MIT Press ESForum

14

"Hot Spots" and Dynamic Coordination in Gestalt Perception

Ilona Kovács

Abstract

How is light that is transduced by retinal receptors and interpreted by neurons converted into visual information? To answer this question, vision science typically employs two types of scientists: those interested in local receptors and neurons that analyze very small pieces of the retinal image, and those interested in global visual information (surfaces, objects, scenes, and events that are meaningful to people). While both types may function well within their own areas of research, "translating" the results between the areas is a problem. Two explicit issues are discussed where strictly local processing stops short: (a) the problem of accumulating local errors and (b) the trade-off between spatial and temporal resolution in pictorial representations. To illustrate the first issue, which is architectural, an old and wonderful architectural mystery, the enigma of the Florence Dome, is used. An example from the history of photography illustrates the second issue, which is representational. Both problems have an important aspect in common: the solutions are both based on global geometry. Both classic examples will be accompanied by visual phenomena demonstrating the relevance of symmetry-based representations in the dynamic coordination of visual perception.

In What Way Is a Gestalt More Than Its Elementary Features?

> If a line forms a closed, or almost closed, figure, we see no longer merely a line on a homogeneous background, but a surface figure bounded by the line. This fact is so familiar that unfortunately it has, to my knowledge, never been made a subject of special investigation. And yet, it is a very startling fact, once we strip it of its familiarity. — Koffka (1935:150)

Gestalt, shape, *prägnanz* constitute the core mysteries of perception. Gestalt theorists considered the formation of perceptual pattern a dynamic process, best demonstrated by the various ambiguous figures with which they have

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

entertained the world. Edgar Rubin's popular face-vase reversal image demonstrates that even boundary ownership is flexible, and the percept can quickly reorganize although the physical stimulus is constant. The compelling multistability of ambiguous images is induced by carefully balanced stimuli, where the two interpretations are equally "salient." One of the simplest multistable stimuli is shown in Figure 14.1. The two superimposed gratings are equiluminant and "orthogonal," both in terms of orientation and color. By simply staring at this stimulus, one observes monocular rivalry, and the two gratings start to alternate spontaneously.

The exciting instability of the perceptual system elicited by such images is not only among the most popular topics within the modern quest for neural correlates of conscious percepts (Crick and Koch 1995; Kovács et al. 1996; Leopold and Logothetis 1996), it also clearly demonstrates that our usual sense of perceptual stability is an illusion, and that the brain has many different ways to assemble new "realities" from competing pieces of concurrent external and internal events. Each competing interpretation entails a certain segmentation of the image into figure and ground.

The Unsolved Problem of Segmentation

Among the many unsolved issues of vision, the issue of segmentation may be one of the toughest. Although it does not sound very difficult to parse an

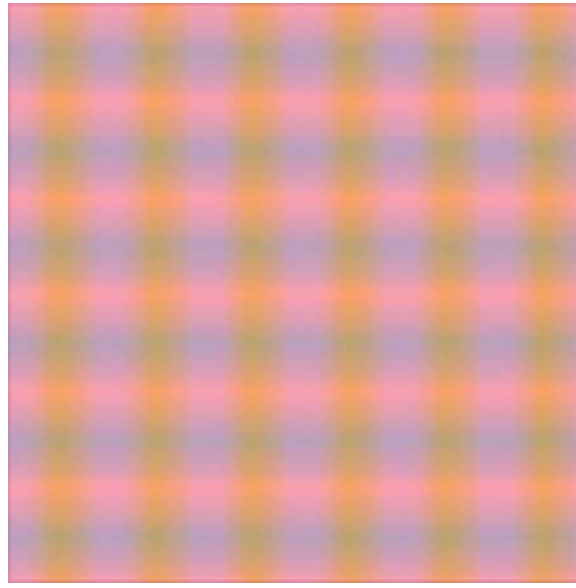

Figure 14.1 By staring at the image, spontaneous alternation between the two gratings is observed and monocular rivalry is stimulated.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

image into different regions that correspond to objects and ground, machine or computer vision systems still cannot match the capabilities of the human visual system, not to mention categorization and image comprehension, which are strongly linked to segmentation. Human segmentation of visual images might depend on acquired knowledge and proceed interactively with object recognition (e.g., Ullman 2007). However, the explicit manner in which these higher-level knowledge systems communicate with low-level feature extractors has not yet been clarified.

There is obviously some higher-order structure, globally organized by the brain. However, there is continuous input from the environment, analyzed by low-level, local feature extractors of some sort. Just how do these computationally very different levels of processing (i.e., global organization vs. local analysis) meet and interact to provide us with our visual world?

According to the "standard" view of visual processing, visual information is first transmitted from the retina through several parallel pathways to the brain in a compressed version, emphasizing edge information at a number of spatial scales. A crucial second step is carried out by cortical area 17 (V1, or primary visual cortex), which is assumed to extract a set of local features based on the retinal input. Although the standard view then proceeds to progressively more complex representations, let us focus on the second step and the cortical "mosaic" generated by the primary visual cortex. It has been known for over forty years that the receptive fields of the primary visual cortex are composed of elongated antagonistic zones (Hubel and Wiesel 1959). The shape and layout of these receptive fields furnish the cells with selectivity for oriented line segments, and receptive field size determines the spatial scale of orientation information.

The primary visual cortex thus provides a neural description of oriented edge primitives and their locations at a number of spatial scales. This can be viewed as an enormous puzzle containing millions of pieces to be put together into figure and ground. A possible candidate for assembling local information already within the primary visual cortex is the plexus of long-range horizontal connections (e.g., Gilbert 1992). These connections are thought to establish connections between neighboring processing units, thereby aiding the segmentation process. The mechanism by which local interactions combine and boundaries of a visual object form is, however, unknown (Figure 14.2). Can all of this be based on local interactions?

The Problem of Accumulating Local Errors: The Puzzle of the Dome

To illustrate the problem of relying on local operators, consider an example from the history of architecture. Construction of the Florence Cathedral started in 1296, and its magnificent dome was completed in 1436. Driven by the urge to surpass Pisa and Siena in the size and decoration of the cathedral, the Florentine cupola is still the largest masonry dome in the world. Unfortunately,

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

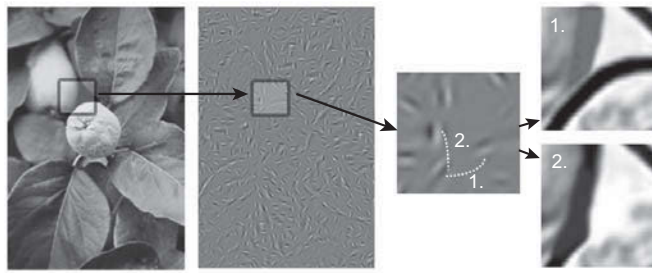

Figure 14.2 The problem of visual segmentation. The primary visual cortex provides a neural description of oriented edge primitives and their locations at a number of spatial scales. The natural image shown in the left panel activates a very large number of cortical filters; however, those that receive input according to their selectivities will be more active. The next panel shows the most activated filters for each location within the image. Neural interactions within the primary visual cortex are assumed to connect the most active filters in an orientation selective, facilitatory manner (e.g., similar-to-similar orientations). However, when viewed locally within the inset, it seems that these connections might be ambiguous. Which one is the better connection? Number 1 is a good choice, as the central object will be well segmented. Number 2 would be a bad choice, as it connects the boundaries of two independent objects. Is there some global reference guiding these local decisions?

only sparse information exists on how this wonderful three-dimensional shape was constructed from bricks and mortar. The dome was built using approximately 4 million specially designed bricks, collectively weighing about 37,000 tons. Filippo Brunelleschi's masonry techniques made a great contribution to architecture (King 2000), and his work is probably the best example of Renaissance engineering. However, he was also very secretive and never revealed his methods in detail. Although the dome is probably the most studied building on the planet, it is still uncertain how it was built. It is not known exactly, for example, what instruction was given to the, assumedly, eight groups of bricklayers as they raised the cupola's eight sides. How were they to know the exact positioning of each brick? Walls are easy to construct vertically, although some quasi-global reference (e.g., a masons' level or a plumb line) is needed from time to time. However, when the wall is curved, there are three dimensions to control: (a) the longitudinal curvature of the dome, (b) the circumferential curvature, and (c) the course by course, precise, and nonuniform change of the inward tilt of the bricks. A simple plumb line, even if it is combined with a level, cannot obviously control these three dimensions.

Brunelleschi was a great geometer, and the plan of the dome was probably perfect. He even took care of designing the shapes of the bricks for each new course himself. However, even with the best planning, and greatest care, bricklayers make slight errors, and if they only use local references (e.g., the previous course of bricks, or the ribs of the dome), these small errors will accumulate. Even an error of a hundredth of a degree in any of the three dimensions would result in a catastrophe in the case of 4 million bricks. In fact, such

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

a catastrophe was observed in Siena after a massive addition to the existing cathedral and dome was undertaken in 1339. Although Sienese ambition was at least as great as Florentine ambition, Brunelleschi's ingenuity only served Florence. Just how did Brunelleschi manage to control the global shape of the dome and achieve this unprecedented and still unequaled construction?

With respect to the global shape of the dome, it may have seemed a good idea to use some central reference during the construction. However, wooden centering or scaffolding, which could have served to support as well as guide the overall arches, was not employed. How was the shape of the dome preserved without any scaffolds to guide it? How does the complicated pattern of bricks fill the spaces between the corners of the dome? Imagine that eight bricklayer groups are working on the wall, and other than along the circumference of the wall, they cannot compare notes. If there is no central pole of any kind (and there was certainly none because the dome is over 80 meters high) to use as a reference, and the masons cannot communicate with each other directly, how are the eight sides going to meet at the top? Perhaps there was some central reference after all, and it was removed after the dome was completed.

Massimo Ricci, a contemporary Italian architect, spent almost as much time trying to figure out the secret of the dome as Brunelleschi did building it. It took Ricci fifteen perplexing years to come up with an idea that might explain the riddle posed by the dome's construction. According to Ricci, the real secret of Santa Maria del Fiore lies in its extremely simple, although, in this context, surprising shape: a flower (Figure 14.3). The eight petals of the flower grew out of a circle, centered within the octagonal base, with a diameter three-fifths of the octagon diameter (the dome is based on a quinto-acuto, four-fifths measure). The flower was probably made of metal, and long ropes were attached to it, traversing the internal space of the dome. The shape of the petal controlled the circumferential curvature (the second dimension); the length of the ropes attached to the petals controlled the longitudinal curvature (the first dimension); and the tilt angle of the ropes controlled the inward tilt of the bricks (the third dimension). Each rope, connecting a certain location of the wall to the petal across the base of the dome, was adjusted to cross the central axis of the cupola. This was achieved by "centering" ropes between the corners of the cupola and the vertices of the flower. When a bricklayer wanted to align a new brick, he would move his rope to the new position, and his apprentice would shift the other end of the rope along the petal until the rope crossed the central axis again in a straight line. Perhaps the procedure was not repeated for each individual brick, but whenever it was done, the brick adjusted in this way would fall into place in accordance to the global reference point, and earlier local errors would not accumulate. This sounds like the solution indeed!

Even if historical evidence for the flower theory is missing, Ricci's scale model of the dome attests to its feasibility. The idea is simple: *when only local operators are given, use axis-based global reference to achieve a global shape and avoid the accumulation of local errors*. According to Ricci, this is the only

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

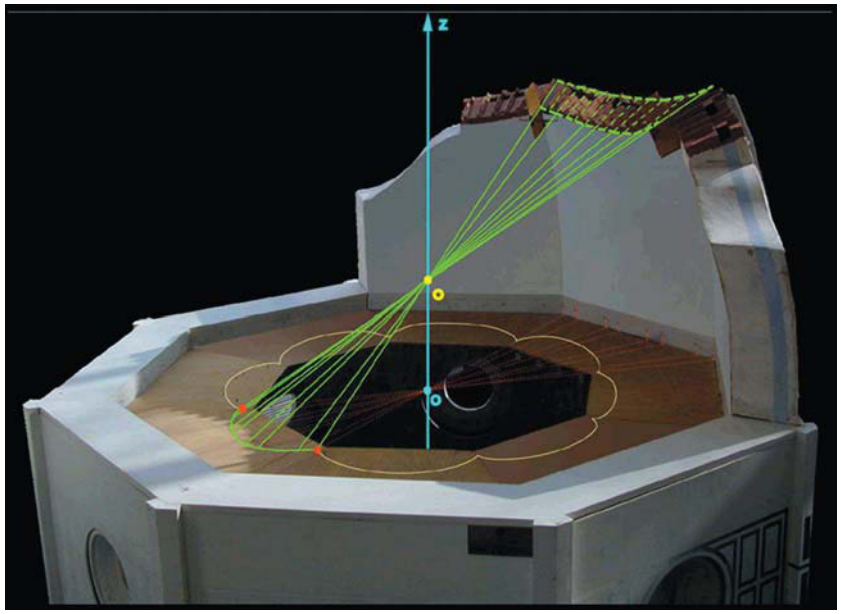

Figure 14.3 The flower theory based on Massimo Ricci Theory (used with kind permission from Luciana Burdi). The flower-shaped "skeleton" and many ropes between the flower and the wall serve to adjust the spatial position of the bricks according to the requirements of longitudinal, circumferential, and inward tilts. It is the symmetry axis of the shape that is employed as a global reference; however, the axis does not have to be there. It is determined with the help of ropes.

way of constructing such a shape. This might appear too ambitious; however, the usefulness of symmetry axes in avoiding the accumulation of local errors is very clearly illustrated with this example.

What Are the "Strings"? The Puzzle of "Closure"[1]

To challenge the local filters and local interactions of the primary visual cortex, and to investigate the type of integration that might be carried out at this "early" cortical level, Kovács and Julesz (1993) designed a psychophysical paradigm. The stimulus used in this paradigm consisted of a closed chain of co-linearly aligned Gabor signals (contour) and a background of randomly oriented and positioned Gabor signals (noise) (Figure 14.4). Gabor signals roughly model the receptive field properties of orientation selective simple cells in the primary visual cortex. Therefore, they are appropriate stimuli for the examination of these small spatial filters and their interactions. Notice that

[1] Closure is an old Gestaltist term used, e.g., by Kurt Koffka (1935:150) in *Principles of Gestalt Psychology*. According to Koffka, the term refers to the superiority of closed contours over open ones.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

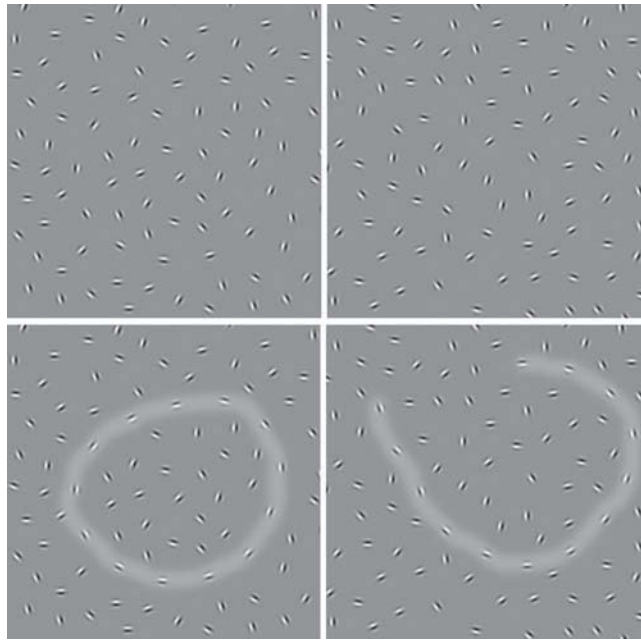

Figure 14.4 The closure superiority effect. The top two panels both have a contour embedded in noise. The solutions are presented in the bottom panels. Most observers find the closed contour in the upper left panel very easy to see, while it is difficult to trace the open contour even in the presence of the solution. (Because of individual variability in noise-tolerance, certain individuals might need different noise levels for the closure effect to appear.)

the contours cannot be detected by purely local filters or by neurons with large receptive field sizes corresponding to the size of the contour. The long-range orientation correlations along the path of the contour can only be found by the integration of local orientation measurements. The noise forces the observer to do these local measurements at the scale of the individual Gabor signals and to rely solely on long-range interactions between local filters while connecting the signals perceptually.

Since the "contour in noise" stimulus was designed to isolate long-range interactions that subserve spatial integration of orientation information in the primary visual cortex, all we expected was that human observers would be able to detect the contours even if noise density is greater than contour density (in other words, when noise elements are closer to each other than contour elements). This, indeed, was observed, but there was another surprising observation: closed contours in these images were much easier to see than open ones (Kovács and Julesz 1993; Mathes and Fahle 2007). We called this a "closure superiority" effect. As described in Kovács and Julesz (1993), closure superiority can be measured at perceptual thresholds. Threshold effects are difficult

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

to illustrate; however, Figure 14.4 is an attempt to show the results of the experiments in an instant demonstration. According to the results, in spite of the locally equivalent parameters (same elements and same spacing parameters), closed contours are perceived differently: they seem to get a kick during the process of segmentation.

If local features are detected by local filters, and their interactions are also local (between neighbors), what causes the elements of a closed contour to jump out, while the same types of elements along an open one blend in with noise? Closure is a global shape property (similar to the shape of the dome). Local filters and local interactions—even if they form long chains—cannot deal with global shape properties. Local errors will accumulate along the chains of local interactions, and the result will be uncertain. Top-down instructions that arrive from higher levels of the cortical hierarchy may also not help. The higher-level "hypotheses" about the shape will meet unsegmented local orientation signals (contour + noise) and, in such a dense field of elements, any global suggestion can take shape, and the result will be uncertain again. Are there more than just local interactions during segmentation in the primary visual cortex? Does some kind of global reference serve segmentation much the same way as the Brunelleschi–Ricci flower served the construction of the Florence Cathedral dome? If there is such a reference system, what are the "strings" (taking the analogy further) in the brain that connect the flower and the bricks?

How Does Neural Activity Signal Gestalts?

The answer to this question might be found in computational models (e.g., Li 2005; Mundhenk and Itti 2005) or in the careful investigation of cortical microcircuits (e.g., Angelucci et al. 2002). It seems to be clear that long-range lateral interactions between neighboring neurons in the primary visual cortex are relevant in co-linearity-based contour grouping; however, these might be insufficient to account for integration beyond that of neighboring neurons (e.g., Loffler 2008). The closure superiority effect might share the underlying neural mechanisms with perceptual phenomena such as surface perception, surface interpolation, or "filling-in" (an excellent review is provided by Komatsu 2006). This Forum has provided a platform to explore these issues: in particular, whether the binding problem as defined by von der Malsburg (1981/1994), temporal correlations in the activity patterns of neurons (e.g., Engel et al. 1992; Phillips and Singer 1997), or a flexible assembly of spatial patterns of coordination (Haken et al. 1990; Kelso 1995) might explain global effects, such as closure superiority. I suggest that it might be essential to consider representational constraints before pointing out actual neural mechanisms. I will define and illustrate one of these constraints in the next section.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

The Problem of Space–Time Resolution: Representing Living Things

Parsing an image into figure and ground is still far removed from the goal(s) vision may have evolved to accomplish. In addition to the role vision plays in guiding locomotion across space, perceiving the complex movements of living things is also essential. Although our visual environment is dominated by artificial objects, the human visual system appears to be fine-tuned to extract effortlessly socially relevant information from the movements of another person—a task that is essential for interpersonal interactions. Considering the nonrigid movements of the body, this requires an efficient and coupled coding of visual shape and motion information.

To illustrate the representational constraints on coding for biological movement information, let us turn to another historical example, this time from photography. A French medical doctor, Etienne-Jules Marey (1830–1904), attempted to capture time and to make all movements of the human body visible and measurable. He invented several devices to track circulation, respiration, and muscle function. In his studies on locomotion, his goal was to generate a description of complex human motion and depict the relationships, both in time and space, between various body parts (e.g., during a walk) using a single representation, within a single image. Marey realized that the two-dimensional graphs, which he used earlier to record the changes of a single parameter in either time or space, would not be useful in this case. While thinking about the appropriate space–time representation, he realized that photography might be an appropriate tool to capture and characterize human movement in time. Marey invented a camera with a fixed photographic plate and a rotating, slotted-disk shutter, which allowed him to overlay multiple exposures on the same plate and to reduce blur that would result when trying to take a shot of a moving subject. However, even with the fixed-plate camera, the problem of spatial blur was not completely solved. In fact, there was a trade-off between acuity in time and in space. If Marey increased the number of exposures (number of slots in the shutter), there were more pictures, and the resulting temporal resolution was better. However, due to contour overlap, spatial resolution was poor, and the images were blurred. Conversely, spatial resolution could be improved by decreasing the number of slots, but only at the expense of temporal resolution. Marey explicitly recognized the trade-off between spatial and temporal resolution (cited in Braun 1995:83):

> In this method of photographic analysis the two elements of movement, time and space, cannot both be estimated in a perfect manner. Knowledge of positions the body occupies in space presumes that complete and distinct images are possessed; yet to have such images, a relatively long temporal interval must be had between two successive photographs. But if it is the notion of time one desires to bring to perfection, the only way of doing so is to augment greatly the frequency of images, and this forces each of them to be reduced to lines.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

Chronophotography provided the final solution for capturing time. To prepare for his chronophotographs, Marey dressed his subject in a black costume and marked the joints with shiny buttons connected by metal bands (Figure 14.5a). The subject moved around in the dark, such that only the movements of the buttons and wires were recorded in the picture. By selecting what he considered the most informative points and lines, he was able to read the successive postures of the body in his plates and follow the important trajectories of motion (Figure 14.5b). The relevance of this solution for human vision was later confirmed by Gunnar Johansson's work on the perception of biological motion in the point-light walker displays (Johansson 1973), and by the motion-capture method employed in psychology (Troje 2002) and in modern animation techniques.

The Role of Symmetry in Optimizing Space–Time Resolution

I began discussion of the closure superiority effect by referring to Kurt Koffka, who emphasized the relevance of surface regions enclosed by closed contours. The conclusions of the "dome" and the "closure" stories might be according to the taste of Gestalt psychologists. Indeed, the wonderful shape of the dome is not simply a collection of arches, but rather a three-dimensional volume; the closed line is not simply a collection of line segments, but rather a two dimensional surface. Ricci's flower theory suggests how the Gestalt of the dome can

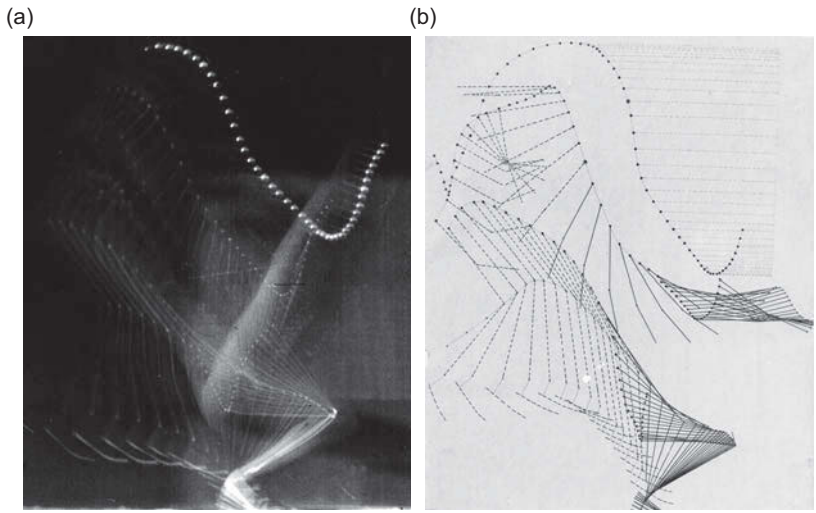

Figure 14.5 Geometric chronophotography by Marey. (a) The subject, dressed in black and photographed against a dark background, has been recorded jumping from a chair. (b) The chronophotograph of a jump is analyzed with graphics. The original glass plates (from 1883) are held by the College de France Archives.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

be assembled from millions of bricks using an ingeniously simple line of reference. This line of reference is the main symmetry axis of the dome.

Marey's space–time diagrams are wonderful examples of the representational issues inherent in capturing the movement of complex bodies. In addition to their beauty, the chronophotographs demonstrate the pertinence of symmetry axes. The question arises whether the brain, when processing similarly complex information, uses similar representational solutions. Do symmetry axes play a crucial role in vision? In a series of experiments (Kovács and Julesz 1994), this question was posed with respect to the representation of simple shapes within the primary visual cortex. How is an extended, closed circle represented in the activity pattern of orientation-tuned neurons that have small receptive fields?

Simultaneous activity of a large number of interacting neural elements can be revealed by tracking the activity of several units simultaneously in search of their higher-order correlations, such as in electrophysiological cross correlation and multiunit studies. An alternative way is to estimate how the activity of one unit is affected in the context of the activity of other units. The latter approach was used in a psychophysical reverse mapping technique, where the activity of one unit is measured as a function of the changing context. Psychophysically measured local contrast sensitivity reflects the local activity of neurons, and the context of the interaction pattern can be manipulated by changing the overall stimulus design. Closed contours (illustrated in Figure 14.4) were employed as context, and single Gabor signals were used to obtain local contrast sensitivity of human observers (Kovács and Julesz 1994). The position of the local signals varied within the closed contours, and by measuring sensitivity for many locations within these contours, a map of contrast sensitivity was obtained for each investigated shape: circle, ellipse, cardioid, and triangle (detailed in Kovács et al. 1996, 1998).

To be able to see if any sensitivity change within these contours is related to the symmetry axes of the shape, we used a medial-axis type transformation (Blum 1967; Siddiqi and Pizer 2007). The D_ε function, as shown in Figure 14.6, is based on an equidistance metric, where the D_ε value of each internal point represents the degree to which this point can be considered as the center of the local boundary segment around it. The transformation provides a nonuniform skeleton of the shape, with one or more peak values. The peaks are very important, and are equidistant from the longest segments of the boundary. In other words, these are the most informative points, and long contour segments can be traded for them. We used the maxima of the D_ε function, which we called the medial-point representation, to predict potential sensitivity changes within the simple shapes mentioned above. If there is any change related to symmetry axes in local contrast sensitivity, it should be around these maxima!

To our great surprise, the D_ε function was a wonderful predictor of psychophysical performance. The prediction worked for simple shapes, circles, and ellipses as well as for shapes with curved and branching symmetry axes

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

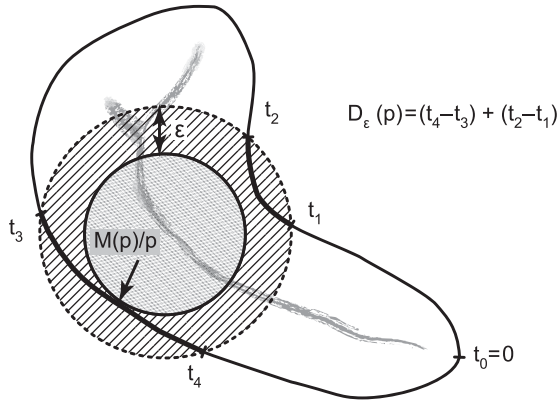

Figure 14.6 The D_ε function: D_ε is defined for each internal point by the percentage of the boundary points that are equidistant from the internal point within a tolerance of ε.

(Kovács et al. 1998). The contrast sensitivity changes were extremely specific, not simply some inside-specific enhancements. The maxima of the sensitivity changes corresponded to the maxima of the D_ε function. Neural correlates of these results have also been found in the modulation profiles of single-cell activity in the primary visual cortex (Lee et al. 1998; Lee 2003), although further confirmation of these would be useful. The psychophysical and neurophysiological data indicate that in addition to being sensitive to global shape properties (e.g., closure, and figure–ground relationships), the primary visual cortex is sensitive to specific shape properties and can host a medial-point type representation. The most provocative possibility is that this early cortical area provides a sparse skeletal code of shape!

Notice the similarities between the chronophotographs (Figure 14.5) and the medial-point representation (Figures 14.6, 14.7). Both provide a very small number of local points that can replace an "infinite" number of points composing a shape. The compactness does not preclude the representation from reflecting global shape properties. Such a sparse shape-coding would be a very desirable tool of communication between low- and high-level cortical visual areas, such as between the primary visual cortex and the inferotemporal area. Using only a handful of "hot-spots" to send information of segmentation in a bottom-up manner, and to send knowledge-based expectations top-down,

Figure 14.7 D_ε for sequential frames of the movements of an animal. The maxima of the function are good candidates as primitives for biological motion computations.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5,
series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.

might provide the brain with a channel that works better than any current artificial image-compressing tool.

The physical world seems to operate along basic laws that exhibit known symmetries (e.g., translation in space, rotation through a fixed angle, etc.). The near-symmetries and broken symmetries might be even more interesting for physicists. A powerful example of a broken symmetry involves the phase change of water with decreasing temperature. As temperature goes down to 0°, liquid begins to solidify. Interestingly, at the same time, rotational symmetry disappears from the structure that H_2O molecules line up into. This might, in fact, be generalized to phase transitions in the domain of visual shape, and breaking those simple symmetries might be more interesting than the symmetries themselves.

Are the above-mentioned contrast sensitivity maps purely epiphenomenal, or are they proof of intelligent image compression in the visual cortex? If the latter, how is this implemented precisely by the cortex? Perhaps (not necessarily oscillatory) synchronous firing of orientation-tuned neurons mediates the compression and provides "hot-spots" in the neural representation of the segmented visual input. In what neural language would these "hot-spots" then be transmitted to more abstract levels of processing to meet with linguistic representations and semantic memory? In addition, how would the enriched information advance thereafter through the feedback pathways to enhance segmentation?

Acknowledgments

The author was supported by the ETOCOM project (TAMOP-4.2.2-08/1/KMR-2008-0007) through the Hungarian National Development Agency in the framework of the Social Renewal Operative Programme supported by the EU and co-financed by the European Social Fund.

From "Dynamic Coordination in the Brain: From Neurons to Mind,"
C. von der Malsburg, W. A. Phillips, and W. Singer, eds. 2010. Strüngmann Forum Report, vol. 5, series ed. J. Lupp. Cambridge, MA: MIT Press. ISBN 978-0-262-01471-7.