1996 SID International Symposium, Digest of Technical Papers Santa Ana, CA, 92705-5421: Society for Information Display. 961-964.

43.1: First Principles of Second-Order Perception

G. Sperling, Z.-L. Lu, C. Chubb
University of California, Irvine, CA

ABSTRACT

The classical perceptual computations, such as direction-of-motion, texture-slant, lateral inhibition (Mach bands, Chevreul illusion) are designated "first-order" because they involve photons directly. Second-order phenomena are revealed by computer-generated stimuli that are transparent to first-order processes. The phenomena of second-order perception perfectly mimic those of first-order and rival it in statistical efficiency. Second-order perceptual computations are the same as first-order, but the elements are features, not photons, and the functional receptors are "texture grabbers" not rods or cones.

Here we briefly review the current status of theories concerning the first level of sensory processing, *first-order* perception, show how these theories can be elaborated to encompass more complex processing, *second-order* perception, and suggest possible applications.

Classical First-order Perception

Overview. The classical theories of first-order perception were developed after the Second World War and have held sway since then. In brief, they were linear theories that modeled the visual system in terms of arrays of parallel linear filters. Subsequent processes were typically encapsulated in a decision component. Initially, the decision component was a simple threshold device: If its input exceeded its threshold it responded "Yes" (to indicate detection or discrimination), otherwise, no. Subsequently, the critical role of internal and external noise was recognized, and matters have become much more complex.

Adaptation. The first stages of visual processing, the receptor and immediate post-receptor processes, essentially normalize the visual input s(x,y) at a point (x,y). Normalization can be represented as follows: Let $\bar{s}(x,y)$ be the average input in a weighted area (the surround) around (x,y). The point-contrast of s is $(s - \bar{s})/\bar{s}$. Thus, point-contrast is simply the fractional amount by which input deviates from its mean value. Nearly all subsequent visual processes are based on point-contrast and have no direct knowledge of absolute luminance.

Receptive Fields. Retinal ganglion cells, which form the primary pathway from retina to brain, occur in an enormous range of sizes. Each point of the visual field is represented by ganglion cells that vary in diameter by a factor of 100. Ganglion cells are modeled as three-dimensional bandpass filters. In space, they transmit only about an octave-wide band of frequencies that is inversely related to cell receptive field size. Ganglion cell receptive fields have been approximated by a difference of Gaussians (DOG), and by a smoothed

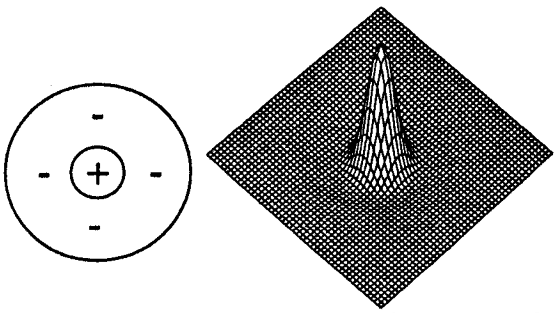

Fig. 1. Receptive fields, input filtering. a) Conventional representation of a circularly symmetric receptive field (e.g., of a retinal ganglion cell). b) 3D representation of (a) as the impulse function of a filter.

---

*On leave from Rutgers University, New Brunswick, New Jersey.

ISSN0097-0966X/96/2701-0961-$1.00 + .00 © 1996 SID

Laplacian. A DOG is illustrated in Fig. 1.

In the time domain, ganglion cells also have a bandpass characteristic. In both space and time, the surround drops out at very low intensities, so that the actual filter shape changes with the level of mean luminance.

In the area V1 (Visual 1) of cortex, receptive fields first exhibit orientation specificity. These can be approximated by Gabor functions. Such an oriented, spatial-frequency-specific filter is illustrated in Fig. 2 and will be referred to here as a V1 filter.

Channels. When vision represents the world perfectly, it is virtually impossible to learn how the visual computation is performed. Errors reveal the underlying processing mechanisms; this suggests why we have such a great interest in visual illusions. We learn about visual channels from an illusion illustrated in Fig. 3. Looking persistently (with only small eye movements) at the high- and low-frequency grating on the left fatigues the neurons that are most tuned to those frequencies in those locations. When we later look at a stimulus that has a medium spatial frequency, the areas with fatigued high-frequency neurons appear to contain lower spatial frequencies, whereas areas with fatigued low spatial-frequency neurons appear to contain higher spatial frequencies. This demonstration shows that the visual system must contain channels. Further variations of procedures like this permit the estimation of the channel bandwidths.

From the retina on, the visual signal is split into spatial-frequency channels. A surprising amount of visual processing is carried out more-or-less separately in each channel before, eventually, information from different channels is recombined. Why spatial frequency channels? Because an object or scene is best represented at a particular scale. Since the same object can occur at many different object scales depending on the viewing distance, visual system matches its processing scale to the object scale.

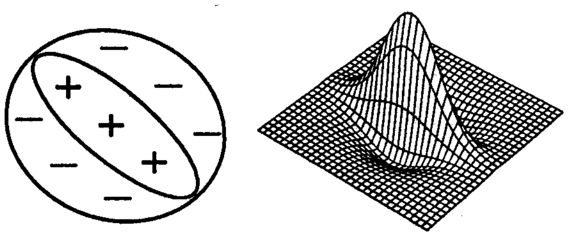

Fig 2. A spatial frequency and orientation specific spatial impulse response function for a V1 filter based on Hubel-Weisel cortical "simple" cells.

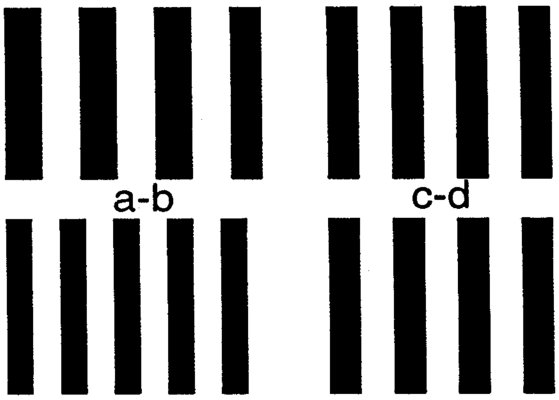

Fig. 3. Blakemore-Sutton paradigm for demonstrating channels. Look alternately at the a-b between the left column of the gratings for about a minute (adaptation) and then look at the c-d between the right pair (test). The illusion is a shift of the apparent frequency of the test away from the adaptation stimulus. (Based on Blakemore & Sutton, 1969).

Second-Order Perception

Basic Phenomena. There are numerous experimental date and some fascinating illusions for which these first-order principles give an excellent account. In particular, in motion perception, an area that we have explored extensively, there are a several counter-intuitive phenomena that have been predicted and then almost perfectly verified by experiment.

A problem arose when computers enabled us to generate motion stimuli that are, on the one hand, invisible to the first-order motion processes and, on the other, perfectly perceived. These stimuli have in common the property that features are substituted for pho-

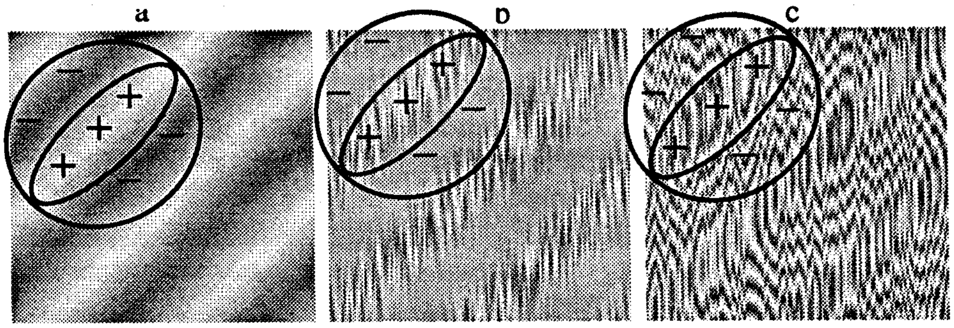

Fig. 4. Grating examples: (a) First-order. A luminance-modulation grating. (b) Second-order. A texture-contrast modulation grating (c) Second-order. A texture-frequency modulation grating. Alternatively, interpreting the vertical axis as a time axis, t, yields an x,t representation of a flicker grating (in which the band of rapid flicker moves across the screen). (Based on Chubb & Sperling, 1988).

tons. Thus, a sine wave grating contains more photons in the light stripes and fewer in the dark stripes (Fig. 4a). One kind of second-order grating contains stripes composed of more texture (e.g., higher texture-contrast) alternating with stripes of less texture (lower texture-contrast, Fig. 4b). In second-order gratings, the expected luminance is exactly the same everywhere. Figure 4 shows a V1 filter superimposed on the gratings, and it is obvious why it fails to respond to the second-order stimulus: The amount of light reaching the "+" areas and the "-" areas of the V1 filter is virtually the same. Indeed, Chubb and Sperling proved that stimuli of this type are *driftbalanced*: no linear filter could discriminate their orientation. These second-order stimuli are also *microbalanced*: They can be viewed through any shape of aperture or window and they retain their driftbalanced property.

Texture contrast is not the only feature that can be used to define second-order stimuli. Flicker (versus nonflicker) is a temporal feature that can be used, and grating spatial frequency and orientation (e.g., high spatial frequency versus low; left slant versus right slant) also can be used. Figure 4c illustrates a microbalanced grating composed of two different binary textures that have exactly the same contrast. The perception of second-order motion is as vivid as that of first-order motion--observers do not have any direct sensory knowledge of which motion system is being stimulated. We have yet to find an observer who fails to perceive second-order motion.

Mechanism. The mechanism for extracting features is a *texture grabber* (Fig. 5). A linear filter, such as exemplified by a specific size of center-surround configuration, or by an oriented, spatial frequency tuned V1 filter, defines the texture. The stimulus is processed in parallel by a field of such selfsimilar filters. The outputs of the filters are then fullwave rectified (Solomon & Sperling, 1994). That is, positive and negative outputs are treated equally, as in an absolute value or a square. Subsequent processes want to know how much of the feature was present at a location, not whether it happened to have a positive or negative phase. Indeed, if outputs were not rectified, the expected output value would be zero. A texture grabber also includes a temporal bandpass filter to sensitize it to the onset of a new texture. In second-order processing, the output of a

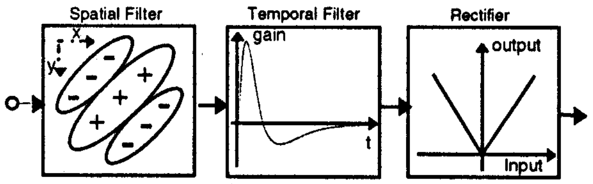

Fig. 5 A second-order mechanism. An input stimulus passes through a spatiotemporal band-pass filter and then is rectified (texture grabbing). Subsequent operations on the outputs of texture grabbers are analogous to operations on raw photon inputs. (Based on Chubb & Sperling, 1988)

field of texture grabbers is then processed exactly the same way as the output of photon grabbers--the rods and cones.

Illusions. The magnitude of second-order illusions tends to be similar to that of the analogous first-order illusions. Two examples are shown: A classical lightness illusion and the Chevreul illusion (Figs. 6,7). In each case, photon intensity (luminance) in the first-order example has been replaced with feature intensity (texture contrast) in the second-order example.

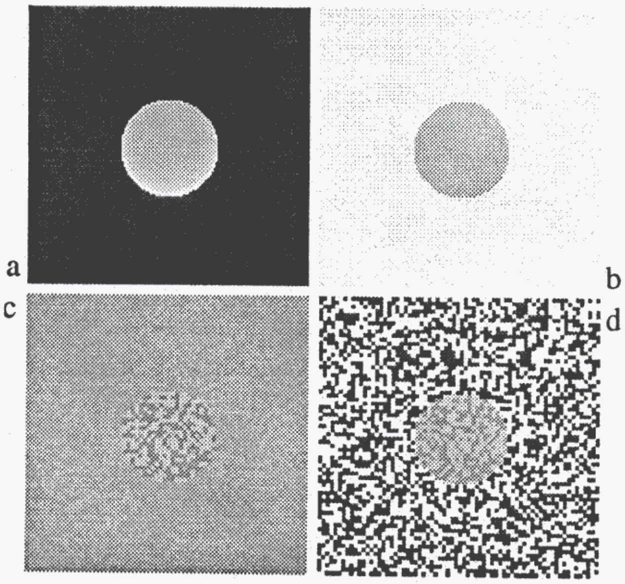

Fig. 6. The classical lightness illusion and the contrast-contrast illusion. The center disks in the left and right panels are identical. (a,b) In the lightness illusion, the (apparent) brightness of the center depends on the surround. (c,d). In the contrast-contrast illusion, the apparent contrast of the center depends on the surround. (Based on Chubb, Sperling, & Solomon, 1989).

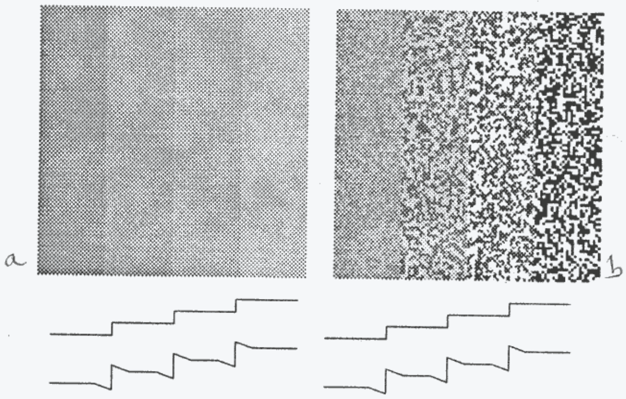

Fig. 7. Chevreul illusion. (a) A rectilinear *luminance* staircase appears to be scalloped (apparent) brightness staircase. (b) In the second-order version, the rectilinear *contrast* staircase appears to be scalloped. (Based on Lu & Sperling, 1995).

Postscript

Obviously, in designing displays, it is useful to have a theory or model of the visual system of the viewer. The most obvious uses for models of the early stages of perception are in the interpretation of displays, i.e., in the realms of image understanding and in the evaluation of image quality. Second-order properties are potentially useful for the interpretation of displays. But they can also be used to develop better schemes for for bandwidth reduction in image transmission, for data compression in the storage of visual information, for making better computer generated dispays, and for uses that are yet to be discovered and exploited.

REFERENCES

Blakemore, C. & Sutton, P. (1969). Size adaptation: A new aftereffect. Science, 166, 245-247.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. J. Opt. Soc. Amer. A, 5, 1986-2006.

Chubb, C. Sperling, G, & Solomon, J.A. (1989). Texture interactions determine perceived contrast. Proc. Natl. Acad. Sci., USA, 86, 9631-9635.

Lu, Z-L, & Sperling, G. (1996). 2nd-order illusions: Mach bands, Chevreul, and Craik-O-Brien-Cornsweet. Vision Research, 36, 559-572.

Solomon, J. A. & Sperling (1994). Full-wave and half-wave rectification in 2nd-order motion perception, Vision Research, 34, 2239-2257.

ACKNOWLEDGMENT

Supported by AFOSR, Life Sciences, Information Processing Program.