Lawrence Berkeley National Laboratory
Recent Work

Title
SPECIAL SESSION ON VISION

Permalink
https://escholarship.org/uc/item/5jk2v6tr

Author
Montalvo, F.S.

Publication Date
1979-05-01
SPECIAL SESSION ON VISION

Fanya S. Montalvo

May 1979

Prepared for the U. S. Department of Energy under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782
SPECIAL SESSION ON VISION

chaired by

Fanya S. Montalvo
Lawrence Berkeley Laboratory
Berkeley CA 94720

Is one picture really worth a thousand words? Why do cleverly designed graphic displays make visual information stand out more clearly with strikingly greater impact than numbers buried in pages of computer printout?

Graphic output devices shift the burden of integrating information generated by computers onto the human vision system: the sensory channel with the highest capacity for distributed parallel processing. The system consists of hundreds of successive two-dimensional arrays of millions of interconnected parallel computers. Perception seems instantaneous because we are not conscious of the massive amounts of computation that occur. What we consciously "see at a glance" is already a highly structured, synthesized, and summarized version of the actual light intensity mosaic that enters the retina.

We will demonstrate some results of the visual structuring that occurs in the human visual system, show why some features stand out instantaneously and others do not, and explain why knowledge of the human input device is crucial to the design of effective computer output devices and displays.

PARTICIPANTS

Stephen R. Levine  Lawrence Livermore Laboratory
Introductory Remarks

Fanya S. Montalvo  Lawrence Berkeley Laboratory
Vision and Computer Graphics

Stuart Anstis  York University
Motion

Bela Julesz  Bell Telephone Laboratories
Preattentive Perception

Richard L. Gregory  University of Bristol
Color
1. INTRODUCTION

Why should computer graphics concern itself with human vision? The very reason for the rise in the use of graphics in the first place is deeply embedded in the way humans perceive. Graphs and pictures convey more information more quickly than pages of numbers. We all know this intuitively. Few of us, however, stop to consider why. What is it about visual perception that allows this seemingly instantaneous communication from computer to human? Conversely, why does the synthesis of concepts or objects from pages of numbers require such time-consuming mental effort? The answer is to be found in the way the visual system is structured and in what it was designed to compute. It cannot and does not compute everything it "sees." Our unawareness of the enormous amounts of parallel computation taking place gives the impression that seeing is simple, that seeing is a one-to-one mapping, and that everything on the physical display is transmitted directly to the user's brain. Nothing could be further from the truth. Tremendous inhomogeneities, distortions and super-acuities exist in the human visual system. The system is rich in special parallel hardware designed to detect specific classes of features. Neural cells code for specific features, such as the width and orientation of bars, and operate in a distributed fashion throughout the two-dimensional visual field. Interactions between cells can make some features stand out more clearly, while obscuring others, and distorting still others. The distinction between preattentive perception, requiring no conscious effort, and that requiring attention or serial inspection helps to identify which features are processed by this special visual hardware and which require higher cognitive processes.
We need to know what characteristics of visual presentation make information stand out quickly, clearly, without distortion and without conscious effort. If we cannot eliminate distortion we should know how to compensate for it. If we cannot duplicate a natural image exactly we should know what approximations are close enough to the perception we wish to transmit. If we expect a pattern recognition program to make automatic decisions about natural images, we need to know about the subliminal decisions made by our brains in processing those images. In generating synthetic images, we need to compensate for distortions such as aliasing and staircasing, produced, respectively, by our approximations to smooth boundaries and curved surfaces. These devastating distortions occur because of the special characteristics of the visual system and not because rastering and faceting are poor approximations. Other approximations do not cause severe problems, for example, distortions of the color space produced by a badly tuned color television. The reasons for these differences in severity can be found in the various units of special hardware the visual system uses to detect specific features and in the interactions between these units that accentuate some features at the expense of others.

2. SPECIAL HARDWARE

Biological visual systems possess many specialized processes that enhance contrast, making boundaries sharper [Ratliff & Hartline 59] and making gradual intensity changes disappear. Other processes detect symmetry within local regions [Julesz 75]. In addition, neural cells have been found that are specifically tuned to line orientation [Hubel & Wiesel 62], size or spatial frequency [Blakemore & Campbell 69], binocular disparity or depth [Barlow et al. 67], color [DeValois et al. 66], and speed and direction of motion [Barlow
et al. 64]. All of these special feature detectors presumably operate in parallel, arranged topographically over the two-dimensional visual field. They interact with each other, enhancing some of their neighbors and inhibiting others.

2.1 LATERAL INHIBITION
Lateral inhibition between cells in the retina causes an exaggeration of intensity differences at boundaries [Ratliff & Hartline 59]. Mach bands, which appear as dark and light bands at positive and negative discontinuities in the first derivative of light intensity, are also caused by lateral inhibition [Ratliff 65]. In computer graphics, contrast enhancement of light intensity is known as staircasing when it occurs at boundaries between flat areas of uniform intensity, and as scalloping when it occurs in 2D polygonal approximations to 3D curved surfaces. Lateral inhibition can also cause the appearance of two areas of equal intensity to differ if the transition between them has a sharp intensity discontinuity completely surrounding one of the areas [Cornsweet 70]. Areas that actually do differ in intensity may not be discriminated if the transition is too gradual.

2.2 FEATURE DETECTORS
The visual system contains neural cells selectively sensitive to specific features in the visual input. Some of these detect line orientation, bar width, spatial frequency, direction and speed of motion, and binocular disparity. For example, pools of cells, each representing a specific degree of tilt, map the continuous dimension of line orientation. Each cell fires maximally to its preferred feature along its continuous feature dimension and drops off in response the more the input feature differs from the optimal one. In the case of orientation, for example, a cell may fire maximally for a
vertical bar and drop off to 50% of its maximal response for 15 degrees of tilt away from vertical. The accuracy with which we can detect differences between bars of different orientations seems to depend on the density of such cells along the feature dimension [Montalvo 76]. The distribution of line orientation cells has clusters at vertical and horizontal, and it can be shown that discrimination excels for differences near these orientations [Mansfield 74]. Time and space interactions between detectors enhance contrast between near neighbors. Space interactions are known as simultaneous contrast and time interactions are known as adaptation. Both refer to the exaggeration of differences between adjacent features along a given dimension. Masking, the degradation of one feature by another, often occurs for features adjacent both in space and in the feature domain [Weisstein 72]. One example is demonstrated by Harmon and Julesz [73]. Face recognition is more effectively impaired by noise that is spectrally adjacent to the picture's spatial frequency spectrum. Noise of the sort introduced by digitization need not interfere with the information content of the picture if it is spectrally far enough away.

2.3 SPECIAL COMPUTATION

Some processes cannot be isolated to single cells using neurophysiological methods. However, using psychophysical methods we can infer that a given computation must occur if humans can identify particular aspects in the visual field. If identification can occur anywhere in the visual field without eye movements or time for inspection then we can assume that computation occurs automatically and in parallel over the entire field. Julesz has clarified the distinction between preattentive perception and the form of perception requiring serial inspection and attention [Julesz et al. 73]. By generating various series of discrimination tasks between textures that vary only in
their Nth-order statistics, or in minor aspects of their local geometry, they were able to isolate crucial factors of preattentive perception.

Two preattentive visual processes that cannot be isolated to computations occurring in single cells are binocular depth perception and a local symmetry operation. Julesz [71] was first to show that these correlation processes occur even in the absence of meaningful form cues. If two identical random dot patterns are generated, one with a central area of dots slightly displaced, and these patterns are projected binocularly, that is, each to one eye, the central portion will appear at a depth different from the background. This will occur even though no edges or meaningful forms are visible to either eye separately. The symmetry perception operation can be demonstrated with a similar random dot pattern. Each quadrant of the field is replicated with axial symmetry about the x and y axes. The pattern will appear symmetric only in a certain zone near the x and y axes on the pattern. Beyond this zone the pattern will appear random. This demonstration implies that the symmetry operation is limited by distance.

The preattentive property of both the binocular and symmetrical correlation operations can be clearly distinguished from other kinds of correlations requiring mental effort. For example, if the four quadrants in a random dot pattern are duplicated four times without being mirrored, the entire array appears totally random [Julesz 75]. Such demonstrations distinguish those operations in the visual system which are "pre-wired" to occur automatically from those that require focusing of attention towards producing the necessary operations at the cognitive level. The best strategy for producing clear displays that transmit information quickly is to distribute the information over the many parallel, automatic processes that exist in the low-level
visual system.

2.4 VERNIER ACUITY AND ALIASING

Aliasing, the perceived jaggedness at high contrast edges, has become a particularly annoying display problem with the growth of computer-generated, raster scan pictures and films. The problem is not necessarily solved by using higher resolution display devices. Individual edge discontinuities may become too small to see, but regular patterns of edge breaks may be much more visible as global patterns. Thus, besides the added expense of higher resolution, these more offensive problems due to sampling may result [Crow 77]. The heuristic solution to aliasing is a smooth, rather than abrupt, intensity transition from one line segment to the next, computed from the degree of overlap of the desired line and the model display grid. The resulting display transforms location information which is finer than the display grid into intensity variations. This solution works because of the visual system's peculiar tolerance for the replacement of location information by intensity information.

The aliasing problem can be better understood by examining the visual system's extreme sensitivity to jagged edges. This sensitivity has been found to be in the same range as vernier acuity [Hamerley & Springer 78]. Vernier acuity is the sensitivity to a slight displacement between parallel, almost aligned contours. It is more than ten times greater than the spatial resolution of the retina. Spatial resolution is limited by receptor spacing and light diffraction on the retina. Receptor separation is about 30 seconds of visual angle, but, remarkably, vernier acuity is on the order of 2 to 3 seconds of visual angle [Berry 48, Riggs 65]. Thus, changes in line width or edge location finer than the retinal mosaic can be detected only as variations in
intensity. Since long contours involve sampling by many receptors along a straight line or edge, some form of signal averaging or comparison over many units must be involved [Riggs 65]. Signal averaging of intensity over long contours dramatically increases the system's ability for localization.

2.5 COLOR

2.5.1 Cones: The First Layer

Color graphics devices can reproduce virtually all colors perceived by humans with only three variables. This is because the human eye has only three types of narrow-band light receptors, called cones, with peak sensitivities at three different wavelengths on the electromagnetic spectrum. Their peak sensitivity corresponds roughly (not strictly) to monochromatic red, green and blue. The appearance of any monochromatic color can be simulated by reproducing the excitation it produces in these three distinct populations of receptors. Similarly, any color composed of a distribution of monochromatic colors across the spectrum can be reproduced with proper proportions of red, green, and blue. [See Smith 78 and Joblove & Greenberg 78 for further discussion of the 3D color space.]

One problem arising from the non-uniformity of the distribution of color receptors over wavelength is a non-uniformity of hue discrimination over wavelength and intensity [Hurvich & Jameson 57, DeValois 65, Gregory 66]. Because of the close proximity of red and green cone cells' peak responses and the sharp drop in response between their peaks, discrimination of hue differences about yellow is much more acute than discrimination near blue. It is because of these non-uniformities that we have to be very cautious about representing data by variations in hue. The greatest differences in the data may not correspond to the greatest perceptual hue differences unless the color
space is corrected for equal hue steps. If we want to be able to pick out distinctive features of the data we wish to display, we need to have some idea of how the perceptual hue mapping corresponds to the color space of the output device.

2.5.2 Opponent-Color Processes

Beyond the retina the three color receptors are linked to cells in the lateral geniculate nucleus (LGN) on the way to the visual cortex [Abramov 68]. Cells at the LGN are organized into four general populations: cells stimulated by red and suppressed by green, cells stimulated by green and suppressed by red, cells stimulated by yellow and suppressed by blue, and cells stimulated by blue and suppressed by yellow [DeValois et al. 66]. This kind of opponent-color process seems to account for color contrast and adaptation [Hurvich & Jameson 57, Bajcsy 75]. Color contrast is a spatial phenomenon. Gray next to red looks much greener than the same gray next to green. Adaptation refers to negative effects over time. After hours of watching a color television image with an overall red component, the colors begin to look perfectly balanced. This is partly due to the automatic gain control performed by the retina on each color channel. Other forms of color aftereffects are believed to occur in the LGN and visual cortex as well [McCollough 65]. In the case of our color TV example above, the perceptual color space is gradually rescaled such that the net red component in the red-green system appears white. Rescaling of a continuous feature space seems to be involved in many other forms of adaptation as well [Over 71, Montalvo 76].

2.5.3 Color Resolution

Spatial resolution for color is lower than the resolution for black and white [VanDerHorst 69]. This fact has implications for anti-aliasing and
anti-staircasing algorithms in color. Sharp borders between colors of equal luminance tend to disappear [Boynton & Kaiser 68]. (Luminance is a term that refers to light intensity corrected for the spectral sensitivity of the visual system in order to produce equal effectiveness in stimulating the retina.) This means that chromatic edges between areas of equal luminance need not be anti-aliased as carefully as achromatic edges. No border enhancement occurs between colors of equal luminance, as is true of the black and white domain [DeValois & Pease 71]. Thus, the degree of anti-staircasing found necessary for black-white transitions need not be as high for purely chromatic transitions. If a color output device is properly balanced to ensure equal luminance between colors at all levels, fewer levels need be devoted to hue transitions than to luminance transitions.

3. CONCLUSION

We have seen that the existence of parallel, automatic populations of cells and operations in the human visual system make visual information stand out quickly and clearly in graphic displays. Exploitation of these properties of the visual system allows us to transmit more information to the user more quickly. We, in effect, utilize the human channels with the highest bandwidths. Doing so shifts the burden of the integration of information onto the visual system and away from higher mental processes requiring more time and attention. However, we need to be aware of the idiosyncrasies of the system in order to best utilize its full capacity and in order to compensate for nonlinearities that can seriously distort the display output.
REFERENCES


Berry, R.N. Quantitative relations among vernier, real depth, and stereoscopic depth acuities. J.Exp.Psychol. 38, 708-721 (1948).


This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.