

This is a preprint of

6057-47 paper in SPIE/IS&T Electronic Imaging Meeting, San Jose, January, 2006

## **High-Dynamic-Range Scene Compression in Humans**

John J. McCann  
McCann Imaging, Belmont, MA 02478 USA

Copyright 2006 Society of Photo-Optical Instrumentation Engineers.  
This paper will be published in the Proceedings of SPIE/IS&T Electronic Imaging, San Jose, CA and is made available as an electronic preprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

# High-Dynamic-Range Scene Compression in Humans

John J. McCann\*  
McCann Imaging  
Belmont, MA 02478, USA

## ABSTRACT

Single pixel dynamic-range compression alters a particular input value to a unique output value - a look-up table. It is used in chemical and most digital photographic systems having S-shaped transforms to render high-range scenes onto low-range media. Post-receptor neural processing is spatial, as shown by the physiological experiments of Dowling, Barlow, Kuffler, and Hubel & Wiesel. Human vision does not render a particular receptor-quanta catch as a unique response. Instead, because of spatial processing, the response to a particular quanta catch can be any color. Visual response is scene dependent.

Stockham proposed an approach to model human range compression using low-spatial frequency filters. Campbell, Ginsberg, Wilson, Watson, Daly and many others have developed spatial-frequency channel models. This paper describes experiments measuring the properties of desirable spatial-frequency filters for a variety of scenes. Given the radiances of each pixel in the scene and the observed appearances of objects in the image, one can calculate the visual mask for that individual image. Here, visual mask is the spatial pattern of changes made by the visual system in processing the input image. It is the spatial signature of human vision. Low-dynamic range images with many white areas need no spatial filtering. High-dynamic-range images with many blacks, or deep shadows, require strong spatial filtering. Sun on the right and shade on the left requires directional filters. These experiments show that variable scene-dependent filters are necessary to mimic human vision. Although spatial-frequency filters can model human appearances, the problem still remains that an analysis of the scene is still needed to calculate the scene-dependent strengths of each of the filters for each frequency.

## 1.0 INTRODUCTION

A major goal of the Polaroid Vision Research Laboratory's work in the 1960's was to study the properties of human vision and develop algorithms that mimic visual processing. One of many important experiments was the Black and White Mondrian: a High Dynamic Range (HDR) image using an array of white, gray and black papers in non-uniform gradient illumination.<sup>1</sup> Here a white paper in dim illumination sent the same luminance to the eye as a black paper in bright illumination. This experiment demonstrated that the sensations of white and black can be generated by the same luminance. In addition, the range of the papers' reflectances in display was about 33:1; the range of the illumination was also 33:1; the total dynamic range of the display was 1000:1. The appearances of the gray areas in the display were very similar to those from the display in uniform light. One goal was to propose a model of vision that could predict sensations in the HDR B&W Mondrian.<sup>2,3,4,5</sup> In short, calculate sensations and write them on film.

Examples of rendering high-dynamic range scenes onto low-dynamic range systems include printers, displays and the human visual system. Although the retinal light receptors have a dynamic range of  $10^{10}$ , the optic nerve cells have a limited range firing range of around 100:1. In general, there are three different approaches to rendering HDR scenes. First, there are tone-scale S-shaped curves used in most chemical and digital photography<sup>6</sup>. Tones scale curves are based on photographic sensitometry developed by Hurter and Driffield<sup>7</sup> and extended by C. K Mees<sup>8,9</sup>. These tone scale functions are the equivalent of a lookup Table (LUT) that transforms input digit to output digit. Such curves have little value in HDR scenes such as the B&W Mondrian. Since both white and black sensations have identical input digits, tied to luminance, tone scale cannot provide a meaningful solution to the problem. In Land's 1968 Ives Medal Lecture, he introduced the Retinex image processing model that calculated sensations and sent sensations to

a display.<sup>1</sup> This algorithm automatically scaled all pixels in the image to the maxima in a highly non-linear manner.<sup>10</sup> The B&W Mondrian experiment, along with a wide variety of others, including experiments on simultaneous contrast, out of focus images, and color images led to three important general conclusions about vision. First, that human visual process was scene dependent. Second, that an auto-normalizing visual system was referenced to the maxima in each channel. Third, that vision used multi-resolution components to achieve distance constancy. The parallel channel-maxima referencing in independent L, M, and S color channels provide a mechanism for color constancy.<sup>11</sup> Further, this mechanism is consistent with experiments measuring departures from perfect constancy with variable illumination.<sup>12,13</sup>

Fergus Campbell and John Robson's 1965 classic paper<sup>14</sup> introduced the sinusoidal spatial displays in vision research. Blakemore and Campbell's experiments showed the existence of independent adaptation channels having different spatial frequencies and different sizes of receptive fields.<sup>15</sup> Tom Stockham of MIT saw the B&W Mondrian demonstration and proposed a spatial-frequency-filter mechanism to compress the dynamic range of the scene luminances.<sup>16</sup> Since then there has been a wide range research using complex images by analyzing their spatial frequency components.<sup>17</sup>

Stockham's example of a small building with an open door and the Black and White Mondrian both require strong spatial filters. However, there are images that require little, or no, spatial filtering. The maxima reset used in Retinex image process has been used to control the extent of spatial processing applied to different images. In this case each resolution image is auto-normalized to the maximum, just as in color, when each color channel is auto-normalized to each channel's maxima. This highly non-linear algorithm has the desirable property that it generates scene dependent changes in images.<sup>10</sup>

This paper looks at three different HDR image-processing systems. First, it studies lightness matching experiments to calculate the visual mask that is the spatial signature of human vision for different targets. Second, it uses the same tool to analyze the spatial signature of a software model of vision. Third, it uses the same tools to analyze the spatial signature of a firmware/hardware processing in a commercial digital camera.<sup>18</sup> In each case we will compare the spatial frequency signature of the visual mask for different images. We find that the three processing techniques have a common property. Each generates different, image dependent masks. Spatial frequency filters are very effective in rendering HDR images into low-dynamic range displays and printers. Nevertheless, there remains a problem for these techniques, namely the algorithm that calculates the filter that is specific for each image. Fixed spatial filters can be shown to be effective for some images, but cannot mimic human vision if one tests the full range of images from a foggy day to sun and shade.

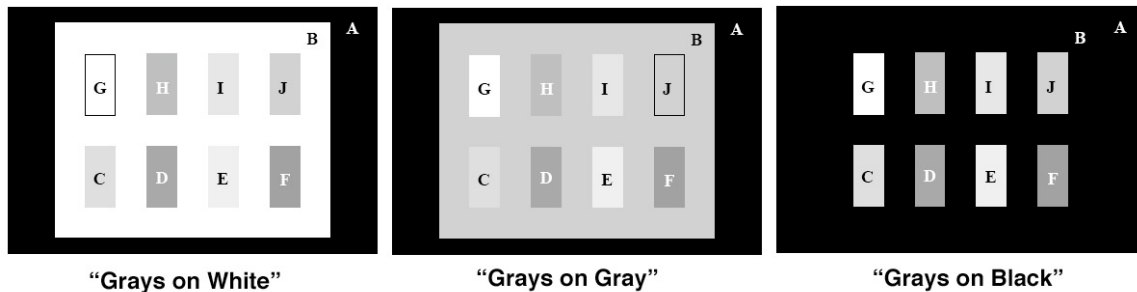


Fig. 1 shows the transparency displays used in the matching experiments. Observers match the same gray transmissions with different surrounds.

## 2.0 MATCHING EXPERIMENTS

The experiment consisted of matching a constant set of gray patches on white, gray and black surrounds (Fig. 1) to a Standard Lightness Display (SLD).<sup>19</sup> The targets were photographic transparencies. The optical density for each area (C through J) made to be a close as possible. The object of the experiment was to measure the change of appearances of the same gray patches with different surrounds in a complex image.<sup>20</sup>

Observers matched each area in Figure 1 to Lightness patches in the SLD (Range 9.0 - 1.0).. The luminance of each matching patch in the SLD was measured with a Gamma Scientific telephotometer.. This calibration luminance vs. lightness curve was fit by a five-degree polynomial so as to be able to calculate the luminance by interpolating Lightness Match values (LM) into luminance. That is, if the average of LMs from the three observers (three trials per observer) is 6.7, then the luminance estimate from the polynomial fit (Equation 1) is 355.7 ft-L.

$$\text{luminance} = -0.1156 * (\text{LM})^5 + 1.6592 * (\text{LM})^4 - 4.1084 * (\text{LM})^3 - 4.1954 * (\text{LM})^2 + 34.021 * \text{LM} - 26.533$$

Table 1 lists the telephotometer measurements of Target Luminances from the display vs. the interpolated SLD luminances of the average match for that area. If human vision acted as a simple photometer, then the Target Luminances should equal Matching Luminances. Any departures from equal luminances are a signature of the human signal processing mechanisms. The data in Table 1 (“Grays on White”) shows matches that are similar to SLD luminances.

Area	Target	Match SLD	Target	Match SLD	Target	Match SLD
	Luminance	Luminance		Luminance		Luminance
<b>G</b>	1003	1005	1000	1009	920	1009
<b>E</b>	765	826	695	1005	655	1007
<b>I</b>	570	690	515	860	470	968
<b>C</b>	431	391	411	741	368	882
<b>J</b>	258	306	222	510	195	741
<b>H</b>	131	94	102	250	89	560
<b>D</b>	61	33	44	59	39	391
<b>F</b>	37	17	24	33	19	287
<b>B</b>	910	1006	198	510	0	3
	<b>"Gray on White"</b>		<b>"Gray on Gray"</b>		<b>"Gray on Black"</b>	

Table 1 lists the measured luminances for each area described in Fig. 1. It also lists the interpolated values for the luminances of the average match chosen by observers in a Standard Lightness Display.

The data in Table 1 “Grays on Gray” show most matches have significantly higher SLD luminances. The data in Table 1 “Grays on Black” shows matches that are even higher than “Grays on Gray” luminances.

Figure 2 plots the difference in luminances (Matching Luminance - Target Luminance) for each area (C through J) sorted with highest luminance on the left and lowest on the right. The “Grays on White” areas have matching luminances that are all close to the actual luminances of the target. Some matches had higher luminances (positive) and others lower luminances (negative), and some were very close to zero differences. We do not know if these differences are associated with spatial differences between the targets and the STD, experimental errors, such as small departures for uniform illumination, or observer variability. In any event there is no evidence of significant systematic differences in matching vs. actual luminances.

### Luminance Difference (Match-Actual)

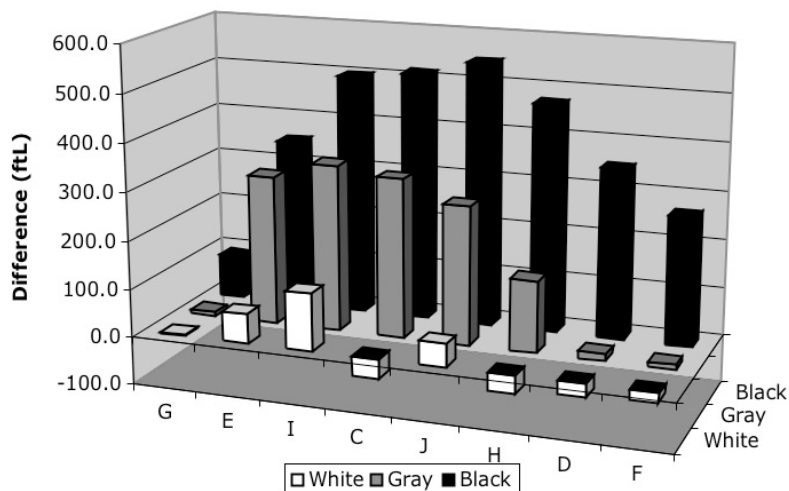


Fig. 2 shows the differences between match and actual luminances for each area shown in Figure 1. The areas are sorted by luminance: Area G on the left has the highest luminance (1003,1000,920 ft-L) and Area F on the right has the lowest luminance (37, 24, 19 ft-L). There is no systematic difference between Match and Actual for “Greys on White”. Some differences are positive and some are negative. “Greys on Gray” have matches that have differences that are 345 ft-L higher for area I. “Greys on Black” have differences that are 545 ft-L higher for area J.

The second row of columns in Fig 2 plots the differences for the “Greys on Gray” target. There is no difference for the lightest area (G). As the luminances of the areas decrease, the differences for Areas E, I, C, J, and H are higher, with the maximum difference for Area I. For the lowest luminances, Areas D and F, the differences are close to zero. The third row of columns in Fig 2 plots the differences for the “Greys on Black” target. The match for the most luminous area G is slightly higher than actual. As the luminances of the areas decrease the matches are higher for all areas, with maximum difference for Area J. For the lowest luminance area, Area F, the difference was 267 ft-L.

### 3.0 HUMAN SPATIAL PROCESSING

The goal of this paper is to evaluate the spatial frequency signature of human vision. As reported by countless other experiments, observers match grays in dark surrounds with higher luminances. Is the spatial influence of white, gray and black surrounds consistent with the hypothesis that human vision incorporates spatial-frequency filters as a mechanism in calculating appearance? This is where data for both the luminance of the target and the match can be used to identify the spatial-frequency signature of human vision. This data can be used to determine if vision uses a fixed set of spatial frequency filters, or instead uses mechanisms that are scene dependent. The issue of fixed processing versus image dependent processing is an important one in selecting models of vision.

If the input to vision is an array of luminances, and the output is a second array of matching luminances, then the signature of the visual process is the change between input and output. Fig. 2 shows one analysis, namely the difference between match and actual luminances. The data describes the signature, but does not help us to understand the underlying mechanism. A better analysis is to calculate the transmissions of each pixel in a spatial mask. This mask is the spatial signature of the visual system for that image. Fig. 3 shows the idea of a spatial mask. In the upper left it shows the luminances of each gray area. In the bottom right it shows the average matching luminances chosen by the observers. In between, it shows the relative transmission of a mask that has the same spatial signature as the human visual system. The values at each pixel are simply the ratios of the output [Match] and the input [Actual] luminances. Imagine a lightbox illuminating the “Greys on Black” transparency target. Superimpose the second transparency that is the spatial record of the human visual system mask. The mask has transformed the actual luminances to the same relative luminances as the matches. Since the ratio for Area F is 14.8, we need to increase the output

of the lightbox by 14.8 times. The resulting luminances for the combined transparencies are the same as those chosen by the observers as matches. The mask described here is the signature for the human visual system for this particular target.

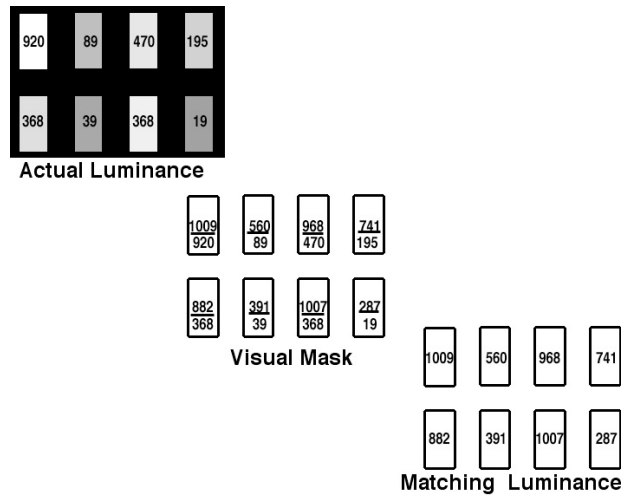


Fig. 3 shows the comparison of the Actual luminances in “Grays on Black”(upper left) and the average Matching luminances (bottom right). The signature of human visual processing is calculated by taking the ratio [Matching/Actual] luminance for each pixel in the image. This visual mask is the spatial array of human image processing for this target. This mask is what vision did to “Grays on Black” input image so as to generate the observer matches.

If we compare the spatial masks for all three displays we see they are very different. Figure 4 plots the ratio of [Matching/Actual] luminances normalized to 1009. The luminance masks are very different. The visual system do not apply a significant mask to the “Gray on White” target; it applied a significant mask to the “Gray on Gray” target; and it applied a very strong mask to the “Gray on Black”.

**Visual Masks (Luminance Ratios)**

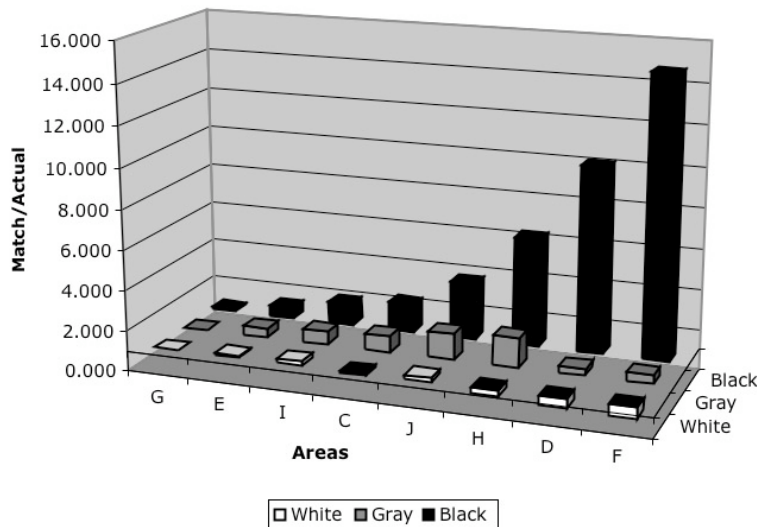


Fig. 4 shows the signature of human visual processing. The visual mask is plotted as the ratio [Match / Actual] luminances for each area are shown in Figure 1. The areas are sorted by luminance. There is no significant mask for [Match and Actual] for “Grays on White”. “Grays on Gray” data show a systematic mask applied to the input image. “Grays on Black” data show a very strong mask. Area F, the lowest luminance area in area is matched in the Standard Lightness Display by an area with 14.8 times the luminance.

In 1972 Stockham proposed an image processing mechanism using spatial filters. This idea using spatial-frequency models is a popular approach in many image processing and vision models. Here we evaluate the present data as a spatial filter. We made images (512 by 512 pixel arrays) for each target for Actual and Matching Luminance data. The 512x512 array of ratios [Matching/Actual] luminances, normalized by 15, was the input to a Matlab program that calculated the shifted 2D Fourier transform of each target. The arrays of ratios describe the visual mask applied by human vision. The shifted 2D FFT are spatial filters that describe human vision. Figure 5 shows the FFTs of the visual masks derived from matching data.

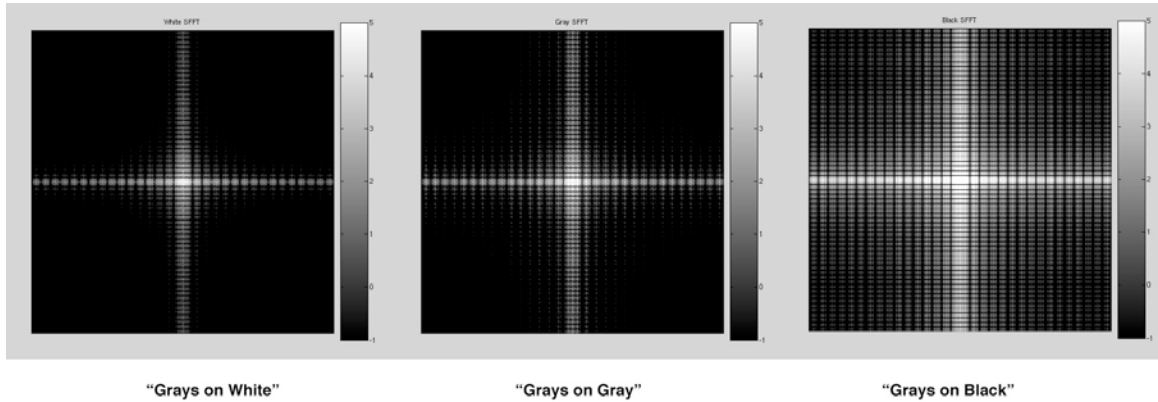


Fig. 5 shows the signature of human visual processing presented as a set of spatial filters. The shifted 2D FFT of visual masks show distinctly different spatial filters for each display. A model of human vision that incorporates spatial filters needs to first calculate an image-dependent spatial filter. One filter does not fit all scenes. Vision models need to be responsive to image content because human vision has that unique imaging property.

#### 4.0 RETINEX PROCESSING

Fig.6 shows a Raw and Retinex Processed image of a pair of Jobo test targets in sun and shade. The Raw image is rendered so that digit is proportional to log luminance. This image is a real-life version of the

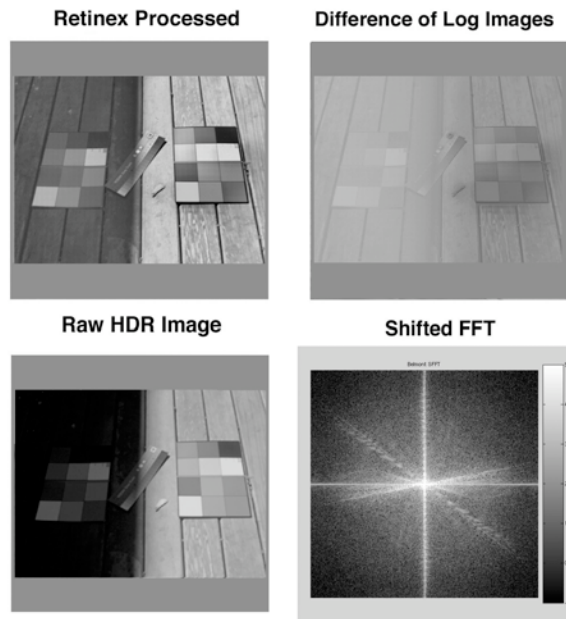


Fig. 6 shows the signature of Retinex processing as a spatial filter (bottom right). This is the shifted 2D FFT of visual masks shown in the upper right. The mask is the ratio image (normalized difference of log luminance) between Retinex Processed image and Raw HDR image. In the Raw HDR image the black square in the sun (top row-right) is the same digit as the white square in the shade (second row-right), namely digit 80. In the Retinex processed image black in the sun is rendered to an output digit of 27, while the white in the shade is rendered as an output digit of 169.

Black and White Mondrian in that the black square in the sun has the same luminance as the white square in the shadow. In the Retinex Processed image (top left) the black in the sun has lower digital values and the white in the shadow has higher values than in the Raw image. The visual mask is calculated by taking the difference of log luminance images. The shifted FFT is a highly directional spatial filter. This Retinex software algorithm has made image changes that are equivalent to an image dependent spatial filter.

### 5.0 DIGITAL CAMERA PROCESSING

We have looked at the equivalent spatial filters made from human vision and image processing algorithms. The third analysis uses images made and processed in a camera.<sup>18</sup> A camera setting activates an option to capture images and apply Bob Sobol's modification of Frankle and McCann Retinex. As well, the processing can be shut off, so as to record a conventional digital image. Color images were converted to grayscale images and scaled to [digit ~ log luminance] with a calibration lookup table. The visual mask equivalent is the normalized difference of log luminances. The shifted 2D FFT was calculated as above.

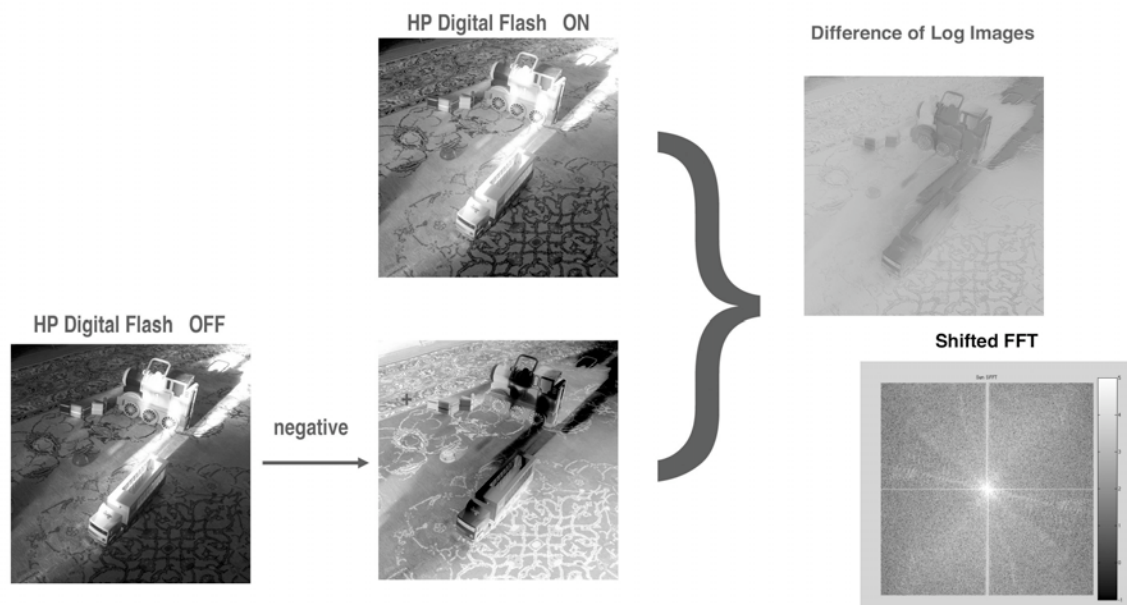


Fig. 7 shows processed and unprocessed images of toys on the floor in high-dynamic range illumination. The “Digital Flash OFF” image is a conventional digital image. The “Digital Flash ON” image is a Retinex processed digital image. The negative of the OFF image is combined with the positive ON image to form the log luminance mask. The shifted FFT of the mask is shown on the bottom right.

The shifted FFT in Figure 7 is a strong oriented filter. The effect of a bright patch of sunlight was to make the Retinex processing in the camera alter the control image significantly, thus make a strong visual mask equivalent.

Figure 8 show the same analysis of the same scene taken a half hour later. The sunlight is gone and the illumination is much more uniform. The visual mask equivalent is nearly uniform and the shifted FFT is a much weaker spatial filter.



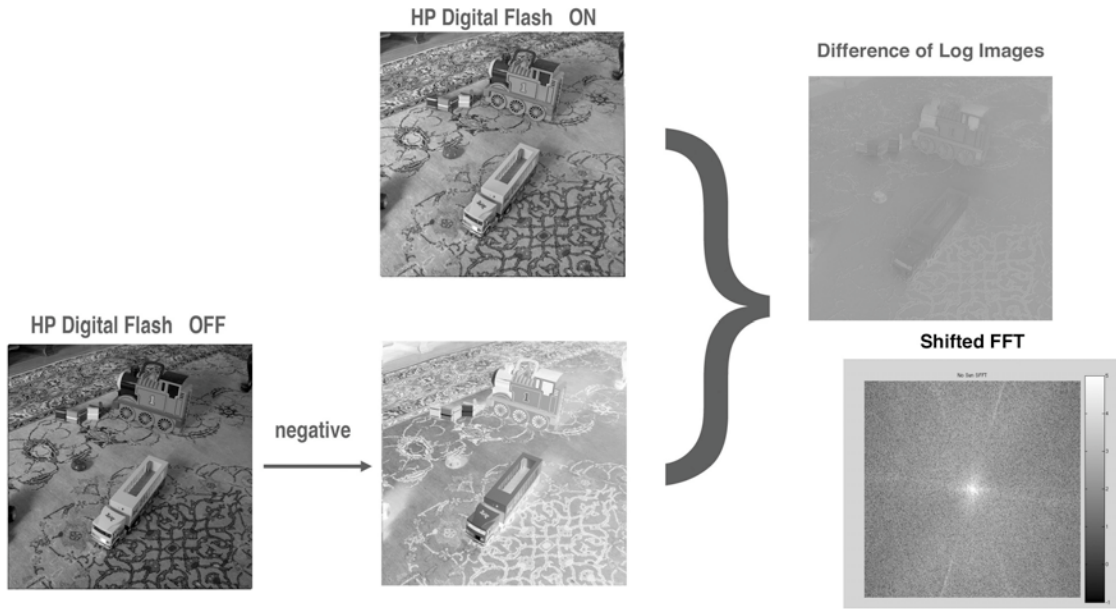


Fig. 8 shows processed and unprocessed images of toys on the floor in more uniform illumination without sunlight. The “Digital Flash OFF” image is a conventional digital image. The “Digital Flash ON” image is a Retinex processed digital image. The negative of the OFF image is combined with the positive ON image to form the log luminance mask. The shifted FFT of the mask is shown on the bottom right.

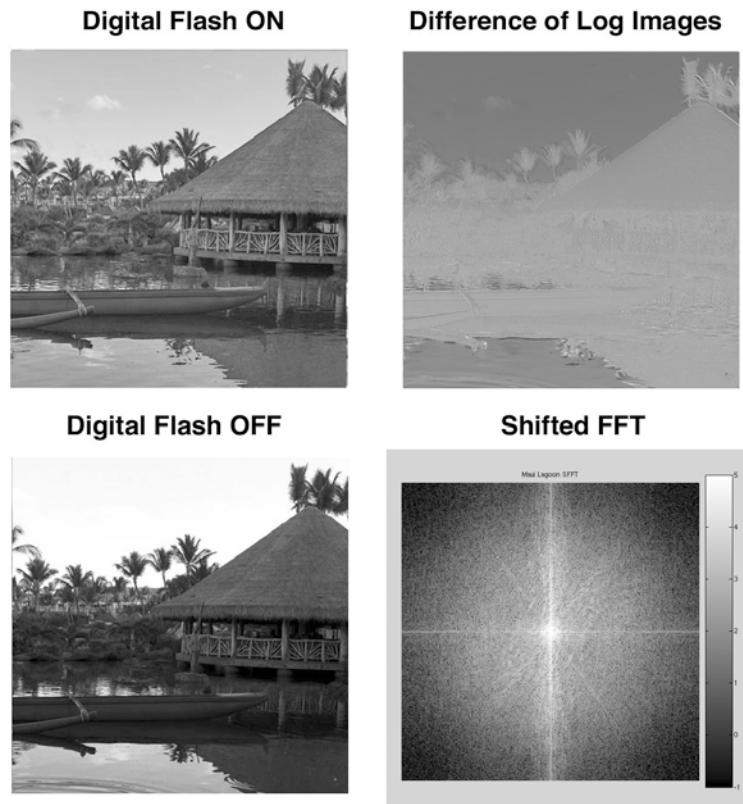


Fig. 9 (left) shows Retinex image and conventional image; (right) log luminance mask and its shifted FFT.

Figures 9 and 10 show the same analysis of two different outdoor scenes. Figure 9 has very high dynamic range when the camera was in the shade looking towards the sun. Figure 10 is taken with the sun behind the camera. As shown above, the visual mask equivalent for high-dynamic range Figure 9 is higher in contrast than that for Figure 10. The shifted FFT in Figure 9 is a strong oriented filter. Figure 10 show a visual mask equivalent that is nearly uniform and the shifted FFT is a much weaker spatial filter.

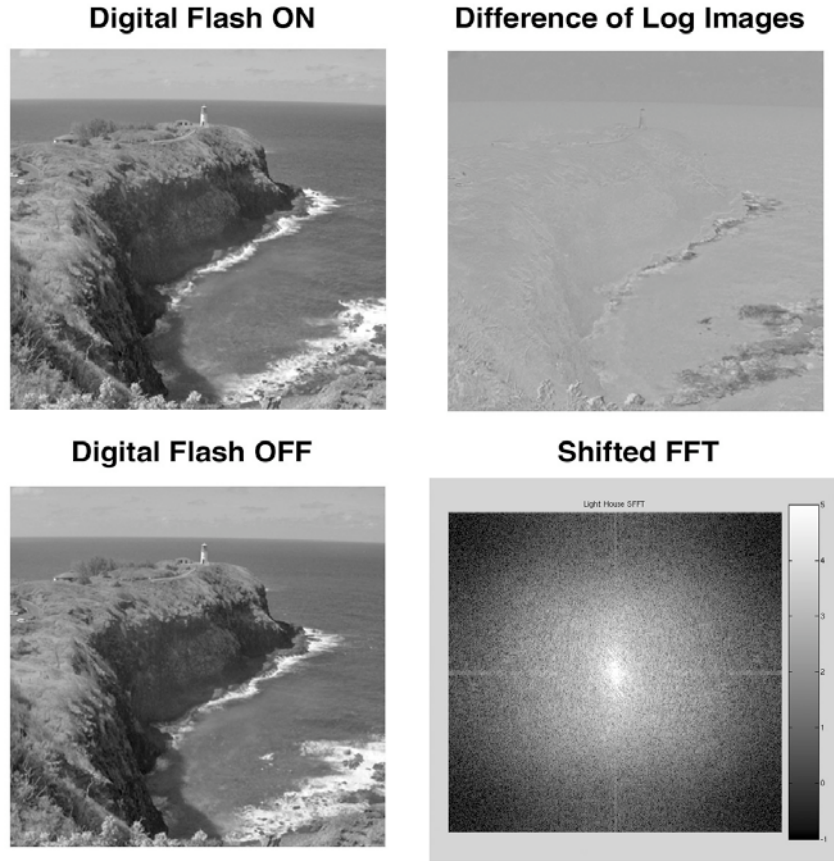


Fig. 10 (left) shows Retinex image and conventional image; (right) log luminance mask and its shifted FFT.

Figure 11 shows the four shifted FFTs from the camera-processed images. They are all different. This shows that the camera Retinex processing generates visual masks and spatial-filter equivalents that are scene dependent. In that regard, this process mimics human vision. As seen in the observer matching data (Table 1) human image processing is image dependent.

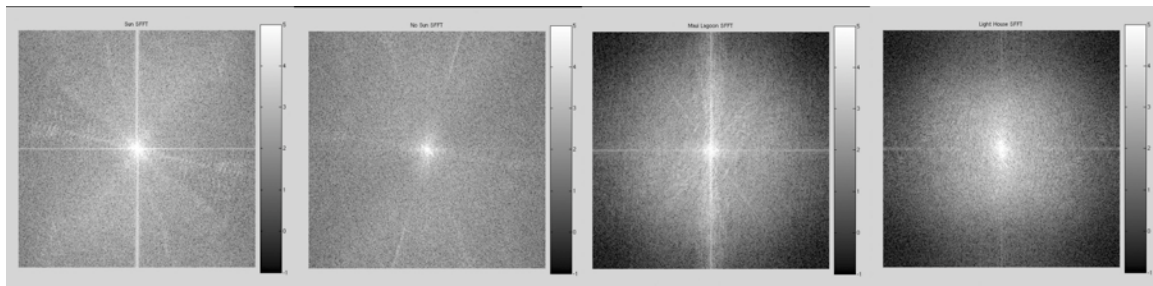


Figure 11 shows the shifted 2D FFTs of the camera image processing for the scenes described in Figures 7,8,9,10.

## 6.0 DISCUSSION

Gatta's recent thesis reviews a wide range of work in HDR imaging.<sup>21</sup> In one section he summarized many tone scale-mapping algorithms. Tone scale cannot solve the problem identified in the B&W Mondrian because white and black sensations are generated by the same input digit. Compressing the digital values near white helps render details near black. As well compressing digital values near black helps render details near white. Tones scales, as employed in all of imaging for the past 100 years, is an attempt to find the best average rendition for all scenes including pictures of objects in fog and HDR scenes. Its fixed tone scale curve is optimal for only one scene dynamic range.

The reset step in the Retinex algorithm provided means to model simultaneous contrast and to provide auto-normalization.<sup>10</sup> Even more important was the idea that reset provided a mechanism for calculating a low-spatial frequency filter equivalent that was image dependent. This was the important differentiation from the work of Stockham<sup>17</sup>, Fergus Campbell<sup>22</sup>, Mari<sup>23</sup>, Horn<sup>24</sup>, Wilson<sup>25</sup>, Watson and Ahumada<sup>26</sup>, Daly<sup>27</sup> as well recent variations by Pattanik et. al.<sup>28</sup> and Fairchild<sup>29,30</sup> They all looked to apply spatial filters to receptor images, but did not have a mechanism to independently adjust the filter coefficients to each scene.

## 7.0 CONCLUSIONS

Human vision generates a scene-dependent spatial filter. Patches in a white surround need no spatial filtering, patches in a gray surround need some spatial filtering, and patches in a black surround need strong spatial filtering. Retinex image processing algorithms and camera firmware show the ability to generate the equivalent of scene-dependent spatial filters. The best image rendering for high-dynamic range images is to calculate the appearance and write the altered image on low-dynamic range media. That means that some scenes need little or no alteration, while other high-dynamic range scenes require significant changes. The Retinex processes described in this paper also show scene-dependent processing.

## 8.0 ACKNOWLEDGEMENTS

The author wishes to thank Mary McCann and Ale Rizzi for many helpful discussions.

## 8.0 REFERENCES

- 
- <sup>1</sup> E. H. Land & J. J. McCann "Lightness and Retinex Theory", *J. Opt. Soc. Am.* **61** 1-11, 1971.
  - <sup>2</sup> E. H. Land & J. J. McCann, "Method and system for reproduction based on significant visual boundaries of original subject", *U.S. Patent* 3,553,360, June 5, 1971.
  - <sup>3</sup> E. H. Land, L. A. Ferrari, S. Kagen & J. J. McCann, "Image Processing system which detects subject by sensing intensity ratios", *U.S. Patent* 3,651,252, Mar. 21, 1972.
  - <sup>4</sup> J. Frankle & J. J. McCann "Method and apparatus of lightness imaging", *U. S Patent* 4384336, May 17, 1983.
  - <sup>5</sup> J. J. McCann "Calculated Color Sensations applied to Color Image Reproduction", in *Image Processing Analysis Measurement and Quality*, Proc. SPIE, Bellingham WA, **901**, 205-214, 1988.
  - <sup>6</sup> John J. McCann, "Color imaging systems and color theory: past, present, and future", Proc. SPIE Vol. 3299, p. 38-46, in *Human Vision and Electronic Imaging III*; B. E. Rogowitz, T. N. Pappas; Eds. 1998.
  - <sup>7</sup> F. Hurter and V. C. Driffield, "The Photographic Resources Ferdinand Hurter & Vero C. Driffield", W. B. Ferguson, Ed., Morgan and Morgan Inc. Dobbs Ferry, 1974.
  - <sup>8</sup> C. E. K. Mees, "An Address to the Senior Staff of the Kodak Research Laboratories", Kodak Research Laboratory, Rochester, 1956.
  - <sup>9</sup> C. E. K. Mees, "Photography", The MacMillan Company, 1937.
  - <sup>10</sup> J.J. McCann, "Lessons Learned from Mondrians Applied to Real Images and Color Gamuts", Proc. IS&T/SID, Seventh Color Imaging Conference, 1-8, 1999.
  - <sup>11</sup> J. J. McCann S. McKee and T. Taylor "Quantitative studies in Retinex theory: A comparison between theoretical predictions and observer responses to 'Color Mondrian' experiments" *Vision Res.* **16** 445-458, 1976.
  - <sup>12</sup> J. J. McCann, *Mechanism of Color Constancy*, Proc. IS&T/SID Color Imaging Conference, IS&T/SID, Scottsdale, Arizona, **12**, 29-36, 2004.
  - <sup>13</sup> J. J. McCann, "Do humans discount the illuminant?", Proc. SPIE Vol. 5666, 9-16, in *Human Vision and Electronic Imaging X*; B. E. Rogowitz, T. N. Pappas, S. J. Daly; Eds., Mar 2005.

- 
- <sup>14</sup> F. W. Campbell & J. G. Robson, "Application of Fourier analysis to the visibility of gratings", *J. Physiol. (Lond.)* **197**, 551-566, 1968.
- <sup>15</sup> Blakemore, C. and F. W. Campbell (1969). "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images." *Journal of Physiology* **213**: 237-260.,
- <sup>16</sup> T. P. Stockham, "Image Processing in the Context of a Visual Model", *Proc. IEEE*, **60**, 828-284, 1972.
- <sup>17</sup> P. G. Barten, "Contrast Sensitivity of the Human Eye and Its Effects on Image Quality", SPIE press, Bellington, 1-232, 1999.
- <sup>18</sup> The camera used in these experiments is an HP 945 with Digital Flash. This camera uses the Frankle and McCann algorithm [J. Frankle and J. J. McCann "Method and apparatus of lightness imaging", *U. S Patent* 4384336, May 17, 1983.] as modified by Sobol [R. Sobol, "Improving the Retinex algorithm for rendering wide dynamic range photographs", *J. Electronic Imaging*, **13**, 65-74, 2001].
- <sup>19</sup> J. J. McCann, E. H. Land and S. M. V. Tatnall, "A Technique for Comparing Human Visual Responses with a Mathematical Model for Lightness", *Am. J. Optometry and Archives of Am. Acad. Optometry*, **47(11)**, 845-855, 1970.
- <sup>20</sup> B. V. Funt, F. Ciurea and J. J. McCann, "Tuning Retinex parameters", in *Human Vision and Electronic Imaging VII*, B. E. Rogowitz and T. N. Pappas, ed., *Proc. SPIE* **4662-43**, 358-366, 2002.
- <sup>21</sup> C. Gatta, "Human Visual System Color Perception Models and Applications to Computer Graphics", Ph.D thesis, Universit'a Degli Studi di Milano, Milano, Italy, 2006.
- <sup>22</sup> F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings", *J. Physiol. (Lond.)* **197**, 551-566, 1968.
- <sup>23</sup> D. Marr, "The computation of lightness by the primate retina", *Vision Res.* **14**, 1377-1388, 1974.
- <sup>24</sup> B. K. P. Horn, "Determining lightness from an image", *Comp. Gr. Img. Proc.* **3**, 277-299, 1974.
- <sup>25</sup> H. R. Wilson and J. R. Bergen, "A four mechanism models for threshold spatial vision", *Vision Res.*, **26**, 19-32, 1979.
- <sup>26</sup> A. B. Watson, & A. J. Ahumada, Jr, "A standard model for foveal detection of spatial contrast", *Journal of Vision*, **5(9)**, 717-740, (2005). <http://journalofvision.org/5/9/6/>
- <sup>27</sup> S. Daly, The visible difference predictor: an algorithm for the assessment of image fidelity, *International Journal of Computer Vision* **6** (1993), 179-206.
- <sup>28</sup> S.N. Pattanik, J. Ferwerda, M. D. Fairchild, and D. P. Greenburg, "A Multiscale Model of adaptation and spatial vision for image display", in *Proc. SIGGRAPH* **98**, 287-298, 1998.
- <sup>29</sup> M. Fairchild and G. M. Johnson, "Meet iCAM: A next generation Color Appearance Model", in *Proc. 10th IS&T/SID Color Imaging Conference*, Scottsdale, Arizona, 33-38, 2002. **2**, 17-36, 1983.
- <sup>30</sup> M.D. Fairchild and G.M. Johnson, "The iCAM framework for image appearance, image differences, and image quality," *Journal of Electronic Imaging*, **13**, 126-138, 2004.