Benary's Cross: Belongingness or Coarse Sampling

John J. McCann*
McCann Imaging, Belmont, MA 02478, USA

ABSTRACT

Late 19th century Gestalt psychologists rebelled against the tenants of traditional epistemology and newly defined psychophysics. They introduced powerful new ideas. These ideas were supported by a myriad of experiments showing that the *whole* is not the *sum of the parts*. These appearance experiments include Simultaneous Contrast, Gelb's doorway, and Benary's Cross. The traditional Gestalt explanations of these experiments call for cognitive top-down inferences, that are difficult to describe as detailed computer models.

Late 20th century ideas of image processing refined bottom-up sensation mechanisms to include low-level spatial interactions and multi-spatial frequency image analysis. As did Gestalt theory, spatial image processing ideas depart from single pixel models, such as Colorimetry. Spatial models of lightness have been shown to account for Simultaneous Contrast, Gelb's doorway and Mondrian experiments.

Benary's Cross used two identical gray triangles, surrounded by one black and two white sides, making them have qualitatively equal surrounds. The Gestalt interpretation is that the two gray triangles look different because one appears to belong to the cross and the other does not ("Belongingness"). This paper analyzes Benary's Cross using spatial-image-processing. This quantitative analysis shows that bottom-up coarse sampling is as good an explanation of apparent lightness as top-down cognition. Sampling is much easier to calculate from image data.

Keywords: Benary's Cross, contrast, multi-resolution imaging, low-spatial-frequency vision

1. INTRODUCTION

Gestalt theorists protested against "the 'new' German psychology of the late nineteenth century, the psychology of Wundt, G. E. Muller and Titchner"[1]. One of their principle weapons was the direct comparisons of *identical* stimuli that did not appear the same. Figure 1 (left) illustrates Simultaneous Contrast. The gray area in the middle of the white area has the same reflectance as the gray in the black surround. The gray surrounded by black looks lighter than the one surrounded by white. 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).

Wilhem Benary's classic 1924 paper, "Beobachtungen zu einem Experiment uber Helligkeitskcontrast"[2] is one of many experiments that supported the idea that lightness is influenced by top-down cognitive processes. Benary's Cross (Figure 1 middle) showed that grays still look different with two black sides and one white side. Benary argued that the different lightnesses were due to "surface appurtenance", namely that we recognize and interpreted the image in terms of what we already known. By this idea, 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's paper, only six pages long, credits M. Wertheimer for suggesting the experiment. He showed four different examples of the effect. He also described contrast matching experiments using spinning disks, and size invariance experiments using masks. In recent usage Benary's word "appurtenance" has been modernized to "belongingness". The hypothesis is the same, namely that these figure are evidence for top-down models of lightness.

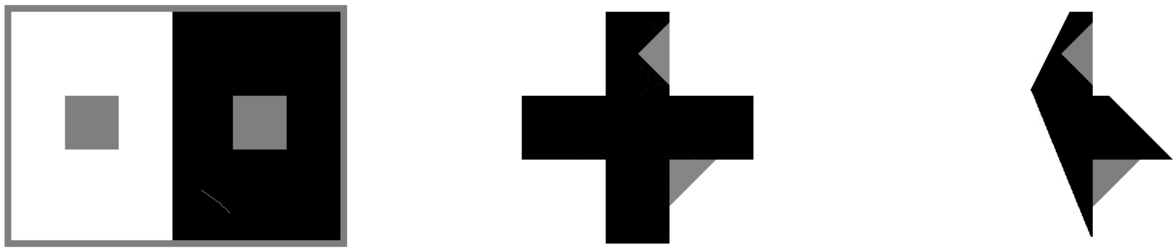

Figure 1 shows three contrast experiments. All gray areas have the same reflectance. Differences in appearance (lightness) are due to spatial processing in the visual system. The left figure is Simultaneous Contrast. The middle figure is Benary's Cross. Note that the grays have the same surrounds, at least on the qualitative basis of identical white and black adjacent borders. The right figure modifies Benary's Cross into an abstract figure lacking the associative qualities of the black cross. Note that the lightness differences are still present.

Figure 1 (right) shows an asymmetric version of the cross; parts of the black cross have been removed. This abstract black shape is much less recognizable than the cross, yet the lightness discrepancy is unchanged. This figure can be used to question the necessity for recognition and implies that other mechanisms might be influencing lightnesses in Benary's Cross.

2. SURROUND VS. RECEPTIVE FIELD

Since Leonardo, artists have known of the importance of surrounding colors on appearance. The definition of surround is sometimes ambiguous. Some concepts build on the presence/absence of a surround stimulus. We can think of this as a qualitative description of the spatial image. Grays with the same adjacent areas, for example, one white and two black edges, have the same surrounds.

Other concepts build on the numerical spatial integrals of quanta catch of all retinal responding to a surround stimulus. We can think of this as a quantitative analysis of the image. Such analysis of the spatial properties of vision began with dark adaptation experiments.[3] The periphery uses larger pools of receptors at threshold than the fovea and the perifovea.[4] In 1953 neurophysiological experiments by Barlow[5] and Kuffler[6] measured the spatial responses of ganglion cells of frog and cat. They showed that small spots of light in the center of the receptive field excited the cell, but that light in the surrounding area inhibited it. Further, they showed that excitation and inhibition antagonistically summed. Uniform light over the entire receptive field caused only weak responses. These finding stimulated a wide range of neurophysiological experiments recording the cellular responses to complex spatial stimuli on the retina. Hubel and Wiesel extended the complexity of receptive fields by reporting the characteristics of simple, complex and hypercomplex cells in visual cortex[7]. Campbell and colleagues introduced grating stimuli and spatial frequency analysis into vision research.[8]

Recent experiments studying White's Effect[9], and Adelson's and Logvinenko's 'Diamond Walls"[10] have suggested that coarse image sampling can explain their effects.[11] This paper analyzes the Benary's Cross image in terms of its coarse-resolution or low spatial frequency properties.

3. SPATIAL AVERAGES

This sections provides a quantitative analysis of Benary's cross using circular receptive field pools of pixels. The pool is defined simply as the average value of all pixels in the circular sample. A series of different pool sizes were chosen to characterize both the top and bottom triangles in Figure 2. The pixel values in the input image were calibrated by making a printed sample and measuring the characteristic curve of reflectance vs. digital value. Average luminance was calculated by summing the relative reflectance of each pixel in the pool and dividing by the number of pixels.

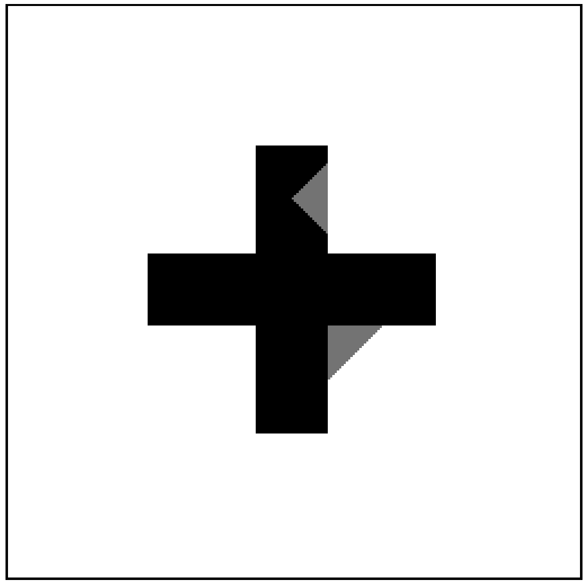

Figure 2 show a 1024 by 1024 pixel image of Benary's Cross. It will be used to calculate the average luminance of different size receptive field pools. All pools are centered on the middle of the gray triangles.

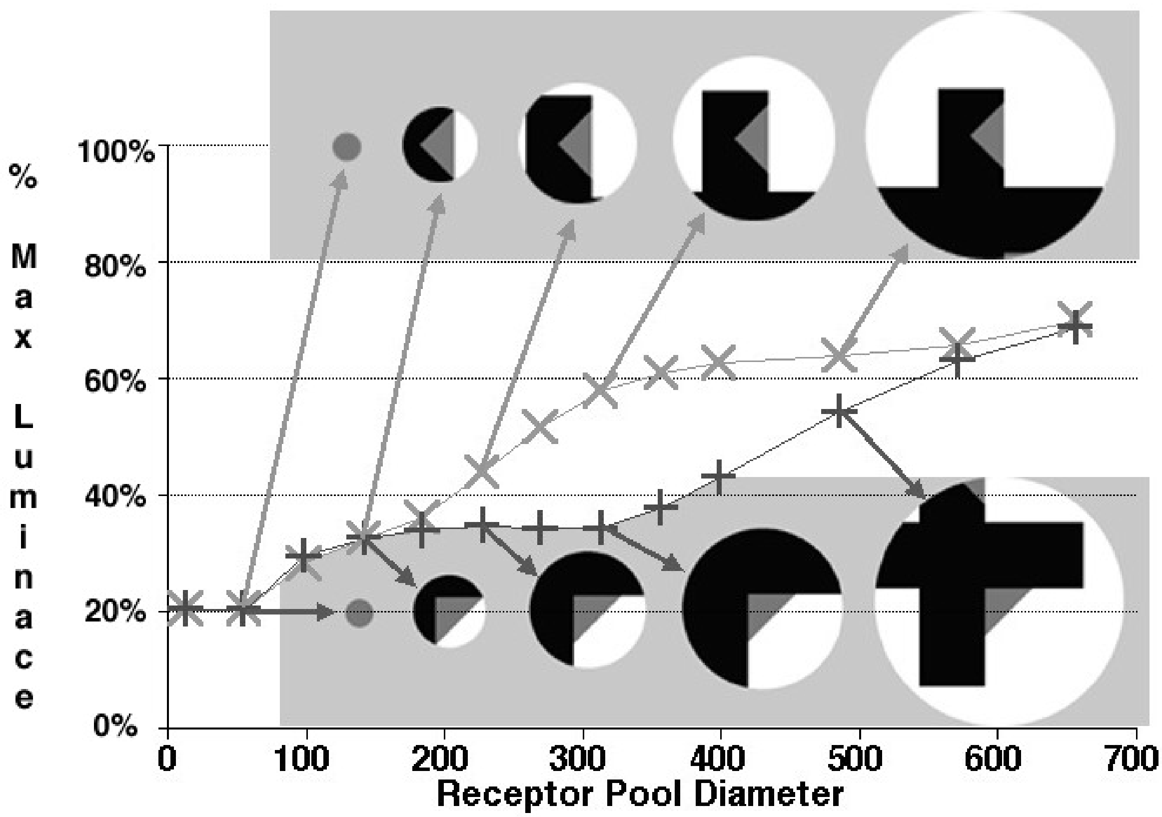

Figure 3 shows a plot of average luminance, relative to maximum luminance, vs. receptive pool diameter in pixels. With a diameter of 12 pixels the top-triangle pool and the bottom-triangle pool fall entirely in their respective gray triangles. At 141 pixel diameter the pool circumference intersects both base vertices of the gray triangles. The images in these receptive fields are identical and the average is 33%. With larger pools the average value of the top is greater than that for the bottom. The difference is largest for 313 pixel diameter. Above that size the pool values converge. The quantitative pool average values around Benary's triangles are not equal.

Figure 3 shows the quantitative measurements of the different surrounds around the gray triangles. With a diameter of 12 pixels the top-triangle pool and the bottom-triangle pool fall entirely in their respective gray triangles.The 12 pixel diameter average luminances are equal to 21% maximum luminance. As well, the 55 pixel diameter pool averages only gray pixels and is 21%. At 98 pixel diameter the pool average includes white and black pixels along with the gray pixels. The equal average rises to 29%. At 141 pixel diameter the pool circumference intersects both base vertices of the gray triangles. The images in these receptive fields are still identical and the average is 33.

Above 141 pixel diameter the pool averages diverge. The insert at 227 pixel diameter clearly illustrates the reason. The bottom triangle has 270 degrees of black pixels while the top has roughly 180 degrees. The pool value is 44% for the top and 35% for the bottom triangle. By the pixel diameter of 313 pixels the poll values are 58% for the top and 34% for the bottom. In other words, the bottom triangle has a difference in average luminance of 24% (58% - 34%) and a ratio of 1.7 (58 / 34) times darker.

4. MULTI-RESOLUTION IMAGES

Quantitative evaluations of human vision go back to Maxwell's color matching functions.[12.] Pixel based Colorimetry was well established the time that computer hardware allowed real time spatial processing of real images.[13] Changing the strategy from

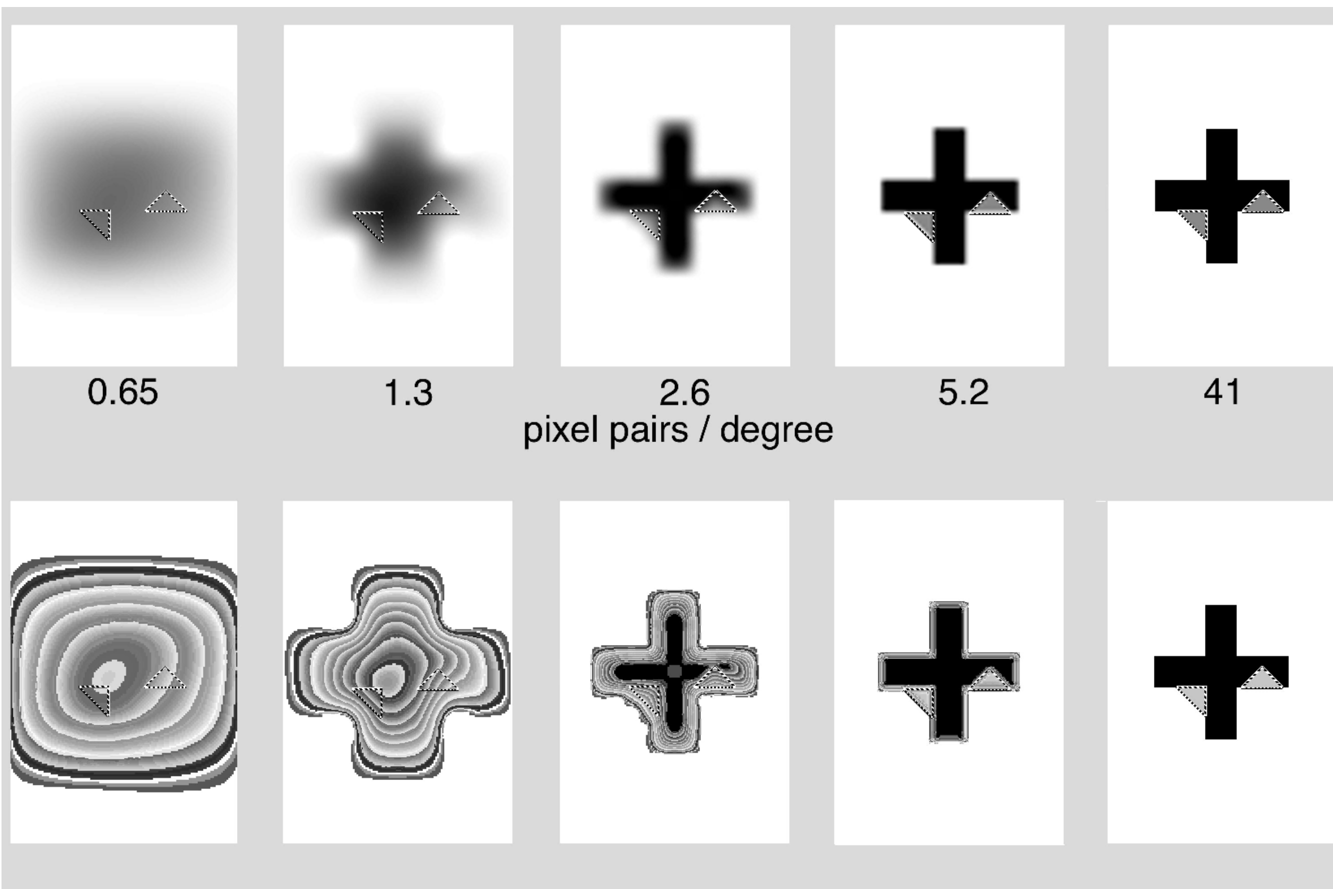

Figure 4 shows a 1024 by 1535 pixel image of Benary's Cross (upper right). The four images to the left are made by averaging the original image such that 1 pixel is the average of different size pixel arrays (left to right- 64 by 64, 32 by 32, 16 by 16, 8 by 8). For comparison each subsampled image is rescaled to the same print size using bicubic expansion. The subsample images have 0.65, 1.3, 2.5, 5.1 and 41 pixel pairs per degree (See text). The dashed lines indicate the area of the triangles in the full resolution image. The bottom half of the image shows grayscale indexed LUT renditions of the images in the top half. These images show that the left (outside) gray triangle have lower digital values than the right (inside) triangle in the 0.65 and 1.3 pixel pairs per degree images . The difference in appearance correlates with the difference in coarse sampling.

Figure 5 shows Benary's Black Triangle vs. Cross experiment. The Black Triangle (left) was made by removing the top and parts of the other three arms of the Cross (right). Benary reports that the addition of white in the surrounding area of the Black Triangle made its gray triangle look lighter. He did the experiment to test between his appurtenance hypothesis and simultaneous contrast. He argued in favor of appurtenance because more white in the vicinity should strengthen simultaneous contrast and make the gray darker. He did not consider that coarse sampling could explain why adding more white made the gray lighter.

evaluating the response of one pixel to include the comparison of all pixels greatly increased the computational burden. New computational techniques were needed to efficiently compare each pixel with all other pixels.[14] One such process generates a series of different resolution images from the original.[15] Each smaller image is one quarter the size of the previous one. It is made by averaging the values of four pixels to make one pixel. The original image can be thought of as the result of an array of single receptors. The next smaller image is the result of an array of 2 by 2 pixel square receptive field pools. The next is 4 by 4, 8 by 8, 16 by 16 and so forth.

Figure 4 shows five of these coarse-resolution images of Benary's Cross. The original display is shown on the top right. The image is 1536 by 1024 pixels. To its left is an 8 by 8 pixel pooled image, the middle is 16 by 16, the left center is 32 by 32 and the far left is 64 by 64 pixel pool. All these subsampled coarse-resolution images have been expanded to the same size on the page by bicubic interpolation.

Many vision scientists think of images in terms of their spatial frequency spectrums. They can compare the image to the contrast response function that plots sensitivity vs. cycles per degree. By way of approximate calibration we can identify each image in Figure 4 as being composed of a specific number of pixel pairs per degree. The original image is 1024 pixels over 3 inches. Viewed at 14 inches it subtends 12.4 degrees and has 41 pixel pairs per degree. The upper limit of the contrast sensitivity function (CSF), or the highest spatial frequency detectable is around 60 cycles per degree. The peak of the CSF is around 3 cycles per degree[8] and the low-frequency cut off depends on the size of the image.[16] Starting from left to right the first image has 0.65 pixel pairs and is near the low- frequency limit of vision; the next image has 1.3 pixel pairs and part way up the CSF; the next image has 1.3 pixel pairs and is half way up the CSF; the next two image has 2.6 and 5.2 pixel pairs and near the top of the CSF.

The effect of the sub-sampling shows the quantitative differences seen in Figure 3. The dashed lines indicate the area of the gray triangles. It is difficult, if not inadvisable, to try to judge the average digital value of the subsampled image by visual inspection. A way to avoid the difficulties of evaluating gradient is to use pseudocolor or gray scale index look-up tables (LUT). The bottom half of Figure 4 shows the image in the top half after an index lookup table. As best seen on the left this LUT cycles from white to black every 32 digital image values from 0 to 255. The image looses it lightness values but presents an easy to interpret display of image intensity. The left (outside) triangle has significantly lower digital values than the (inner) one in both the 0.65 and 1.3 pixel pairs per degree images. At higher resolutions the differences are small.

4. DISCUSSION

Benary's explanation of appurtenance would be difficult to model in a modern computer. All of the steps are possible. The model could perform image segmentation to isolate figure from background, perform object recognition to identify the cross and the gray triangles, calculate the property of "insideness" and "outsideness", calculate the imposition of simultaneous contrast rules to modify input image. A particular example could be implemented. The problem becomes more complex if the demand of size invariance is also required. It is still possible. This model's problem is that has many steps. By comparison parallel comparison of low-frequency channels requires much less computation.

Benary's second experiment[2] is even more interesting than the first. He masked off parts of the Cross to make a Black Triangle. He removed all of the top and parts of the other three arms (Fig 5). He compared the gray in the Black Triangle with the gray in the Cross without masks. He reported that the gray in the Black Triangle looked lighter. He argued that it should have looked darker, because there was more white in the vicinity, and hence more simultaneous contrast. Since it looked lighter he concluded that it was caused by appurtenance. This paper was written in 1924, the year that Fergus Campbell was born. Benary did not consider low-spatial frequency channels that would report the effect of more white in the region was to make grays look lighter.

Late 20th century ideas of image processing refined bottom-up sensation mechanisms to include low-level spatial interactions and multi-spatial frequency image analysis. As did Gestalt theory, spatial image processing ideas depart from single pixel models, such as Colorimetry. Spatial models of lightness have been shown to account for Simultaneous Contrast, Gelb's doorway, the Black and White and Color Mondrians and real life scenes.[16]. All these experiments as well as Logvinenko's "Diamond Wall" gradients and Adelson's "Checker Shadow" and "Snake" have been modeled using the same Ratio-Product-Reset-Average model.[11] It 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 found in a recent review paper.[17]) The image processing models can account for the observed lightness in all the above flat images without depth or apparent depth information. Such image processing models can be used as an argument that lightness is a complex bottom-up process.

Balanced Contrast Effects[18] such as Benary's Cross, White's Effect and Rizzi's Effect do not follow the simple rules of simultaneous contrast. Their grays do not look lighter adjacent to black. These images have symmetrical local contrasts; both white and black areas are adjacent to the gray test patch. One can explain all of these effects by combining in parallel the assessment of each spatial-frequency channel or coarse-resolution sampling. This has been described for Whites effect.[19] Similar results and conclusions are found in the analysis of the of Adelson's "Straight Edge Effect".[11]

5. CONCLUSIONS

The gray triangles in Benary's Cross have the same qualitative surround. They have one white and two black boundaries. As with other Balanced Contrast examples, it fails to follow the simple hypothesis that grays look lighter when adjacent to black. Quantitative analysis of receptor pooling and coarse-sampling of Benary's Cross produces subimages in which the area corresponding to the gray at full resolution has higher luminance. The difference in appearance correlates with the difference in intensity in coarse-resolution sampling. Any multi-channel model can account for Benary's Cross by parallel combination of coarse-resolution or low-spatial frequency information. All such models would not require the calculation of image segmentation, object recognition, object relations, imposition of simultaneous contrast principle to modify input image. The gray triangle outside the cross looks darker because the average luminance in the vicinity is lower.

6. ACKNOWLEDGMENTS

The author wants to thank Mary McCann for very helpful discussions.

7. REFERENCES

*mccanns@tiac.net
1. E. Boring, "History of Experimental Psychology", Appleton Century Croft , 587, (1957).
2. W. Benary, "The influence of Form on Brightness", in A Source Book of Gestalt Psychology, W, Ellis trans. Kegan, Paul, Trench, Truber & Co. Ltd., London, 104-108, (1950).

[3.] M. H. Pirenne, “Dark Adaptation and Night Vision”, in The Eye Vol. 2, H. Davson, ed., 93-122, (1962).
[4.] S. Hecht, C. Haig and G. Wald, “The dark adaptation of retinal fields of different size and location”, J. gen. Physiol. **19**, 321-327, (1935).
[5.] H. Barlow, “Summation and inhibition in the frog’s retina”, J. Physiol. **119**, 69-88, (1953)
[6.] S. W, Kuffler, “Discharge patterns and functional organization of mammalian retina”, J. Neurophysiol. **16**, 37-68, (1953).
[7.] D. Hubel and T. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat’s striate cortex”, J. Physiol. **148**, 574-591, (1959);
D. Hubel and T. Wiesel, Receptive fields of single neurons in the cat’s striate cortex”, J. Physiol. **160**, 106-154, (1962);
D. Hubel and T. Wiesel, Receptive fields and functional architecture of monkey striate cortex”, J. Physiol. **195**, 215-243, (1968).
[8] F. W. Campbell and J. G. Robson, “The application of Fourier analysis to the visibility of gratings”, J. Physiol. (Lond.), **197**, 551-556, (1968);
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, 3-38, (1978);
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).
[9.] 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).
[10.] A. D. Logvinenko, “Lightness Induction Revisited”, Perception, **28**, 803-816, (1999).
[11.] 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”, in Human Vision and Electronic Imaging V, B. Rogowitz and T. Pappas, eds., SPIE Proc., **3959**, 243-253, (2000).
[12..] J. C. Maxwell, “VII. Experiments on colour, as perceived by the Eye”, with remarks on Colour-Blindness”, Vision, in The Scientific papers of James Cleck Maxwell, W. D. Niven ed., Dover, NewYork, Vol.1, 126-154, (1965).
[13.] J. J. McCann and K. L Houston, Color Sensation, Color Perception and Mathematical Models of Color Vision, in: Colour Vision , J. D. Mollon, and L. T. Sharpe, ed., Academic Press, London, 535-544, (1983);
J. J. McCann and K. L. Houston, Calculating color sensations from arrays of physical stimuli, IEEE Transactions on Systems, Man, and Cybernetics, VOL. SMC 13, 1000-1007, (1983).
[14.] J. Frankle and J. McCann, “Method and apparatus of lightness imaging”, U.S. Patent 4,384,336, (1983).
[15.] J. J. McCann, “Lessons Learned from Mondrians Applied to Real Images and Color Gamuts”, Proc.IS&T/SID Seventh Color Imaging Conference, **7**, 1-8, (1999);
J. J. McCann, “Color Theory and Color Imaging Systems: Past, Present and Future”, J. Imaging. Sci. and Technol., **42**, 70, (1998).
[16] J. J. McCann, R. L. Savoy, J. Hall, Jr., and J. J. Scarpetti, “Visibility of Continuous Luminance Gradients”, Vision Res., **18**, 891-894 (1974);
R. L. Savoy and J. J. McCann, Visibility of low spatial-frequency sine-wave targets: dependence on number of cycles, J. opt. Soc. Am., **65**, No. 3 , 343-350, (1975).
J. J. McCann, R. L. Savoy and J. A. Hall, Jr. Visibility of Low Frequency Sine-Wave Targets, Dependence on Number of Cycles and Surround Parameters, Vision Research, **18**, No. 7, 891-894, (1978);
J. J. McCann, “Visibility of gradients and low-spatial frequency sinusoids: Evidence for a distance constancy mechanism”, J. Photogr. Sci. Eng., **22**, 64-68, (1978).
J. J. McCann and J. A. Hall, Jr. Effects of average-luminance surrounds on the visibility of sine-wave gratings, J. opt. Soc. Am., **70**, No. 2, . 212-219, (1980).
[17] B. Funt, F. Ciurea and J. McCann, “ Retinex in Matlab” in Proc. 8th IS&T/SID Color Imaging Conference, Scottsdale, Arizona, in 112-121, (2000).
[18] J. J. McCann, “Gestalt Vision Experiments from an Image Processing Perspective”, Proc. IS&T PICS, **4**, in press, (2001).
[19.] J. J. McCann, “Spatial Contrast and Scatter: Opposing Partners in Sensations”, in Human Vision and Electronic Imaging IV, B. Rogowitz and T. Pappas, Eds., SPIE Proc., Vol. 3644, pp. 97-104, (1999).