psychology*. New York, NY: Harcourt, Brace and Company.
- Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: monkey and human fMRI studies. *Neuron*, *37*, 333–346, [https://doi.org/10.1016/s0896-6273\(02\)01174-1](https://doi.org/10.1016/s0896-6273(02)01174-1).
- Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effects of closure in figure–ground segmentation. *Proceedings of the National Academy of Sciences of the USA*, *90*, 7495–7497, <https://doi.org/10.1073/pnas.90.16.7495>.
- Lamme, A. F., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*, 529–535, [https://doi.org/10.1016/s0959-4388\(98\)80042-1](https://doi.org/10.1016/s0959-4388(98)80042-1).
- Lamme, V. A. F. (1995). The neurophysiology of figure–ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615, <https://doi.org/10.1523/jneurosci.15-02-01605.1995>.
- Li, W., Pièch, V., & Gilbert, C. D. (2006). Learning to link visual contours. *Neuron*, *57*, 442–451, <https://doi.org/10.1016/j.neuron.2007.12.011>.
- Li, W., Pièch, V., & Gilbert, C. D. (2008). Contour saliency in primary visual cortex. *Neuron*, *50*, 951–962, <https://doi.org/10.1016/j.neuron.2006.04.035>.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, *10*, 903–940, <https://doi.org/10.1162/089976698300017557>.
- Liu, T., Slotnick, S., & Yantis, S. (2004). Human MT+ mediates perceptual filling-in during apparent motion. *NeuroImage*, *21*, 1772–1780, <https://doi.org/10.1016/j.neuroimage.2003.12.025>.
- Lorenz, K. (1951). The role of Gestalt perception in animal and human behaviour. In L. Whyte (Ed.), *Aspects of form* (pp. 157–178). London, UK: Lund Humphries.
- Lu, Y., Yin, J., Chen, Z., Gong, H., Liu, Y., Qian, L., . . . Wang, W. (2018). Revealing detail along the visual hierarchy: neural clustering preserves acuity from V1 to V4. *Neuron*, *98*, 417–428.
- Luo, J., He, K., Andolina, I. M., Li, X., Yin, J., Chen, Z., . . . Wang, W. (2019). Going with the flow: The neural mechanisms underlying illusions of complex-flow motion. *Journal of Neuroscience*, *39*, 2664–2685, <https://doi.org/10.1523/jneurosci.2112-18.2019>.
- Mach, E. (1865). Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf der Retina [On the effect of the spatial distribution of the light stimulus on the retina]. *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, *52*, 303–322.
- Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4), 91–98, <https://doi.org/10.1038/scientificamerican0474-90>.
- Metzger, W. (1936). *Gesetze des Sehens* (1st ed., 2nd ed. 1953). Frankfurt am Main: W. Kramer. [Translated by L. Spillmann, S. Lehar, M., Stromeyer, & M. Wertheimer (2006), *Laws of Seeing*. Cambridge, MA: MIT Press], doi:10.1162/leon.2008.41.3.297.
- Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2006). Primary visual cortex activity along the apparent motion trace reflects illusory perception. *PLoS Biology*, *3*(8), e265, <https://doi.org/10.1371/journal.pbio.0030265>.
- Nelson, J. I., & Frost, B. J. (1985). Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex. *Experimental Brain Research*, *61*, 54–61, <https://doi.org/10.1007/BF00235620>.
- Nieder, A. (2002). Seeing more than meets the eye: Processing of illusory contours in animals. *Journal of Comparative Physiology A*, *188*, 249–260, <https://doi.org/10.1007/s00359-002-0306-x>.
- Nothdurft, H., Gallant, J., & Van Essen, D. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience*, *17*, 421–436, <https://doi.org/10.1017/S0952523800173092>.
- Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to “biological motion” stimuli. *Journal of Cognitive Neuroscience*, *6*, 99–116, <https://doi.org/10.1162/jocn.1994.6.2.99>.
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*, 1040–1042, <https://doi.org/10.1038/35059085>.
- Pan, X. Y., Chen, M., Yin, J., An, X., Zhang, X., Lu, Y., . . . Wang, W. (2012). Equivalent representation of real and illusory contours in macaque V4. *The Journal of Neuroscience*, *32*, 6760–6770, <https://doi.org/10.1523/jneurosci.6140-11.2012>.
- Pan, X. Y., Wang, L. Y., Wang, Z. W., Xu, C., Yu, W. W., Spillmann, L., . . . Wang, W. (2016). Representation of illusory and physical rotations in human MST: A cortical site for the Pinna illusion. *Human Brain Mapping*, *37*, 2097–2113.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342, <https://doi.org/10.1007/BF00239352>.
- Peterhans, E., Heider, B., & Baumann, R. (2005). Neurons in monkey visual cortex detect

- lines defined by coherent motion of dots. *European Journal of Neuroscience*, *21*, 1091–1100, <https://doi.org/10.1111/j.1460-9568.2005.03919.x>.
- Peterhans, E., & von der Heydt, R. (1989). Mechanism of contour perception in monkey visual cortex: II. Contours bridging gaps. *Journal of Neuroscience*, *9*, 1749–1763, <https://doi.org/10.1523/jneurosci.09-05-01749.1989>.
- Peterhans, E., & von der Heydt, R. (1991). Subjective contours bridging the gap between psychophysics and physiology. *Trends in Neuroscience*, *14*, 112–119, [https://doi.org/10.1016/0166-2236\(91\)90072-3](https://doi.org/10.1016/0166-2236(91)90072-3).
- Pinna, B. (2009). Pinna illusion, *Scholarpedia*, *4*(2), 6656.
- Pinna, B., & Brelstaff, G. J. (2000). A new visual illusion of relative motion. *Vision Research*, *40*, 2091–2096, [https://doi.org/10.1016/S0042-6989\(00\)00072-9](https://doi.org/10.1016/S0042-6989(00)00072-9).
- Roelfsema, P. R., & Singer, W. (1998). Detecting connectedness. *Cerebral Cortex*, *8*, 385–396, <https://doi.org/10.1093/cercor/8.5.385>.
- Rubin, E. (1915). *Synsoplevede Figurer*. Kopenhagen: Glydendalske.
- Self, M. W., Mookhoek, A., Tjalma, N., & Roelfsema, P. R. (2015). Contextual effects on perceived contrast: Figure–ground assignment and orientation contrast. *Journal of Vision*, *15*(2), <https://doi.org/10.1167/15.2.2>.
- Shipley, T. F., & Kellman, P. J. (1994). Spatiotemporal boundary formation: Boundary, form, and motion perception from transformations of surface elements. *Journal of Experimental Psychology: General*, *123*, 3–20, <https://doi.org/10.1037/0096-3445.123.1.3>.
- Sigman, M., Cecchi, G.A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences of the USA*, *98*, 1935–1940, <https://doi.org/10.1073/pnas.98.4.1935>.
- Singer, W. (1989). Search for coherence: A basic principle for cortical self-organization. *Concepts in Neuroscience*, *1*, 1–26.
- Spillmann, L. (1997). Guest editorial. Colour in a larger perspective: the rebirth of Gestalt psychology. *Perception*, *26*, 1341–1352, <https://doi.org/10.1068/p261341>.
- Spillmann, L. (1999). From elements to perception: Local and global processing in visual neurons. *Perception*, *28*, 1461–1492, <https://doi.org/10.1068/p2763>.
- Spillmann, L. (2006). From perceptive fields to Gestalt. *Progress in Brain Research*, *155*, 67–92, [https://doi.org/10.1016/S0079-6123\(06\)55005-8](https://doi.org/10.1016/S0079-6123(06)55005-8).
- Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research*, *49*, 1507–1521, <https://doi.org/10.1016/j.visres.2009.02.022>.
- Spillmann, L. (2012). The current status of Gestalt rules in perceptual research: Psychophysics and neurophysiology. In L. Spillmann, V. Sarris, & V. M. Wertheimer (Eds.), *Max Wertheimer on perceived motion and figural organization* (pp. 191–242). Cambridge, MA: MIT Press, <https://doi.org/10.7551/mitpress/9222.003.0008>.
- Spillmann, L. (2014). Receptive fields of visual neurons: The early years. *Perception*, *43*, 1145–1176, <https://doi.org/10.1068/p7721>.
- Spillmann, L., Dresch-Langley, B., & Tseng, C. (2015). Beyond the classical receptive field: The effect of contextual stimuli. *Journal of Vision*, *15*, 1–23, <https://doi.org/10.1167/15.9.7>.
- Spillmann, L., & Ehrenstein, W. (1996). From neuron to Gestalt—Mechanisms of visual perception. In R. Greger, & U. Windhorst (Eds.), *Comprehensive human physiology: From cellular mechanisms to integration* (pp. 861–893). Berlin, Germany: Springer-Verlag, https://doi.org/10.1007/978-3-642-60946-6_44.
- Spillmann, L., & Ehrenstein, W. (2004). Gestalt factors in the visual neurosciences. In L. M. Chalupa, & J. S. Werner (Eds.), *The visual neurosciences (Vol. 2)*, pp. 1573–1589. Cambridge, MA: MIT Press.
- Spillmann, L., Laskowski, W., Lange, K. W., Kasper, E., & Schmidt, D. (2000). Stroke-blind for colors, faces and locations: Partial recovery after three years. *Restorative Neurology and Neuroscience*, *17*, 89–103.
- Spillmann, L., & Werner, J. S. (1996). Long-range interactions in visual perception. *Trends in Neurosciences*, *19*, 428–434, [https://doi.org/10.1016/0166-2236\(96\)10038-2](https://doi.org/10.1016/0166-2236(96)10038-2).
- Sterzer, P., Haynes, J.-D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *Neuroimage*, *32*, 1308–1316, <https://doi.org/10.1016/j.neuroimage.2006.05.029>.
- Stürzel, F., & Spillmann, L. (2004). Perceptual limits of common fate. *Vision Research*, *44*, 1565–1573, <https://doi.org/10.1016/j.visres.2003.12.019>.
- Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*, 626–641, <https://doi.org/10.1152/JN.1989.62.3.626>.

- Todorović, D. (2008). Gestalt principles. *Scholarpedia*, 3, 5345, <https://doi.org/10.4249/scholarpedia.5345>.
- Tseng, C. H., & Jingling, L. (2015). A salient and task-irrelevant collinear structure hurts visual search. *PLoS One*, 10(4), e0124190, <https://doi.org/10.1371/journal.pone.0124190>.
- Tseng, C. H., Chow, H. M., Liang, J., Shioiri, S., & Chen, C. C. (2021). Collinear search impairment is luminance contrast invariant. *Scientific Reports*, 11(1), 1–8, <https://doi.org/10.1038/s41598-021-90909-w>.
- Veras, C., Quang-Cuong, P., & Maus, G. W. (2017). The silhouette zoetrope: A new blend of motion, mirroring, depth, and size illusions. *i-Perception*, 8, 1–8, <https://doi.org/10.1371/journal.pone.0124190>.
- von der Heydt, R., Peterhans, E., & Baumgartner, C. (1984). Illusory contours and cortical neuron responses. *Science*, 224, 1260–1262, <https://doi.org/10.1126/science.6539501>.
- von der Heydt, R., & Zhang, N. R. (2018). Figure and ground: How the visual cortex integrates local cues for global organization. *Journal of Neurophysiology*, 120, 3085–3098, <https://doi.org/10.1152/jn.00125.2018>.
- Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, 20, 325–380, <https://doi.org/10.1007/BF02409790>.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61, 161–265. [Translated in: L. Spillmann, V. Sarris, & M. Wertheimer (2012): Max Wertheimer on Perceived Motion and Figural Organization (pp. 191–242). Cambridge, MA: MIT Press].
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, 4, 301–350.
- Westheimer, G. (1999). Gestalt theory reconfigured: Max Wertheimer's anticipation of recent developments in visual neuroscience. *Perception*, 28, 5–15, <https://doi.org/10.1068/p2883>.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, 20, 6594–6611, <https://doi.org/10.1523/JNEUROSCI.20-17-06594.2000>.