Analysis of Spatial Image Rendering

John McCann, McCann Imaging, Belmont, USA; Vassilios Vonikakis, Democritus University of Thrace, Greece ; Carinna Parraman, University of the West of England; Bristol, (UK); and Alessandro Rizzi, University of Milan, Italy

Abstract

Spatial image processing, such as Retinex, ACE, spatial-frequency, or bilateral, filtering, use the entire image in rendering scenes. These algorithms process captured scene radiances as input; then use the spatial information to synthesize a new image for rendition to a display or print. Spatial algorithms have different properties from pixel-processing algorithms. Pixel processes apply the same transform to all image pixels, so that all pixels with the same input value (i) have the same output value (o). However, spatial algorithms can convert identical input values into different output values. We discuss techniques most appropriate for measuring the success of spatial algorithms.

We would like a simple figure-of-merit calculation for our favorite algorithm. We found that goal impractical. Spatial color algorithms are in the middle of the imaging chain, and their success is affected by pre- and post-processing. There are a variety of distinct goals for different spatial algorithms: one is to find the objects' reflectance; one is to find the illumination; one is to make the best HDR picture; another is to model human vision. As well, there are different ground truth goals for each type of algorithm. Instead of presenting a universal solution to evaluate all types of algorithms, we describe a number of steps measuring scene characteristics that evaluate spatial processes. We describe examples of a number of control and test experiments that are useful in quantitative evaluation of portions of the imaging chain. This paper provides test images, measurements of scene characteristics, and examples of a set of flexible tools for quantitative evaluations of spatial color algorithms. Quantitative measurements of spatial algorithms evaluate the true performance of the central spatial process. This paper works in parallel with <http://sites.google.com/site/3dmondrians/> that provides appendices for detailed data.

Introduction

We use spatial processing for modeling human vision, and rendering high-dynamic-range (HDR) scenes. Spatial processes are needed to render images that cannot be processed using single pixel (Tone Scale) approaches. In human vision we see that color appearance does not correlate with the quanta catch of the receptors. Appearance does not correlate with pixel value. [1,2] In HDR photography we find that spatial rendering onto lower range display media avoids truncating scene information by converting HDR scenes into LDR images appropriate for the media. In both cases, we use spatial information because pixel processes cannot solve the problem.

The common point of the spatial algorithms is that computations build up the output from the spatial information in the scene. They start with the quanta catch of the sensor pixel and apply spatial computations to make new image renditions, as does human vision. The spatial color family includes all the various Retinex implementations that can differ quite remarkably in the way they transform the image and apply spatial processing. They include alternative spatial algorithms (ACE or RACE). They include image domain ratio-products, frequency based spatial filters, some tone rendering algorithms and bilateral filters. They all are nonlinear spatial transforms of the receptor quanta catch. In general spatial algorithms are applied to the captured scene radiance so as to render the scene data with improved image quality. Spatial color algorithms have been studied by many authors. < http://web.me.com/mccanns/Spatial/Processing.html.>

Types of Evaluation

There are both subjective and objective measurements of images. Subjective measurements ask observers to select the preferred rendition of an image. Objective measurements compare the algorithms output pixel values in the processed image with ground truth. Ground truth is the goal of the algorithm. There are a great many spatial algorithms, and they have a wide variety of goals. The ground truth for each algorithm is defined by its author, so one ground truth does not apply to all algorithms.

A second choice about evaluation techniques is whether to use a single metric for all attributes, or a series that measure individual image characteristics. Although efficient, observer preference experiments cannot discriminate between the many attributes that contribute to the judgement and measure their relative importance. Although more difficult, a series of objective measurements can help our understanding.

Beauty Contest

An effective subjective technique for evaluating scene rendering is to ask a number of observers to select the preferred image among different candidates. This is how photographic film response functions were determined. The weakness of this technique is that it provides little feedback on the underlying principles of why the algorithm works. We can identify the most preferred image, but learn little about why it is preferred.

Moreover the judgement can be affected by the display technique and setup. Large displays in a dark room has different visual stimuli than a small print. The monitor, its color gamut, its profile, the display luminance, the ambient light on the screen, the viewing angle, and the image's visual angle all influence the appearance of the array of calculated digits.

Departures from Ground Truth

If reproductions actually replaced the light coming from the scene with an identical stimulus, then scene radiance would be the ground truth of image reproductions. Error metrics, such as the mean-squared-error comparing light from the scene, and that from the reproduction would be simple and effective. The problem is that reproduction media have response functions that transform the scene into a preferred rendition. Photographs do not reproduce scenes accurately. The preferred rendering is very different from scene radiances.[3] In order to perform a mean-squared-error calculation, model output vs. ground truth goal, we need information about the model's goal. For human vision, ground truth is the appearance of the objects in the scene. For spatial algorithms that calculate objects' reflectance, or for those that calculate the illumination on a scene, there are measurable ground truths, namely, the set of physical measurements of reflectance and illumination. The ground truth for the best-preferred reproduction of scenes has no universal definition.

Rendition Quality Metric

Can we find an objective analysis of image rendering using error metric analysis? First, we would need to measure the error for each pixel in a complex image. The error is a distance between the spatial processed output and the ideal ground truth. We also need to compare these errors in a uniform color space. In such a space, apparent changes in hue, lightness and chroma are all equal to numerical distances in the 3-D space. In uniform spaces, such as Munsell, the distance in the space represents the size of the change in appearance, while in XYZ, RGB, and sRGB spaces distance does not equal change in appearance. In addition, camera digits follow sRGB guidelines, but do not always follow the standard in regions near the limits of their color space along the color gamut. Built-in color enhancement firmware distorts these near-gamut regions of color space. In order to accurately convert camera digit to colorimetric XYZ, one needs detailed proprietary information of the signal processing in each camera. It is impractical to assume that we can transform the rendered camera response back into scene XYZ values and then convert them into an accurate, uniform color space coordinates.

The second problem is that we need a goal image; we need an array of perfectly rendered pixels. How does one find the ideal rendered image of the scene? Algorithms that calculate physical quantities, such as reflectance and illumination, have well-defined ground truths. Algorithms that calculate appearance, or seek to make the best, most preferred reproduction has to find a quantitative description of appearance, or preference, to use objective analysis. We can use the beauty contest techniques for finding most preferred individual image, but in such experiments we find that the conclusions are image dependent. [4]

Although it would be desirable to analyze images using error metrics, we see that the camera digit rendition of the scene is not equally spaced and that ground truth depends on the goal of the algorithm. Consequently, the single, universal objective measurement of combined properties is not practical. Nevertheless, we need objective evaluation techniques to measure the effectiveness of algorithms. Instead of combining all the properties of algorithms into a single metric value we can learn a great deal from the study of individual image-processing characteristics. For this we need scenes with known radiometric values.

3-D Mondrians

This work is based on a series of experiments from the CREATE project.[5,6] The set of experiments used a scene with a Low-Dynamic-Range portion next to a High-Dynamic-Range portion in the same room at the same time. Both LDR and HDR parts were made of wooden blocks painted with 11 different paints. The LDR blocks were placed inside an illumination cube so as to be as uniform as possible. The HDR blocks had two highly directional lights.

This 3-D test target has been measured with a wide variety of techniques: measurements of objects (reflectances); the light coming from 104 facets (XYZ); multiple exposure photographs using a number of different cameras; magnitude estimates of appearance of block facets; and watercolor paintings of the entire scene as a measure of appearance. We employed a number of these scene measurements to discuss possible evaluation techniques of spatial image processing.

Spatial Color Examples

Figure 1 is examples of images from the 3-D Mondrian experiments: normal digital images, spatial processed images, and Carinna Parraman's watercolor paintings.

Figure 1 shows LDR (top row) and HDR (bottom row) parts of the scene. The columns show normal digital photographs (left); the Vonikakis spatial image processing (center); and watercolor rendition of appearance (right).

Comparison 1 (Figure 1) shows different renditions of LDR and HDR CREATE scenes (rows). The left column shows control photographs taken with a Panasonic DMC FZ5 digital camera (top, LDR; bottom HDR). The middle column shows the LDR and HDR outputs of a spatial algorithm (VV). The right column shows the Carinna Parraman watercolor painting of the scenes (rendition of scene appearance).

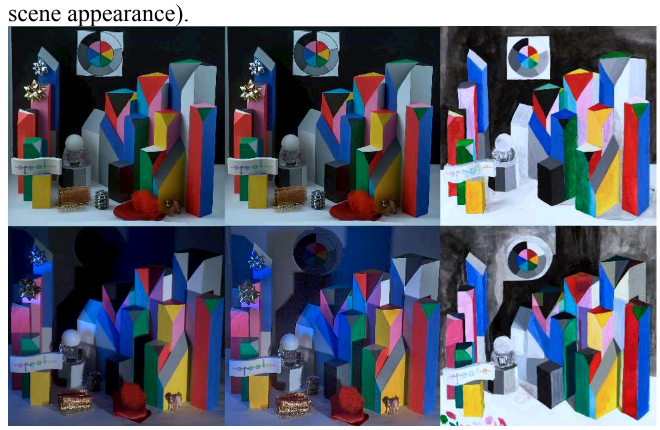

Figure 2 show the normal digital photographs (left); the HP 945 Retinex spatial image processing (middle); and watercolor appearance (right).

VV is a center-surround image-processing algorithm, which employs both local and global parameters.[7] The local parameters, which significantly affect the new value of a pixel, are its intensity (center) and the intensity of its surround. The global parameters that affect the overall appearance of the image are extracted from image statistics. The surround is calculated using a diffusion filter, similar to the biological filling-in mechanism, which blurs uniform areas, preserves strong intensity transitions and permits partial diffusion in weaker edges. New pixel values combining local and global parameters, are inspired by the shunting characteristics of the ganglion cells of the human visual system. The algorithm is applied only to the Y component. <http://sites.google.com/site/vonikakis/software>

Comparison 2 (Figure 2) shows LDR and HDR control images captured by an HP945 digital camera, their spatial processed image along with the Watercolor painting of the scenes. (HP945). The retinex algorithm is a menu selectable part of the image processing firmware (Digital Flash) in the HP945 camera. It is a multi-resolution retinex process with ratio limits, described by Sobol.[8,9] Both the VV and HP945 Retinex algorithms belong to a subset of spatial algorithms called Spatial Color Synthesis Algorithm (SCSA). These algorithms attempt to mimic vision.

Measurements of LDR & HDR 3-D Mondrians

The scene has 2 identical sets of 3-D painted color blocks using only 11 paints on 100 facets. The LDR half is in nearly uniform illumination and the HDR half is in highly directional illumination. Calibration measurements of the CREATE 3-D Color Mondrians are available at <http://sites.google.com/site/3dmondrians/> Table 1 list the measurements. Camera images are multi-exposures (jpeg) of the LDR & HDR portions of the scene. Measurements of appearances are spectral reflectance measurements of paints in LDR and HDR watercolors. The artist recorded the appearances of the scene in non-uniform illumination on the watercolor painting made in uniform illumination. The reflectances of the watercolor are a measure of the scene appearance.[5]

Data Available	Format	Source
Paint reflectance	spectra, XYZ	Spectrolino
LDR radiances	LDR XYZ	Konica Minolta CS100
HDR radiances	HDR XYZ	Konica Minolta CS100
LDR camera	digits (sRGB)	Multiple exposures
HDR camera	digits s(sRGB)	Multiple exposures
LDR appearances	spectra, XYZ	Spectrolino
HDR appearances	spectra, XYZ	Spectrolino

Table 1 lists the data available on the web ***<https://sites.google.com/site/3dmondrians/measurements>*** from the CREATE experiment.

Scene Characteristic Analysis

Instead of looking for a single, universal metric value, we need to break the problem down into a number of practical questions that are possible to implement, with objective measurements. We can use the above data sets to perform a number of different analyses that help us to understand the many characteristics of spatial image processing. Table 2 lists seven different comparisons that characterize the properties of the rendered image.

Test	Segment	Data A	Data B	Quantification
1		LDR control photo	measure paint	effect of profiles
2		LDR control photo	HDR control photo	effect of illumination
3		LDR control photo	LDR spatial process	unwanted compression
4		HDR control photo	HDR spatial process	wanted compression
5	 	LDR watercolor	HDR watercolor	change in appearance
6		HDR watercolor	HDR spatial process	appearance vs. process

Table 2 lists six characteristic tests of the image processing chain.

The first comparison identified in Table 2 provides information about how the camera transforms the scene in image capture. The second shows how the HDR illuminations changed the camera response to the paints. The third analysis shows the effect of the spatial algorithm on the LDR image. This is a very important measurement because it differentiates Spatial Color Synthesis from Tone Mapping algorithms. A Tone Mapping process that significantly improves the HDR rendition will also significantly alter the LDR rendition. A successful spatial algorithm will have no effect on the LDR rendition. This component test looks for unwanted range compression of LDR images.

The next component test measures the range compression of the HDR images by the algorithm. Here, the goal is to measure the amount of dynamic range compression for the circular target in the shadow in the HDR image. The effect of the spatial algorithm can be seen by comparing the normal photograph with the spatial-processed rendition.

The next test compares the compression found in the spatial-processed output with appearance measured by the watercolor. If the goal of the algorithm rendition is to mimic human vision, then we should measure compression similar that found in the watercolor painting. The final tests listed use selected areas in the target to evaluate the effects of the processing on chroma and colors near white and black. These following examples evaluate the 11 paints in uniform illumination using the circular test target. We used this limited number of facets for simplicity of explanation. In most cases it would be appropriate to evaluate all the facets.

Results of Scene Characteristic Analysis

Figure 3 illustrates the steps in image processing between the scene and display. Starting at the scene, cameras alter scene radiances with camera glare, sensor sensitivity, color filters and signal processing, such as anti-blooming, noise reduction, digitization, de-mosaic, color enhancement and preferred tone-scale shaping. We have grouped together all these transforms in the red box, identified as pre-processing.

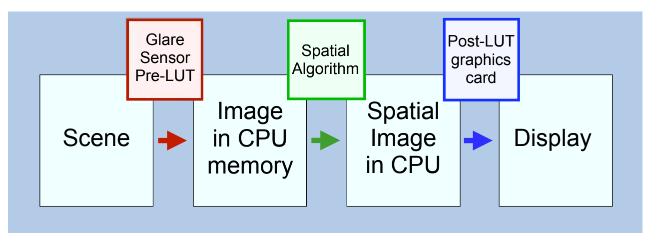

Figure 3 shows the steps in image processing. The middle two blocks illustrate the input and output stages of the spatial image processing (green box). The red box identifies the transforms found in cameras that convert scene radiance to camera digits. The blue box identifies the transforms found in display systems that convert rendered digits into display radiances.

At the other end of the processing chain we show a blue box that includes the post-spatial-processing transformation of rendered digit values into display radiances for viewing. These changes include the post-LUTS (device profiles), graphics card display systems, and unwanted device artifacts.[10,11] In summary, there are a great many pre- and post-processing image modifications that complicate the analysis of spatial algorithms.

Test 1 - Effect of Profiles and Camera Firmware

We can see the effects of pre-processing by comparing camera digits with the original scene. We know the reflectances (XYZ) of the paints on the blocks and on the circular test target on the back of the scene. We can use the sRGB standard to convert XYZ reflectances into expected sRGB digits. In Figure 4 we plot the sRGB values of the paint (scaled to 255 = 100% reflectance) to the HP 945 LDR camera response. Figure 4 shows that the normal HP945 photograph is darker for white, but lighter for grays and black, than the scene. This indicates that the normal image should have had more exposure, and that the camera has changed the achromatic tone scale of the scene. The larger changes are seen in the chroma of the camera rendition. The sR responses to red and yellow paints are significantly boosted, while the sR response to green and cyan are significantly reduced. The sB responses to green, cyan and blue are boosted. These color transform effects are well known to be camera specific, and exposure dependent. This analysis becomes important when one uses images from different cameras.

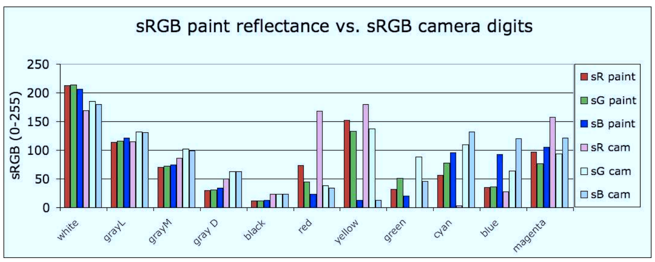

Figure 4 shows the comparison of the circular test targets of sRGB paint reflectance vs. camera digit.

Any algorithm that attempts to calculate reflectances of objects from scene radiances can use this technique to objectively measure its success. The camera's sRGB response makes this difficult.

Test 2 - Effect of Illumination

The LDR circular test target has 11 paints is in maximum uniform illumination. In the HDR image, the circular target is in a shadow created by the box around it. Table 3 lists the changes in illumination as the difference in log X, log Y and log Z between the radiances measured with the KM100 telephotometer.

	delta logX	delta logY	delta logZ
white	0.96	0.97	0.70
grayL	1.38	1.38	1.15
grayM	1.37	1.36	1.15
gray D	1.36	1.33	0.97
black	1.31	1.31	1.12
red	1.28	1.30	1.18
yellow	1.41	1.40	0.98
green	1.40	1.40	1.18
cyan	1.27	1.29	1.16
blue	1.17	1.25	0.90
magenta	1.38	1.38	1.18
average	**1.30**	**1.31**	**1.06**
	5%	**5%**	**9%**

Table 3. Changes in radiance for the 11 paints in the circular test target.

Compared to the LDR illumination, Table 3 shows the average HDR illumination on the circular target is 5% X, 5% Y and 9% Z (linear) for this portion of the scene. The painted-circles image portion is easy, because the light is uniform. However, it is highly variable in the rest of the scene. Any spatial algorithm that attempts to calculate the illumination falling on objects in all parts of the scene could use this technique to objectively measure its success for all 104 facets. The 3-D Mondrian HDR is a particularly difficult target for illumination detection algorithms, yet provides a good challenge.

Test 3 - Unwanted Range Compression

The comparison of the normal LDR photograph with the spatial rendition measures the amount the image processing has compressed the spatial relationships in the image. Figure 5 shows the results for the HP 945 images. Here we see that the average camera digits for all 11 paints in the circular test target are almost identical. This means that the spatial processing has not compressed the LDR image. If we had used a Tone Scale map to compress the range of the HDR image, then that map would alter the LDR image. Appropriate spatial processing does not affect the LDR rendition. The spatial process used in Figure 5 shows no unwanted range compression of the LDR control image.

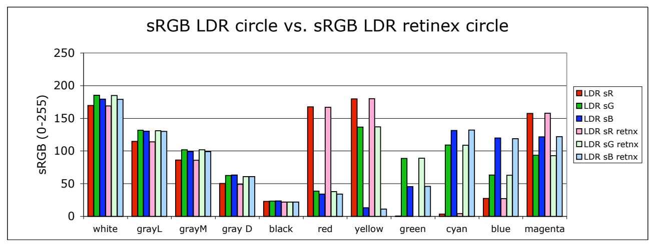

Figure 5 shows the effect of spatial processing on the LDR part of the scene.

Test 4 - Wanted Range Compression

Table 4 compares the outputs of the HP945 with its input image for the darkest regions of the HDR scene. It subtracts the input sRGB digits from the processed values.

Effect of Compression

	HDR sR	HDR sG	HDR sB	HDR sR SCSA	HDR sG SCSA	HDR sB SCSA	delta sR	delta sR	delta sR
white	6.9	12.6	25.5	40.8	49.9	73.2	**33.9**	**37.3**	**47.7**
grayL	3.8	4.8	12.9	29.0	34.7	56.6	**25.2**	**29.9**	**43.7**
grayM	2.6	2.8	6.5	22.6	25.5	40.1	**20.0**	**22.7**	**33.6**
grayD	1.2	1.2	2.9	14.1	12.3	20.6	**12.9**	**11.0**	**17.7**
black	0.6	0.5	1.7	8.4	7.0	14.0	**7.8**	**6.4**	**12.3**
red	7.5	1.2	2.6	44.8	10.8	15.0	**37.3**	**9.6**	**12.4**
yellow	9.2	5.1	2.6	50.1	36.2	21.3	**40.8**	**31.1**	**18.8**
green	1.3	2.6	4.4	9.2	25.2	24.1	**8.0**	**22.6**	**19.6**
cyan	0.9	3.4	9.7	9.7	28.4	49.5	**8.8**	**25.0**	**39.8**
blue	1.0	1.4	7.6	10.3	14.7	46.1	**9.3**	**13.4**	**38.4**
magenta	5.9	2.7	10.7	43.9	23.1	50.5	**38.0**	**20.4**	**39.8**
						average	**22.0**	**20.8**	**29.4**

Table 4 lists the sRGB values of the 11 paint circular target in normal HDR photograph, the processed HP945 values and their differences.

On average, this algorithm rendered the circular test target 22 sR, 21 sG and 29 sB digits lighter (Table 4). These numbers give us a quantitative measure of the scene's spatial range compression for this particular algorithm. This pair of tests is essential to evaluating the success of spatial processing: leave the LDR input unchanged (unwanted compression), while maximizing the wanted range compressing for the HDR input.

Test 5 - Appearance LDR vs HDR

Table 5 lists the changes in appearance measured by the watercolor painting. We converted the reflectance spectra to XYZ, and then to sRGB for the LDR and HDR portions of the watercolor On average, the HDR watercolor was darker than the LDR watercolor by 20 sR, 24 sG, and 19 sB calculated digits. These numbers give us a quantitative measure of the human dynamic range compression. Spatial algorithms that model human vision can use the technique to objectively measure success.

Change in Appearance LDR-HDR (sRGB)
LDR-HDR watercolor paintings

	LDR sR 255	LDR sG 255	LDR sB 255	HDR sR 255	HDR sG 255	HDR sB 255	delta sR	delta sR	delta sR
white	253	254	246	233	237	234	**-20**	**-17**	**-12**
grayL	192	195	195	151	153	155	**-41**	**-42**	**-39**
grayM	156	159	160	136	139	142	**-20**	**-20**	**-18**
gray D	124	126	128	89	89	92	**-36**	**-37**	**-36**
black	61	57	56	62	59	57	**1**	**1**	**1**
red	184	61	56	165	59	63	**-18**	**-1**	**7**
yellow	258	178	0	246	168	0	**-11**	**-10**	**0**
green	16	139	70	63	118	66	**47**	**-21**	**-4**
cyan	132	207	219	88	178	181	**-44**	**-30**	**-39**
blue	73	139	225	21	105	186	**-52**	**-34**	**-39**
magenta	237	207	233	209	156	204	**-28**	**-51**	**-28**
						average	**-20**	**-24**	**-19**

Table 5 lists circular test target sRGB in the LDR /HDR watercolor painting.

Test 6 - Appearance vs. Algorithm Rendering

We see from the results in Table 4 that the spatial processing in the HP945 makes the images 20 to 30 units lighter processed vs. unprocessed. Since we used the CREATE scenes as test targets we can use the variety of data to make further comparisons. For example, we can compare the spatial algorithms with human vision. If we want the spatial algorithm to mimic vision, then the sRGBs of the processed images should equal those of the watercolor painting. In Table 6 we list the sRGBs of the Carinna Parraman (CPHDR) painting. We also list the difference between the painting and the two spatial processes: HP945, and VV.

Compare Appearance (watercolor) with HP945 and VV HDR rendering

	CPHDR sR	CPHDR sG	CPHDR sB	HP945-CPsR	HP945-CPsG	HP945-CPsB	V V-CPsR	V V-CPsG	V V-CPsB
white	253.4	254.2	245.7	-245.0	-247.2	-231.7	-167.1	-170.9	-134.9
grayL	192.0	194.8	194.8	-151.2	-144.8	-121.6	-135.8	-135.1	-119.8
grayM	156.1	158.5	160.3	-112.2	-135.4	-109.8	-117.9	-119.1	-114.9
gray D	124.3	125.7	127.9	-79.6	-115.0	-112.9	-105.9	-105.7	-104.3
black	60.8	57.3	55.9	-46.6	-45.0	-35.3	-52.5	-46.7	-44.4
red	183.5	60.7	56.0	-173.2	-46.0	-9.9	-83.5	-49.4	-33.2
yellow	257.9	177.9	0.0	-207.9	-141.7	21.3	-143.6	-115.9	17.2
green	16.4	139.0	70.4	12.6	-104.3	-13.8	-9.0	-89.3	-50.5
cyan	131.7	207.4	219.4	-122.0	-179.0	-169.9	-116.7	-145.0	-135.7
blue	73.4	138.7	224.9	-64.2	-113.5	-200.8	-59.8	-111.9	-137.7
magenta	237.0	207.2	232.7	-214.4	-181.7	-192.6	-137.3	-166.7	-125.6
			average	**-127.6**	**-132.1**	**-107.0**	**-102.6**	**-114.2**	**-89.4**

Table 6 compares appearance measured by the HDR watercolor painting (CPHDR) with spatial color renditions. The center columns are sRGB values of [HP945-CPHDR]; the right columns are [V V - CPHDR].

We see that the 20 to 30 unit increases in HP945 image would need to be 115 more to accurately mimic vision, as measured by the watercolor. (If the set of HP945 images had more exposure the effects of processing would give slightly larger values. (The white sRGB in the painting is [253, 254, 246]; while the HP945 is [200, 203, 207]). We see that the values in VV algorithm in Figure 1 would need to be about 100 units lighter.

Most photographic algorithms are not designed to strictly match appearance. A model of vision must predict appearance. However, quantitative comparison of photographic images with vision is very helpful in establishing numerical goals of the model. Figure 6 plots the HDR sRGB values of the CP watercolor painting, the HP945 and the VV image renderings. While we see in that the VV is closer to the watercolor (Figure 6), it does not approach the image range compression in humans. Any spatial algorithm that attempts to make the best reproduction needs a different ground truth for optimizing that subjective analysis.

We can combine the results of these individual characteristics measurements. We see that the HP945 camera firmware and pre-spatial processing profiles modified scene radiances. We see in Figure 4 an increase in chroma and a compression of gray-scale in the image without spatial processing. These are distortions of the scene information by the cameras conversion of light to sRGB values. We also see a need for more camera exposure for these images. We see that this spatial processing did not introduce unwanted range compression of the LDR scene (Figure 5). We see that the HP945 spatial processing increased the sRGB values of the circular target in shade by about 20 units. The VV spatial processing increased those same values about 20 units more than the HP945. Both processes were considerably lower than the appearance values from the watercolor painting (Figure 6, Table 6). This comparison benefits from the advantage that they avoid the post-spatial-processing transformations that occurs in display and printing.

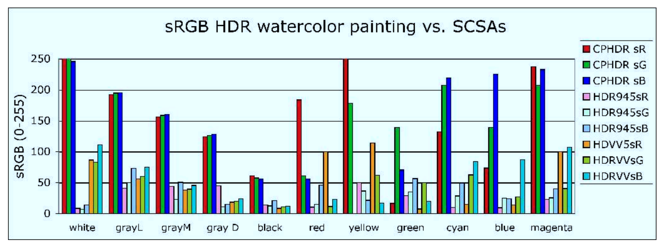

Figure 6 compares the HDR watercolor reflectance sRGBs with those of the two HDR processed images.

Discussion

This paper describes three parts of the puzzle of analyzing spatial algorithms:

First, it provides the download source of multiple exposures of digital photos of LDR/HDR CREATE 3-D Mondrian test target.

Second, it provides the measurements of paint reflectances, scene radiances and appearances of the LDR/HDR test target.

Third, it describes a few (6 tests) of the many possible analyses of image processing characteristics. Once one defines goal of the calculation, one can use these source images and calibration measurements to make objective evaluations of the characteristics of an algorithm.

With these examples in mind, one can measure the influence of camera pre-processing on all 104 colored facets for LDR and HDR images for each exposure. One can test an algorithm for undesirable range compression of LDR renditions. One can measure the amount of range compression for bright and shadow detail for HDR images. One can compare the algorithm's range compression with that found in humans. The examples of spatial algorithms, shown here, are not intended to be examples of optimal processes. Rather, the intent is to illustrate the tools that identify departure from optimal. This paper is not about the success of these algorithms, rather the means to quantitatively assess possible improvements.

The combination of photographs and measured data for the 104 block facets makes it possible to use spatial color metrics for these images. In the above examples, we report on the input/output values of single facets. The analysis of images is not limited to pixel comparisons. We plan to add spatial evaluation techniques, such as the comparison of ratios of pixel values, and the comparisons of multi-resolution segments of these images before and after spatial processing. Spatial evaluation metrics is an ongoing part of this project. Before presenting this approach we are working to reduce the distortions in camera capture, namely improve the quality of the input image, based of the KM meter scene radiance measurements.

We began by looking for the best objective metric for spatial image processing. We described how the signal processing has three categories, each with many steps. The image capture (pre-spatial processing) identifies all the operations between the light from the scene and the input for the spatial algorithm. Image display (post-processing) includes all the operations and hardware artifacts between the computer's digital output and the light reaching the observers' eyes. Profiles, or standards, are used to control these pre- and post-processing operations. They effect the process in two fundamental ways: First the distort the scene information, and second affect the appearance of displayed images. These two large effect need to be isolated from the analysis of spatial image processes. Using different, or inappropriate, pre- and post-processing profiles invalidates comparisons of spatial algorithms.

Quantitative measurements of the spatial characteristics of the imaging chain are necessary because they can measure the effects of pre-processing and remove the effects of post-processing. Looking at an image on a display, or a print, includes the unwanted influence of the post processes. We begin with a scene. We can think of it as a set of three-dimensional scene radiances. They are a distribution in color space. These radiances are transformed by the pre-processing into a new, camera-specific distribution in color space. The pre-processing has changed the relationship of these colors. We applied two spatial algorithms, each of which uniquely transforms the relationships of colors. When we view the processed images we add still another transform of color relationships with post-processing.

Comparing results of different spatial algorithms by looking at the pictures just measures how well the algorithm's color-space rendition anticipates the display's preset post processing. Each spatial algorithm creates a unique 3-D set of color outputs. That space may, or may not, be the one that the display is expecting. Without individual optimal post-processing of each algorithm, visual comparisons are meaningless. Objective numerical comparisons, not visual comparisons, give a better evaluation of spatial image processing. These numerical comparisons avoid the post-processing transformations.

Summary

Although we would like to have a simple objective program that gives us a reliable figure of merit for our favorite algorithm, we found this goal was impractical. The spatial color algorithm is in the middle of the imaging chain and its success is affected by pre- and post-processing. The sRGB input and output values are not in a uniform color space, so sRGB distances between actual and goal have variable differences in appearance. Mean-squared-distance calculations in a nonuniform 3-D space have questionable value. Post-processing properties of individual devices add unknown distortions to the analysis.

There are a variety of goals for different spatial algorithms: one is to find the objects reflectance; one is to find the illumination; one is to make the best HDR picture; another is to model human vision. There are different ground truth goals for each type of algorithm.

Instead of presenting a universal solution to evaluate all types of algorithms, we describe a number of steps that evaluate spatial processes. We describe examples of a number of control and test experiments that are useful in quantitative evaluation of portions of the imaging chain. The goal here is to provide test images, measurements of scene characteristics, and examples of a set of flexible tools for quantitative evaluations of spatial color synthesis algorithms. Whereas subjective selection of preferred pictures is confounded by the effect of post-processing, quantitative measurements of spatial algorithms evaluates the true performance of the central spatial process.

References

<http://sites.google.com/site/3dmondrians/>

[1] Land, E. H., & McCann, J. J. (1971) Lightness and retinex theory, JOSA, 61, 1-11.
[2] Wyszecki, G. (1973) Colorimetry, in Color Theory and Imaging Systems, SPSE, R. Eynard, Ed., Washington, 24-49.
[3] Fedorovskaya, E., de Ridder, H. & Blommaert, F. J. J., Chroma variations and perceived quality of color images of natural scenes, Color Research & Application, 22, 96–110.
[4]Kuang, J., Yamaguchi, H, Liu, C., Johnson, G, & Fairchild, M. (2007) Evaluating HDR rendering algorithms, ACM Transactions on Applied Perception (TAPI), 4 (2), 9.
[5] Parraman, C., McCann, J. J. & Rizzi, A, (2010) Artist's colour rendering of HDR scenes in 3-D Mondrian colour-constancy experiments, in Proc. IS&T/SPIE Electronic Imaging, San Jose, 7528-1.
[6] McCann, J. J., Parraman, C. E., & Rizzi, A. (2009b) Reflectance, Illumination and Edges, Proc. CIC, Albuquerque, CIC 17, 2-7.
[7] Vonikakis, V. Andreadis, I., & Gasteratos, A. (2008). Fast centre-surround contrast modification. IET Image processing, 2(1), 19-34.
[8] Sobol, R. & McCann, J. J. (2002) Method for variable contrast mapping of digital images, U.S. Patent, Oct, 24.
[9] Sobol, R. (2004). Improving the Retinex algorithm for rendering wide dynamic range photographs, J. Electronic Imaging 13(01), 65-74.
[10] Feng, X. & Daly, S. C. (2005). Improving color characteristics of LCD. in Proc. IS&T/SPIE EI, Color Imaging X, 5667, 328-335.
[11] Feng, X. & Yoshida, Y, (2008). Improving the Gray Tracking Performance of LCD, Proc. CGIV, 327-330.

Author Biography

John McCann, McCann Imaging; Vassilios Vonikakis, Democritus University of Thrace, Greece; Carinna Parraman, University of the West of England; and Alessandro Rizzi, University of Milan, are active participants in CREATE <http://www.create.uwe.ac.uk/>.