Gestalt Vision Experiments from an Image Processing Perspective

John McCann
McCann Imaging
Belmont, MA

Abstract

In the late 19th century, the Gestalt Psychology rebelled against the popular new science of Psychophysics. The Gestalt revolution used many fascinating visual examples to illustrate that the whole is greater than the sum of all the parts. The physical interpretation of sensations and their quantification by JND's and Weber fractions were met with innumerable examples in which two identical physical stimuli did not look the same.

The debate continues today with proponents of both physical, pixel-based thinking and perceptual, image based thinking. Image processing concepts can provide a new way of analyzing famous Gestalt displays. By this way of thinking, simple multi-resolution spatial processes can account for the various appearances of identical stimuli by analyzing the spatial properties of the entire image. Benary's Cross, Adelson's diamonds and Logvinenko's gradients can all be modeled by the low-, middle- and high-spatial-frequency components of the image. Pixel based analogs of physics and object cognition can be replace by familiar image processing concepts, such as multi-resolution spatial comparisons.

Introduction

E. Boring, in his "History of Experimental Psychology" describes that both Gestalt and Behavioral Psychology as reactions against the idea of quantifying psychology by adopting the procedures of physics.. He wrote, "Each is a protest of the 'new' German psychology of the late nineteenth century, the psychology of Wundt, G.E. Muller and Titchner, but the two are different protests. Gestalt psychology primarily protested against the analysis of consciousness into elements and the exclusion of values from the data of consciousness, whereas behaviorism mostly protested against the inclusion of the data of consciousness in psychology".[1]

One of the principle weapons used by Gestalt psychologists was the direct comparisons of "identical" stimuli that did not appear the same. Figure 1 illustrates this point because the gray area in the middle of the white area has the same reflectance as the gray in the black surround. They do not appear the same. The gray surrounded by white looks darker than the one surrounded by black. Equal stimuli do not generate equal appearances. The Gestalt psychologists argued that this was evidence that the whole (appearance) was not equal to the sum of the parts (reflectance of the gray pixels).

In a recent editorial Kindom[2] described the century-long debate on visual appearance conducted by disciples of Helmholtz and Hering. Helmholtz[3] is harder to understand because of the duality of his thinking. On the one hand, he popularized Thomas Young's idea that color required only three types of retinal receptors. Young invented, Maxwell quantified and Helmholtz established trichromatic color theory. Helmholtz and Konig extended Maxwell's work of measuring human color matching functions. Helmholtz did much to establish the importance of physics-based "pixel" thinking of colorimetry. On the other hand, he introduced the idea of "unconscious inference" to explain color constancy. This top-down concept is in sharp contrast with the bottom-up ideas of 20th century colorimetry.

Hering[4] began with a need to explain the uniqueness of red, yellow, green and blue hues. This led to the idea of opponent processing. Ladd Franklin[5] combined these apparently conflicting approaches by proposing a sequential zone theory. Kuffler[6] discovered that ganglion cells had an excitatory response to the center of its visual field and an inhibitory re-

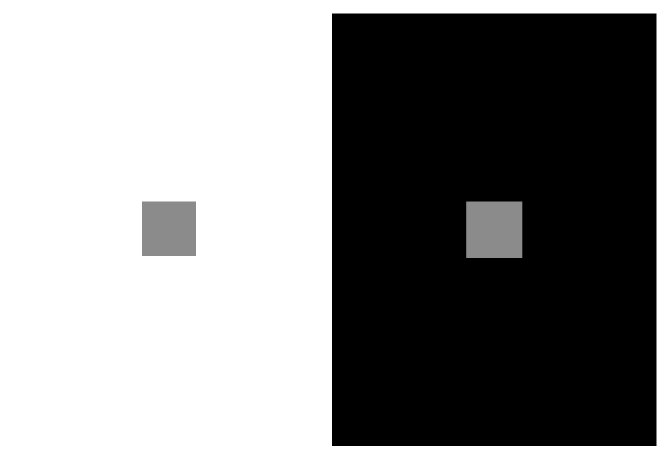

Figure 1 illustrates simultaneous contrast. The gray central squares have the same reflectance. The gray surrounded by white looks daker.

sponse to the periphery. A great many neurophysiological experiments lead to today's spatial opponency concepts of human visual processing.

The theme of this paper is that future research in vision will not resolve the debate of whether Helmholtz was right, or Hering was right . In the 21st century we should take the example of Ladd-Franklin and search for the integration of the best parts of both traditions. The theme of this paper is that common ground can be built around concepts of image processing. The Young-Helmholtz ideas of trichromatic spectral response to light are universally accepted. The Hering-Kuffler ideas are not in competition, and are accepted as well. What is needed next in visual theory is further understanding of spatial integration of the entire field of view in vision. These ideas come from experiments by Land[7] , Campbell[8] and Zeki[9]. They clearly demonstrate that human vision employs extensive spatial interaction. The resolution of Helmholtz vs. Hering debate will come from a better understanding of spatial vision.

Large Contrast Effects

As Kingdom[2], points out, Helmholtz's top-down inference ideas have been supported by many different experiments by Alan Gilchrist[10]. Along with Adelson[11,12], Logvinenko[13], and Ross and Pessoa[14], they all suggested mechanisms in which apparent depth and/or apparent illumination altered the appearance of lightness. In this paper we will restrict the discussion to flat displays, that look flat. This class of image is very important because it includes many complex phenomena that require more complex visual models than single pixel based algorithms. A wide variety of different experiments fall in this category. Gelb's experiment[15], B&W and Color Mondrians and real life scenes have demonstrated that nearly the total range of appearance can be generated by identical stimuli[16].

A number of image processing models have been proposed for these images. Land and McCann[17] described a spatial comparison extending Wallach's[18] observation about the importance of edge ratios. Stockham[19] suggested low-spatial frequency filtering. Ratio-Product-Reset-Average models have been used calculate the observed lightness in Simultaneous Contrast (Figure 1) since 1970.[20] The Black and White Mondrian was modeled in the original Land and McCann [17] article. Color Mondrians were modeled by the same technique.[21] In the early 1980's Frankle and McCann[22] extended the ration-product-reset-average operation to highly efficient multi-resolution image processing. A variety of outdoor images including a flat, real-life equivalent to the B&W Mondrian.[16]

All these experiments have been modeled using the same Ratio-Product-Reset-Average model which takes the information from the entire field of view and calculates an imperfect global normalization of the image. It is sensitive to the separation and enclosure of this maximum. It works with quanta catch at the retina as the only input. (A detailed discussion of the steps in this calculation, including MatLab code can be fount in a recent review paper.[23]) The model can account for the observed lightness in all the above flat images without depth or apparent depth information. Recent experiments by Logvinenko[13] and Adelson[12] integrate gradients and abrupt edges in computer displays. Logvinenko reports that the gradient variant of diamond walls generates twice the effect of other non-gradient Diamond Wall experiments.

All of these large contrast effects can be modeled by image processing programs that take local ratios and uses products to integrate across the field of view. A nonlinear reset removes the gradients. The general conclusion was that the model evolved from the study of Mondrians can as well calculate appearances of both real life scenes , simultaneous contrast, including the successful prediction of Logvinenko gradient "Diamond Wall" and Adelson's "Checkered Shadow" images.[16]

As originally described such spatial models do not provide a robust explanation of Balanced Contrast Effects such as Benary's Cross and White's Effect and Rizzi's Effect. These image have been created to have symmetrical local contrasts; both white and black areas are adjacent to the gray test patch.

Balanced Contrast Images

There are two interesting common properties of Balanced Contrast images. First, observers report appearances opposite those of Simultaneous Contrast. Second, these displays have popular cognitive, "unconscious inference" explanations.

Benary's Cross

Benary Cross (Figure 2 top) showed that grays still look different with two black sides and one white side. Benary argued that the different appearances were due to appurtenance, namely that we recognize and interpreted the image in terms of what we already known. We expect that the gray should look darker when surrounded by white, and we recognize that the gray triangle on the left is in the white surround, while the gray triangle on the right is in part of the black cross. Benary showed four different examples of appurtenance.[24] (Recently the word "appurtenance" has been modernized to "belongedness".) This is one of the oldest and most influential arguments for top-down modelsof vision.

White's Effect

Figure 2 middle.shows White's Effect.[25] The grays are surrounded by both white and black adjacent areas. Observers report the opposite to Simultaneous Contrast. The gray bars on the left are darker (with more adjacent black) than those on the right (with more adjacent white).

Rizzi's Effect

Recently Alessandro Rizzi showed me the effect of checkerboards on gray patches (Figure 2 bottom). It is easy to assume that because the gray on the left has four black squares adjacent to the square gray patch that it should behave the same as having a black surround. It has four white squares next to the black ones, but they are adjacent to the corners of the gray area. A similar argument applies to the gray with adjacent white edges on the right side of the image. Again the human visual system did not follow the rules of simultaneous contrast. Vision generates appearances that are again opposite to the effect seen in Figure 1.

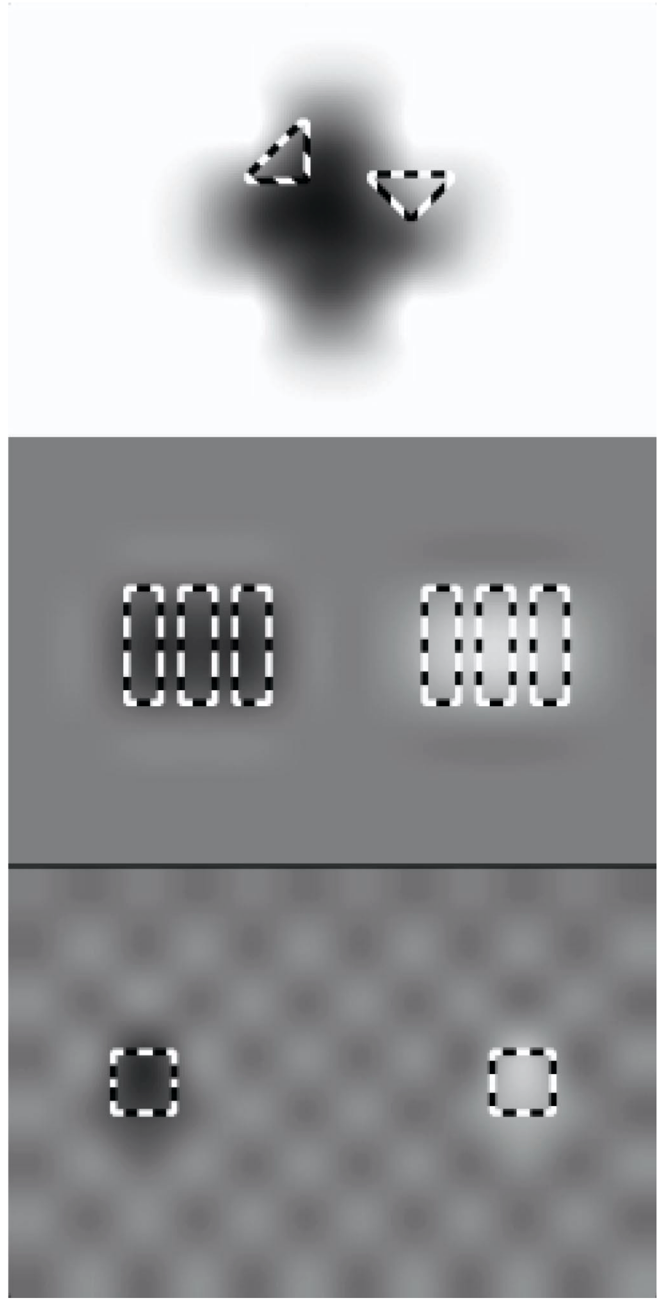

Figure 2 shows three different examples of Balanced Contast images. Each gray area is adjacent to both white and black. Each image exhibits the opposite appearance shift from Simultaneous Contrast. The top image is Benary's Cross. The triangle on the left appears darker than the one of the right. Benary's explanation was based on appurtenance, or recognition. The middle image is White's Effect. Unlike Simultaneous Contast, the grays on the left look darker, despite the fact that there is more black surround adjacent to the grays. The bottom image is Rizzi's Effect. As well, the grays on the left look darker despite the fact that there are four black areas adjacent to the grays.

Figure 3 show the images from Figure 2 after receptive field pooling (See text). It illustrates the idea that all three effects can be explained by human low-spatial frequency sampling. The common feature of all three displays is that the areas of interest on the left (top-triange, middle-rectangles, bottom-square) are darker than those on the right. The three different effects look darker on the left because the images are darker in low-spatial frequency sampling. The application of simple, human visual-image-procssing ideas can provide a single explanation of historicaly different phenomena. Top-down interpretations requiring cognition are not necessary.

Image Processing Ideas

As describe above, Ratio-Product-Reset models, that do very well with contrast and gradient images, do not explain these Balanced Contrast images. However, other spatial concepts provide an even simpler explanation of these effects. Campbell and Blakemore[26] demonstrated the existence of spatial frequency channels. Frankle and McCann[22] used multiresolution images and Burt and Adelson[27] used pyramid processing for efficient image computation. Almost any multi-resolution model can account for these results by sampling the output from each resolution independent of the others.

If the original image is three inches high and was viewed at 14 inches. The image subtends 12.3 degrees. The analysis of the image consists of averaging four adjacent image pixels to make a new image of one-quarter the size. The process is repeated each time making a new image one-quarter the size of the previous sampled image. In the image shown here the full resolution images were 256 pixels high. The pictures in Figure 3 were made by averaging 32 by 32 pixel blocks to make low resolution image (0.32 pixel pairs / degree). For analysis of the gray areas the low resolution image was rescaled to 256 pixels high with bicubic interpolation. The process described here is similar to making a low-spatial-frequency channel by pooling the responses of many pixels. The expansion allows us to compare this sampled image with the original. A black and white dashed line was drawn around the position of the gray areas in the original image.

In all three images the grays on the left are darker because the local sampling of the gray and the area around the grays are darker. The reason that the areas on the left look darker is that they are darker in a very coarse sampling of the image. Models of vision that combine independent multi-resolution versions of the image can predict the appearances of these Balanced Contrast images.

Benary's Cross

In Figure 3 top we see a resized (bilinear expansion) image of a 32x32 block average of Figure 2 top. The average sampling to generate very-coarse spatial information has made an image that is darker in the region of the left triangle than in the region of the right triangle. The average value inside the right triangle is 101 ± 28 while the average value inside the left triange is 168 ± 30. McCann[28] has analyzed all four of Benary's original images. He found that in each case has a lower coarse resolution average lightnesses in the areas that looks darker in the image. All of Benary's experiments can be explained by any model that combines low-resolution sampling information with high-resolution sampling.

White's Effect

Since Michael White's first paper in 1979, there have been a great many studies interpreting the cause of White's effect. Recent papers has provided a very wide range of interpretations of the effect. Spehar et. al. look to High Vision mechanisms.[29] Todorovic[30] and Zaidi et. al[31]. described T junctions reminiscent of Lettvin's triads[32] White's 1981 and Moulden & Kingdom papers[33] demonstrated the need for two different low-level mechanisms for Simultaneous Contrast and White's effect. Blakeslee and McCourt's[34] shows that an oriented difference of gaussian (DOG) model can account for both White's experiments and the White and White 1985 experiments. J. du Buf et al discusses spatial frequency filters.[35]

McCann[16] argues that the gray bars in White's effect are controlled by the values in very-low-spatial frequent channels. Bars look lighter because they have higher values in the very-low-spatial frequency channels.

In Figure 3 middle we see a resized (bilinear expansion) image of a 32x32 block average of Figure 2 middle. The average sampling to generate very coarse spatial information has made an image that is darker in the region of the left rectangles than in the region of the right rectangles. This simple demonstration and the spatial filtering experiments by White, Moulden & Kingdom, Blakeslee & McCourt all provide evidence of the important role of very-low-spatial frequencies in the appearance of equal grays in different surrounds. McCann's analysis of White's effect showed that in each case has a lower coarse resolution average lightnesses in the areas that looks darker in the image.[16]

Rizzi's Effect

In Figure 3 bottom we see a resized (bilinear expansion) image of a 32x32 block average of Figure 2 bottom. The average sampling to generate very coarse spatial information has made an image that is darker in the region of the left square than in the region of the right square.

Similar results are found in experiments that measure the effect on gray appearance with variable enclosure by white.[36]. Six displays have a 1.5 degree gray square, a fixed area of white area on a background of black. The variable is the placement of the white. When white surrounded the gray with an equal width on four sides, observers matched the gray to Munsell value 1.5. When all the white was on one side, lightness matched 3.5. Four other, intermediate placements gave intermediate lightness matches. These experiments along with others that measure the effects of angular separation from whites make the same point. Spatial positions of test patches relative to whites have large effects on lightness. Matches varied from 1.5 to 3.5 Munsell units. That is 28% of the lightness difference between white and black.

Other Experiments

A different "Diamond Wall" experiment described by Adelson in 1993 and measured using the Munsell Scale by Logvinenko. (Logvinenko's Figure 1)[13]. Logvinenko measured the differences in lightness between the gray diamonds with light surrounds and dark surrounds with both corrugated edges and straight edges. Observers reported that the difference in lightness in with straight white-dark boundaries was 1.0 Munsell unit greater than in with corrugated boundaries. Logvinenko argued that perceived illumination is responsible for this effect. The pixel values of the four adjacent areas are the same. Analysis of the images showed different low-spatial frequency content consistent with observer matches. [16]

Further, these experiments and other in the paper demonstrate that "Early-Vision" lightness mechanisms can account for lightness appearance of "Diamond Wall" without the need for feedback from higher-level mental processes.[16]

Conclusions

Full-resolution models that can successfully predict lightness in Simultaneous Contrast images cannot predict Balanced Contrast displays, such as Benary's Cross, White's Effect and Rizzi's Effect. Nevertheless, almost any multi-resolution model can account for these results by sampling the output from each resolution channel independently and then combining in parallel the coarse-resolution information with the high-resolution information.

Clearly pixel based analogs of physics apply to the early stages of vision at the receptors. Hering's idea of opponentcy and Kuffler's observation of spatial opponentcy have become an important concept bin our understanding of visual processes. More complex, bottom-up processes have been demonstrated to exist in human vision and these spatial image processing ideas are all this necessary to explain certain images that were previously attributed to top-down, cognitive processes. In these cases, cognition explanations can be replaced by familiar, image-processing concepts, such as parallel combination of multi-resolution spatial sampling.

Acknowledgments

I want to thank Mary McCann for her thoughtfully discussions and suggestions.

References

1. E. Boring, "History of Experimental Psychology", Appleton Century Croft , 587, (1957).
2. F. A. A. Kingdom, "Old wine in new bottles? Some thoughts on Logvinenko's Lightness induction revisited", *Perception*, **28**, 929-934, (1999);
3. H. von Helmholtz, "Physiological Optics", J.P.C. Southall, Ed., Optical Society of America, Washington, D.C., (1924).
4. E. Hering, "Outline of A Theory of the Light Sense", trans. by L.M. Hurvich and D. H. Jameson, Harvard University Press, Cambridge, MA, (1964).
5. C. Ladd Franklin, "Color and Color Theories", Harcourt, Brace & Co. NewYork, (1929)
6. S. W, Kuffler, "Discharge patterns and functional organization of mammalian retina", *J. Neurophysiol.* **16**, 37-68, (1953).
7. E. H. Land, "Color Vision", in Edwin H. Land's Essays, Vol.III, M.McCann ed., IS&T, Springfield, VA, 1-188, (1993).
8. O. Braddick, F. W. Campbell and J. Atkinson, "Channels in Vision: Basic Aspects", in *Handbook of Sensory Physiology VII,* R. Held, H. W. Leibowitz and H. Teuber, pp. 3-38, (1978).
9. S. Zeki, *A Vision of the Brain*, Blackwell Scientific Publishers, Cambridge, (1993)
10. A. Gilchrist, "Perceived lightness depends on perceived spatial arrangement", *Science*, **195**, 185-187, (1975).
11. A. Gilchrist, "The perception of surface blacks and whites", *Scientific American*, **24,**(3), 88-97, (1979).
    A. Gilchrist, S. Delman, and A. Jacobsen, "The classification and integration of edges as critical to the perception of reflectance and illumination", *Perception and Psychophysics*, **33**(5), 425- 436, (1983).
    A. Gilchrist, "Lightness contrast and failures of lightness constancy: a common explanation.", *Perception and Psychophysics*, **43** (5), 415-424, (1988).
    A. Gilchrist, "Lightness, Brightness and Transparency, Hillsdale:Lawrence Erlbaum Associates, (1994).
    J. Cataliotti, & A. Gilchrist, "Local and global processes in lightness perception" *Perception and Psychophysics*, (2), 125-135, (1995).
    E. H. Adelson and A. P. Pentland, "The perception of shading and reflectance, in Channels" in *Visual Nervous System: Neurophysiology, Psychophysics and Models*, Ed.. Blum, London: Freund, 195-208, (1991).
    E. H. Adelson, "Perceptual Organization and the Judgment of Brightness", *Science*, **262**, 2042-2044 (1993).
    P. Sinha, and E. H. Adelson, "Recovering Reflectance in a World of Painted Polyhedra", *Proc. of Fourth International Conference on Computer Vision*, **4**, 156-163, Berlin; May 11-14 (1993).
    P. Sinha, and E. H. Adelson, "Verifying the Consistency of Shading Patterns and 3-D Structures", *Proc. IEEE Workshop on Qualitative Vision*, 71-81, New York, NY; June 14 (1993).
12. E. H. Adelson, "Lightness Perception and Lightness Illusions" in *The Cognitive Neurosciences*, M. Gazzaniga, ed., Cambridge, MA:3644, 97-104, (1999).
13. A.D. Logvinenko, "Lightness Induction Revisited", *Perception*, **28**, 803-816, (1999).
14. W.D. Ross and L. Pessoa, "Lightnss from contrast: A selective integrative model", *Perception and Psychophysics,* **62**, 1160-1181, 2000.
15. A. Gelb, "Colour Constancy", in *A Scource Book of Gestalt Psychology,* W, Ellis trans. Kegan, Paul, Trench, Truber & Co. Ltd., London, 196-209, (1950).
16. J. J. McCann, "Calculating The Lightness Of Areas In A Single Plane", *J. Electron. Imaging,* **10,** Jan. in press, (2001)
    J. J. McCann, "Calculating The Lightness Of Areas In A Single Plane", B. Rogowitz and T. Pappas, Eds., *SPIE Proc.*, 243-253, (2000);
17. E. H. Land and J. J. McCann, "Lightness and Retinex Theory", *J. Opt. Soc. Am.* **61**, 1, (1971).
18. H. Wallach, "Brightness constancy and the nature of achromatic colors", *J. Exptl. Psychol,* **38**, 310-324, (1948).,
19. T.P. Stockham, "Image Processing in the Context of a Visual Model", *Proc. IEEE*, **60**, 828-284, (1972).
20. J. J. McCann, E. H. Land and S. M. V. Tatnall, , "A Technique for Comparing Human Visual Responses with a Mathematical Model for Lightness", *Am. J. Optometry and Archives of Am. Acad. Optometry*, **47**(11), 845-855, (1970).
21. J. McCann, S. McKee and T. Taylor, "Quantitative Studies in Retinex theory, A comparison between theoretical predictions and observer responses to Color Mondrian experiments", *Vision Res.*, **16,** 445-458, (1976).
22. J. Frankle and J.J. McCann, Method and apparatus of lightness imaging, U.S. Patent 4,384,336, (1983).
23. B. Funt, F. Ciurea and J. J. McCann, " Retinex in Matlab" in *Proc. 8th IS&T/SID Color Imaging Conferenc*e, Scottsdale, Arizona, **8**, 112-121, (2000).
24. W. Benary, "The influence of Form on Brightness", in A Scource Book of Gestalt Psychology, W, Ellis trans. Kegan, Paul, Trench,

Truber & Co. Ltd., London, 104-108,(1950).

25. M. White, "A new effect of pattern lightness", *Perception*, 8, 413- 416, (1979);
M. White, "The effect of the nature of the surround on the perceived lightness", *Perception*, 10, 215-230, (1981).
M. White and T. White, "Counterphase lightness induction", *Vision Res.*, 25, 1331-1335, (1999);
26. C, Blakemore and F.W. Campbell, "On the existence of neurons in the human visual systems selectively sensitive to the orientation and size of retinal images", *J. Physiol.* (Lond.), **203**, 237-260, (1969).
27. Peter J. Burt and Edward H. Adelson, "A Multiresolution Spline with Application to Image Mosaics", *ACM Transactions on Graphics*, **2**, 2-17-236, (1983).
28. J. J. McCann, "Benary's Cross: Belongedness or Low-Spatial Frequency Sampling", B. Rogowitz and T. Pappas, Eds., *SPIE Proc.*, **4299**, in press, (2001);
29. B. Spehar, A, Gilchrist, and L. Arend, "The critical role of relative luminance relations in White's and grating induction" , *Vision Res*., 35, 2603-2614, (1995).
30. D. Todorovic, "Lightness and Junctions", *Perception*, 26, 379-395, (1097).
31. Q. Zaida, B. Spehar, and M. Shy, "Induced effects of backgrounds in three-dimensional configurations: the role of T-junctions", *Perception*, 26, 395-408, (1997).
32. J. Lettvin, "The Colors of Colored Things", MIT Research Laboratory of Electronics Quarterly Progress Reports, **87,** 193-225, (1967).
33. B. Moulden and F. A. A. Kingdom, "White's effect: a dual mechanism", *Vision Res.*, 29, (1989);
F. A. A. Kingdom and B. Moulden, "A multi-channel approach to brightness coding", *Vision Res.*, **32**,1565-1582, (1992);
34. B. Blakeslee and M. E. McCourt, "A muliscale spatial filtering account of White effect, simultaneous brightness contrast and grating induction", *Vision Res.*, 4361-2377, (1999);
B. Blakeslee and M. E. McCourt, "Similar mechanisms underlie simultaneous brightness contrast and grating induction", *Vision Research*, 37, 2849-2869, (1997);
35. J. M. Hans du Buf and S. Fischer, "Modeling brightness perception and syntactical image coding", *Optical Eng.,* 1900-1905, (1995);
36. J. J. McCann and R. L. Savoy, "Measurement of Lightness: Dependence on the Position of a White in the Field of View ", B. Rogowitz, Ed., *SPIE Proc.*, 1453, 402-411, (1991).

Biography

John McCann received his B.A. degree in Biology from Harvard University in 1964. He managed the Vision Research Laboratory at Polaroid from 1961 to 1996. His work concentrated on research in human color vision, large format instant photography and the reproduction of fine art. He is a Fellow of the IS&T. He is a past President of IS&T and the Artists Foundation, Boston. He is currently consulting and continuing his research on lightness and color vision.