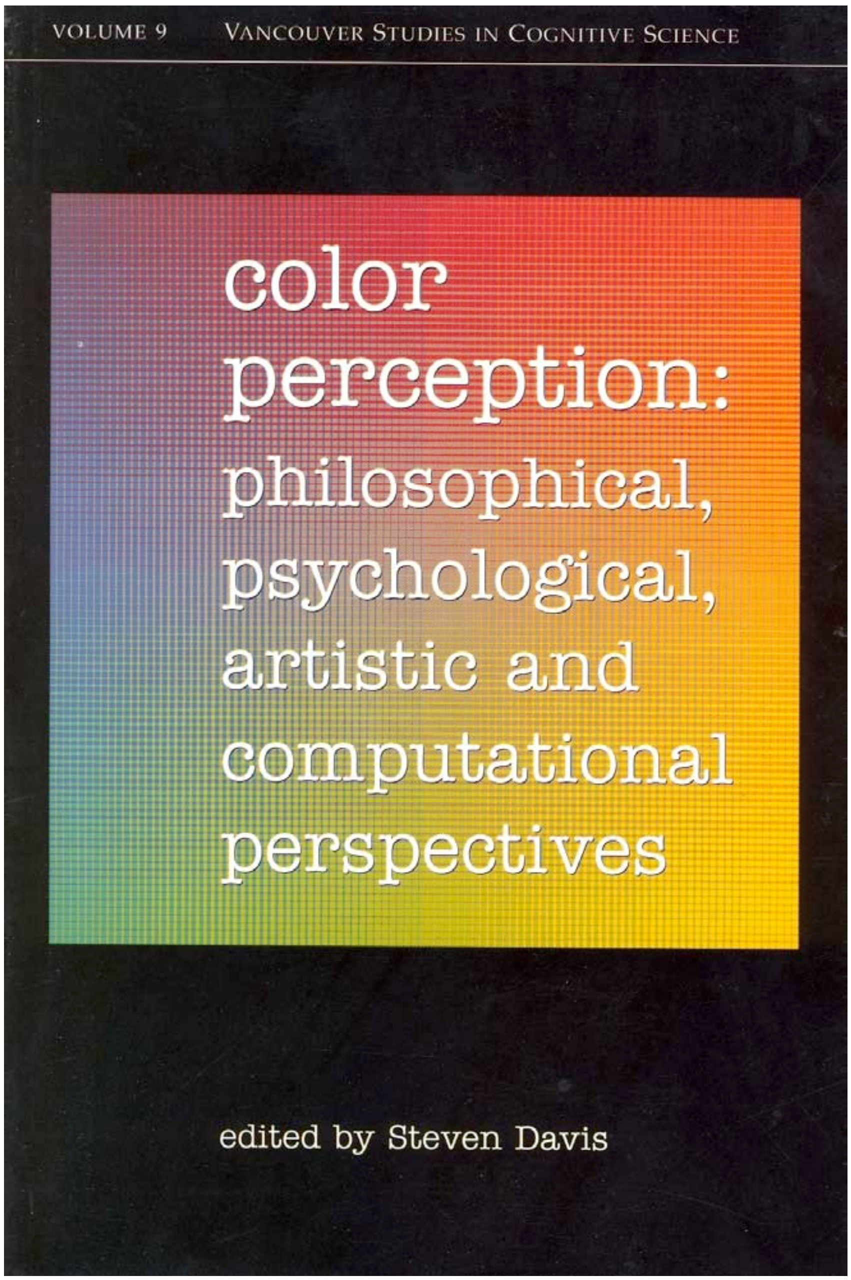
J. J. McCann, "Simultaneous Contrast and Color Constancy: Signatures of Human Image Processing" Chapter 6 in *Color Perception: Philosophical, Psychological, Artistic, and Computational Perspectives*, Vancouver Studies in Cognitive Studies, pp. 87-101, S. Davis, Ed, Oxford University Press, USA, 2000.



Colour Perception:
Philosophical, Psychological, Artistic and Computational Aspects
9 th. Annual
Vancouver Cognitive Science Conference

Simultaneous Contrast and Color Constancy: Signatures of Human Image Processing

John J. McCann
McCann Imaging

Rarely does a conference include all the relevant aspects of a subject. This conference is the exception in that it has embraced the philosophical, artistic, physical, physiological, psychophysical and computational dimensions of color. The problem is where to begin. The following paper is a discussion of computational models that mimic human visual image processing. The objective is to describe a range of visual phenomena that are unique to human color vision. The common theme is that psychophysical experiments describe certain visual properties or signatures that can be used to test whether image processing algorithms are acting as the eye does. This paper is an attempt to discuss the scope of human color, and the limitations that separate human color processing from other mechanisms.

Four Categories of Human Color Models

For the sake of discussion, we can categorize models of human color vision in four types (Figure 1). The first is the physical analysis of the scene by the science of colorimetry. It models the response of the cone-shaped, light-sensitive cells in the retina. Although based on psychophysical measurements,[1] the spectral sensitivities are in close agreement with physical measurements of the cone cell absorption spectra.[2] In short, the psychophysical experiment of matching colors is successfully modeled by a purely physical mechanism of quanta catch by retinal receptors.

[1]W. D. Wright, The Measurement of Colour, 3 rd Ed, Hilger & Watts, London, 1964 and CIE Proceedings 1931, p. 19, Cambridge University Press, Cambridge, 1932.
[2]V.. Smith and J. Pokorney, Spectral Sensitivity of the Foveal Cone Photopigments between 400 and 500 nm, Vision Res., **15,** p. 161, 1975.

The second type of color model is the psychophysical model of sensation. The curious property of human color vision is that the physics of colorimetry will always predict matches, but cannot predict colors. The same quanta catch at the retina can appear gray, red, yellow, green or blue.[3] Human vision calculates color by comparing responses over the field of view: color is relative to other colors in the scene. Sensation models use the quanta catch at all pixels in the field of view as the input. The output is the color appearance.

The third type color model is the cognitive model of recognition. The perception model of human color vision combines prior experience with sensation and quanta catch models. The output is the color recognition.

The fourth type of color model is one the attempts to calculate emotion. It is based on the study of fine art and uses the modern tools of digital imaging. This field is simply a set of questions of how well the rules of aesthetics can be quantified to predict emotional responses to images.

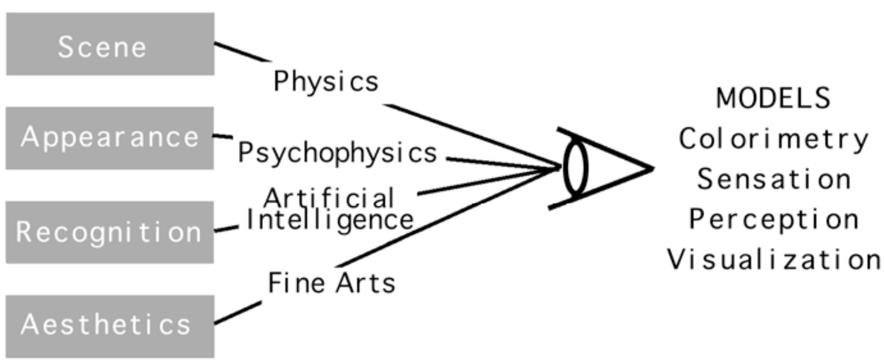


Figure 1. The left-hand column of boxes lists the four categories of problems: Scene, Appearance, Recognition, Aesthetics. The four arrows list the disciplines that address these categories: physics, psychophysics, artificial intelligence and fine arts. The list to the right of the eye is the name of the models used to predict human behavior: colorimetry, sensation, perception and visualization.

Sensation & Perception

Often we see the models of colorimetry, sensation and perception used in a manner that makes them seem interchangeable. We see colorimetry

[3]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**, pp. 445-458, 1976.

diagrams, such as the CIE u,v color space that have been painted the colors of the spectrum. These diagrams may help us remember that long-wave light is plotted in the lower right. However, it implies that color appearance is correlated with location on the u,v graph, when it is not.[4] Likewise, people often use color sensation and color perception interchangeably. The Optical Society of America uses a pair of definitions that are similar to those of the Scottish philosopher Thomas Reid(1710-1796).

> Sensation: Mode of mental functioning that is directly associated with the stimulation of the organism.[5]
>
> Perception: Mode of mental functioning that includes the combination of different sensations and the utilization of past experience in recognizing the objects and facts from which the present situation arises.[6]

There are many nuances and variations found in the very frequent use of these terms. The common feature of them all is that perception is more complex than sensation, and involves past experience.

It is helpful to compare and contrast these terms in a single image to stabilize our vocabulary as we progress from 16th century psychology to 21st image processing. A good example is a photograph of a raft, --a swimming float-- in the middle of a lake[7]. The photograph was taken in early morning: the sunlight falls on one face of the raft while the other face is illuminated by skylight. Although standard daylight is 6500 degrees Kelvin, it is in fact the mixture of a very yellow sunlight and a very blue skylight. Estimates of the sunlit face of the raft are about 3000 degrees Kelvin, whereas estimates of the skylit face are 20,000 degrees Kelvin. The sunlit side reflects about 10 times more light that skylit side. In summary, the two faces have very different quanta catches and hence very different colorimetric values.

---

[4]G. Wyszecki, "Colorimetry", in Color Theory and Imaging Systems, Society of Photographic Scientists and Engineers, R. Eynard, ed., Washington, p. 24, 1973.

[5]OSA,[5]The Science of Color, OSA, Washington, DC, p381, 1953.

[6]OSA, The Science of Color , OSA, Washington DC, p377, 1953.

[7]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).

Fig 2. Photograph of raft.

To measure sensations we need to ask observers to select the colors they see from a lexicon of color samples, such as the Munsell Book or the catalog of paint samples from the hardware store. The question we ask the observer is to find the paint sample that a fine-arts painter would use to make a realistic rendition of the scene. Observers select bright white with a touch of yellow for the sunlit side and a light gray with a touch of blue for the skylit side. The sensation matches render the two faces as similar, but slightly different.

To measure perceptions we need to ask observers to select the colors from the same catalog of paint samples, but with a different question. The perception question we ask is to find the paint sample that a house painter would use to repaint the raft. The observers select white. They recognize that the paint on the raft is the same despite the illumination. The perception matches render the two faces identical.

The raft faces are very different, similar or identical depending on whether the experimenter is measuring colorimetry, sensation or perception. These terms cannot be used interchangeably. They embody completely different kinds of image processing. Colorimetry models tell the story of the receptors, sensation the story of the color appearance -- a spatial calculation, and perception -- the story of cognition.

What is the Computational Goal?

The definitions of sensation and perception are important because they have a very large effect on the desired calculation. You cannot write a single algorithm because, as we saw above, the raft's sensation answer is different from its perception answer. So first, you need to select the your goal. Do you want the calculation to have the sensation, the fine arts painter's rendition? Or do you want the cognitive, house painter's rendition? One can combine the two algorithms by saying that humans compute sensations first. This step would handle color constancy, compress the range of the image, emphasize edges and minimize gradients. This sensation image could be the input to the perceptual image that undertakes the difficult task of recognition.

There is an entirely different sequence of ideas that starts with the array of light in the scene and attempts to calculate the reflectance of objects. Such techniques are successful in scenes with single illuminants. It becomes more difficult with many illuminants in a single scene. This topic will be discussed in detail later in this conference.

The most important idea here is the eventual use of the calculated image. What is it for? A successful "reflectance" rendition of a scene will be the same as a photograph of that scene in perfectly uniform light. There will be no shadows, no light modeling. It would be a meticulously accurate "paint-by-numbers" portrait of the objects. This image will be stripped of the light shading we associate with three dimensional objects. Such an image is a good object map.

A successful sensation image would be a fine-arts painting of the scene. The reflectance range of the rendition is limited to 30:1. The color shifts due to illumination are almost removed. The rendition looks like the original scene. The effects caused by lighting are still visible. Only their magnitude has been altered. This is an ideal image to write on paper as a properly exposed photograph. The photographs in Figures 3, 4, and 5 are an excellent example of an image that does not fit the dynamic range of a conventional photograph. The original scene was a clear day in Yosemite Valley that included sun and shade. In the foreground is a Macbeth ColorChecker©. In the background John is holding a white card. The range of reflected light from the ColorChecker© is 30:1. The ratio of sunlight to shade is 30:1. The white card in the shade sends the same light to the camera as the black patch in the sun.

Fig. 3. An accurate rendition of the scene in sunlight.

The problem with a photographic print of this scene is that the print has the same range as the ColorChecker©, i.e., 30:1. Figure 3 is an accurate reproduction of the sunlit scene. It correctly portrays the ColorChecker©. White is white and black is black.

Fig. 4. An accurate rendition of the scene in shade.

Figure 4 is an accurate portrayal of the shade image. The white card in John's hand is white, but the black patch in the ColorChecker© is also white. Figures 3 and 4 have accurate renditions of two different 30:1 portions of the 1000:1 dynamic range scene.

Fig. 5. A Retinex-processed image, with a 30:1 dynamic range replacing the 1000:1 dynamic range in the original scene.

Figure 5 is an image-processed rendition of the scene[8]. Here the total dynamic range has been reduced from 1000:1 to 30:1. The image processing had compared each pixel with each other pixel in the scene. In the input image, the white in the shade and the black in the sun send the same light to the camera, and hence have the same input pixel value, namely 126. In the output image the pixels in the white card in the shade have an average value of 238, while those in the black on the ColorChecker© have a value of 15. The image processing has done what the human eye does. It has reassigned identical input values to very different output values. In fact, the outputs are almost as different as possible in a 0 to 255 digital image.

In chemical photography there is unique, monotonic tone curve, sometimes called a characteristic or H&D curve. It records the input/output function of the system. In chemical photography each pixel operates independently from all the other pixels. One characteristic curve -- exposure in verses density out -- describes all the pixels. If the input digit is 128, and the output digit for one pixel is 150, then the output pixel for all 128 inputs will be 150. This approach is necessary for chemical systems and efficient for digital systems, but very limiting in image processing. Vision is different. A 128 input value can become any output value, depending on the values of other pixels in the scene. Since each input value is associated with a variable output value, there is no tone curve for vision.

---

[8]J. Frankle and J.J. McCann, Method and apparatus of lightness imaging, US. Patent 4,384,336.

Signatures of Human Vision

Here we take a closer look at models of color sensations. We will look at the unique properties of human image processing that are the signatures of the way we see. These signatures are of critical value in selecting the best models for vision. Color constancy, for example, can be accounted for by any number of schemes. The best model is not one that can account for only one example, but rather the one that is consistent with the experimental conditions when color constancy breaks down. The successful model is one that includes the limits of the phenomenon. The following section describes a series of psychophysical limits of human vision. They are useful as signatures of human image processing and can be used to differentiate models of vision.

Absolute Radiance

We often see in popular scientific articles the idea that the eye is like the camera. There are two important facts about human vision that don't fit the camera-film model. First, films with compensating exposures give essentially the same print with different levels of illumination. Human vision has a whiter white in high illumination. Of greater interest is the fact that blacks are blacker in high illumination than in low illumination. This signature of vision is a subtle, but important characteristic that distinguishes models of vision from camera mechanisms.[9]

The second important difference is that human vision independently normalizes to the maxima in long-, middle- and short-wave light.[10] Cameras and Grayworld artificial intelligence mechanisms use the average of the scene.

Simultaneous Contrast

The effects of colors surrounding the color of interest has been studied since Leonardo.[11] Hering's extensive study[12] of the surround led to many

---

[9]J. C, Bartleson and E. J. Breneman, Brightness perception in complex fields, J. Opt. Soc. Am. **47**, 953-957, 1967, and D. Jameson and L. M. Hurvich, Theory of brightness and color contrast in human vision, Vision Res. , **4**, 135-154, 1964

[10]J.J. McCann, Rules for Color Constancy, Opthal. Physiol. Opt., **12:,** pp. 175-177,1992. and J.J. McCann, Color Constancy: Small overall and large local changes, in Human Vision, Visual Processing, and Digital Display III, B. Rogowitz ed., Proc SPIE, 1666, pp. 310-321, 1992.

[11]J. McCann, "Human Color Perception" in Color Theory and Imaging Systems, Society of Photographic Scientists and Engineers, R. Eynard, ed., Washington, p. 1, 1973.

[12]E. Hering, Outline of A Theory of the Light Sense, trans. by L.. Hurvich and D. H. Jameson, Harvard University Press, Cambridge, MA, 1964.

different opponent-processing mechanisms both in psychology[13] and neurophysiology[14]. Alber's book "The Interaction of Colors"[15] is a lexicon of contrast effects. It shows many different examples of identical papers that appear different sensations when viewed in different surrounds.

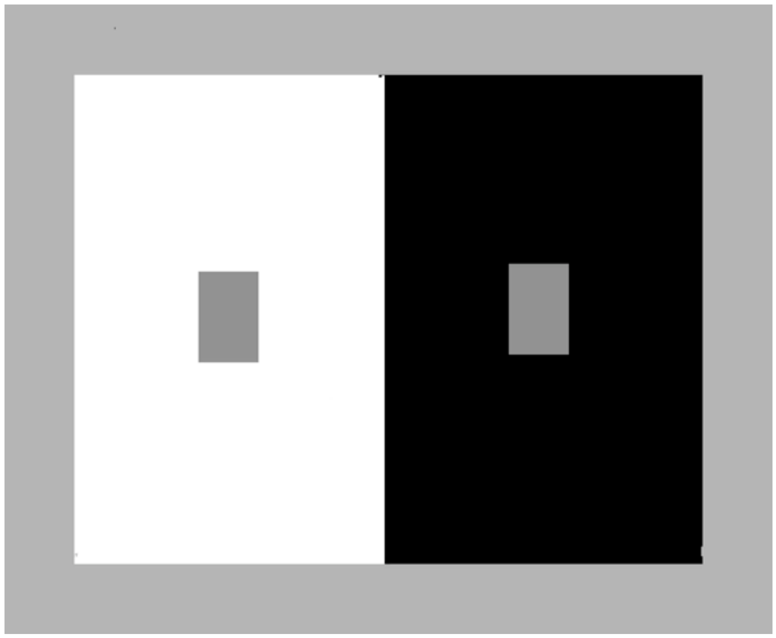


Figure 6. Simultaneous Contrast

The simplest example is shown in Figure 6. Here two identical gray patches are surrounded by white in one case, and black in the other. The one in the black surround looks lighter than the one in the white. This simple observation is an essential signature of human vision. Any mechanism that successfully finds the reflectance of objects in an input image will not have this essential human vision signature. Such a model will report that the two grays are identical. A model that mimics human vision must recreate this failure to render equal reflectances as equals.

Color Constancy

This phenomenon has the very simple roots that red objects look red in cool skylight, neutral daylight, warm sunlight and red firelight. These illuminants have a significant effect on the quanta catch of the cones, but has very little effect on the color sensations. Nevertheless, experiments have measured the signature of human vision's departure from perfect

---

[13]L. M. Hurvich, Color Vision, Sinauer Assoc. , Sunderland, MA, 1981.
[14]R. L. DeValois, Color Vision Mechanisms in Monkey, J. Gen. Physiol. 43, 115-128, 1960.
[15] J. Albers, "The Interaction of Colors" Yale University Press, New Haven, CT, 1963.

color constancy. [3,16] Here again, mechanisms that actually remove all traces of illumination will remove this important signature of human vision.

Discounting the Illuminant

This idea was popular in nineteenth century European psychology. The notion is that the eye can calculate the illuminant and discount it. This idea could work well with color constancy, but has serious problems with real three dimensional scenes. Mechanisms that can remove all traces of illumination would also remove the gradients so valuable in recognizing three dimensional shapes.

No unique tone curve

If we recall the images in Figures 3, 4, and 5, we see that the input scene had two identical values for a white and a black in the scene. Photographic film can only have a single tone curve. It is very carefully crafted to be the best for the widest variety of scenes. Nevertheless, one input has only one output value. Humans are different. Both white and black are possible outputs from exactly the same input at a pixel. This property is probable the best litmus test of human image processing techniques. If the objective of a technique is to mimic human vision then it must have spatial interactions capable of transforming input values into the whole range of outputs. In other words, it has no unique tone curve.

In summary, human vision has six distinctive signatures that can be used as a test of models for vision.

1. Dependence on overall radiance
2. Independent normalization of long-, middle-, and short-wave images.
3. Color is influenced by surround.
4. Color constancy is not perfect.
5. Gradients in illumination are not discounted on 3-D objects.
6. Any output, i.e., white or black, can be generated from a single input value.

---

[16]J.J. McCann, Magnitude of Color Shifts from Average-Quanta Catch Adaptation Proceedings of 5th IS&T and SID Color Imaging Conference , vol. 5, pp. 215,1997 and J.J. McCann, Color Mondrian Experiments without Adaptation, Vol. **1**, pp. 159-162, AIC Color 97 Proceedings, Kyoto, 1997.

Model Design

The above list of signatures comes from a series of different experiments. Although a complete model of sensation must include all these signatures the model need not be tied directly to the experimental source. For example a well-chosen model can simplify the problem by combining different signatures into the same mechanism.

Simultaneous Contrast

As noted above, the most familiar example of simultaneous contrast is two identical pieces of gray paper; one placed on a white background, one placed on a black background. The gray-on-black papers look about 10% lighter than the gray-on-white paper.

Usually these experiments are used to illustrate the ideas that;

1. The color of objects is determined by the areas around them.
2. Simultaneous contrast is a local mechanism.

However on closer inspection we find that simultaneous contrast experiments are:

1. Dependent on normalization to maxima
2. Normalization is sensitive to distance and enclosure.

These conclusions are evident when we look at a wide variety of center-surround experiments. White and black will change the appearance of grays. However, different gray surrounds will not change the appearance of whites or blacks. The phenomenon is best explained by the idea that the entire image is normalized to the maxima.

The biological normalization process is very different from the mathematical one. If we normalize a numerical array of numbers to the maxima, the result is the same regardless of the position in the array of the maximum value. Vision is different. The distance and the degree of enclosure influence the extent of normalization. Areas contiguous with the maxima are effected more than those at a distance. Areas surrounded by maxima are darker than those with maxima on one side.

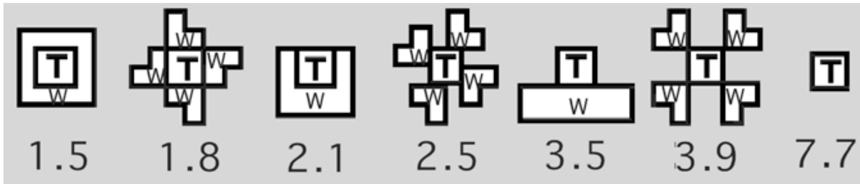


Figure 7 shows a series of gray patches (T) and a white surrounding element (W). Each white element has the same area. The numbers below the diagram report the matching gray value for each target.

Figure 7 shows the matches to a gray scale with white =9.0 and black =1.0. The results show that the gray sensation can vary from 1.5 to 3.9 simply by varying the placement of the same area of white surround. The last gray patch on the right shows that it has a lightness of 7.7 when there is no white in the field of view. Figure 7 illustrates the nature of normalization to maxima in human vision[17].

In summary, is it essential that we keep in mind the unique properties of vision. When we use the word normalize in a mathematical sense, we divide each value in the array by a single number. Visual normalization divides each number in the array by a different number depending on the distance and the degree of enclosure of the maxima.

Color Constancy

The most familiar example of color constancy is two identical complex displays of papers; one placed in a long-wave-rich illumination, one placed in a short-wave-rich illumination. The paper looks essentially the same in both, regardless of the light coming to the eye. In color constancy, the observer ignores the illuminant. McCann, McKee and Taylor's quantitative study of the range of colors from the same quanta catch is an example. Usually these experiment are used to illustrate the ideas that;

1. The color of objects is determined by the objects' reflectance.
2. The color of the surround has little influence.
3. Color constancy is a global mechanism.

However, on closer inspection we find that color constancy experiments are:

1. Dependent on normalization to maxima
2. Normalization is sensitive to distance and enclosure.

---

[17]J.J. McCann and R. L. Savoy, Measurement of lightness: Dependence on the position of a white in the field of view, in Human Vision, Visual Processing, and Digital Display II, B. Rogowitz ed., Proc SPIE, 1453, pp. 402-411,1991.

The detail matches in color constancy experiments show that average mechanisms and adaptation do not explain color constancy. Normalization to the maxima in the scene is the only explanation consistent with the matching data.[3,16] Furthermore, other experiments demonstrate that this normalization is independent for each of the long-, middle-, and short-wave mechanisms.[10]

The quantitative data on color constancy reinforces the general observation that the surround does not matter in this context. The reason is simple. The 18-Area Mondrians are well populated with maxima in all wavebands. With such displays the surround effects are minimal.

Historically many authors consider simultaneous contrast and color constancy as different psychological phenomena. Nevertheless, they have precisely the same roots, or underlying mechanisms. The normalization to maxima shows itself as lighter gray in a black surround and as a color correction in color constancy. Both experiments show a greater than local, but less than global scale of interaction.

A successful model of vision will incorporate unifying principles such as normalization to maxima. It combines several different signatures into a single mechanism. Other mechanisms will be necessary to incorporate absolute radiance. This however will contribute substantially to the departures from perfect color constancy.[3] Finally, the mechanism for normalization is the basis for the lack of the unique tone curve and the proper renditions of gradients.

Color Aesthetics

Painters have great skill in synthesizing emotion. They create the mood in our minds by their choice of tone, color, contrast, brush strokes and scale. The tools they use are the same for the full range of emotions. The selections create the desired emotion. All of these tools are today under computer control. In principle, one could write a program that takes a scanned picture and then reconstructs the common image into a series of different renditions that emote fear, happiness, calm, anger, etc. If you combine these color and tone tools with facial expressions tools, we could imagine a simplistic computer equivalent of an artist's ability to synthesize emotion.

A particularly good example of an artist's visualization is Ansel Adams' account of his invention of visualization for photography. Adams describes[18] climbing to the top of the valley walls of Yosemite, he had just taking an 8x10 glass negative of Half Dome. He had one negative plate left. He paused to study the scene he had just captured. He visualized what the real life colored scene into the black and white print he was about to make. He realized that the blue sky and the gray face of Halfdome would have about the same tonal value. He then took out a deep red filter, that would change the tonal value of the sky from light gray to black in the print. Adams visualized that the print would have the light face of Halfdome dramatically portrayed against the darkened sky. This was the image he wanted, rather than the first exposure. Adams describes this experience as the first time he visualized the printed image at the time of capturing the scene. Adams developed a wide variety of photographic techniques to capture, control and modify gray-scale values, so as to create the visualized image. Digital image processing techniques available today are much more powerful and convenient than those used by Adams.

Summary

This paper has been a review of the different goals of image processing programs that mimic human vision. It provides a list of different signatures of vision that can be used to distinguish vision models from machine vision algorithms. It describes how some of the historically familiar psychological principles can be simplified by being restated as more general biological image processing mechanisms. As we return to the list of models in Figure 1, we can observe an uneven distribution of effort. By far, the greatest scientific effort has been devoted to colorimetry or the elucidation of receptor mechanisms.

Sensation models have been discussed at length in this paper. True perception models, involving recognition of objects and light sources requires cognitive modes that are difficult and computationally intensive. A large body of work has avoided this approach for a more attainable goal of machine vision: the search for the reflectance of objects, often accompanied by simplifying assumptions about the illuminant, and a departure from the goal of mimicking human vision.

---

[18] A. Adams, "Examples The Making of 40 Photographs", Little Brown and Company, Boston, pp. 2-5, 1983

The fourth type of model is in its infancy. Here we use digital image processing to calculate emotion. The idea is that a rule-based analysis of images could analyze any image and predict human emotional response. Artists have familiar rules that they use to create pictures that convey: happiness, sadness, joy, terror, peace, anger, tranquillity, etc. The artists' tools are color selection, contrast, facial expression, and visual environment. These are the same elements that are easily within the reach of today's image processing tools. Today, feature length movies are being made with computer synthesized images controlled by artists. In the future rule based programs created by artists might synthesize the emotional message.