Image data's hidden assumptions: Consequences in Colorimetry and Natural Scenes

John J. McCann, McCann Imaging, Arlington, MA 02474, USA

Abstract

Electronic Imaging is a major source of input data in today's wealth of computational models. Using models of Color, as an example, we can see that photographs as input data can be problematic. Valid input data requires caution about the actual quantitative information recorded in digital images. Attention to quantitative numbers, such as their calibration and linearity is needed when we use camera data as the input information for computations. As well, does a single pixel's data provide all the information needed to predict Color? Sometimes in Color Appearance models we use single spots of light in a no-light background, while at other times we use HDR Natural scenes. Is it appropriate to assume that a model, established using a single spot of light for the stimulus, is transferable to the Natural Scene's Appearances? The problems of hidden assumptions in Electronic Imaging Models is discussed using Color as an example.

Introduction

In many scientific journals, it has become common practice to disclose any possible commercial conflicts of interest. This practice helps to address issues of possible, and actual problems of funding sources influencing scientifiv results. While this is almost never an issue in Electronic Imaging, there is another kind of full disclosure that could help our field. We should consider a Full Disclosure of a Computational Model's Hidden Assumptions (FDCMHA).

Alessandro Rizzi and Stephen Westland organized a Special Issue in Coloration Technology, "Challenges and Open Problems in Colorimetry". Their call-for-papers aimed at "*attracting provocative commentaries on colorimetry, addressing old and new challenges*". This EI talk extends my Coloration Technology contribution called "What scene information is needed for Models of Color Appearance in the Natural World?" [1]

Over my career I have witnessed a revolution in models of imaging. When I began, all imaging models predicted the light-matter interactions of a small *spot of light*. Silver halide (AgX) Photography and Animal Vision were modeled by the physical chemistry at an atomic and molecular scale. AgX cameras were the cell phones of today. Everyone knew the basics of how a camera worked, and even a college level course taught by a Nobel Laureate described human vision's mechanisms by comparing them to an AgX film camera.

In the 1960's very successful imaging film models predicted the response of sensors to the light falling on them. The input to *spot of light* color models was a constant, uniform spot of light calibrated in watts per meter^2 at every wavelength. Each unique film response was generated by a unique quanta catch of input light.

Today's AI models uses trillions of images, made of millions of pixels from tens of thousands of data bases. We live in a tsunami of digital images: photos, frames in videos, slides, Facebook, Instagram, and thousands of other databases. Electronic Imaging research and development is the framework that has guided this 60 year transformation from AgX's molecular scale spots to digital imaging's arrays of millions of pixels. These digital arrays have opened the door to AI object recognition and autonomous vehicles.

Today's imaging models have replaced the small *spot of light* with arrays of millions of pixels reporting uncalibrated digital values. These digital values are entirely different from light's Watts/meter^2. These digits are records of the interactions of light falling on a camera's lens, unwanted camera glare, sensor response, proprietary camera firmware and software (dead pixel, black level, demosaic, tone-scale and chroma enhancements, data compression, fusions of multiple exposures, etc.). Unlike the calibrated *spot of light* inputs, there are no physics-based units of measure [2] in digital camera values. Glare is unique in each scene, image processing is variable with camera settings, and camera responses vary with the manufacturer, and the model. Digital images look good; however they are substantial transformations of the light coming from the scene. Camera digits do not have a physics-based unit, and are an undefined nonlinear quantity. Performing calculations using these digits requires assumptions. For example, an average of digit values assumes linear data. The average of pixel values need not be equal to the average of scene radiances, because cameras have nonlinear responses to light.

Without calibration measurements the quantity described by a pixel value is undefined. There is an often cited assumption that digital camera sensors have a linear response function because the sensor literally counts photons. This assumption is accurate. Sensor response is linear. Using sensor linearity as an argument about camera response, requires a hidden assumption; namely, that the light falling on the lens is linearly proportional to the light falling on the sensor. That assumption is a serious unforced error. The camera's optical glare is a substantial spatial transformation of scene radiances.[3] After that, the camera's image processing introduces an additional sequence of nonlinear transformations of light from the scene. Camera response is a cascade of highly nonlinear functions of camera sensor responses.[4] Both glare and the camera's digital processing spatially transform scene radiances.

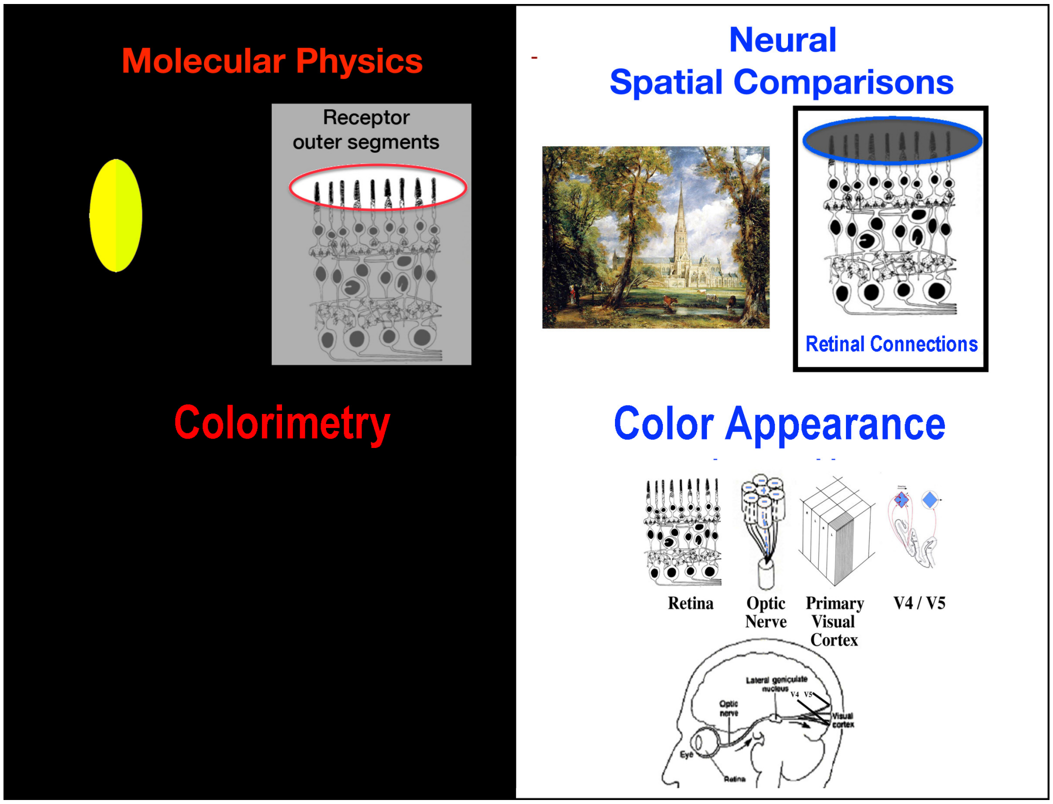

Figure 1 (left) Wright's Colorimetry uses the quanta catch of receptors to predict Color Matches. (right) Models of Neural Spatial Comparison are needed to predict Color Appearances in the Natural Scene.

Colorimetry's *Spot of Light* vs. the Natural Scene's Array

The consequences of the selection of input data for computational models are obvious in Color. Color is the response to light from a scene. The description of the scene is the model's input. Hidden assumptions play a fundamental role in Color models. In fact, they lead to different theories because they have distinctly different inputs. Colorimetry's input is the spectral radiance of a *spot of light*. As David Wright pointed out *"colorimetry ends once the light has been absorbed by the colour receptors in the retina and that appearance science begins as the signals from the receptors start their journey to the visual cortex."* [5]

David Wright splits CIE Color Standards into its two topics: Colorimetry, and Appearance. Colorimetry is the study of the Molecular Physical Chemistry of light /matter interactions at an atomic level. The spectral sensitivities of the receptors are the input to vision (Figure 1, left). After that, *Appearance* is the study of the cascade of neural interactions that begin with the signals at the opposite end of the rod and cone receptors (neural junctions), and travels down the optic nerve and throughout the brain (Figure 2, right). Wright's clear dichotomy between Colorimetry (light/matter interactions) and Appearance (comparison of neuron responses) forms the framework of distinct Color Vision Models.

Two Scenes leads to Two Color Models

All human light/matter interactions take place in the retinal rod and cone outer segments (and ganglion cells). The red ellipse at the top of the left side identifies the only site of light/matter cone interactions. The only scene radiance data allowed in all CIE calculations are the X,Y, Z values of a single small spot of light on these receptors. Colorimetry predicts MATCHES from the spectral radiances of single *spots of light*.

Spots of light on a no-light background are unique stimuli. There is no glare from surrounding scene segments. The only glare comes from the spot itself.

David Wright stated that: "*colorimetry ends once the light has been absorbed by the colour receptors in the retina*". That red ellipse on the left illustration is Wright's Colorimetry stop sign.

Figure 1 (right) illustration begins at all receptors' neural junctions. The neural input is the complete array of all receptors' responses to quanta catch. Neural spatial processing uses all outputs from receptors in the retina. Unlike Colorimetry, real Natural Scenes have abundant optical veiling glare in both vision and cameras.[6] Glare transforms the HDR scene radiances into a substantially different image on the retina.[7] While scene radiance is the appropriate input to Colorimetry's model of Matches, we cannot use scene radiances as the input to neural spatial processing. A *spot of light* from a Natural Scene does not include the substantial light scattered from all the other segments of the scene.

The right side of the above figure uses a different Scene icon for Color Appearance, namely, John Constable's HDR painting Salisbury Cathedral from the Bishop's Garden, (1825). It illustrates the Bishop, his cathedral in sunlight, and his garden in shade. It is a rendition of what Constable saw - his Appearances - made from this typical real-world HDR scene.

Figure 1 (right), the white half of the figure, illustrates models of appearance, as described by Wright as the neural *"journey to the visual cortex"* and beyond. Color Appearance models use receptor's response as input to the retina's complex spatial comparisons. Dowling's retinal connections are magnified at the top and shown again below in the map of the entire visual pathway. (From retina to optic nerve, to primary visual cortex and beyond (V4,V5 etc.). The bottom of the right illustration shows the many different types of neural spatial comparisons, and their location along the visual pathway. It illustrates the work of Dowling, Kuffler, Barlow, Daw, Hubel, Wiesel, and Zeki. Neurophysiologists have shown that vision uses neural spatial comparisons at every stage along the visual pathway. They have shown that the visual pathway is a cascade of spatial comparisons starting with receptor's output synapses and continuing at every stage throughout the brain. [8]

Understanding the neurophysiology of vision, so as to know the mechanisms of neural spatial comparisons, is an essential question in color appearance. These models require input data from all parts of the scene.

Two Models of Color Vision

Colorimetry measures the light reaching the front surface of the eye. It then calculates the quanta catch of receptors to predict color matches. The hidden assumptions in different models of color vision lead to different consequences. Colorimetry limits the scene input data to a single spot of light. The consequence is that it can predict if two different spectra will match, with great accuracy. The additional consequence is that it cannot predict the color appearance of either spectrum.[9, 10]

CIE Appearance models (CIELAB, CIELUV, CIECAM) limit scene input to the XYZ values from a single scene segment. This XYZ triplet is the only input data from the scene. The CIE Appearance models require assumptions about illumination and parameter information that is not specifically calculated from the scene, or from the digital image. The model's programmer assigns this information. CIE Appearance models make the hidden assumption that the quanta catches of all other receptors in the observers field of view are not relevant to Color Appearance.

However, "Visual Illusions" contradict this assumption. Since daVinci, there are thousands of examples of two identical scene segments that have dramatically different appearances (Simultaneous Contrast, paintings by Albers [11], Land's Color Mondrians [9]). In the Natural Scene a particular quanta catch can have any color appearance.[9] Spatial comparison models require measurements of the entire scene as input. It compares all receptors' quanta catches along the entire visual pathway. The consequence is that it can calculate Appearances in Natural Scenes.

FDCMHA!

It is essential that we look for, and articulate, our assumptions when we design a model, particularly a Color model. We need to be highly aware of both calibration and hidden assumptions.

It is so easy when we pick up and use digital data from cameras, or from databases on the web. We can address almost any problem. We can do it so fast that we do not have to stop and think about things like calibration, radiometric units, linearities, etc. When we copy and use data-base digits how do we describe that quantity? When we cite that database we document the source, but who owns the responsibility that this data is actually the most appropriate input for our model? When we try a new idea we can see its effect instantly. If we see what we wanted to see on the display, we feel reassured that it is accurate, because it worked the way we wanted. We do it so fast that there was no time to examine, and articulate, our hidden assumptions. The unanticipated consequences do not occur to us. That is why digital imaging needs FDCMHA! Namely, Full Disclosure of the Computational Model's Hidden Assumptions. In particular we need to give considerable thought to our input data. We need to think about its calibration, standard units of measure, linearity, and whether it has excluded essential input information relevant to the definition and purpose of the model.

Human Color vision's response to the Natural Scene has two powerful spatial transformation mechanisms. The first is glare; the second is neural comparisons. Glare causes a major reduction in image contrast in the image on the retina. Neural comparisons introduce variable contrast mechanisms that counteract glare. Maximum glare scenes have the highest slope visual response function. Minimum glare scenes have the lowest slope visual response function. Neural processing tends to compensate for glare.[7] Models of Color in the Natural Scene require the data from all scene elements to predict the effects of glare, receptor quanta catch and neural spatial processing.

Colorimetry's success depends on the minimal glare in its restricted scene. With minimal glare, and its single spot/background edge, it isolates quanta-catch information in its matching measurements. But it cannot not embrace glare's transformation of the scenes radiance falling on receptors, and the neural spatial processing to counteract glare without input data from the entire field of view.

Summary

Recall that camera digits do not record scene radiances. Recall that Colorimetry calculates Matches, while spatial Natural Scene models calculate Appearances. Restricting input data to a single triplet of X,Y,Z values prevents CIE Appearance models from calculating sensations in the Natural Scene.[1] Consider the hidden assumptions in computational models, and their input data.

Acknowledgements

I want to thank Alessandro Rizzi for his work on Coloration Technology's, "Challenges and Open Problems in Colorimetry", and organizing this session. His work and his conversations on the future of Colorimetry is important, and have been very helpful.

References

[1] J. J. McCann, "What scene information is needed for Models of Color Appearance in the Natural World?"Coloration Technology, Nov 22, 2020. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cote.12502>

[2] "A unit of measurement is a definite magnitude of a quantity, defined and adopted by convention or by law, that is used as a standard for measurement of the same kind of quantity. Any other quantity of that kind can be expressed as a multiple of the unit of measurement. For example, a length is a physical quantity." Wikipedia <https://en.wikipedia.org/wiki/Unit_of_measurement>

[3] J.J. McCann, V. Vonikakis, and A. Rizzi,"HDR Scene Capture and Appearance"SPIE Spotlight Tutorial, Chapters 1-15, 2018. <http://spie.org/Publications/Book/2315540?&origin_id=x109925&SSO=1>

[4] J.J. McCann and V. Vonikakis, "Accurate Information vs. Looks Good:Scientific vs. Preferred Rendering", Proc. CGIV, 6, 2012.

[5] W.G. Wright, "A plea to Edwin Land" in Talking about color, Color Research & Applications 12(3), 1987.

[6] J.J. McCann and A Rizzi, "Camera and visual veiling glare in HDR images," J Soc Info Display 15/9, 721-730, 2007.

[7] J.J. McCann and V.Vonikakis, "Calculating Retinal Contrast from Scene Content:A Program", Front. Psychol. 2017 <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02079/full>

[8] J. J. McCann, Retinex Algorithms: "Many spatial processes used to solve many different problems".Electronic Imaging, in Retinex at 50, pp. 1-10(10), 2016. <https://doi.org/10.2352/ISSN.2470-1173.2016.6.RETINEX-017>

[9] E.H. Land and J.J. McCann, "Lightness and retinex theory" J.Opt. Soc. Am., 61, 1-11, 1971.

[10] G. Wyszecki, "Colorimetry", in Color Theory and Imaging Systems. Eynard R, ed., Washington: Soc Photographic Sci & Eng; 24-49, 1973.

[11] J. Albers, Interaction of Color, New Haven ; London, Yale University Press, 1963.

Author Biography

John McCann worked in, and later managed, the Vision Research Laboratory at Polaroid from 1961 to 1996. He currently consults and continues to do research on color. He has studied human color vision, digital image processing, large format instant photography and the reproduction of fine art. His publications have studied Retinex theory, color from rod/Lcone interactions at low light levels, appearance and intraocular scatter, and HDR imaging. He is a Fellow of the Society of Imaging Science and Technology (IS&T) and the Optical Society of America (OSA). He received the SID Certificate of Commendation. He is the IS&T/OSA 2002 Edwin H. Land Medalist, and IS&T 2005 Honorary Member. He is past President of IS&T and the Artists Foundation, Boston. He served as Secretary of the Inter-Society Color Council, the USA Member body of AIC. He has spoken at Electronic Imaging meetings since 1988.