Measuring Constancy of Contrast Targets in Different Luminances in Complex Scenes

John J. McCann, McCann Imaging, Belmont, MA/USA

Abstract

Objects in complex images appear almost constant. This is true with changes of the overall level of illumination, and somewhat true with changes of the surround around a test area. The small departures from perfect constancy, however, provide important evidence on the underlying mechanisms of constancy. First, this paper measures the departures from constancy with changes in overall luminance. Second, it measures the effects of contrast using white, gray, and black surrounds. Third, it compares the results from flat-2D transparent displays with those using 3D shapes. With 3-D paper targets, illuminated in direct light and shadow, we find the same small decrease in matching value with large decreases in illumination level (low-slope behavior) found in flat-2D transparent displays. Within the direct light and in the shade parts of the targets, the matches showed the same high-slope contrast behavior. Here, the arrangement of reflectances, illumination and depth did not affect the appearance matches made by observers. In both flat-2D transparent and complex 3-D data, observers' matches fit the simple two-step physical description. The local maxima are dependent on luminance, and other, darker areas, are dependent on spatial contrast.

Introduction

One of the first reported psychophysical experiments was the classification of stellar magnitude by Hipparchus of Nicea in the 2nd century B.C. Although the original manuscripts have been lost, the results were documented by Ptolomy. After many centuries, stellar magnitude is in common use today. It has been modified to be a photometric measurement starting with Pogson in 1856. The stellar magnitude changes by 100:1 when the measured luminance changes by 100,000:1. In other words, stellar magnitudes have a slope of 0.4 in a plot of log luminance vs. stellar magnitude.[1]

Flat-2D transparent Displays

In our *flat-2D-transparent* study, observers matched the appearance of eight white, gray, and black patches using different illuminances. The experiment asked seven observers to match uniform luminance patches to a standard display. This experiment included 3,150 observations. Both test and standard displays were transparent photographic films viewed on two high-luminance Aristo lightboxes. The targets subtended 25 by 30 degrees. The two light boxes were mounted at right angles so that the observers turned their heads back and forth to make matches. The left eye was used for observing the contrast targets, and the right eye for observing the standard.

The standard was a series of 9 patches each with a different luminance transmission. The nine patches were surrounded by white (1000 ft-L). The patches were selected to have equal differences in appearance between white [9.0] and black [1.0]. In the experiments described below, the observers were asked to report on the mixture of colors on a palette that would match what they saw.[2] Observers were asked to interpolate between the 9 reference values seen in the standard display.

The dynamic range of the test targets covered 2.3 log units; the dynamic range of the illumination covered 3.3 log units. The resulting dynamic range of the experiment covered 5.6 log units. One advantage of this apparatus is the luminance uniformity across the display and within each test patch. The luminances reported in this paper were measured with a Gamma Scientific telephotometer. Another advantage of using photographic transparencies is that we can be certain that the contrast displays are exactly constant at each illumination level. The ranges of the transparent contrast targets are larger because they are not limited by paper surface reflectances. The constancy of image content is far superior to what is possible in simulations on monitors and LCD displays.

This study investigates the role of luminance and contrast in complex images. It uses eight different transmission patches (2.7° by 5.5°) test patches. in a white, a gray and a black surrounds. These patches are viewed at 5 different luminances and matched to a constant standard. The goal of the study is to understand the limits of constancy with overall changes in luminance. The experimental data will be compared to various hypotheses such as "discounting the illuminant", simultaneous contrast, and the ancient low-slope change of appearance with luminance found in stellar magnitude. A further goal is to expand the range of luminances matched in related studies.[3,4,5]

Figure 1 plots the observer matches for 8 test areas in a white surround. The lines connect the average match for 7 observers for the same area in 5 different illuminations

We used a least squares regression linear fit to the data from each line in a white surround to calculate the linear slopes of the eight patches. The average slope for a white surround is 0.500 ± 0.108. When analyzing the data for the same luminance test areas in a gray surround the average linear fit for the eight patches in gray surround was 0.503 ± 0.108. The average linear fit for the black surround was 0.612 ± 0.066. These average slopes are plotted in Figure 2. The average linear fit for all lines in all surrounds is 0.539 ± 0.103.
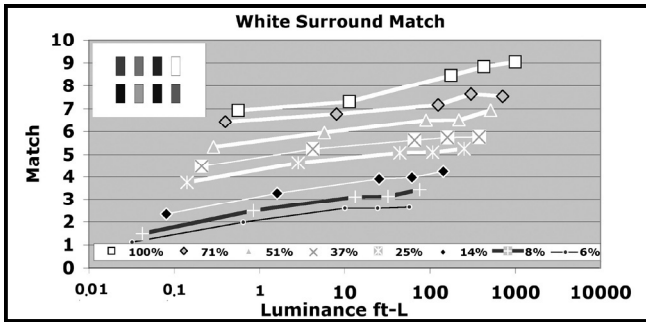
Figure 1 plots the match for 8 test areas in a white surround. Here, the different lines plot the matches for each patch in the contrast display. The eight lines for each target area are parallel lines with a very low slope
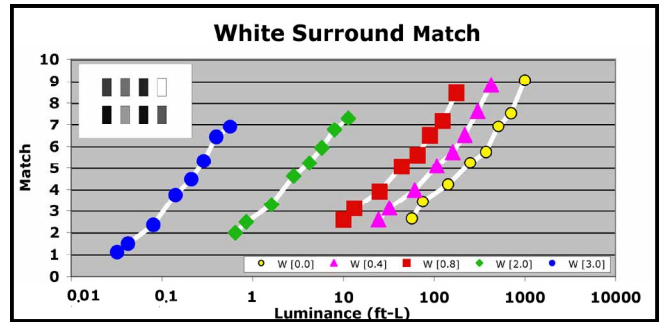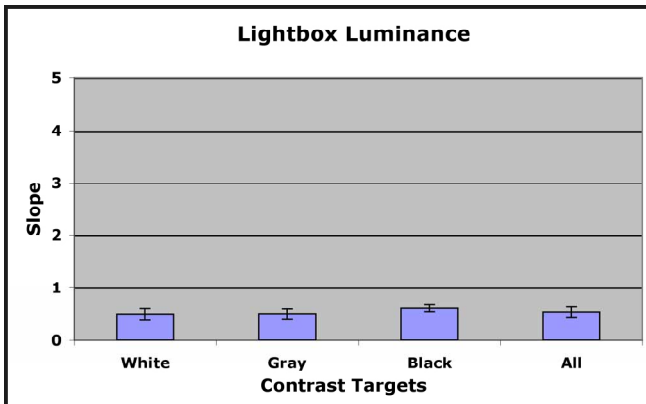

Figure 2 plots the average slope for each of 8 patches in 5 illuminants. Here we see the same behavior for white, gray, and black surrounds.

The fact that matches in different surrounds exhibited the same low-slope rate of change with luminance give an important insight into the effect of overall illumination. Decreasing illumination departs from perfect constancy, but at a slow rate. That rate is the same regardless of the effects of spatial contrast. That rate is the same for all test patches in all surrounds. Changes in overall luminance adjust all parts of the image so that the maxima in the scene fall on the 0.54 slope "Hipparchus line".[6]

Figure 3 replots the same data used in Figure 1. Here the lines connect the average match for 8 test areas in each illumination. These are the contrast lines. The contrast plots of observer data show 5 parallel lines, one for each illumination level. We fit the data from each line in Figure 3 with linear least square regression to calculate the contrast slopes. The values were: 4.70, 4.75, 4.46, 4.34, 4.82. The average slope for a white surround is 4.615 ± 0.206.

Figure 3 plots the match for 8 test areas with a white surround. The lines show the matches for the 5 levels of lightbox luminance. The five contrast plots are parallel.

We repeated the analysis with the gray surround. The contrast slopes for the gray surround in were: 3.82, 3.77, 3.77, 3.83, 4.03 (Average = 3.846 ± 0.107). The contrast slopes for the black surround were: 2.59, 2.66, 2.32, 2.45, 2.83 (Average = 2.569 ± 0.195). These average slopes are plotted in Figure 4.
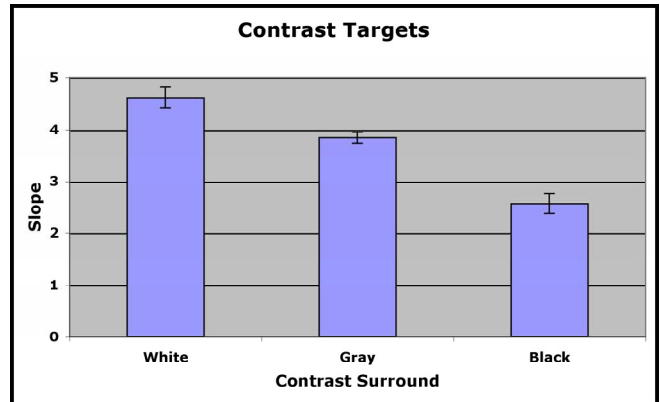

Figure 4 plots the average slopes and standard deviations for white, gray and black surrounds. The effect of the surround produces different rates of change of match with luminance.

The 3,150 matches over a luminance range of nearly six log units showed that all data fit two simple physical rules. First, the match for the highest luminance in the field of view decreased at a slope of 0.54 with luminance ("Hipparchus line"). Second, contrast matches for darker areas start at the "Hipparchus line". Contrast lines have higher, background dependent, slopes (4.6 for white, 3.8 for gray and 2.6 for black) (See Figure 5).
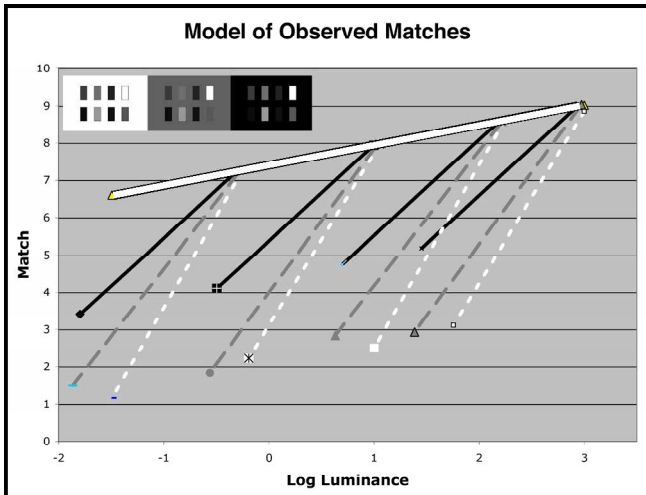
Figure 5 *summarizes observer matches. It is the physical description for matches in overall changes and contrast changes in luminance. The upper (solid white) low-slope "Hipparchus line" has slope 0.54 (triangles). All luminance maxima in the field of view fall on this line. All areas with less luminance fall on background-dependent contrast lines. The (dotted) white lines have slope 4.6; the (dashed) gray lines have slope 3.8; and the (solid) black lines have slope 2.6.*

Figures 6, 7, and 8 show quite good fit to the simple physical description of luminance and contrast. Unlike film and electronic sensors, human visual responses are scene dependent. There is no dependence on the quanta catch of the receptors. Observers do not select the same luminance in the standard as a match for the test area (Solid white lines). If humans "discounted the illumination", then the match for each test area should be constant. The data shows that the maximum in each target falls on the Hipparchus low-slope line. In all experiments with lowered illumination the matches are significantly lower than those predicted by "discounting the illumination" constancy. Matches to maxima are dependent on absolute luminance and darker areas are dependent on spatial contrast interactions.
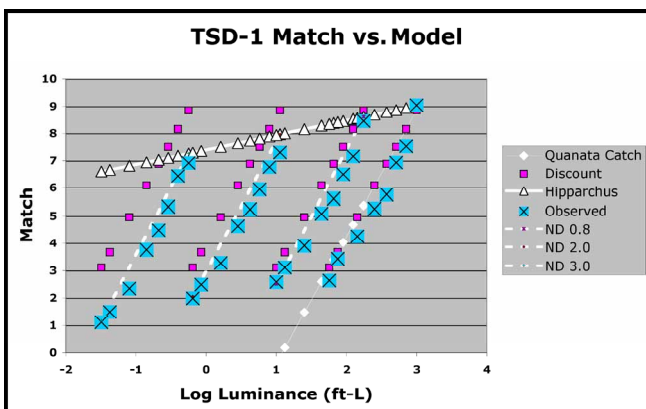

Figure 6 *plots the white-surround physical description vs. observer data. The slope 4.6 predictions are the dotted white lines. The Xs plot the observer data for white surround in four illumination levels. The fit is quite good. The triangles plot the low-slope Hipparchus line. The squares plot the expected matches for perfect constancy. The solid white line plots the predictions for a luminance match (Quanta Catch) to the standard display.*
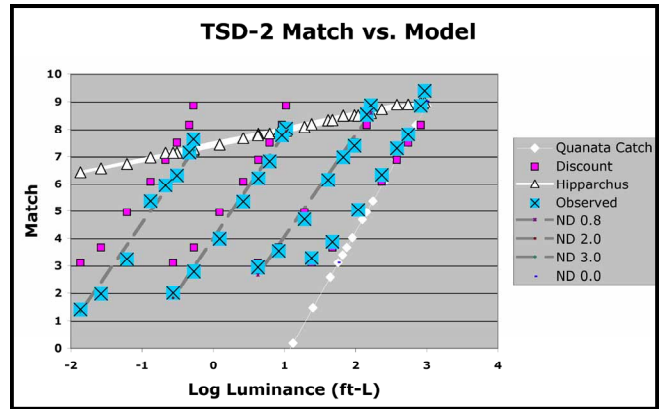

Figure 7 *plots the gray-surround physical description vs. observer data. The slope 3.8 predictions are the dashed gray lines.. The Xs plot the observer data for gray surround in four illumination levels. The fit is quite good. The triangles plot the low-slope Hipparchus line. The squares plot the expected matches for perfect constancy. The perfect constancy predictions are higher than actual matches. The solid white line plots the expected match for a luminance match (Quanta Catch) to the standard display.*
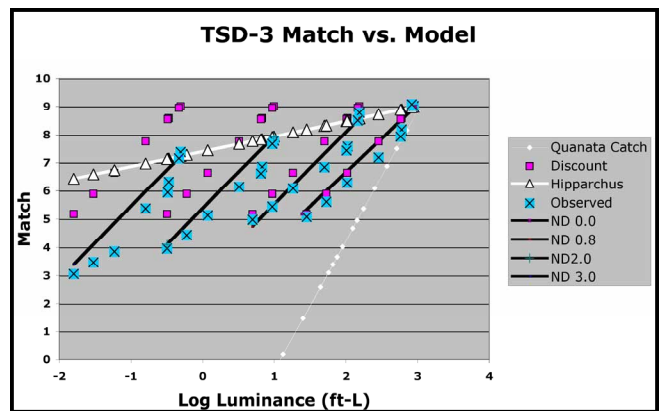

Figure 8 *plots the black-surround physical description vs. observer data. The slope 3.8 predictions are the solid black lines. The Xs plot the observer data in four illumination levels. The fit is quite good. The triangles plot the low-slope Hipparchus line.*

A model that predicts matches must use spatial comparisons to be responsive to the surround. All white surround matches decrease from the maximum at a slope of 4.6 from the maximum (Figure 6). All gray surround matches decrease from the maximum at a slope of 3.8 (Figure 7). All black surround matches decrease from the maximum at a slope of 2.6 (Figure 8).

These experiments require that all edge ratios in the entire scene are perfectly constant for all illumination levels. These experiments used transparent components to achieve very reliable luminance displays with an extraordinary dynamic range. It is very difficult to achieve the same control of dark luminances with papers because of surface reflections. It is even more difficult to manage display devices over such large ranges.

`An important result of these experiments is the measurement of the value of the Hipparchus line slope for complex image matching. The second experiment in this set is a study of matches in *complex 3D* scenes. Here identical reflectances in bright light and shadow are matched to a standard. The "Hipparchus line" gives us a quantitative estimate of the effect of the shadow as local maxima. Further, this simple physical description lets us examine issues of perceived illumination and other suggested higher-level mechanisms with a new quantitative base. As shown above, matches do not correlate with constancy implied by discounting the illumination. Matches do not correlate with luminance. We can use the simple two-step physical description of maxima and spatial contrast to evaluate complex scenes using direct and shadow illumination.

Complex 3-D Scenes

The 3D experiments study the influence of uniform direct illumination and shade in a real 3D scene. The experiments use printed-paper targets that are folded. The images on each side of the fold are the same. One side was in direct illumination and the other in shade. A separate card has standard reference reflectances printed on the direct illumination side (Figure 9). Observers matched these patches in both direct and shade illumination. This data can be used to assess whether real 3D illuminations of scenes influence the appearance of matches. If humans "discount the illumination", then these matches will differ from those reported in *flat-2D transparent* experiments. If the observers behave the same as with 2D, then we will fail to find evidence for "discounting the illumination" in complex scenes.

In trying to tease apart the mechanism of appearance it is essential to control the question presented to observers. McCann[2] used a photograph of a float to illustrate that observers give different answers to different questions about the same stimulus. When asked to recognize the paint on the sides of the raft, observers say it is white, even though one side is in sunlight and the other is in shade (perception). When observers are asked to hypothetically mix paints to render the appearance in a painting of the scene, they select a whiter, more yellow match for the sunny side, and a darker, more blue match for the shady side (sensation). Arend and Goldstein[7] documented this observation in experiments using display Mondrians. Hurlbert also makes this distinction in the analysis of Mach Card experiments.[8] In the experiments described below the observers were asked to report on the mixture of colors on a painters palette that would match what they saw. They were instructed not to try to guess the reflectance of the patch they were matching.

Experiments

A 4.4 by 2.8 inch folded card was placed on a table on a black cloth. The light falling on the paper came from 2 two-foot fluorescent lamps (distance = 18 inches). Care was taken to make the illumination of each side of the card uniform. The side directly illuminated had a 99.3 ft-L luminance from the white part of the card. The other side had 4.86 ft-L from white. The shade was 4.9% of the bright side. The standard card was slightly wider than the test targets. The 13 standard patches covering the range of 9.0 to 3.0 were calibrated to have the same % reflectance as the % transmission in the standard in the *flat-2D transparent* study. The

reflectance standard had lower luminance and lower dynamic range than the transmission standard. The luminance values (ft-L) of the standard were: 99.7 = [9.0], 92.5 = [8.5], 81.6 = [8.0], 69.1 = [7.5], 54.7 = [7.0], 44.6 = [6.5], 34.8 = [6.0], 24.6 = [5.5], 14.8 = [5.0], 10.0= [4.5], 8.36 = [4.0], 4.93 = [3.5], 4.47 = [3.0].
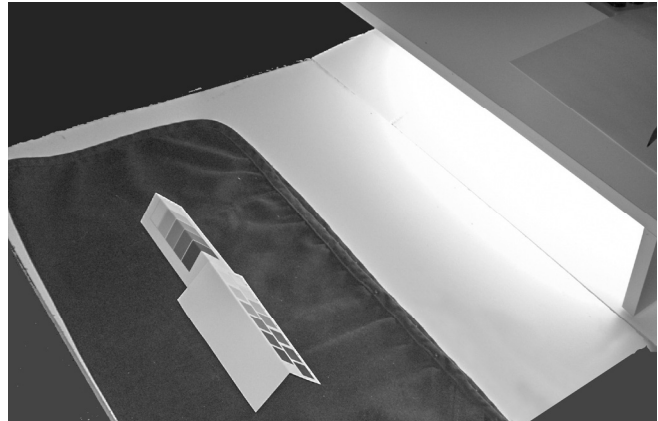
Figure 9 shows the 3-D display. Two fluorescent lamps were mounted in a white box placed on a white table. The standard set of gray patches and folded test card were placed on a black cloth. The distance between the cards and the lamps controlled the direct illumination on the cards. To the left of the photograph (out of the image) was a large white reflector. The distances between the lights, the cards and the reflectors were adjusted to control the level of the shade illumination. All cards were measured with a telephotometer to insure uniform direct and shade illuminations.

Measurements of the luminances from the card showed that the illumination was uniform over the entire card. The black cloth under the paper plays an important role in uniform illumination. Without a black cloth, it is very difficult to control unwanted light reflected from the table surface, leading to higher luminances along the bottom of the card. A large white surface (2 by 2 feet) was parallel to the lamps on the far side of the folded paper target. It acted as a reflector to control the uniformity and illuminance falling on the paper on the darker side of the fold. A black card (3 by 3 feet) was placed perpendicular to the lamps to reduce reflected light onto the card from that direction. The observer viewed the folded card and the matching palette from a distance of 21 inches using both eyes. Binocular vision prevented perceived reversal of depth for the cards.

The experiments used three different folded cards shown in Figure 10. Observers were asked to match test areas in the Rectangles, Circular Spots and Mondrian targets. In addition, they were asked to make additional matches within areas of uniform reflectance of the Rectangles target.
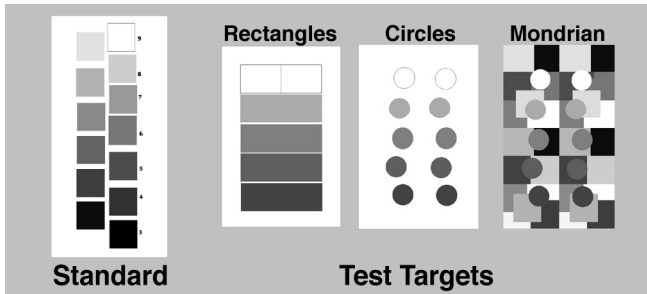
Figure 10 shows the Standard, the three test targets. The test targets were folded along the vertical center line. Matches for the five test areas were measured using the Standard as a reference.
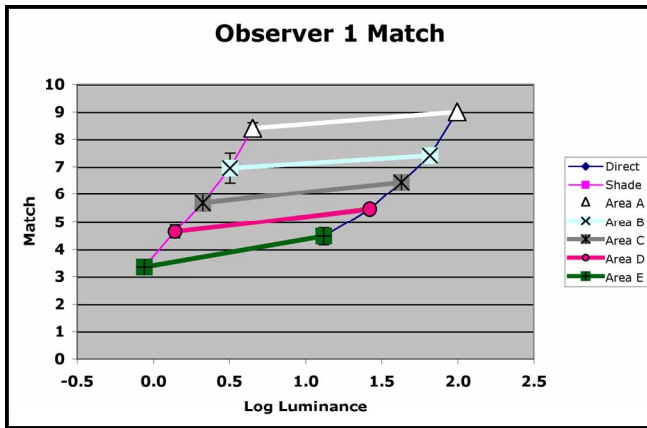

Figure 11a plots the average of four trials ± the standard deviation for the rectangles target for observer MAM. The legend identifies plots match vs. log luminance for the gray scale in direct light (slope =5.0) and in shade (slope =6.9). In addition, change in match for each area is plotted. Areas A through E show matches for white through dark gray. The average slope is 0.59 ± 0.23.

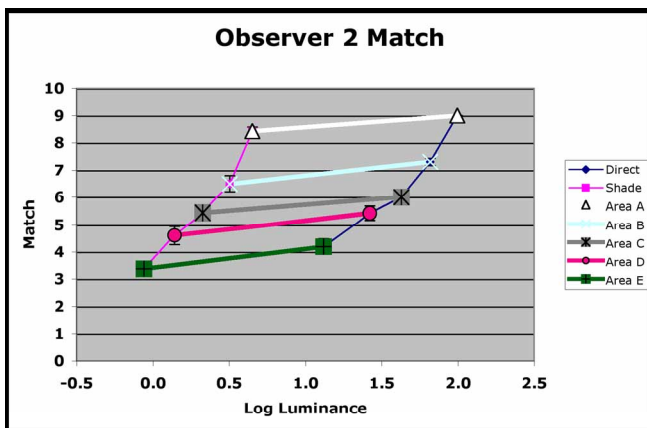Figure 11b plots the average of four trials ± the standard deviation for the rectangles target for observer JMC. The legend identifies plots match vs. log luminance for the gray scale in direct light (slope = 5.2) and in shade (slope = 6.7). In addition, the change in match for each area is plotted. The average slope is 0.56 ± 0.13.

Results

Figure 11 plots the observer matches vs. luminance for the Rectangles test patch card. Figure 11a plots the matches for one observer; Figure 11b plots the results for a second observer. Both observers made very similar matches. The average slopes for illumination change here were 0.59 and 0.56 (0.54 for *flat-2D-transparent* study). The white surround Rectangle slopes were 5.0 and 5.2 (4.6 for *flat-2D-transparent* study). The shade Rectangle slopes were 6.9 and 6.7.

Figure 12 plots the average of the data in Figure 11 and compares it with the average data from the *flat-2D transparent* study with uniform illumination. In addition, Figure 12 plots average matches for two other folded cards with different backgrounds: circular spots on a white background; and the same spots on a Mondrian background.

Figure 12 show that there is very little difference between matches with the center of the Rectangles, the Circular Spots, (both in white surrounds), and a Mondrian. Further, the data fits very well the physical relationship described for the *flat-2D transparent* experiments. The whites in the direct illumination have matches averaging 9.0. The average match in shade was 8.4 and falls on the Hipparchus (slope 0.54 line). There is very little difference between matches made in *flat-2D transparent* targets and real *complex 3D* scenes.

We find no evidence for discounting the illumination. Observer matches are the same as in uniform illumination. The observer seems to follow the simple physical rule that the local maxima respond to luminance. Areas darker than the local maxima in white surrounds are controlled by the same high-slope contrast mechanism.
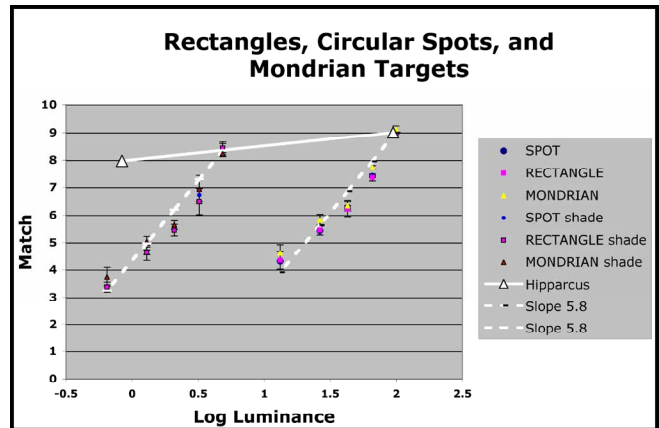

Figure 12 plots the average of eight trials (2 observers) ± the standard deviation for three different backgrounds around the matching patches. The linear regression fit for circular spots with a white surround has a slope of 5.2 for direct and 6.2 for shade. The fit for the Mondrian surround has a slope of 5.0 for direct and 6.1 for shade. That is not very different from the slope of 5.1 for direct and 6.8 for shade in the Rectangles experiment (Figure 11). The average slope is 5.8. The lines are the predictions made in simple physical description described in the flat-2D transparent study. The dot-dash line plots the Hipparchus line that predicts the matches for whites or local maxima. The solid white line plots the high-slope behavior found for white surrounds.

It is interesting to note that the Mondrian background, despite its lower average luminance, has the same slope as the other two white backgrounds (Circular Spot and Rectangular patches). This fact is important because it shows that standard matching targets should use white surrounds. Surrounds with gray surrounds will not behave the same as complex images.

The targets in Figure 10 are binocular variants of the Mach Card demonstration. These demonstrations have considerable variability with spatial content.[9] The appearance matching experiments in this paper differ in three important properties from Mach perceptual experiment. First, the displays have three different complex spatial patterns printed on them. Second, the displays were seen in binocular vision that inhibited the depth reversal necessary for change in the perceived reflectances. After all matches had been completed, observers were asked if they could make the card reverse using binocular vision. They said they could not. Third, the question asked of observers was the artist's palette problem of finding a value in the standard that looked like the patch. Again, after all matches had been made, observers were asked to use monocular vision to see if they could make the card reverse in depth. They could. They were then asked if the reversal changed the appearance of the gray patches on the card. They said that the reversal did not. Although there are physical similarities to Mach card, there are also important differences. These results are not in conflict with perceptual studies of Mach's original perceptual experiment, they are different answers to different questions using different stimuli.

Do uniform stimuli appear uniform?

Additional matches were made along the rectangular patches as illustrated in Figure 13. Observers were asked to match all along the gray stripe that started near the table at the bottom of the card, traveled to the top and continued down the shade side. At the top the observer was asked to fixate on the edge created by the shadow at the very top. Here the observers were asked to match the edge they observed at the transition.
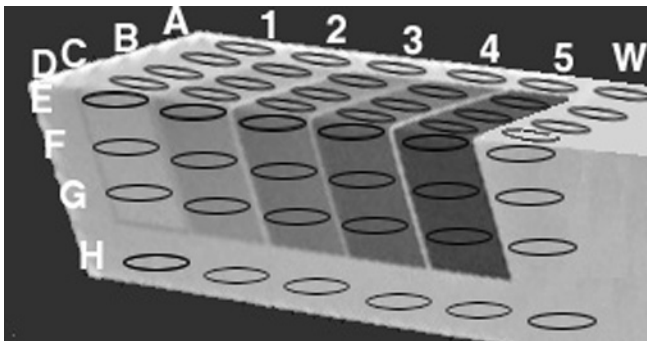

Figure 13 shows the diagram given observers to identify the test patch segment for matching. Here, A through H identify position along the patch.

Figure 14 plots the average matches for image segments along the rectangular patches. The plot starts at the bottom of the card in the shade (Position 0). The first matches are in the white surround below the rectangles (position 0.15); the next match was one-third the way up the patch (position 0.50); the next match was two-thirds the way up (position 0.75); the next match was at the

very top on the shade side (position 0.95); the next match was at the very top on the direct-light side (position 1.05); the next match was one-third the way down the patch (position 1.25); the next match was two-thirds the way down (position 1.67); last match are in the white surround below the rectangles (position 1.85).

The results in Figure 14 show that uniform luminances do not always appear uniform. The right side of the graph shows that the gray rectangles appear nearly uniform in direct light. At the edge created across the rectangle by the change from direct to shade illumination there is a significant decrease in match followed by higher values further along the strip. The match at the center of the shade portion of the strip falls on the Hipparchus line. The spatial comparisons at the illumination edge report a much darker patch. All the rest of the spatial comparisons report a lighter patch consistent with the spatial comparisons to the local maxima. This non-uniform spatial appearance is an important piece of data for any computational model using multi-resolution computations. This data can be used to identify the contributions of different size spatial components.
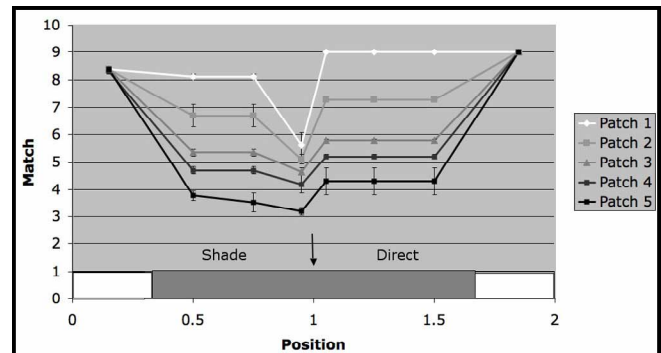

Figure 14 plots the matches along the rectangular patch for 5 target patches. The horizontal axis is the distance along the paper from white surround in shade to fold at the center and on to the white surround on the direct illumination side.

Discussion

The principle goal of these experiments was to measure the extent of constancy with changes in luminance. We studied *flat-2D transparent* targets and *complex 3D* shapes in direct light and shade. To make these measurements we must use an experimental design that accurately quantifies constancy. We chose to instruct the observer to make the artist's palette judgment to match the appearance of the test patch. We told the observer not to guess the reflectance of the patches. Guessing the patch's reflectance does not accurately quantify the appearance. Two patches that look different could be assigned the same apparent reflectance[2]. Such matches are appropriate for measuring how well human do at reflectometry,[5,10,] but do not help quantify the accuracy of appearance constancy match.

These experiments were designed to measure the magnitude of the phenomenon called "discounting the illuminant." We performed extensive experiments to measure the changes in match

with overall, uniform changes in illumination (*flat-2D transparent* targets). In 3D, we measured identical reflectance patches in direct and shadow illumination. We found no difference in observer matches in uniform illumination and real 3D scenes. We found no evidence for perfect constancy. We found no evidence for high-level recognition of illuminants and mechanisms that discount those illuminants. Instead, we found that that the maxima in the field of view and the local maxima follow the low-slope decrease in match with luminance. This is the simplest form of response to light. Response is proportional to log luminance. Areas darker than the maxima show a rapid change in match due to contrast [slope 5.4 (*flat-2D transparent* targets) and slope 5.8 (*complex 3D* targets)]. These simple low-level mechanisms are all that are necessary to explain luminance constancy in *flat-2D* and *complex 3D* targets.

Human vision does not exhibit perfect constancy in luminance and in color. The departures in both are small enough so that they are easily overlooked in everyday observations. Careful matching experiments can quantify the departures from perfect constancy. These measurements are the signatures of the visual mechanisms that approach perfect constancy. Human vision normalizes images using the maxima in each color channel. Experiments matching color in a Mondrian showed a slow change in appearance in L, M, and S channels from shifts in overall illumination.[11] Other experiments have shown that departures from perfect color constancy correlate with crosstalk between color channels.[12] Such crosstalk is the result of spatial comparisons within a single set of receptors.[13] Human vision normalizes images using maxima. This behavior is seen in color constancy, rod-Lcone color[14], and here in achromatic experiments.

Conclusions

Observers matched six contrast targets to standards. The effect of illumination was to decrease the match along the low-slope "Hipparchus line". Observers matched the maximum in the field of view to this line. Image areas, less than the luminance maxima, decrease with luminance at a much higher slope. In the *flat-2D transparent* targets, observer matches had slopes of: 4.6 for white; 3.8 for gray; and 2.6 for black surrounds. In the *complex 3D* study, the average slope was 5.8 for the Mondrian targets, the Rectangle, and Spot targets in a white surrounds.

The matches were not consistent with the predictions of constancy implied by "discounting the illuminant". The observers' matches fit the simple two-step physical description of local maxima, that is dependent on luminance, and other darker areas, are dependent on spatial contrast. This simple summary of matches fit both *flat-2D transparent* and *complex 3D* data.•

Acknowledgements

References

[1] R. E. Zissell, "Evolution of the "Real" Visual Magnitude System", JAAVSO **26**, 151, 1998.

[2] 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.

[3] D. H. Jameson and L. M. Hurvich, "Complexities of perceived brightness", Science, **133**,174-179, 1961.

[4] C. J. Bartelson and E. J. Breneman, "Brightness perception in complex fields", J. Opt. Soc. (A), **57**, 953-957, 1967.

[5] A. Gilchrist, Ed. Lightness, Brightness and Transparency, Lawrence Erlbaum Associates, Hillsdale, 1994.

[6] J. J. McCann, "Rendering High-Dynamic Range Images: Algorithms that Mimic Human Vision", in Proc. AMOS Technical Conference, Maui, p 19-28, 2005

[7] L. Arend and R. Goldstein, "Simultaneous Constancy, Lightness and Brightness", J. Opt. Soc. (A), **4**, 2281-2285, (1987).

[8] M. G. Bloj and A. C. Hurlburt, "An empirical study of the traditional Mach card effect", Perception, **31**, 233-246, 2002.

[9] A, C, Hurlburt, "Editorial: Illusions and reality-checking on the small screen", Perception, 27, 1998,

[10] R. Robilotto and Q. Zaidi, " Limits of lightness identification for real objects under natural viewing conditions", J. Vision, **4**, 779-797, 2004.

[11] J. 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.

[12] J. J. McCann, Mechanism of Color Constancy, Proc. IS&T/SID Color Imaging Conference, IS&T/SID, Scottsdale, Arizona, 12, 29- 36, 2004.

[13] J. J. McCann, "Do humans discount the illuminant?", In Proc. SPIE 5666, Human Vision and Electronic Imaging X; B. E. Rogowitz, T. N. Pappas, S. J. Daly; Eds., San Jose, p. 9-16, 2005.

[14] J.J. McCann, J. L. Benton, S. P. McKee. "Red/white projections and rod/long-wave cone color: an annotated bibliography", J. Electronic Img, 13, 8-14, 2004.

Author Biography

John McCann received an A.B. in Biology from Harvard in 1964. He managed the Polaroid's Vision Research Laboratory from 1961 to 1996. He worked on human color vision, large-format instant photography, the reproduction of fine arts and digital imaging (since 1975). He is a Fellow of the IS&T and has received SID and IS&T service awards for the initiation of their Color Imaging Conference. He is a Past-President of IS&T and the Artists Foundation, Boston. He has chaired numerous meetings, including the IS&T/SPIE Electronic Imaging 2000. He is an Edwin Land Medalist and an IS&T Honorary Member. He consults, continues research on color/HDR imaging and explores the history of color photography.