Title
Why We Should Report Colorimetric Data In Every Paper

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Veale, Richard E
Kawashima, Takahiko
Okada, Tomohisa
et al.

Publication Date
2021

Peer reviewed

Why We Should Report Colorimetric Data In Every Paper

Richard Veale[1], Takahiko Kawashima[2], Tomohisa Okada[3], and Jun Miyata[2]
[1]Department of Neurobiology, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[2]Department of Psychiatry, Graduate School of Medicine, Kyoto University, Kyoto, Japan
[3]Human Brain Research Center, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Abstract

This is a modern horror story about an innocently misbehaving projector, and why we beseech everyone to report minimal colorimetric data about stimulus displays. We present anecdotal experience of configuring a projector to display video stimuli in a high-tesla MRI room, along with all the gotchas, (broken) technical assumptions, and theoretical rehashings that should be considered by every scientist who uses computers to display visual stimuli. The moral of our story is: (1) check that your monitor/projector is actually showing the colors and luminances that you think it is, (2) make explicit assumptions regarding the physical/perceptual space of your stimuli and how they relate to any model analysis you will perform. This is especially important when modeling non-human animals, since most equipment and data formats implicitly assume human perception. We show that innocent changes to display settings such as brightness reduction can cause dangerously unexpected results. Understanding and reporting colorimetric data in scientific publications is important for two reasons: (1) reproducibility, and (2) model fidelity.

Keywords: color space; saliency; animal perception;

Introduction

Humans have the amazing ability to percieve visual properties such as color in an invariant manner. We innocently accept significantly different physical stimuli as identical. For example, you immediately recognize the picture displayed on your laptop's screen and the same picture displayed on your mobile phone screen, even though each different device displays the image using different combinations of wavelengths of light and physical fluxes of light. These practical differences raise issues when stimuli are used for psychophysical or perceptual tasks, and when comparing human responses to the behavior of other animals such as marmoset or macaque monkeys (Chen et al., 2021). Assuming without evidence that other animals will perceive visual stimuli in the same fashion as humans is a common error (Stevens & Cuthill, 2005). Indeed, the technical fields of monitor engineering, calibration, measurement of luminance, and translation between color spaces are predicated on theories and models specifically created for *human* perception (Broadbent, 2004; Stiles & Burch, 1959).

Thus, for reproducibility in both human and animal studies, it is important to ensure that stimuli are similar along *relevant task dimensions* to the original setup. This is difficult to quantify, and the best option – using identical equipment – is often untenable as equipment manufacturers change technology over time. There is no great solution to this issue, we argue that it is of key importance for a scientist to be aware that the visible spectrum and properties can differ between animal species – and account for it in any attempted experimental reproductions or modeling.

(a) Target: sRGB (b) Orig. bright=-24 (c) Calibrated

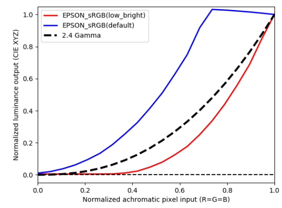
Figure 1: Video frame before/after calibration.

We first motivate our argument using an anecdote: the projector in a university-administered fMRI setup has been used for years under "default" settings. These settings turned out to be degenerate. We briefly review the theory behind how color and luminance are defined for visual stimulus displays. This is relevant for scientists building models to predict behavior using visual stimuli (in both humans and non-humans). An example of this modeling is presented in the final section: we present an example from recent work ((Chen et al., 2021)) in which we applied saliency map model analysis to video stimuli shown to different animal species (human, macaque monkey, marmoset monkey) and predicted gaze behavior. In this paper, we show how mistaken assumptions about the color space of the stimuli can lead to different modeling results.

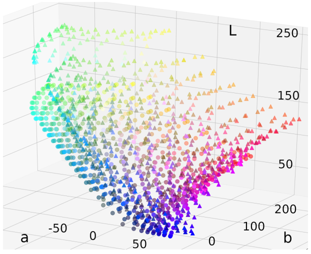
In light of this, we propose minimal procedures for verifying the sanity of one's visual stimulus display configuration. We also recommend reporting a minimal set of display calibration data along with behavioral data. We hope that this will lead to an improvement of reproducibility both in human and animal studies, as well as more correct application of image processing and perceptual models, by encouraging full knowledge of the assumptions and limitations of the data and machines we use daily.

1. Computer displays are not always well-behaved, and may transform identical computer image data in unexpected ways. Solution: report visual display (colorimetric) calibration data. Colorimeters operate on assumptions predicated on human-biased color spaces, but are cheap, quick, and force one to consider problems with one's visual stimuli.

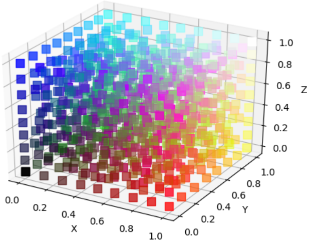
2. Colorimetric data is sufficient to guarantee reproducibility in most human perceptual studies, but fails to capture stimulus properties that are relevant in some non-human animal species or humans with types of color-blindness. To achieve a species-agnostic method of reporting visual stimuli, one needs to measure the spectral power distribution (SPD) over a wide visual spectrum. For example,



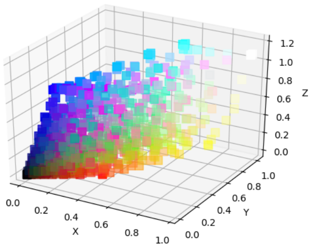
(a) Luminance transfer functions for two default projector settings.



(b) Absolute CIE Lab values of projector gamut before calibration (circles) versus reference laptop (triangles).



(c) Input RGB values for measuring projector transfer function.



(d) Output (normalized) CIE XYZ for calibrated projector.

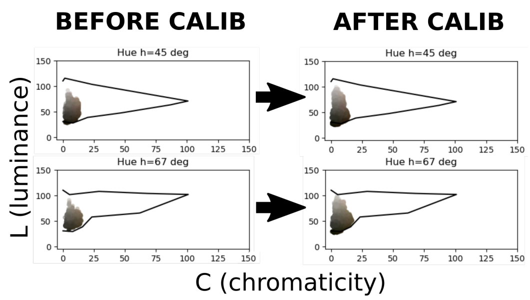
Figure 2: Projector input-output mapping

Gagin et al. (2014) used a DP-655 spectroradiometer in a exemplary human and macaque study. Spectroradiometers are expensive (10000 USD) and time-consuming. An option is to request that manufacturers provide measurements of the spectral power distributions of their products. Solution: there is no good solution for how to exactly reproduce the correct spectral power distribution using a different display for animal experiments other than measuring the color-matching function of that species and finding a monitor that can cover it. The easiest option is to ensure that the SPD of the reproducing display closely matches the SPD of the original display.

3. Computational models that operate on image or video files often interpret the image/video data incorrectly. This is caused by confusion regarding the relationship between computer data and perceptual/physical properties that the data represents such as SPD, brightness and/or color. Modelers often do not state what the input of a model is intended to represent. Solution: Explicitly address the intended meaning of visual inputs to models.

Motivation: A Misbehaving fMRI Projector

We use a 7-Tesla fMRI scanner (Siemens Healthcare, Erlangen, Germany). Visual stimuli are back-projected onto a grey acrylic screen inside of the MR scanner bore from outside the MRI room using an Epson EB-G5100 (Tokyo, Japan) projector with a long focus zoom lens (ELPLL06). We presented video stimuli to subjects while recording eye position using an EyeLink 1000 eye tracking system system (long distance mount – SR Research, Ottawa, Canada). Pilot subjects reported that some video stimuli (Figure 1(b)) were too dark. Previous subjects on the same machine viewing achromatic (white/black) fixation and text stimuli did not report any such issues. The projector was confirmed to be in default sRGB mode using low lamp power setting, with only one divergence from default values: "brightness" parameter was set to its lowest possible value (-24) because the default brightness level (0) caused discomfort.



(a) Distribution of pixels in CIE Lch space for Figure 1(a) (before calibration) and (Figure 1(c)) (after). Black lines estimate projector gamut borders.

Figure 3: Gamut of uncalibrated project/target.

After resetting the the projector to default settings (sRGB mode), we hypothesized that the dark stimuli were due to properties of the gray acrylic screen, incorrect transformations in the stimulus presenting computer, or color gamut limitations of the projector itself. We used a datacolor SpyderX colorimeter (New Jersey, USA) to enumerate the input-output mapping of the projector, i.e. between input pixel byte value ($[0,255]$) and the output luminance (cd/m^2).

We used "dispcal" software (`https://www.argyllcms.com/`) to enumerate the input-output mapping of the projector at high granularity. Figure 2(c)-Figure 2(d) shows normalized CIE XYZ output values for a tiling of the 3-dimensional input domain (R,G,B). Peak output luminance was unchanged (820 cd/m^2) regardless of projector "brightness" setting, indicating that the curvature or offset of the transfer function was changing, not the minimum and maximum outputs.

We further visualized the marginal input-output functions for achromatic and monochromatic stimuli.[1] The normalized curves are plotted for the original setting, as well as for the projector's default sRGB setting, in Figure 2(a) against a target (scaled sRGB) curve. Observe that under default settings, the output is nearly zero for almost the entire lower half of input values – the left half of the horizontal axis. This explains the "dark" videos reported by subjects.

Disabling sRGB mode revealed additional configuration options in the projector's software menu, such as color temperature and component channel (R,G,B) weighting. Using "dispcal", we adaptively modified settings until the input-

[1] achromatic: $R = G = B$ i.e. grayscale; monochromatic: one component channel is stepped from minimum to maximum while the other two channels locked at zero.

output mapping approached the sRGB standard (white point correlated color temperature identical to black body at 6500 K (D65), effective gamma of 2.4). Best-fit settings were: photo mode, 0 brightness, -24 contrast, 5000 K color temperature, -10 red weight, -4 green weight, +12 blue weight.[2]

After calibration, we confirmed that visual stimuli displayed as expected (Figure 1(c)). Pilot subjects reported videos on the projector looked similar to the same video displayed on a reference screen (razer blade stealth 13 laptop (2019), Figure 1(a)). Figure 3(a) visualizes how the distribution of pixel values changed after calibration. Pixels are plotted in CIE Lch space (Lightness, Chromaticity, Hue), with two hue angle slices shown. Black lines indicate SMBGD estimation of the projector's gamut border (Morovič & Luo, 2000). The example image was one of the least visible stimuli under the original (dark) projector settings.

The moral of our anecdote: check colorimetric properties of your stimulus display device. Monitors and projectors have default modes named "standard RGB", but these can be degenerate and cause unreproducible experimental results if one does not take care. Modern colorimeters are cheap and automated with free software, and a display need only be calibrated once. For dissemination, we recommend enumerating the colorimetric data for displays used in previous experiments along with visual stimulus image/video files, so that posterity can estimate what physical stimuli subjects were actually seeing. Calibration data can be reported as a table of input RGB values (sampled in steps of e.g. 4% of the input space) and output (unnormalized) CIE XYZ values in cd/m^2. We recommend *against* reporting only "gamma" values and minimum/peak luminance, since the effective gamma reported by colorimeters can be misleading if the input-output curve is degenerate such as the red or blue lines in Figure 2(a).

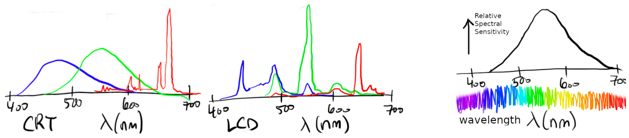
In addition to being important for reproducing experimental results, proper handling of the color space of visual stimuli is of vital importance to accurate comparative animal research as well as computational modeling. We next review the history of color spaces and visual perception for these purposes.

Background: A Review of Color And Light

A visual scene in the natural world is a spatio-temporal pattern of propogating electromagnetic waves, which can also be modeled as quanta (photons). A photon has energy proportional to its wavelength. A visual scene will naturally be comprised of many photons of many different wavelengths. Certain wavelengths ($370 - 700$ nanometers for humans) interact with molecules (opsins and retinol) packed in receptor cells in human retina, causing nerve signals to propogate when photons are absorbed. Humans have four receptor types: L-cone, M-cone, and S-cone cells, and Rod cells (Figure 5(a), "human") (Jacobs, 2009). The names are based on their shape (cone-shaped or rod-shaped) and the relative wavelength of visible light to which they are most sensitive. Rods are engaged primarily in scotopic (night) vision. In contrast, cone cells are active during photopic and mesopic (daytime, dusk) vision. L-cones are most sensitive to long wavelength light (which looks red when viewed monospectrally), M-cones to medium (green), and S-cones to short (blue).[3]

Other animals (or humans with genetic mutantations – "color blindness") express different sets or numbers of receptors with different spectral sensitivities (see "Predator" in Figure 5(a)), and thus different spectral distributions will be ambiguous or discriminable to different species.

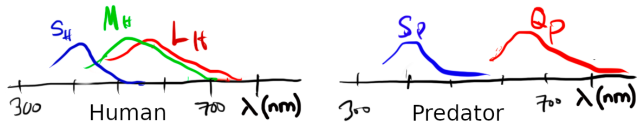
[2]Additionally: PC operating system may weight component (RGB) channels. In Ubuntu 18.04, we confirmed no weighting using "xrandr" – R, G, B weights and gammas were all 1.0. Furthermore, xrandr confirmed "Expanded (full RGB)" mode for the HDMI connection to the projector. Under the alternative "TV (limited RGB)" mode, the PC will clip RGB byte inputs to range $[16, 235]$ (instead of expected $[0, 255]$) before sending them to the display.



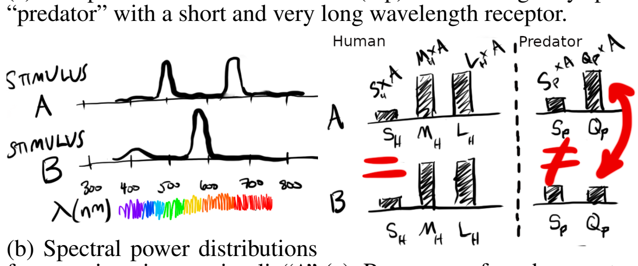
(a) Spectral power distribution emitted by each of the 3 basis elements of a CRT and LCD monitor.

(b) Sketch of human luminosity function

Figure 4: Monitor spectra, human luminosity function



(a) Receptor sensitivities for human (top) and an imaginary species "predator" with a short and very long wavelength receptor.



(b) Spectral power distributions for two imaginary stimuli "A" and "B".

(c) Response of each receptor type of human and the predator to stimulus A and B.

Figure 5: Two stimuli can be distinguishable to one species (the predator) but look identical to another (human).

The intensity of a receptor's response to a stimulus is calculated by convolving the spectral power distribution of the beam of light with the spectral sensitivity function of the receptor. An example is given for two example SPDs "A" and "B" in Figure 5(b). The corresponding neural response of receptor cells with sensitivities defined in Figure 5(a) is cartooned in Figure 5(c). Even though "A" and "B" are mixtures of entirely different wavelengths of light, both stim-

[3]A receptor class responds with different sensitivity to a range of photon energies. Sensitivity distribution is determined by quantum properties of the opsin molecules expressed by the receptor.

uli cause identical responses in human receptor types (Figure 5(c), left). This phenomenon is "metamerism", a type of anti-aliasing resulting from the visual system sampling a high-dimensional space using only three basis functions. Different basis functions change which spectral distributions look the same and which look different (A and B look different to "Predator", Figure 5(c)).

Most extant technology (CRT, LCD monitors) takes advantage of metamerism. Display devices use three light-emitting/transmitting components (e.g. Figure 4(a)), each of which emits a a spectral distribution of light that peaks in a unique region of the visual spectrum. These primaries can be mixed together in appropriate proportions to produce perceptual stimuli that look the same as any desired target spectral distribution.

Luminance, Color Spaces, Color Matching

Full enumeration of the spectral power distribution of a stimulus includes wavelengths not visible to humans (radio, infrared, ultraviolet). To standardize a function of spectrum that matches human perception of luminance, the international commission on illumination (CIE) published a "luminosity function" (photopic function sketched in Figure 4(b)) which measures *candelas* (cd). For visual displays, we report the luminous intensity per unit surface area, i.e. cd/m^2.

The luminosity function is related to the CIE RGB and CIE XYZ color spaces. Scientists defined human "color-matching function" by having subjects mix three mono-spectral "primary" light sources in different proportions to match mono-spectral light targets (Guild, 1931; Wright, 1929). This effectively quantified the metameric partition function for human. CIE standardized three mono-spectral primaries: r at 700 nm, g at 546.1 nm, and b at 435.8 nm and published color-matching functions for those primaries: $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, $\bar{b}(\lambda)$, obtained via transformation from the empirical color-matching results of different primaries. The color matching functions define the proportions of the three primaries necessary to match any given mono-spectral light of wavelength λ. CIE standardized the the response of the "average" participant, creating the "CIE 2° standard observer".[4]

Using these color matching functions, arbitrary spectral power distributions $S(\lambda)$ can be reproduced using a mixture of only the three primaries in different proportions. The three linear gains R, G, B that match SPD $S(\lambda)$ are achieved by $R = \int_0^\infty S(\lambda)\bar{r}(\lambda)\mathrm{d}\lambda$, $G = \int_0^\infty S(\lambda)\bar{g}(\lambda)\mathrm{d}\lambda$, $B = \int_0^\infty S(\lambda)\bar{b}(\lambda)\mathrm{d}\lambda$.

This is CIE RGB. However, $\bar{r}(\lambda)$ included negative values[5]. So, CIE developed CIE XYZ, a color space formed using imaginary primaries X, Y, and Z and color-matching functions such that certain desirable properties are satisfied. Specifically, the color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ are never negative, and the tristimulus value Y corresponds to the luminous intensity of the stimulus.

[4] Colors were matched within a circle two degrees of visual angle in the center of the visual field.

[5] I.e. target wavelength unmatchable without adding red primary to the target.

Illuminants and White Points

White light can be formed of many different mixtures of wavelengths. CIE published "standard" illuminants, based on defined physical systems such as a specific type of metal (tungsten, mercury) at a specific temperature. It is also common to use the spectrum emitted by an idealized blackbody radiator, whose SPD is defined by Planck's law. CIE published standards to simulate outdoor illumination. For example "D65" (used for sRGB standard below) is the spectral power distribution of average direct and diffuse sunlight at noon in Western/Northern Europe. D65's correlated color temperature[6] is 6500 Kelvin. The triplet of primaries (e.g. $[R,G,B]$, $[X,Y,Z]$) that color-match a chosen white illuminant are called the "white point" or "reference white". Defining a reference white is necessary to convert between color spaces, since it defines the "pivot" around which rotations or scalings occur. An absolute white point (in cd/m^2) is necessary to map from a color space to physical luminances.

Visual Stimuli for Computers: sRGB

The combination of a perceptual color space, a fixed white point, an absolute luminosity, and an environment with controlled ambient illumination is sufficient to uniquely determine the perceptual response that a color stimulus will elicit. An example is "standard" RGB (*sRGB* – not CIE RGB). sRGB defines white point (D65), absolute luminance (80 cd/m^2), environment (CRT monitor in dim office environment). sRGB primaries (r, g, b) are defined via CIE XYZ. Most modern computer monitors are calibrated to display sRGB or similar, with varying absolute luminance. This means that image and video data stored on computers (JPEG, PNG, TIFF) is usually implicitly encoded in sRGB.

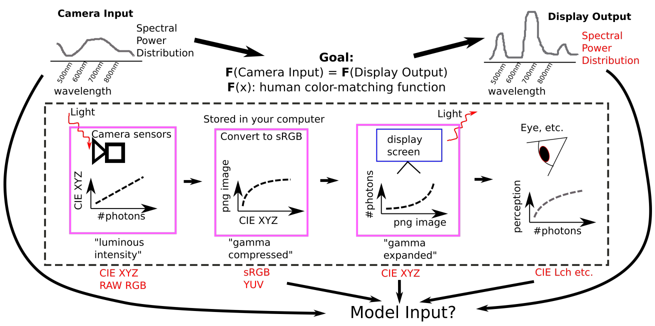


Figure 6: Image data in modern computers. In cameras, light is detected by (three) sensors and converted to internal color space via color-matching functions. Internally, luminance is compressed to more efficiently cover space of human discriminable luminances. Displays use three light-producing elements to emit mixtures of light perceptually indistinguishable from the original, despite comprising physically different distribution of spectral powers.

[6] temperature of blackbody radiator it color-matches

Gamma (non-linear encoding)

For CIE XYZ or CIE RGB, a linear increase in the tristimulus value corresponds to a linear increase in the luminous intensity of the SPD generated by that tristimulus point. Other color spaces do not share this property. For example, sRGB is gamma-compressed to take advantage of the fact that humans are more sensitive to contrast at low luminance than at high luminance. sRGB transforms the linear physical intensity represented by a raw tristimulus value (of arbitrary physical linear primaries, e.g. from a camera image sensor) via a function like $f_{save}(x) = x^{\frac{1}{\gamma}}$. γ is usually $\in [1.8, 2.4]$. It then stores the the "gamma-compressed" value. The reason for this is that devices on which the image data will eventually be displayed are expected to behave such that the mapping between input values and emitted light intensity will implement the inverse of that function (i.e. $f_{display}(x) = x^{\gamma}$ ("gamma expansion"). In this way, fewer bits of information can be used to span the range of brightness in perceptually evenly spaced steps. This works because human perception of luminance is non-linear, following a power law (Stevens's power law) relatinship between physical luminous intensity and apparent magnitude with exponent $1/3$. Gamma-compression and gamma-expansion cancel each other out – but ensure that the *representable space* of luminances of a given data format (e.g. byte) fall at perceptually pleasing, equal intervals.

sRGB is Not a Perceptual Color Space

Thus, gamma correction only influences the distribution of values that are represented by a signal with limited bandwidth. sRGB is not intended to represent the non-linear perceptual response to stimuli. Rather, it is a standard to ensure that the set of representable colors and brightnesses is standard across physical displays (Figure 6). An image in sRGB color space is guaranteed to look identical *to standard humans* displayed on an sRGB calibrated monitor in the proper environment. Note that it will only be perceptually identical to a human standard observer – no such guarantee is made for animals or humans who diverge significantly from the standard observer on which sRGB is defined.

There do exist "perceptual" color spaces which purport to model the subject human perceptual response to colors. For example, CIE Lch, CIE L*a*b*, and CIE Luv represent "Luminance" (L) of a stimulus as a non-linear mapping of linear luminance Y (of CIE XYZ), to account for the non-linearity of human perception. The goal of these color spaces is to model visual stimuli along natural perceptual dimensions. For example, one should be able to only change the perceived hue of a stimulus without modifying its brightness, chromaticity, or saturation, simply by moving along one axis.

Summary: Common Mistakes with Computer Data

1. Non-linearities in data representation – image and video files are usually gamma-encoded and must be passed through an inverse exponential function to recover the linear physical luminances that would be produced by a computer monitor displaying the image or video.

2. Incorrect assumptions about the brightness of stimuli – modelers often misunderstand the relationship of tristimulus primaries (R, G, B) to both perceived and absolute luminance (Nguyen & Brown, 2017).

3. Incorrect assumptions about the meaning of stimulus color spaces – modelers mistakenly assume that commonly used color spaces (RGB, XYZ, Lab, Luv, YUV, CMYK, etc.) uniquely indicate physical visual stimuli. These color spaces are specifically designed to model *only the set of visual stimuli consciously discriminable to humans*. They incorrectly predict the perceptual response of organisms with different photoreceptors or visual processing neural circuits. Corollary: human technology (computer display screens, printed materials, etc.) is incapable of displaying a consistent visual perceptual space of either color or brightness to organisms with perceptual systems different from the standard human.[7]

Results: Saliency Model Input Assumptions

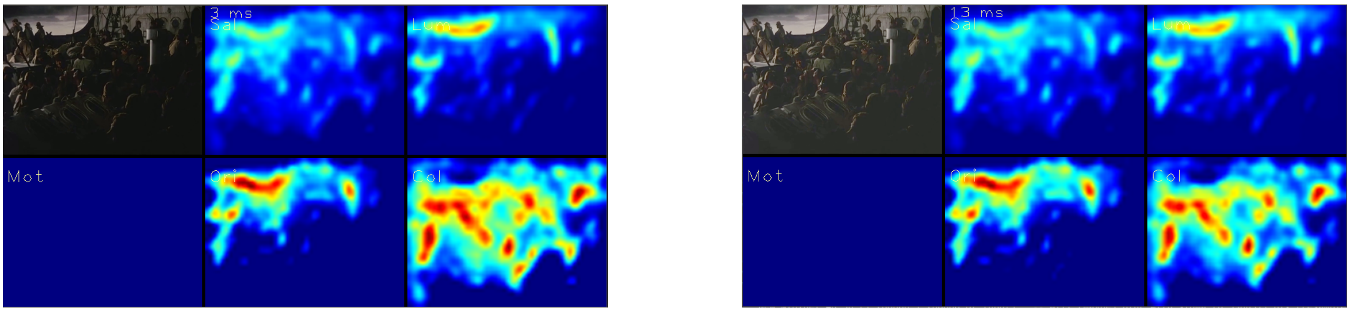
Different assumptions about the input image format lead to different model performance in a model of visual attention we recently reported to predict looking behavior of humans, marmosets, and macaque monkeys (Chen et al., 2021). This is meant to be an example of how an incorrect assumption on the part of the modeler can be responsible for different results, damaging reproducibility and explanatory power.

The saliency map model (Itti & Koch, 2000) is a well-worn model of bottom-up visual attention operating on pixel images. The color space of the images are never made explicit. The saliency map is used as a surrogate of bottom-up visual attention in humans and non-human primates. Recently, correlates of saliency have been identified in parts of the brain related to visuo-ocular control (superior colliculus (White et al., 2017; Veale, Hafed, & Yoshida, 2017)).

The saliency map model extracts visual features of the input such as luminance, color, motion, and oriented lines, and identifies parts of the input where those visual features differ from their surroundings at many spatial scales. It uses center-surround differences of the features extracted at different spatial scales, as well as within-map competition and normalization. Finally, the feature maps (orientation, color, motion, etc.) are combined into a feature-agnostic saliency map.

An overview of the saliency map model and how it would process an example input image (Figure 1(a)), is shown in Figure 7(a). In this example, the default implementation and assumptions of the saliency map model are used for computing color and luminance Itti and Koch (2000). The model is

[7] The representable perceptual gamut will be uneven and include gaps of unrepresentable colors and brightnesses. This can occur even if an organism has the same *number* of photoreceptors, so long as their spectral sensitivity profiles diverge significantly from the standard human.



(a) Input sRGB gamma-compressed pixels (default).

(b) Input pixels are assumed to be linear.

Figure 7: Analysis of saliency map response to different assumptions about input.

implemented via the "salmap_rv" library.[8] Image data pixel values are used to directly compute color and luminance, without gamma-expansion nor color channel weighting. A first question: is it appropriate to use the gamma compressed values, or should they be linearized first? In addition to being gamma-compressed, pixel values implicitly incorporate sRGB channel weightings. We know that the sRGB color space is defined in perceptual linear CIE XYZ space, a space specifically designed using human judgement. Is it appropriate to use this human-biased input for fitting the behavior of marmoset monkeys, for example? We will investigate how model output changes given different input assumptions.

The second image (Figure 7(b)) shows the response of the model to linear RGB input (i.e. with gamma removed and channel weightings removed for luminance calculation). Small differences are visible in the luminance (top right) and color (bottom right) channels.

Our primary purpose in this paper is to reiterate the importance of explicitizing assumptions about the format of inputs to one's models, especially when those models and experiments relate to non-human animals. For the saliency map model, different assumptions regarding the input format lead to different outputs (Figure 7(a) versus Figure 7(b)), although broad qualities of the model output were unchanged. This may be because the test image is an abnormal outlier – it was initially selected for being too dark to discriminate before calibration but visible afterwards (Figure 3(a) – mostly low chromaticity). One can imagine that more extreme changes, such as changing the color space of the input, or using a model which takes raw spectral power distributions as input, would cause more extreme divergences. For example, one might argue that the saliency map should take perceptual-space stimuli as input. Inputs would be transformed to e.g. CIE Lch color space, and an appropriate feature detector added to identify local differences in hue or chromaticity. The nonlinear transforms implicit in the conversion to Lch can be expected to likewise result in different model output, since color and luminances (which form the basis of orientation and motion detectors) will be distorted. On the other hand, one may not see a large change, since the original input, sRGB, is already specialized to represent human perception, and conversion to Lch will not distort as much as expected. A more extreme conversion, for example to a color space predicated on the color-matching function of a different animal species or a color-blind subject, may result in clearer divergence of model output. Macaque monkeys may have a color-matching function almost identical to human ((Gagin et al., 2014)). This is not true for the common marmoset, which is usually dichromatic (Freitag & Pessoa, 2012).

[8] `https://github.com/flyingfalling/salmap_rv`

Conclusion: Report Colorimetry, Be Aware

Image data does not usually contain sufficient information to reproduce a physical stimulus exactly. To do so would require knowledge of the spectral power distribution for every point of the stimulus. Rather, our cameras sample physical stimuli using low-dimensional basis functions (primaries). This turns out to be sufficient to elicit a perceptual experience indistinguishable from the original stimulus using the standard human color matching function. Thus, for a human, a picture of a forest looks the same as if one were actually looking at the forest. However, this is not guaranteed for other animal species unless one tailors the color space of the image to the characteristic color-matching function of that species (which is usually unknown).

The transfer function of a visual stimulus display device for experiments can be quantified via colorimeter or spectrometer. This is a reproducible starting point with minimal assumptions, and will greatly improve the reproducibility of human visual experiments. While spectroradiometer data is preferable since it can be used for animal experiments, colorimetry data may be sufficient in the future as the color-matching functions of more species are quantified. Unfortunately, there is no good way to reconcile the fact that display parameters will be biased towards human perception. Our best hope is to be careful to ensure that visual stimuli look as expected to the visual systems of subjects, be they humans or animals. Finally, for modeling, we caution scientists to be wary of the format of visual stimulus data, and to make explicit decisions regarding the meaning of data used as input to computational models.

Acknowledgements

Research supported by JST COI, AMED Brains/MINDS, JSPS Grant-in-Aid for Scientific Research (KAKENHI) (21K15609) to RV, and Grants-in-Aid for Scientific Research (KAKENHI) 20H05064 (MEXT) and 17H04248 (JSPS) to JM.

References

Broadbent, A. D. (2004). A critical review of the development of the cie1931 rgb color-matching functions. *Color Research & Application*, *29*(4), 267–272.

Chen, C.-Y., Matrov, D., Veale, R., Onoe, H., Yoshida, M., Miura, K., & Isa, T. (2021). Properties of visually guided saccadic behavior and bottom-up attention in marmoset, macaque, and human. *Journal of Neurophysiology*, *125*(2), 437-457.

Freitag, F. B., & Pessoa, D. M. A. (2012). Effect of luminosity on color discrimination of dichromatic marmosets (callithrix jacchus). *JOSA A*, *29*(2), A216–A222.

Gagin, G., Bohon, K. S., Butensky, A., Gates, M. A., Hu, J.-Y., Lafer-Sousa, R., … others (2014). Color-detection thresholds in rhesus macaque monkeys and humans. *Journal of Vision*, *14*(8), 12–12.

Guild, J. (1931). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *230*(681-693), 149–187.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*, *40*(10-12), 1489–1506.

Jacobs, G. H. (2009). Evolution of colour vision in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1531), 2957–2967.

Morovič, J., & Luo, M. R. (2000). Calculating medium and image gamut boundaries for gamut mapping. *Color Research & Application*, *25*(6), 394–401.

Nguyen, R. M., & Brown, M. S. (2017). Why you should forget luminance conversion and do something better. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 6750–6758).

Stevens, M., & Cuthill, I. C. (2005). The unsuitability of html-based colour charts for estimating animal colours–a comment on berggren and merilä (2004). *Frontiers in Zoology*, *2*(1), 1–9.

Stiles, W. S., & Burch, J. M. (1959). Npl colour-matching investigation: final report (1958). *Optica Acta: International Journal of Optics*, *6*(1), 1–26.

Veale, R., Hafed, Z. M., & Yoshida, M. (2017). How is visual salience computed in the brain? insights from behaviour, neurobiology and modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714), 20160113.

White, B. J., Berg, D. J., Kan, J. Y., Marino, R. A., Itti, L., & Munoz, D. P. (2017). Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nature communications*, *8*(1), 1–9.

Wright, W. D. (1929). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, *30*(4), 141.