A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Scenes

Daniel J. Jobson, *Member, IEEE,* Zia-ur Rahman, *Member, IEEE,* and Glenn A. Woodell

***Abstract*— Direct observation and recorded color images of the same scenes are often strikingly different because human visual perception computes the conscious representation with vivid color and detail in shadows, and with resistance to spectral shifts in the scene illuminant. A computation for color images that approaches fidelity to scene observation *must* combine dynamic range compression, color consistency—a computational analog for human vision color constancy—and color and lightness tonal rendition. In this paper, we extend a previously designed single-scale center/surround retinex to a multiscale version that achieves simultaneous dynamic range compression/color consistency/lightness rendition. This extension fails to produce good color rendition for a class of images that contain violations of the gray-world assumption implicit to the theoretical foundation of the retinex. Therefore, we define a method of color restoration that corrects for this deficiency at the cost of a modest dilution in color consistency. Extensive testing of the multiscale retinex with color restoration on several test scenes and over a hundred images did not reveal any pathological behavior.**

I. Introduction

A COMMON (and often serious) discrepancy exists between recorded color images and the direct observation of scenes (see Fig. 1). Human perception excels at constructing a visual representation with vivid color and detail across the wide ranging photometric levels due to lighting variations. In addition, human vision computes color so as to be relatively independent of spectral variations in illumination [1]; i.e., it is color constant. The recorded images of film and electronic cameras suffer, by comparison, from a loss in clarity of detail and color as light levels drop within shadows, or as distance from a lighting source increases. Likewise, the appearance of color in recorded images is strongly influenced by spectral shifts in the scene illuminant. We refer to the computational analog to human vision color constancy as color consistency. When the dynamic range of a scene exceeds the dynamic range of the recording medium, there is an irrevocable loss of visual information at the extremes of the scene dynamic range. Therefore, improved fidelity of color images to human observation demands i) a computation that synthetically combines dynamic range compression, color consistency, and color and lightness rendition, and ii) wide dynamic range color imaging systems. The multiscale retinex (MSR) approaches the first of these goals. The design of the computation is tailored to visual perception by comparing the measured photometry of scenes with the performance of visual perception. This provides a rough quantitative measure of human vision's dynamic range compression—approaching $1000:1$ for strong illumination variations of bright sun to deep shade.

The idea of the retinex was conceived by Land [2] as a model of the lightness and color perception of human vision. Through the years, Land evolved the concept from a random walk computation [3] to its last form as a center/surround spatially opponent operation [4], which is related to the neurophysiological functions of individual neurons in the primate retina, lateral geniculate nucleus, and cerebral cortex. Subsequently, Hurlbert [5]–[7] studied the properties of this form of retinex and other lightness theories and found that they share a common mathematical foundation but cannot actually compute reflectance for arbitrary scenes. Certain scenes violate the "gray-world" assumption—the requirement that the average reflectances in the surround be equal in the three spectral color bands. For example, scenes that are dominated by one color—"monochromes"—clearly violate this assumption and are forced to be gray by the retinex computation. Hurlbert further studied the lightness problem as a learning problem for artificial neural networks and found that the solution had a center/surround spatial form. This suggests the possibility that the spatial opponency of the center/surround is, in some sense, a general solution to estimating relative reflectances for arbitrary lighting conditions. At the same time, it is equally clear that human vision does not determine relative reflectance, but rather a context-dependent relative reflectance since the same surfaces in shadow and light do not appear to be the same. Moore *et al.* [8], [9] took up the retinex problem as a natural implementation for analog very large scale integration (VLSI) resistive networks and found that color rendition was dependent on scene content—whereas some scenes worked well, others did not. These studies also pointed out the problems that occur due to color Mach bands and the graying-out of large uniform zones of color.

We have previously defined a single-scale retinex [10] (SSR) that can either provide dynamic range compression (small scale), or tonal rendition (large scale), but not both simultaneously. The multiscale retinex with color restoration (MSRCR) combines the dynamic range compression of the small-scale retinex and the tonal rendition of the large scale

Manuscript received March 21, 1996; revised January 28, 1997. The work of Z. Rahman was supported by NASA Contract NAS1–1903 and NASA Grant NAG1–1847. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. H. Joel Trussell.

D. J. Jobson and G. A. Woodell are with the NASA Langley Research Center, Hampton, VA 23681-0001 USA (e-mail: d.j.jobson@larc.nasa.gov).

Z. Rahman was with Science and Technology Corporation, Hampton, VA USA. He is now with the Department of Computer Science, College of William and Mary, Williamsburg, VA 23187 USA.

Publisher Item Identifier S 1057-7149(97)04726-X.

Fig. 1. Illustration of the discrepancy between color images and and perception. The right image is a much closer representation of the visual impression of the scene.

retinex with a universally applied color restoration. This color restoration is necessary to overcome the problems that the MSR has in the rendition of scenes that contain gray-world violations. It merges all the necessary ingredients to approximate the performance of human vision with a computation that is quite automatic and reasonably simple. These attributes make the MSRCR attractive for smart camera applications, in particular for wide dynamic range color imaging systems. For more conventional applications, the MSRCR is useful for enhancing 8-b color images that suffer from lighting deficiencies commonly encountered in architectural interiors and exteriors, landscapes, and nonstudio portraiture.

Most of the emphasis in previous studies has been on the color constancy property of the retinex, but its dynamic range compression is visually even more dramatic. Since we want to design the retinex to perform in a *functionally* similar manner to human visual perception, we begin with a comparison of the photometry of scenes to their perception. This defines (at least in some gross sense) the performance goal for the retinex dynamic range compression.

An apparent paradox has been brought to our attention by a colleague as well as a reviewer. This paradox is so fundamental that it requires careful consideration before proceeding. The question, simply stated, is why should recorded images need dynamic range compression, since the compression of visual perception will be performed when the recorded image is observed? First we must state categorically that recorded images with significant shadows and lighting variations *do need compression*. This has been our experience in comparing the perception of recorded images with direct observation for numerous scenes. Therefore, we have to conclude that the dynamic range compression for perception of the recorded images is substantially weaker than for the scene itself. Fig. 1 is a case in point. There is no linear representation of this image, such as the viewing of the image on a gamma-corrected cathode ray tube (CRT) display, which even comes close to the dynamic compression occurring during scene observation. The same is true for all scenes we have studied with major lighting variations. We offer the possible explanation that weak dynamic range compression can result from the major differences in angular extent between scene and image viewing. Image frames are typically about 40° in angular extent for a 50 mm film camera. These same frames are usually viewed with about a 10° display or photographic print. Furthermore, the original 40° frame is taken out of the larger context, which would be present when observing the scene directly. The dynamic range compression of human vision is strongly dependent upon the angular extent of visual phenomena. Specifically, compression is much stronger for large shadow zones than for smaller ones. We feel that this a plausible resolution for this apparent paradox, and are certainly convinced by considerable experience that recorded images do need computational dynamic range compression for scenes that contain significant lighting variations. Likewise, this explanation applies to color consistency.

Since the nonlinear nature of the MSR makes it almost impossible to prove its generality, we provide the results of processing many test images as a measure of confidence in its general utility and efficacy. Results obtained with test scenes—i.e., where direct observation of the subject of the image is possible—are given more weight because the performance of the computation can be compared directly to observation of the scene.

II. The Photometry of Scenes Compared to Perception

We approached learning more about the dynamic range compression in human vision by exploring the perceptual and photometric limits. We did this by selecting and measuring

TABLE I
PHOTOMETRY OF SCENES

	cd/m^2
Visual saturation (white clouds near sun)	49,000
Just below saturation (clouds further from sun)	18,000–37,000
Outdoor building facade—bright sun	7000–13,000
Blue sky—morning	4600
Concrete sidewalk in	
sun	3200
shadow	570
deep shadow	290
Interior conference room-fluorescent lighting	
Floor/walls	36–140
Shadows	4–18
Interior conference room-unlit but by window	
Walls	29
Shadows	6
Inside open closet	1

scenes with increasingly emphatic lighting variations and then examining the point at which dynamic range compression gives way to loss of visual information. In other words, we looked for the dynamic range extremes at which human vision either saturates or clips the signals from very dark zones in a scene. We used a photographic spotmeter for the photometric measurements. In addition, we attempted to calibrate the perceptual lightness difference that occurs when the same surface is viewed in direct sunlight and in shadow. To quantify this difference, we compared the perceived lightness under both conditions to a reference gray-scale in direct sun and asked the question: Which gray scales match the surface in sun and shadow? Whereas the extreme measurements provide information about where dynamic range compression becomes lossy, the sun/shadow/gray-scale matches give some measure of the dynamic range compression taking place within more restricted lighting changes.

The results of the photometric measurements are given in Table I. The conditions shown are representative of the wide dynamic range encountered in many everyday scenes. Scene visibility is good except under the most extreme lighting conditions. On the low end, visibility is quite poor at 1 candles/m^2 (cd/m^2) luminance but improves rapidly as light levels approach 10 cd/m^2. Detail and color are quite easily visible across the range of 10–10 000 cd/m^2, even when all occur together in a scene. We can therefore conclude that dynamic range compression within a scene can approach 1000 : 1, but becomes lossy for wider ranges. For low luminance, color and detail are perceptually hazy with a loss of clarity; and for extremely low levels of luminance (approaching 10 000 : 1 when compared with direct sunlight), all perception of color and detail is lost.

We can also quantitatively estimate from this data the difference between perception and photometry for a very commonly encountered case: objects in sun and shadow. The drop in light level usually associated with a shadow is between 10–20% of the sunlit value, depending on the depth of the shadow. We compared the perceived drop in lightness to a reflectance gray-scale and concluded that the perceptual decrease is only about 50% of the sunlit lightness value. This clearly demonstrates the large discrepancy between recorded images and perception, even for conditions that do not encompass a very wide dynamic range. This data implies that for 10 : 1 changes in lighting, the perception of these changes is about 3–5 : 1 to minimize the impact of lighting on the scene representations formed by consciousness. Hence, as simple and ubiquitous an event as a shadow immediately introduces a major discrepancy between recorded images and visual perception of the same scene. This sets a performance goal derived from human visual perception with which to test the retinex. Clearly, a very strong nonlinearity exists in human vision, although our experiments can not define the exact form of this neural computation.

III. CONSTRUCTION OF A MULTISCALE CENTER/SURROUND RETINEX

The single-scale retinex [10]–[12] is given by

$$R_i(x,y) = \log\ I_i(x,y) - \log\left[F(x,y) * I_i(x,y)\right] \quad (1)$$

where $R_i(x,y)$ is the retinex output, $I_i(x,y)$ is the image distribution in the ith spectral band, "*" denotes the convolution operation, and $F(x,y)$ is the surround function

$$F(x,y) = Ke^{-r^2/c^2}$$

where c is the Gaussian surround space constant, and K is selected such that

$$\iint F(x,y)\,dx\,dy = 1.$$

The MSR output is then simply a weighted sum of the outputs of several different SSR outputs. Mathematically,

$$R_{\mathrm{MSR}_i} = \sum_{n=1}^{N} w_n R_{n_i} \quad (2)$$

where N is the number of scales, R_{n_i} is the ith component of the nth scale, R_{MSR_i} is the ith spectral component of the MSR output, and w_n is the weight associated with the nth scale. The only difference between $R(x,y)$ and $R_n(x,y)$ is that the surround function is now given by

$$F_n(x,y) = Ke^{-r^2/c_n^2}.$$

A new set of design issues emerges for the design of the MSR in addition to those for the SSR [10]. This has primarily to do with the number of scales to be used for a given application, and how these realizations at different scales should be combined. Because experimentation is our only guide in resolving these issues, we conducted a series of tests starting with only two scales and adding further scales as needed. After experimenting with one small scale $(c_n < 20)$ and one large scale $(c_n > 200)$, the need for a third intermediate scale was immediately apparent in order to produce a graceful rendition without visible "halo" artifacts near strong edges. Experimentation showed that equal weighting of the scales—$w_n = 1/3, n = 1,2,3$—was sufficient for most

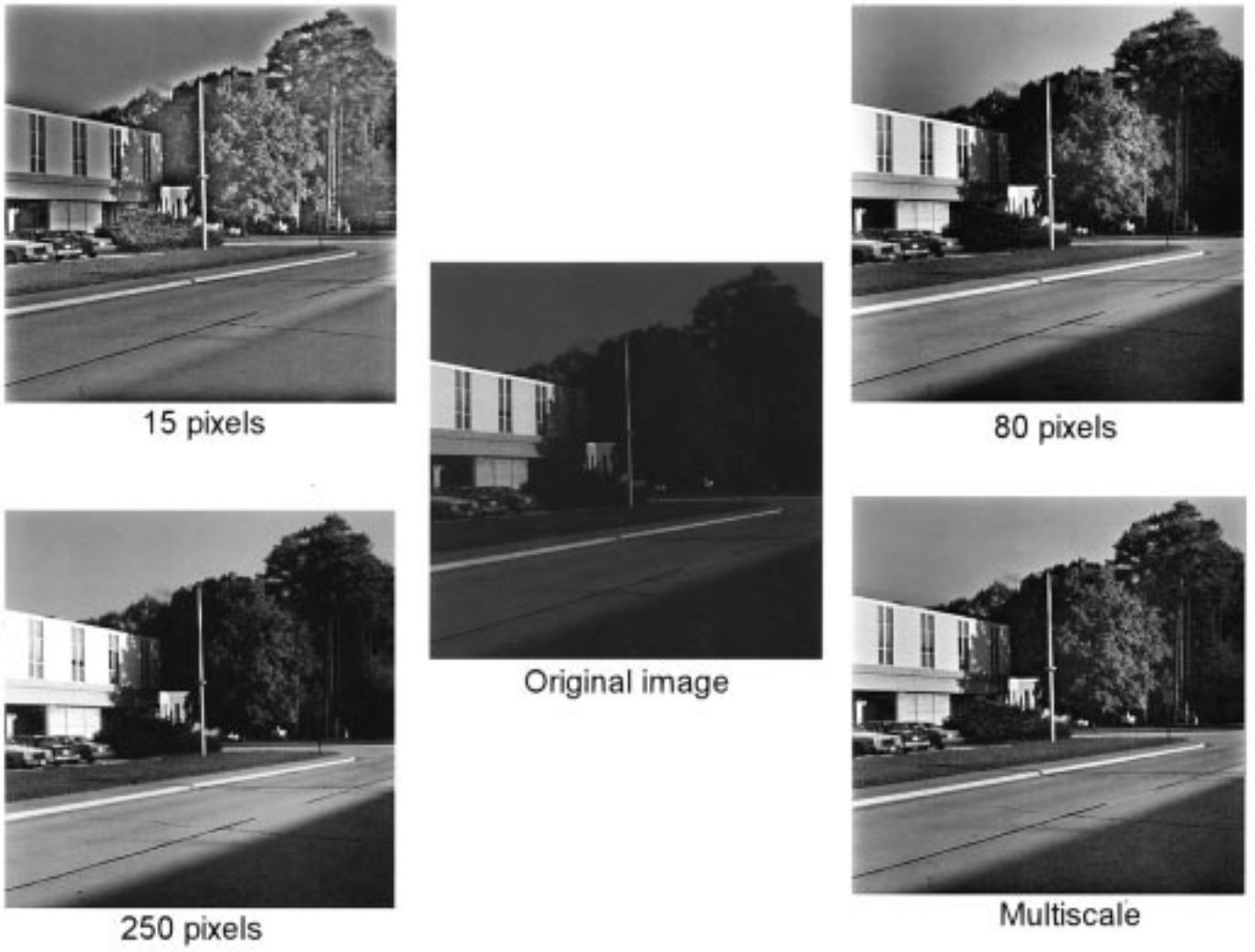

Fig. 2. Components of the multiscale retinex that show their complementary information content. The smallest scale is strong on detail and dynamic range compression and weak on tonal and color rendition. The reverse is true for the largest spatial scale. The multiscale retinex combines the strengths of each scale and mitigates the weaknesses of each.

applications. Weighting the smallest scale heavily to achieve the strongest dynamic range compression in the rendition leads to ungraceful edge artifacts and some graying of uniform color zones.

To test whether the dynamic range compression of the MSR approaches that of human vision, we used test *scenes* that we had observed in addition to test images that we had obtained from other test sources. The former allowed us to readily compare the processed image to the direct observation of the scene. Fig. 2 illustrates the complementary strengths and weaknesses of each scale taken separately and the strength of the multiscale synthesis. This image is representative of a number of test scenes (see Fig. 3) where for conciseness we show only the multiscale result.

The comparison of the unprocessed images to the perception of the scene produced some striking and unexpected results. When direct viewing was compared with the recorded image, the details and color were far more vivid for direct viewing not only in shadowed regions, but also in the bright zones of the scene! This suggests that human vision is doing even more image enhancement than just strong dynamic range compression, and the MSR may ultimately need to be modified to capture the realism of direct viewing. Initially, we tackle the dynamic range compression, color consistency, and tonal/color rendition problems, while keeping in mind that further work may be necessary to achieve full realism.

A sample of image data for surfaces in both sun and shadow indicates a dynamic range compression of $2:1$ for the MSR compared to the 3–$5:1$ measured in our perceptual tests. For the SSR $(c_1 = 80)$ this value is $1.5:1$ or less. These levels of dynamic range compression are for outdoor scenes where shadows have large spatial extent. Shadows of small spatial extent tend to appear "darker" and are more likely to be clipped in recorded images. Fig. 3 shows a high dynamic range indoor/outdoor scene. The foreground orange book on the gray-scale is compressed by approximately $5:1$ for the MSR while compression for the SSR is only about $3:1$, both relative to the bright building facade in the background.

Fig. 3. Examples of test scenes processed with the multiscale retinex prior to color restoration. While color rendition of the left image is good, the other two are "grayed" to some extent. Dynamic range compression and tonal rendition are good for all and compare well with scene observation. Top row: Original. Bottom row: Multiscale retinex.

The compression for human vision is difficult to estimate in this case, since both the color and texture of the two surfaces are quite different. Our impression from this analysis is that the MSR is approaching human vision's performance in dynamic range compression but not quite achieving it. For scenes with even greater lighting dynamics than these, we can anticipate an even higher compression for the MSR to match human vision. However, we are currently unable to test this hypothesis because the conventional 8-b analog-to-digital converters of both our solid-state camera and slide film/optical scanner digitizer restrict the dynamic range with which the image data for such scenes can be acquired. Solid state cameras with 12-b dynamic range and thermoelectrically cooled detector arrays with 14-b dynamic range are, however, commercially available, and can be used for examining the MSR performance on the wider dynamic range natural scenes. Even for the restricted dynamic range shown in Fig. 3 (left), it is obvious that limiting noise has been reached, and that much wider dynamic range image acquisition is essential for realizing a sensor/processing system capable of approximating human color vision.

For the conventional 8-b digital image range, the MSR performs well in terms of dynamic range compression, but its performance on the pathological classes of images examined in previous SSR research [10] must still be examined. Fig. 4 shows a set of images that contain a variety of regional and global gray-world violations. The MSR, as expected, fails to handle them effectively—all images possessing notable, and often serious, defects in color rendition (see Fig. 4, middle row). We only provide these results as a baseline for comparison with the color restoration scheme, presented in the next section, that overcomes these deficiencies of the MSR.

IV. A Color Restoration Method for the Multiscale Retinex

The general effect of retinex processing on images with regional or global gray-world violations is a "graying out" of the image, either globally or in specific regions. This desaturation of color can, in some cases, be severe (see Fig. 4, middle). More rarely, the gray-world violations can simply produce an unexpected color distortion (see Fig. 4, top left). Therefore, we consider a color restoration scheme that provides good color rendition for images that contain gray-world violations. We, of course, require the restoration to preserve a reasonable degree of color consistency, since that is one of the prime objectives of the retinex. Color constancy is known to be imperfect in human visual perception, so some level of illuminant color dependency is acceptable, provided it is much lower than the physical spectrophotometric variations. Ultimately, this is a matter of image quality, and

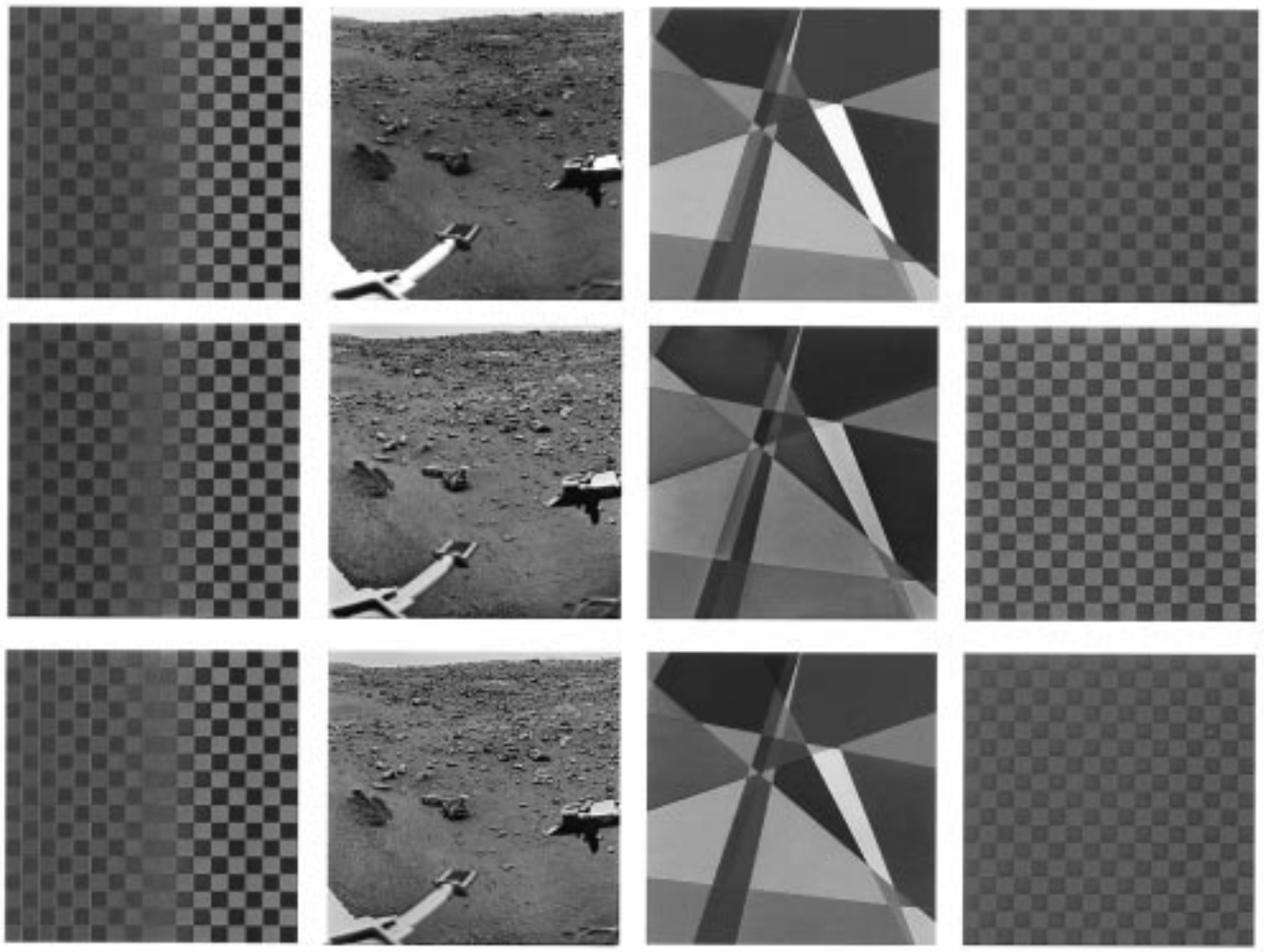

Fig. 4. Pathological "gray-world" violations are not handled well by the multiscale retinex alone (middle row), but are treated successfully when color restoration is added (lower row). Top row: Original.

color dependency is tolerable to the extent that the visual defect is not visually too strong.

We begin by considering a simple colorimetric transform [13], even though it is often considered to be in direct opposition to color constancy models. It is also felt to describe only the so-called "aperture mode" of color perception, i.e., restricted to the perception of color lights rather than color surfaces [14]. The reason for this choice is simply that it is a method for creating a relative color space, and in so doing becomes less dependent than raw spectrophotometry on illuminant spectral distributions. This starting point is analogous to the computation of chromaticity coordinates where

$$I_i'(x,y) = I_i(x,y) / \sum_{i=1}^{S} I_i(x,y) \tag{3}$$

for the ith color band, and S is the number of spectral channels. Generally, $S = 3$, using the red–green–blue (RGB) color space. The modified MSR that results is given by

$$R_{\mathrm{MSRCR}_i}(x,y) = C_i(x,y) R_{\mathrm{MSR}_i}(x,y) \tag{4}$$

where

$$C_i(x,y) = f[I_i'(x,y)]$$

is the ith band of the color restoration function (CRF) in the chromaticity space, and R_{MSRCR_i} is the ith spectral band of the multiscale retinex with color restoration. In a purely empirical manner, we tried several linear and nonlinear color restoration functions on a range of test images. The function that provided the best overall color restoration was

$$\begin{aligned} C_i(x,y) &= \beta \log[\alpha I_i'(x,y)] \\ &= \beta \left\{ \log[\alpha I_i(x,y)] - \log\left[\sum_{i=1}^{S} I_i(x,y)\right] \right\} \end{aligned} \tag{5}$$

where β is a gain constant, and α controls the strength of the nonlinearity. In the spirit of a preserving a canonical computation, we determined that a single set of values for β and α worked for all spectral channels. The final MSRCR output is obtained by using a "canonical" gain/offset to transition between the logarithmic domain and the display domain. Looking at the forms of the CRF of (5) and the SSR of (1), we conjecture that the CRF represents a spectral analog to the spatial retinex. This mathematical and philosophical

TABLE II
LIST OF CONSTANTS USED FOR ONE PARTICULAR IMPLEMENTATION OF THE MSRCR ON A DEC ALPHA 3000, USING THE VMS F77 COMPILER

Constant	N	c_1	c_2	c_3	G	b	α	β	w_n
Value	3	15	80	250	192	-30	125	46	1/3

symmetry is intriguing, since it suggests that there may be a unifying principle at work. Both computations are nonlinear, contextual, and highly relative. We can speculate that the visual representation of wide dynamic range scenes must be a compressed mesh of contextual relationships for lightness and color representation. This sort of information representation would certainly be expected at more abstract levels of visual processing such as form information composed of edges, links, and the like, but is surprising for a representation so closely related to the raw image. Perhaps in some way this front-end computation can serve later stages in a presumed hierarchy of machine vision operations that would ultimately need to be capable of such elusive goals as resilient object recognition.

The bottom row in Fig. 4 shows the results of applying the CRF to the MSR output for pathological images. The MSRCR provides the necessary color restoration, eliminating the color distortions and gray zones evident in the MSR output. The challenge now is to prove the generality of this computation. Since there is not a mathematical way to do this, we have tested the computation on several hundred highly diverse images without discovering exceptions. Unfortunately, space considerations allow us to present only a very small subset of all the images that we have tested.

V. SELECTED RESULTS FOR DIVERSE TEST CASES

Extensive testing indicates that the gain constant α for the CRF and the final gain/offset adjustment required to transition from the logarithmic to the display domain are independent of the spectral channel and the image content. This implies that the method is general or "canonical," and can be applied automatically to most (if not all) images without either interactive adjustments by humans or internal adjustments such as an auto-gain. This final version of the MSRCR can then be written as

$$R_{\mathrm{MSRCR}_i}(x,y) = G[C_i(x,y)\{\log I_i(x,y) - \log[I_i(x,y) * F_n(x,y)]\} + b] \quad (6)$$

where G and b are the final gain and offset values, respectively. The constants G and b intrinsically depend upon the implementation of the algorithm in software. Table II gives a list of the constants used to produce all the outputs in this paper.

We must again emphasize that the choice of the all constants merely represents a particular implementation that works well for a wide variety of images. In no way do we mean to imply that these constants are optimal or "best case" for all possible implementations of this algorithm. The choice of the surround space constants, c_ns, in particular does not seem to be critical. Instead, the choice seems to only need to provide reasonable coverage from local to near global. Likewise, the choice of using three scales was made empirically to provide the minimum number of scales necessary for acceptable performance.

The test images presented here begin with some test scenes since we feel it is fundamental to refer the processed images back to the direct observation of scenes. This is necessary to establish how well the computation represents an observation. Clearly, we cannot duplicate human vision's peripheral vision which spans almost 180°, but within the narrower angle of most image frames, we would like to demonstrate that the computation achieves the clarity of color and detail in shadows, reasonable color constancy and lightness and color rendition that is present in direct observation of scenes. The test scenes (see Fig. 5) compare the degree with which the MSRCR approaches human visual performance. All four of the MSRCR outputs shown in Fig. 5 are quite "true to life" compared to direct observation, except for the leftmost, which seems to require even more compression to duplicate scene perception. This image was scanned from a slide and digitized to 8-b/color. The other three images were taken with a Kodak DCS200C CCD detector array camera. In none of the cases could a gamma correction produce a result consistent with direct observation. Therefore, we conclude that the MSRCR is not correcting simply for a CRT display nonlinearity, and that far stronger compression than gamma correction is necessary to approach fidelity to visual perception of scenes with strong lighting variations. We did not match camera spatial resolution to observation very carefully, so some difference in perceived detail is expected and observed. However, overall color, lightness, and detail rendering for the MSRCR is a good approximation to human visual perception.

The rest of the selected test images (Figs. 6–8) were acquired from a variety of sources (see acknowledgments) and provide as wide a range of visual phenomena as we felt could be presented within the framework of this paper. Little comment is necessary and we will leave the ultimate judgment to the reader. Some images with familiar colors and no strong lighting defects are included to show that the MSRCR does not introduce significant visual distortions into images that are without lighting variations. The white stripes of the American flag in Fig. 6(a) show a shift toward blue-green in the MSRCR output. This is, perhaps, analogous to the simultaneous color contrast phenomena of human perception. Moore *et al.* [8] noted a similar effect in their implementation of a different form of the retinex. The Paul Klee painting in Fig. 7(b) is included as a test of the subtlety of tonal and color rendition. Some of the test images with strong shadows zones where one or two color channels are preferentially clipped do exhibit a color distortion. This is due to the rather limited dynamic range of the "front-end" imaging/digitization, and is not an artifact of the computation. Even for these cases, the MSRCR produces far more visual information and is more "true-to-life" than the unprocessed image. The set of space images are included to show the application of the MSRCR to both space operations imagery and remote sensing applications.

A further test is worthwhile in assessing the impact of the CRF on color consistency. The CRF, as expected, dilutes color consistency, as shown in Fig. 9. However, the residual color dependency is fairly weak and the visual impression of color shift is minimal especially in comparison with the dramatic shifts present in the unprocessed images.

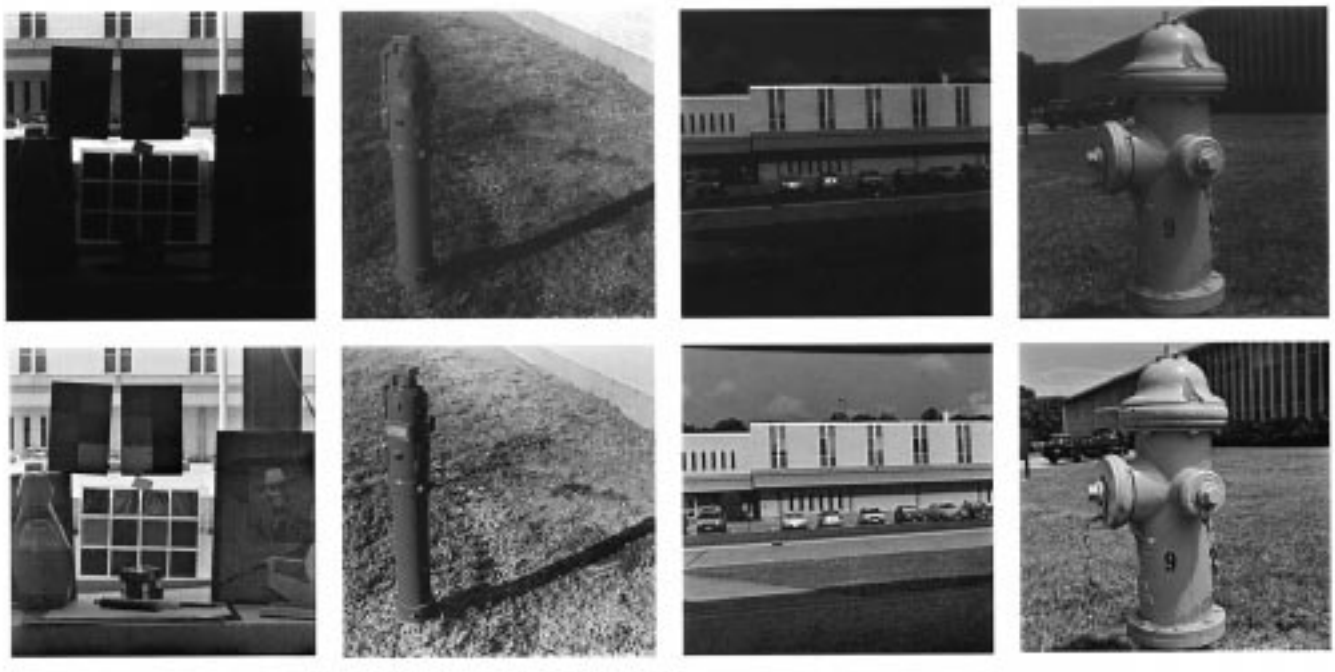

Fig. 5. Test scenes illustrating dynamic range compression, color, and tonal rendition, and automatic exposure correction. All processed images compare favorably with direct scene observation with the possible exception of leftmost image, which is even lighter and clearer for observation. This scene has the widest dynamic range and suggests that even stronger dynamic range compression may be needed for this case. Top row: Original. Bottom row: Multiscale retinex.

Fig. 6. Photographic examples further illustrating graceful dynamic range compression together with tonal and color rendition. The rightmost image shows the processing scheme handling saturated colors quite well and not distorting an image that is quite good in its original form. Top row: Original. Bottom row: Multiscale retinex.

VI. Discussion

While we have not yet conducted an extensive performance comparison of the MSRCR to other image enhancement methods, we have done some preliminary tests of the MSRCR relative to the simpler image enhancement methods—histogram equalization, gamma correction, and gain/offset manipulation [15], and point logarithmic nonlinearity [16]. Overall, the performance of the retinex is consistently good, while performance for the others is quite variable. In particular, the retinex excels when there are major zones of both high and low light levels. The traditional methods that we have compared against are all point operations on the image,

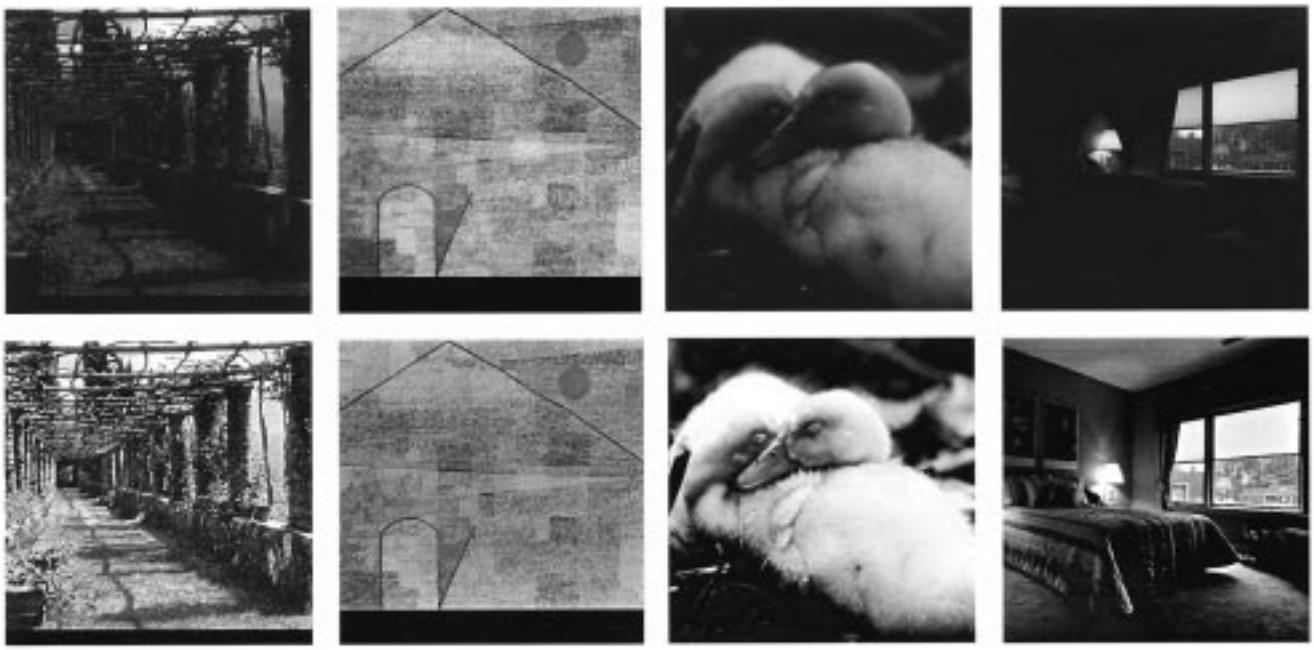

Fig. 7. Miscellaneous examples illustrating fairly dramatic dynamic range compression as well one for subtlety of color rendition (second from leftmost—painting by Paul Klee). Top row: Original. Bottom row: Multiscale retinex.

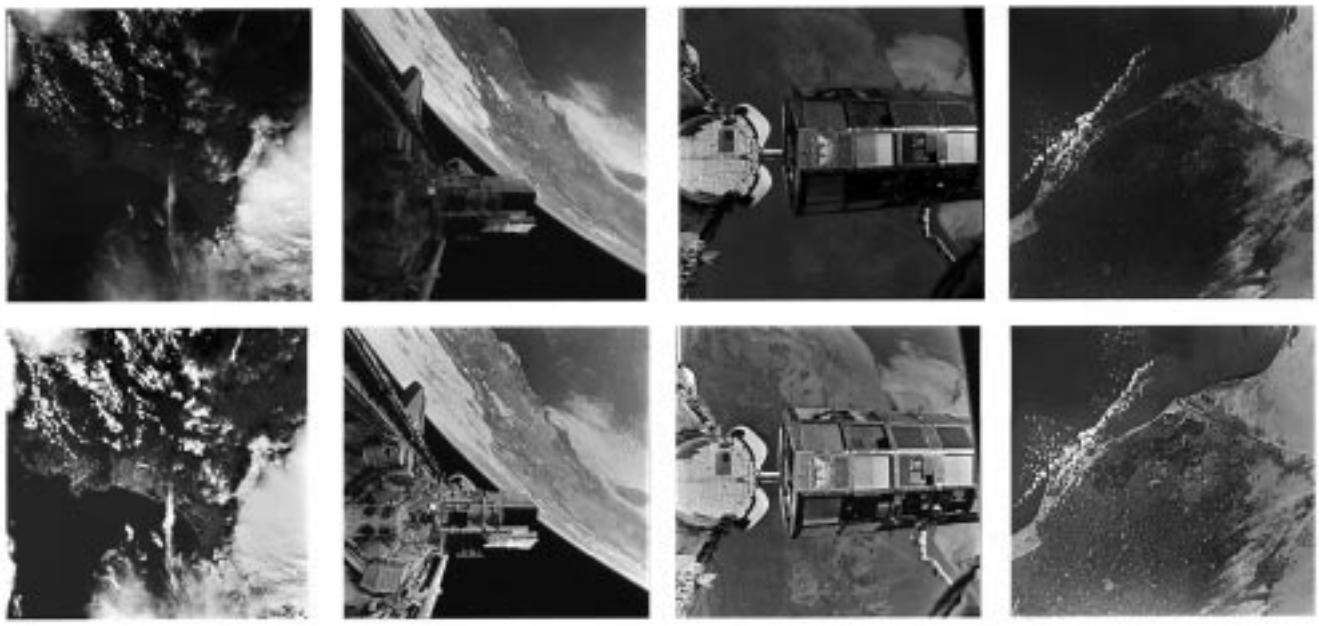

Fig. 8. Selection of space images to show enhancement of space operations imagery and remote sensing data. Top row: Original. Bottom row: Multiscale retinex.

whereas unsharp masking [17] and homomorphic filtering [17], [18] are spatial operations more mathematically akin to center/surround operation of the retinex. Unsharp masking is a linear subtraction of a blurred version of the image from the original and is generally applied using slight amounts of blurring. For a given space constant for the surround, we would expect the retinex to be much more compressive. It is not clear that unsharp masking would have any color constancy property, since the subtraction process in the linear domain is essentially a highpass filtering operation and not a ratio that provides the color constancy of the retinex.

Homomorphic filtering is perhaps the closest computation to the MSRCR and in one derivation [19] has been applied to color vision. Both its original form and the color form rely upon a highpass filtering operation that takes place after the dynamic range of the image is compressed with a point log-

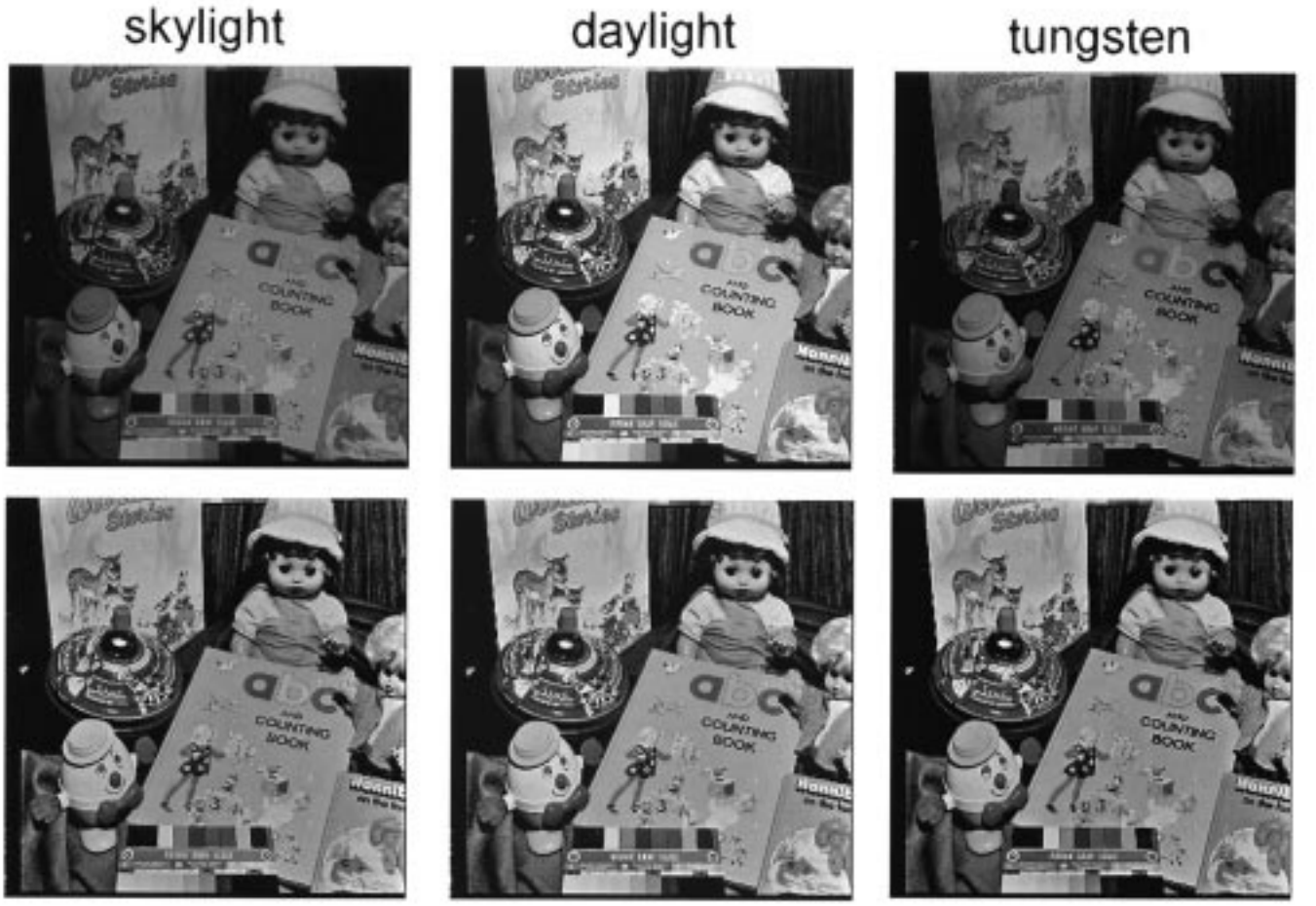

Fig. 9. Toy scene revisited. A test of the dilution of color consistency by the color restoration. While color consistency was shown previously to be near perfect for the SSR and MSR, some sacrifice of this was necessary to achieve color rendition. While slight changes in color can be seen, color consistency is still quite strong relative to the spectrophotometric changes seen in the original images (top row). The blues and yellows are in the color restored multiscale retinex (bottom row) are the most affected by the computer simulated spectral lighting shifts, but the effect is visually weak and most colors are not visibly affected.

arithmic nonlinearity. An inverse exponentiation then restores the dynamic range to the original display space. The color vision version adds an an opponent-color/achromatic transformation after the application of the logarithmic nonlinearity. We have found that the application of the logarithmic nonlinearity before spatial processing gives rise to emphatic "halo" artifacts and have also shown that it is quite different visually and mathematically from the application of the log after the formation of the surround signal [10]. Because of the nonlinearities in both the MSRCR and homomorphic filtering, a straightforward mathematical comparison is not possible. We do, however, anticipate significant performance differences between the two in terms of dynamic range compression, rendition, and, for the color vision case, color consistency. Another major difference between the MSRCR and homomorphic filtering is in the application of the inverse function in homomorphic filtering. The analogous operation in the MSRCR is the application of the final gain/offset. Obviously, the two schemes use quite different techniques in going from the nonlinear logarithmic to the display domain. We conjecture that the application of the inverse log function in the retinex computation would undo some of the compression it achieves.

One of the most basic issues for the use of this retinex is the trade-off between the advantages versus the introduction of context dependency on local color and lightness values. Our experience is that the gains in visual quality, which can be quite substantial, outweigh the relatively small context dependency. The context dependencies are perhaps of most concern in remote sensing applications. The strongest context dependencies occur for the dark regions that are low because of low scene reflectances—for example, large water areas in remote sensing data adjacent to bright land areas. The large zones of water are greatly enhanced and subtle patterns in them emerge. The retinex clearly distorts radiometric fidelity in favor of visual fidelity. The gains in visual information, we hope, have been demonstrated adequately in our results. Even for specific remote sensing experiments where radiometric fidelity is required, the retinex may be a necessary auxiliary tool for the visualization of overall patterns in low signal zones. Visual information in darker zones that may not be detected with linear representations which preserve radiometry will "pop out" with a clarity limited only by the dynamic range of the sensor front-end and any intervening digitization scheme employed prior to the retinex. This may be especially useful

in visualizing patterns in remote sensing images covering land and water. Water has a much lower reflectance than land especially for false-color images including a near-infrared channel. The ability of the MSRCR to visualize features within both land and water zones simultaneously should be useful in coastal zone remote sensing.

The retinex computation can be applied ex post facto on 8-b color images and all of the results presented here represent this application. We have noticed only one problem with this—that the retinex can and will enhance artifacts introduced by lossy coding schemes, most notably lossy JPEG. Hence, the retinex is best applied prior to lossy image coding. One obvious advantage that the MSRCR provides for image compression is its ability to compress wider dynamic ranges to 8-bit or less per band color output, while preserving, and even enhancing, the details in the scene. The overall effect then is a significant reduction in the number of bits (especially in cases where the original color resolution is higher than 8-b/band) required to transmit the original without a substantial loss in spatial resolution or contrast quality.

The greatest power and advantage of the retinex is as a front-end computation, especially if the camera is also capable of wider than 8-b dynamic range. We have seen from scene photometry that 10–12-b dynamic ranges are required to encompass everyday scenes. Obviously, the retinex is most powerful as a front-end computation if it can be implemented within a sensor or between the sensor and coding/archival storage. We have not tested this retinex on wide dynamic range images, since we do not yet have access to an appropriate camera, therefore for wider dynamic range images some modifications in the processing may be anticipated. This may involve adding more scales, especially smaller ones, to provide a greater but still graceful dynamic range compression.

We have encountered many digital images in our testing that are underexposed. Apparently even with modern photographic autoexposure controls, exposure errors can and do occur. An additional benefit of the MSRCR is it capacity for exposure correction. Again, this is especially beneficial if it is performed as a front-end computation.

We do have the sense from our extensive testing thus far that the MSRCR approaches the high degree of dynamic range compression of human vision but may not quite achieve a truly comparable level of compression. Our impressions of the test scene cases is that direct observation is still more vivid in terms of color and detail than the processed images. This could be due to limitations in display/print media, or it could be that the processing scheme should be further designed to produce an even more emphatic enhancement. Further experimentation comparing test scenes to processed images and an accounting for display/print transfer characteristics will be necessary to resolve this remaining question and refine the method if necessary in the direction of greater enhancement of detail and color intensity. The transfer characteristics of print/display media deserve further investigation since most CRT's and print media have pronounced nonlinear properties. Most CRT's have an inverse "gamma" response [17] and the specific printer that we have used (Kodak XLT7720 thermal process) has a nonlinear response. For the printed results shown, we used a modest gamma correction ($\gamma = 1.2$). While this does not represent an accurate inverse that linearizes the printer transfer function, it does capture the the visual information with a reasonable good and consistent representation. Obviously no matter how general purpose the MSRCR is, highest quality results will still need to account for the specifics of print/display media especially since these are so often nonlinear.

VII. Conclusions

The MSR, comprised of three scales (small, intermediate, and large), was found to synthesize dynamic range compression, color consistency, and tonal rendition, and to produce results that compare favorably with human visual perception, except for scenes that contain violations of the gray-world assumption. Even when the gray-world violations were not dramatic, some desaturation of color was found to occur. A color restoration scheme was defined that produced good color rendition even for severe gray-world violations, but at the expense of a slight sacrifice in color consistency. In retrospect, the form of the color restoration is a virtual spectral analog to the spatial processing of the retinex. This may reflect some underlying principle at work in the neural computations of consciousness; perhaps, even that the visual representation of lightness, color, and detail is a highly compressed mesh of contextual relationships, a world of relativity and relatedness that is more often associated with higher levels of visual processing such as form analysis and pattern recognition.

While there is no firm theoretical or mathematical basis for proving the generality of this color restored MSR, we have tested it successfully on numerous diverse scenes and images, including some known to contain severe gray-world violations. No pathologies have yet been observed. Our tests were, however, confined to the conventional 8-b dynamic range images, and we expect that some refinements may be necessary when the wider dynamic range world of 10–12-b images is engaged.

Acknowledgment

The following World Wide Web sites provided the test images used for evaluating the performance of the MSRCR: Kodak Digital Image Offering at www.kodak.com/digitalImaging/samples/imageIntro. shtml; Monash University, Australia, DIVA Library at www.arts.monash.edu.au/visarts/diva/gardens.html; NASA Langley Research Center, LISAR Image Library at lisar.larc.nasa.gov/LISAR/BROWSE/ldef.html; and NASA Lyndon B. Johnson Space, Center Digital Image Collection at images.jsc.nasa.gov/html/shuttle.htm; Webmuseum, Paris at sunsite.unc.edu/louvre. The toy scene image is available from numerous sources.

References

[1] T. Cornsweet, *Visual Perception*. Orlando, FL: Academic, 1970.
[2] E. Land, "An alternative technique for the computation of the designator in the retinex theory of color vision," in *Proc. Nat. Acad. Sci.*, vol. 83, pp. 3078–3080, 1986.

[3] ——, "Recent advances in retinex theory and some implications for cortical computations," *Proc. Nat. Acad. Sci.*, vol. 80, pp. 5163–5169, 1983.
[4] ——, "Recent advances in retinex theory," *Vis. Res.*, vol. 26, pp. 7–21, 1986.
[5] A. C. Hurlbert, "The computation of color," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, Sept. 1989.
[6] ——, "Formal connections between lightness algorithms," *J. Opt. Soc. Amer. A*, vol. 3, pp. 1684–1693, 1986.
[7] A. C. Hurlbert and T. Poggio, "Synthesizing a color algorithm from examples," *Science*, vol. 239, pp. 482–485, 1988.
[8] A. Moore, J. Allman, and R. M. Goodman, "A real-time neural system for color constancy," *IEEE Trans. Neural Networks*, vol. 2, pp. 237–247, Mar. 1991.
[9] A. Moore, G. Fox, J. Allman, and R. M. Goodman, "A VLSI neural network for color constancy," in *Advances in Neural Information Processing 3*, D. S. Touretzky and R. Lippman, Eds. San Mateo, CA: Morgan Kaufmann, 1991, pp. 370–376.
[10] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Processing*, vol. 6, pp. 451–462, Mar. 1997.
[11] Z. Rahman, "Properties of a center/surround retinex, part 1: Signal processing design," *NASA Contractor Rep. 198194*, 1995.
[12] D. J. Jobson and G. A. Woodell, "Properties of a center/surround retinex, part 2: Surround design," NASA Tech. Memo. 110188, 1995.
[13] P. K. Kaiser and R. M. Boynton, *Human Color Vision*, 2nd ed. Washington, DC: Opt. Soc. Amer., 1996.
[14] P. Lennie and M. D. D'Zmura, "Mechanisms of color vision," *CRC Crit. Rev. Neurobiol.*, vol. 3, pp. 333–400, 1988.
[15] Z. Rahman, D. Jobson, and G. A. Woodell, "Multiscale retinex for color rendition and dynamic range compression," in *Proc. SPIE 2847, Applications of Digital Image Processing XIX*, A. G. Tescher, Ed., 1996.
[16] ——, "Multiscale retinex for color image enhancement," in *Proc. IEEE Int. Conf. Image Processing*, 1996.
[17] J. C. Russ, Ed., *The Image Processing Handbook*. Boca Raton, FL: CRC, 1995.
[18] A. Oppenheim, R. Schafer, and T. Stockham, Jr., "Non-linear flitering of multiplied and convolved signals," *Proc. IEEE*, vol. 56, pp. 1264–1291, Aug. 1968.
[19] O. D. Faugeras, "Digital color image processing within the framework of a human vision model," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 380–393, Aug. 1979.

Daniel J. Jobson (M'97) received the B.S. degree in physics from the University of Alabama, Tuscaloosa, in 1969.

He is a Senior Research Scientist at NASA Langley Research Center, Hampton, VA. His research has spanned topics including the design and calibration of the Viking/Mars lander camera, the colorimetric and spectrometric characterization of the two lander sites, the design and testing of multispectral sensors, and the analysis of coastal and ocean properties from remotely sensed data. For the past several years, his research interest has been in visual information processing with emphasis on machine vision analogs for natural vision, focal-plane processing technology, and nonlinear methods that mimic the dynamic-range compression/lightness constancy of human vision.

Zia-ur Rahman (M'87) received the B.A. degree in physics from Ripon College, Ripon, WI, in 1984 and the M.S. and Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in 1986 and 1989, respectively. His graduate research focused on using neural networks and image-processing techniques for motion detection and target tracking.

He is a Research Assistant Professor with the Department of Computer Science, College of William and Mary, Williamsburg, VA. Prior to that, he was a research scientist with the Science and Technology Corporation, and worked under contract to NASA Langley Research Center, Hampton, VA, on advanced concepts in information processing for high-resolution imaging and imaging spectrometry. Currently, he is involved in conducting research in multidimensional signal processing, with an emphasis in data compression and feature extraction methods. This work supports a NASA project for providing readily accessible, inexpensive remote-sensing data.

Dr. Rahman is a member of SPIE and INNS.

Glenn A. Woodell graduated from the NASA apprentice school in 1987 in materials processing.

He is a Research Technician at NASA Langley Research Center, Hampton, VA. His work has included semiconductor crystal growth experiments flown aboard the Space Shuttle in 1985 to study the effect of gravity-induced convection. His research has included demarcation, calculation, and visualization of crystal growth rates and real-time gamma ray visualization of the melt-solid interface and the solidification process. He has recently become involved in research on nonlinear image processing methods as analogs of human vision.