Visual Photons

Encoding the Formation and Sensation of Retinal Images

author: Jonathan D. Lettvin

written: Apr 01 2015
fetched: May 07 2015

1 Introduction

Whither and whence visual photons needs must uncertain be yet by lawful pattern ruled. Photons are produced, scattered, reflected, refracted, diffracted, incident, and aborbed by a retina in "bleaching" events. Daylight photons result from solar blackbody emission with notches from atmospheric absorption. Surviving photons are scattered by atmosphere, reflect off objects, refract through a cornea, diffract through a pupil, and are absorbed by plural species of photoreceptors. Each of these processes contribute to spectral transforms.

Laws governing perceptual invariants such as contrast and color invariance must arise from quantitative retinal photon absorption differentials from a typical scene. Historically, image synthesis depends on use of RGB components usually produced by LEDs. Notwithstanding LMS retinal cone species having similar spectral peaks, images produced as if by a digital camera achromatic lens with an Airy disk smaller than one pixel are not suitable for analyzing images produced by a chromatic natural lens and Airy disk always larger than a human photoreceptor. The subject of this paper is production of synthetic image encoding suitable for broad spectrum analysis of human visual invariants.

Two transforms dominate retinal image formation; corneal refraction and pupillary diffraction. Corneal refraction is modelled by an affine expansion $\propto 1/\lambda$ where short wavelengths are expanded more than long wavelengths. Pupillary diffraction is modelled by an Airy kernel convolution where the kernel size is $\propto \lambda$. In other words, for a white object on a black background oberve a blue tinge ringing the object as seen through a magnifying glass and a red tinge ringing it as seen in the Airy disk from a pinhole.

Both of these transforms apply to mathematical models of images formed from photons counting in the millions or billions. The affine transform from refraction is modelled simply enough. Analysis of diffraction and synthesis of suitable diffraction images from very low to very high photon counts will be the primary focus of this paper. Images produced and encoded using these methods will enable further research into how retinal processing achieves invariance.

Point Source Diffraction

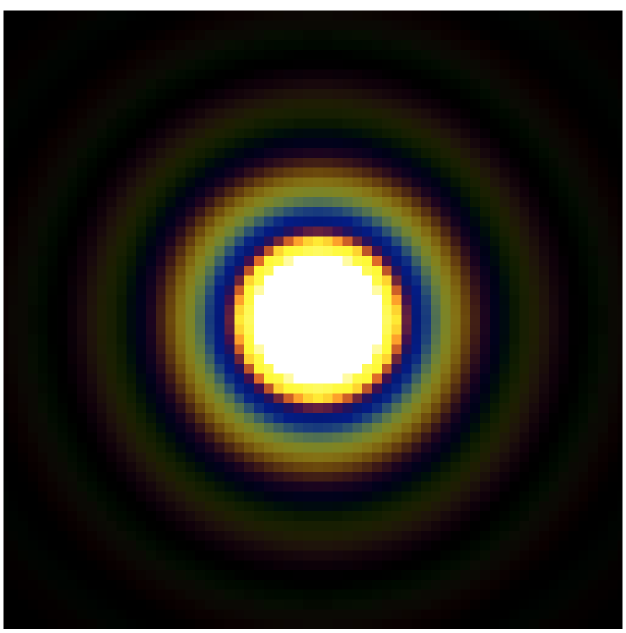


figure 1: Airy pattern

When a very large number of photons from a point source pass through an aperture and then fall on a surface they accumulate in a diffraction pattern called an Airy pattern figure[1]. This intensity pattern is the square of the photon wave function equation[6]. Photon densities are properly synthesized by summing the wave equations for all points over all wavelengths and squaring the result.

This intensity map is ruled by equation[4]. This formula calculates idealized continuous distribution from which deviation is unmeasurable at high photon counts. Discrete photons do not form a continuous set. Nor does a grid of sensors in a focal plane in an eye or camera. The zoomed "pixels" in figure[1] illustrate a grid for billions of RGB source photons for a centrally located point source.

Close examination reveals that different wavelengths contribute differentially to a pattern at different radii. Photographic and visual colors are typically thought to operate using RGB. The names (R)ed, (G)reen, and (B)lue, as wavelengths, are typically thought to be related to color perception. In fact, color perception is independent of wavelength Lettvin et al[3]. Scientific labelling for RGB in retinal photoreceptor absorption is LMS for Long, Medium, and Short. However, since photography and displays use RGB, they will be used interchangeably.

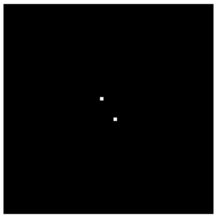


figure 2:
2 white points

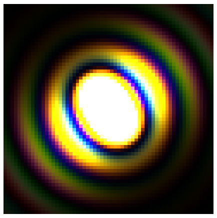


figure 3:
Retinal image with a 1mm pupil

The retinal image of two white RGB point sources incident on a human retina after traversing a 1mm pupil appears as an inchoate blur incapable of being resolved by Rayleigh criterion, Sparrow limit, or even Nyquist sampling limit in a Fourier transform of intensity.

Wavelength dependence in figure[3] is even more apparent than for figure[1]. This is the character of image for narrow band RGB illumination. This paper will explore the problem of forming retinal equivalent incident photons using broadband illumination and methods for producing suitable sensed images for analyzing how sensory data for vision is presented. Even in figure[3], enough photons are present to give the image a seemingly continuous character.

Low Photon Count Diffraction

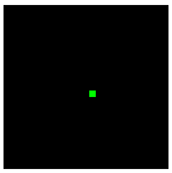


figure 4:
1 photon

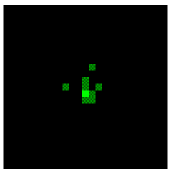


figure 5:
10 photons

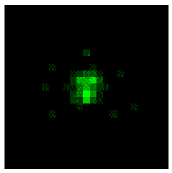


figure 6:
100 photons

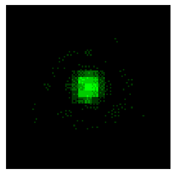


figure 7:
1000 photons

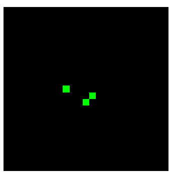


figure 8:
3 photons

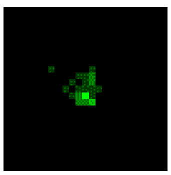


figure 9:
30 photons

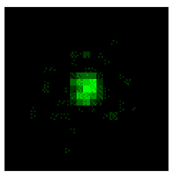


figure 10:
300 photons

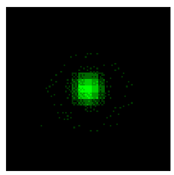


figure 11:
3000 photons

Suppose one views a 0.1mm emerald on a black velvet background pad on a brightly lit white counter. The photon count contributed by the emerald is very low. The observer's pupil is constricted by high ambient light and the act of scrutinizing for detail.

When photon flux is low, the Airy pattern becomes more difficult to discern as can be seen in the models of accumulations of single monochromatic photon events in figure[4] to figure[11].

As can be seen, even at 3000 photons from a point source the Airy pattern is yet to assume the continuous character achieved at much higher photon fluxes experienced when observing less carefully constructed scenes in bright light.

Vision still works fairly well at low light levels so whatever explanation is made for vision must account for the ability to convert patterns such as these to good enough quality internal reconstructions of the scene.

It is important to note that, when point sources are eccentric from the center, different wavelength photons refract and diffract to different centers aligned radially from the center; yet the perception of a polychromatic point is of a single point source.

Photons, Spectra, and Blackbodies

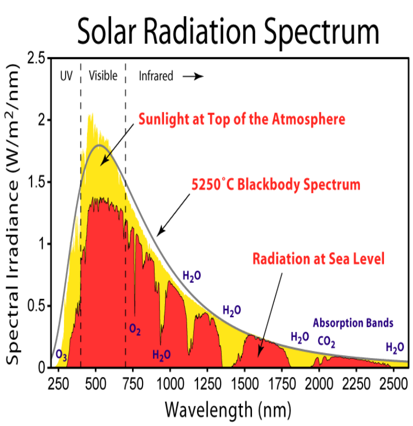


figure 12: Solar Spectrum (wikimedia commons)

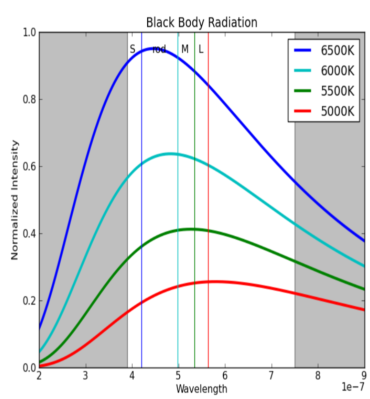


figure 13: Blackbody radiation (visible wavelengths are within the white region of the figure. The peak absorption of typical human photoreceptors are shown as colored vertical lines)

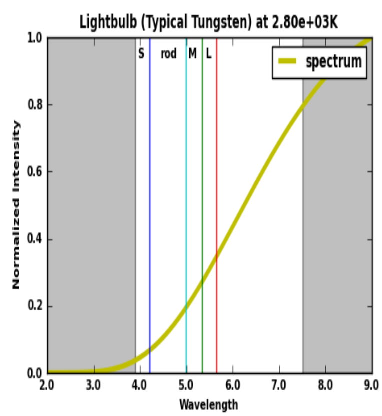


figure 14: Light bulb (at a lower blackbody temperature of 2800K, longer wavelengths dominate unlike sunlight where shorter wavelengths dominate)

Natural lighting is not divided into photons having 3 or 4 energy bands discussed in robotic vision literature. Daylight spectra are considered continuous in that no energies or wavelengths are forbidden.

Typically, ambient light is broad spectrum as in figure[13]. Sunlight resembles the 6500K curve with atmospheric absorption. An incandescent lightbulb peak temperature is 3695K which yields a slope that appears to rise monotonically from the blue end to the red end. The spectrum of figure[14] is typical for a light bulb at 2800K. When moving from sunlight into incandescent light one does not perceive a change in colors. The change of spectrum does not change the colors seen but the size, shape, and spectral composition have changed dramatically by the change in light source.

The two spectra of figure[13] and figure[14] were produced with equation[2]. Blackbody radiation is specified with two parameters, T (degrees Kelvin), and visual angle (or amplitude).

$$I(\lambda, T) = \frac{2hc^3}{\lambda^5} \frac{1}{exp\left(\frac{hc}{\lambda kT}\right) - 1} \quad \textit{Blackbody equation in } \lambda \quad [1]$$

$$I(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{exp\left(\frac{h\nu}{kT}\right) - 1} \quad \textit{Blackbody equation in } \nu \quad [2]$$

where:
h = *Planck constant*
c = *speed of light in a vacuum*
k = *Boltzmann constant*
ν = *photon frequency*
T = *body temperature in Kelvin*

Other spectra like figure[15] and figure[16] are generated using short lists of coefficient triples (a,b,c). Curves in both figures took one or two triples each. This equation is used for convenient curve-fitting.

$$Gauss(x) = a\ exp\left(-\frac{(x-b)^2}{2c^2}\right) \quad \textit{Gaussian function} \quad [3]$$

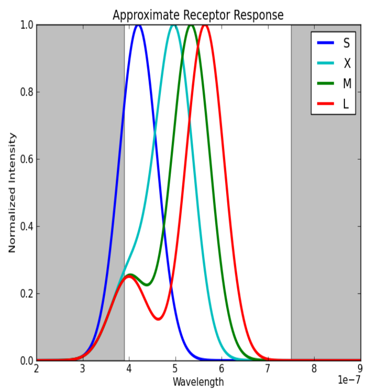


figure 15: Retinal receptor absorption

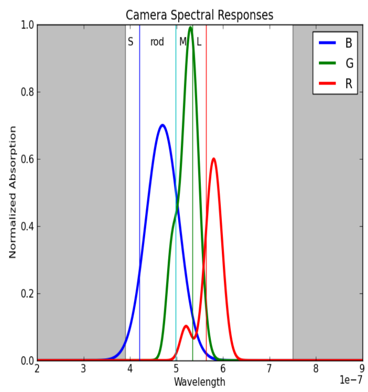


figure 16: Camera pixel absorption

A typical trichromat human has three cone photoreceptors L, M, and S and one rod photoreceptor (X in figure[15]). Human photoreceptors absorb photons having short wavelengths. Absorptions of long wavelength photons have sharp relative cutoffs. S cones absorb almost no photons at wavelengths above ~550nm. M cones cut off at ~625nm, and L cone cuts off at ~710nm. Note a curious increase of L and M cone absorption at shorter wavelengths. All this indicates that wavelengths in a scene are not as sharply divided into color planes as is assumed in standard image processing.

Vision identifies readiness of a resource by its color. Regardless of ambient light, a ripe apple should look red. In fact, it does when viewed under a blue sky, a green forest canopy, a red sunset, or by sodium light having only two yellow and one blue spectral line. Redness of an apple does not depend on presence of long wavelength photons. This is color constancy as described for perception under figure[13] and figure[14].

$$I = I_0\left(\frac{2J_1(u)}{u}\right)^2 \qquad \textit{Airy equation} \qquad [4]$$

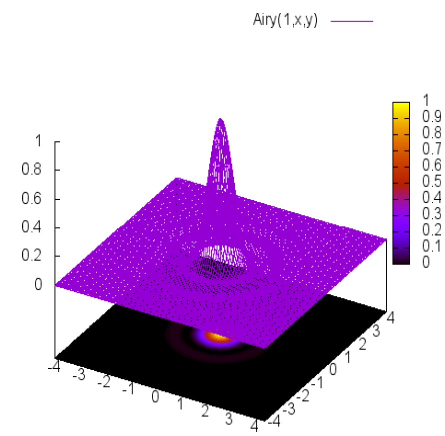
$$u = \frac{2\pi aq}{\lambda R} \qquad \textit{Airy parameter} \qquad [5]$$

where:
J_1 = *Bessel function of the first kind*
a = *aperture radius*
q = *focal plane offset*
λ = *photon wavelength*
R = *distance from aperture to focal plane*

$$A = A_0\frac{2J_1(u)}{u} \qquad \textit{Airy wave equation} \qquad [6]$$

where:
$A_0 = \sqrt{I_0}$

Wave superposition rules when multiple points are being projected onto an image plane. Waves from all points are summed and the result is squared to yield an intensity map (image).



gnuplot 1: Airy(x,y)

The conditions under which this occurs are special. An achromatic lens is required.

The straight lines predicted by optics between this origin identify the center of a probability distribution at the destination. The relative probability of a photon of specific energy from a specific source being absorbed within a specific tile in a grid of a tiled surface is calculated by integrating the infinitesimal probabilities over the tile.

Retinal Photoreceptor Species Mosaic

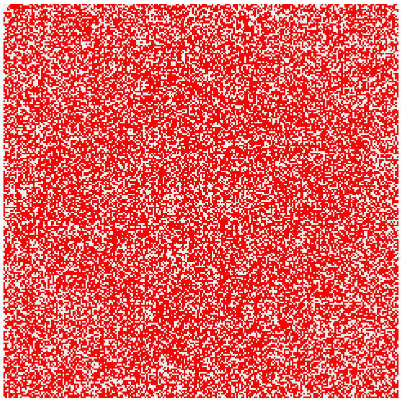


figure 17: Mosaic L cones

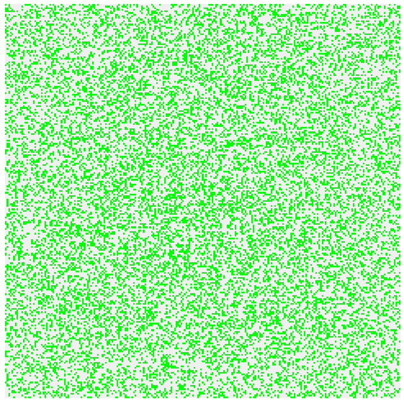


figure 18: Mosaic M cones

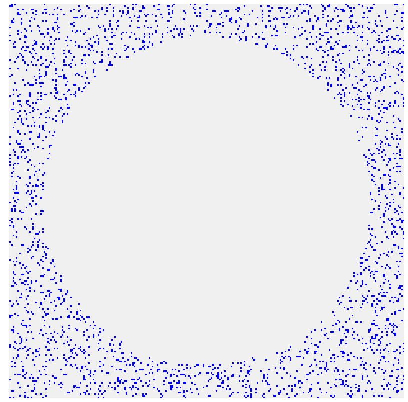


figure 19: Mosaic S cones

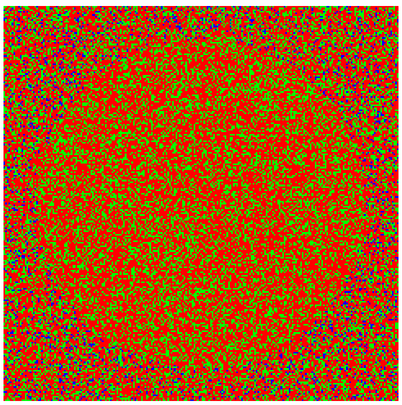


figure 20: Species Mosaic

A human trichromat has four photoreceptor species whose proportion and distribution vary in individuals (Christine A. Curcio and Kenneth R. Sloan[1], David J. Calkins[2]). Our synthetic images are consistent with in vivo data but use a rectangular grid.

Either L cones or M cones dominate a macula. Scarce S cones don't appear within 100μM of center. Rods appear in the outer macula where they increase to 100% population beyond.

As shown in figure[17]-figure[19], effective images for color planes have lacunae. This paper predicts absorption events in full-spectrum scenes.

Mosaic images are generated by PHP code listing[1]. table[1] is typical in human trichromats.

```
<?php
error_reporting(E_ALL | E_STRICT);

function radius($x,$y) { return sqrt($x*$x + $y*$y); }

function species() {
    list($r0, $x0, $y0) = array_fill(0, 3, 120);
    list(     $X , $Y ) = array_fill(0, 2, range(-$r0, +$r0));
    list($P, $M, $D, $rs, $Rmax) = [[], 255, count($X), 100, radius($x0, $y0)];
    $image = ['L' => ''       , 'M' => ''       , 'S' => ''       , 'LMS' => '' ];
    $count = ['L' => 600      , 'M' => 300      , 'S' => 100      , 'I'   => 0  ];
    $cones = ['L' => [$M,0,0], 'M' => [0,$M,0], 'S' => [0,0,$M]               ];
    foreach (array_keys($image) as $k)
        ${'im'.$k} = $image[$k] = @imagecreatetruecolor($D, $D)
        or die("{$k} mosaic fail");
    foreach ($image as $im)
        imagefill($im, 0, 0, imagecolorallocate($im, 240, 240, 240));
    foreach ($cones as $cone => $channel) {
        list($r, $g, $b)     = $channel;
        $$cone               = imagecolorallocate($imLMS, $r, $g, $b);
        ${$cone.'L'}         = imagecolorallocate($imL  , $r, $g, $b);
        ${$cone.'M'}         = imagecolorallocate($imM  , $r, $g, $b);
        ${$cone.'S'}         = imagecolorallocate($imS  , $r, $g, $b);
        if (($N=$count[$cone])) $P = array_merge($P, array_fill(0, $N, $cone));
    }
    foreach ($X as $x1) {
        $x = $x1+$x0;
        foreach ($Y as $y1) {
            $y   = $y1+$y0;
            $r   = radius($x1, $y1);
            $LMS = ($LM  = $count['L'] + $count['M']) + $count['S'];
            $c   = $P[rand(0, (($r < $rs) ? $LM : $LMS) - 1)];
            imagesetpixel($imLMS    , $x, $y, $$c); // Combined mosaic
            imagesetpixel($image[$c], $x, $y, $$c); // Species  mosaic
        }
    }
    foreach ($image as $key => $im) imagepng(    $im, "species{$key}.png");
    foreach ($image as         $im) imagedestroy($im                     );
}

species();
?>
```

Listing 1: Generating the mosaic

Table 1: Photoreceptor Species

L photoreceptors	(cones shown as red)	~566nm peak absorption
M photoreceptors	(cones shown as green)	~533nm peak absorption
I photoreceptors	(rods shown as cyan)	~498nm peak absorption
S photoreceptors	(cones shown as blue)	~433nm peak absorption

Spectral Encoding Constraints

A spectrum, as conceived of here, is a curve in λ to which discrete probability data points may be fitted. Important levels and inflection points must be preserved by conversion back and forth between continuous curve and discrete vector form. An image maps relative probabilities of photon events. In the proposed system, probabilities are calculated from vectors of equidistant values of λ. Each probability value of the vector represents the combined probability at all values of λ along the binning interval from halfway points between values specified in the vector.

The relative probability of a photon event at wavelength λ at point (x1,y1) due to one at point (x0,y0) is the product of all probabilities along all probable pathways. Calculation involves converting all continuous forms of spectra to identically binned λ intervals and multiplying all the probabilities found in a bin.

A new method of encoding spectra for a given retinal pixel is needed. The encoding must support calculation of relative probabilities of absorption. It must support expression of blackbody radiation, monochromatic sources, polynomial spectral shapes, and simple discrete vectors. It must serve for use with any spectral operation used to synthesize an image.

Retinal photoreceptors respond to photons across the entire visual spectrum. The retinal response to a scene lit by sunlight is modelled by computing the dot product of the discrete spectrum presented to a photoreceptor with its absorption spectrum. The point of this paper is to compute that dot product and show the effect this has on each of the cone species mosaics and on the combined mosaics. No knowledge of wavelength survives.

For the vast majority of spectra, it is sufficient to specify a source as either a set of blackbodies or as a set of gaussians. Usually a single blackbody is sufficiently specified as an amplitude and a temperature (two parameters). Usually a single gaussian is sufficiently specified as an amplitude, a peak λ, and a width (three parameters).

Synthetic images are produced by summing wave equations and squaring the result, so it is convenient to use the square root of peak intensity as the amplitude. An image can be fully specified with spectral precision from pixels composed from their spectral components.

Pixels locations may be either specified on pixel groups randomly scattered over the image, or as raster images where the location is implied by its position in the raster line set.

The former permits a full lower-quality image to be delivered early with improving fidelity as more pixels arrive. The latter requires full delivery of all pixels before the full image may be seen. In the former case, a spectral specification is given followed by a sequence of (x,y) pixel coordinates. A 4096x4096 pixel image can use 12 bits per dimension or 3 bytes per pixel. First priority of presentation should be given to generating a neutral background. Second priority should be given to larger differences across boundaries. Last priority should be given to field filling operations.

Digital cameras are unable to produce images of this type since they are restricted to sampling 3 narrow spectral bands. Older photographs are also unable to produce these images for similar reasons. The only feasible manner of production of test images suitable for analysis of vision is synthesis.

As will be shown, acceptable synthetic images can be produced and their effects on retinal photreceptors properly modelled.

2 Methods

Full spectrum image file representation.

Currently, there is no popular stored image format capable of storing a full spectrum image of the sort needed for this analysis. It is possible to specify a format for use in this paper. The format for a given pixel may be given as a list of specifications. Each specification has 2 bits to identify the coefficients as either blackbody figure[13] equation[2], gaussian equation[7] (a variation on equation[3]), polynomial (complex curve-fit), or absolute (data points for a curve fit). If identified as blackbody, a pair of (I_0,T) coefficients or (a,b) follows. If identified as gaussian, an $(I_0, \lambda_{peak}, RMSwidth)$ or (a,b,c) coefficient triple follows. If identified as polynomial, a vector (a,b,c,...) follows where the coefficient sign bit for 'b' is on to distinguish it from blackbody. If identified as absolute,

Historically, images are represented by 3 color values (R, G, B) for each pixel. For efficiency reasons, both in storage and display time, this encoding has been dominant. With the availability of gpgpu processing display time is no longer impacted significantly by encoding complexity or by calculated rendering complexity. Data pathways in DMA, bus width, and bus speed have improved, eliminating encoding size bottlenecks.

$$I(\lambda) = I_0 \; exp\left(-\frac{(\lambda-\lambda_{peak})^2}{2RMSwidth^2}\right) \qquad \textit{Gaussian function in } \lambda \quad [7]$$

encoding reminder ...
SPECIFICATION: first byte, two high bits encode N-byte datawidth (1,2,3,4 bytes) for (x,y), next two high bits encode spectral encoding type (A,B,G,P), leaving 4 TBD bits. 2nd byte gives length of coefficient stream from 1-256. Next two N byte fields specify (x,y) of pixel in image (0,0) in center. Next 16 bits gives half-float amplitude of wave responsible for intensity. Subsequent byte pairs are aligned half-float coefficients. Spectrum is given as normalized (i.e. sum(coeff)/count(coeff) = 1).

The spectrum of a pixel may be fully described as a list of components: (B:((a1,b1),(a2,b2)...), G:((a3,b3,c3),(a4,b4,c4)...), P:((an,bn,...))). Each coefficient is a floating point number limited to a small range. Using the "half" float (16 bits), it is possible for blackbody specifications to occupy 32 bits and gaussian to occupy 48 bits. I_0 Cannot be negative, so the sign bit of coefficient a may be used to indicate whether it is a G (gaussian) or B (blackbody). For shorthand, uppercase letters will indicate Blackbody coefficients when the first coefficient is negative. Here is a fanciful pixel spectrum: (A1,B1,a2,b2,c2,a3,b3,c3,A4,B4). In most cases, only one or two components will be used. In no case shown in this paper will blackbody and gaussian components be mixed.

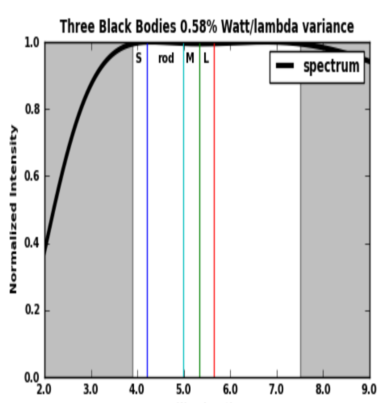


figure 21: White Body Spectrum

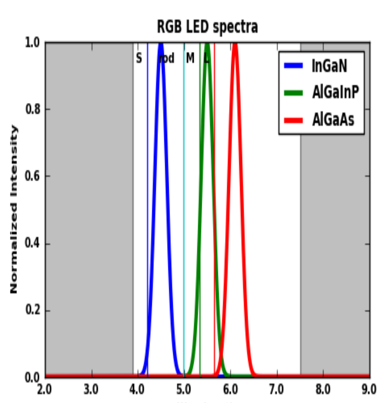


figure 22: RGB LED Spectrum

An example is the white body spectrum figure[21] where intensity variation over the entire visible spectrum is less than 1%. This is expressable as the 96 bit or 12 byte (A1,B1,A2,B2,A3,B3) with values (-1e-2,8.7e+3,-3.226e-2,5.6e+3,-1.732e+0,3e+3). Note the negative values for A1, A2, and A3 indicating black body coefficients.

Another example is the spectral composition of a white LED pixel when using LEDs of types R:AlGaAs, G:AlGaInP, B:InGaN. The b coefficient is given in nanometers. figure[22]. In this figure, each component is given by a single gaussian. (I_R,610,0.18e-7,I_G,550,0.18e-7,I_B,450,1.7e-7,0.18e-7)

Typically, images are laid out as either raster sequences of pixels or as vector graphics. Another alternative, used here, is to give the coordinates of a pixel at which a spectrum is expressed. With (x,y) expressed as 12 bits for each coordinate (24 bits) a pixel is expressed as (x,y,a1,b1,c1,a2,b2,c2...). 3 bytes for coordinates and 6 bytes for each gaussian or 4 bytes for each blackbody component. This pixel spectrum may be used for spectral calculations such as multiplying the probabilities of capture by a photoreceptor by the probabilities of photons arriving from a scene. An S photoreceptor figure[15] takes 9 bytes. An M and L photoreceptor each take 15 bytes for complete spectral representation. For instance with positive a1, and a2 for gaussian coefficients:

List 1: Spectra for human cones

- `L cone spectrum (x1,y1,I0,564,0.58e-7,0.25*I_0,400,0.58e-7)`
- `M cone spectrum (x2,y2,I0,534,0.58e-7,0.25*I_0,400,0.58e-7)`
- `I rods spectrum (x3,y3,I0,498,0.58e-7,0.25*I_0,400,0.58e-7)`
- `S cone spectrum (x4,y4,I0,420,0.58e-7)`

The spectra of figure[15] were generated using these coefficients and they may be compared favorably with published absorption spectra for human photoreceptors, notably Bowmaker J.K. and Dartnall H.J.A.[4].

While the stored form of a full photograph would be expanded by about a factor of 4 with use of this representation, an image composed of photons from point sources may actually be far smaller than with traditional RGB digital images yet capable of expressing the full spectra required for analyzing vision in natural light settings.

Photon probability chain

Probabilities for photon capture are products of probabilities for every step between emission and absorption. The spectrum of blackbody emission E_λ is a probability function of wavelength. Objects have a surface reflectance spectrum S_λ. The cornea and lens have a refraction spectrum L_λ at each retinal photoreceptor. The pupil has a diffraction spectrum P_λ at each retinal photoreceptor where it is absorbed. Finally, each species of photoreceptor has an absorption spectrum A_λ. The relative probability of a photon from the point source being absorbed by the particular photoreceptor is shown in equation[8].

$$E_\lambda * S_\lambda * L_\lambda * P_\lambda * A_\lambda \qquad \textit{Probability chain} \qquad [8]$$

graph 1: Probability chain

In effect, each spectrum between emission and absorption modifies the image as absorbed. No single step in the chain stands alone as dominating the result. Isotropic emission from a blackbody doesn't alter the course of a photon to the photoreceptor. Isotropic reflection, also, doesn't differentially change photon pathways. Refraction performs an Affine transform with the pattern seen for short wavelength photons displaced further to the outside. Diffraction performs two function: convolution with a spatially limited discrete Airy kernel followed by an affine transform with long wavelength photons displaced further to the outside.

The resulting function of wavelength forms a probability pattern which may be converted to a wave equation for superposition. Since the final step in the chain is multiplication by the absorption spectrum of the photoreceptor, the apparent image incident on each species of receptor differs substantially from that for other species.

RGB test pattern

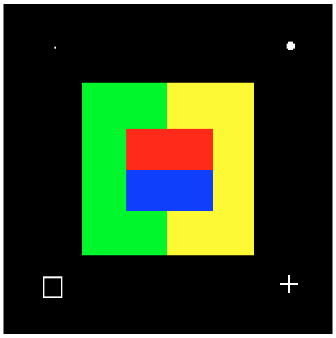


figure 23: RGB test pattern

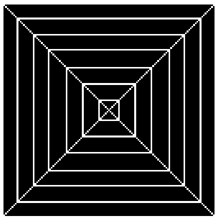


figure 24: Another RGB pattern

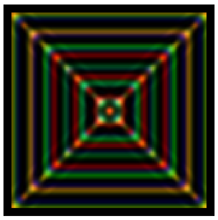


figure 25: As incident on retina

3 Results

4 Discussion

section reminder ...

Remember to fill in this section

equation service documentation ...

`equation.$id|$label|$code`

https://www.tug.org/texlive/doc/texlive-en/texlive-en.pdf

$id: a keyword to use when referencing this equation;
$label: a title for this equation;
$code: TeX instructions to generate this equation.

This service uses TeX to generate an image of the equation. The image is converted to base64 format and inserted directly in the <img> tag instead of referenced through an URL.

/Users/jlettvin/Desktop/Ellipsis/...service.equation.php

equation defaults ...

No defaults ...

/Users/jlettvin/Desktop/Ellipsis/...service.equation.php

5 Acknowledgments

6 Literature

1. Christine A. Curcio, Christine A. Curcio and Kenneth R. Sloan, Visual Neuroscience, 1992, August, 9, 02, 169-180
2. Seeing with S Cones, David J. Calkins, Progress in Retinal and Eye Research, 2001, 20, 3, 255-287
3. The Colors Of Things, Jerome Y. Lettvin et al, Lettvin et al, Scientific American, 1986, September, 255.3, 84-92
4. Visual pigments of rods and cones in a human retina, Bowmaker J.K. and Dartnall H.J.A., J. Physiol., 1980, 298, 501-511

7 Appendices

8 Missing