Journal of Electronic Imaging 13(1), 100–110 (January 2004).

Retinex processing for automatic image enhancement

Zia-ur Rahman
College of William & Mary
Department of Applied Science
Williamsburg, Virginia 23187
E-mail: zrahman@as.wm.edu

Daniel J. Jobson
Glenn A. Woodell
NASA Langley Research Center
Hampton, Virginia 23681

Abstract. *There has been a revivification of interest in the Retinex computation in the last six or seven years, especially in its use for image enhancement. In his last published concept (1986) for a Retinex computation, Land introduced a center/surround spatial form, which was inspired by the receptive field structures of neurophysiology. With this as our starting point, we develop the Retinex concept into a full scale automatic image enhancement algorithm—the multiscale Retinex with color restoration (MSRCR)—which combines color constancy with local contrast/lightness enhancement to transform digital images into renditions that approach the realism of direct scene observation. Recently, we have been exploring the fundamental scientific questions raised by this form of image processing. 1. Is the linear representation of digital images adequate in visual terms in capturing the wide scene dynamic range? 2. Can visual quality measures using the MSRCR be developed? 3. Is there a canonical, i.e., statistically ideal, visual image? The answers to these questions can serve as the basis for automating visual assessment schemes, which, in turn, are a primitive first step in bringing visual intelligence to computers.*
[DOI: 10.1117/1.1636183]

1 Introduction

A common problem with color imagery—digital or analog—is that of successful capture of the dynamic range and colors seen through the viewfinder onto the acquired image. More often than not, this image is a poor rendition of the actual observed scene. The idea of the Retinex was conceived by Land[1] as a model of the lightness and color perception of human vision. Through the years, Land evolved the concept from a random walk computation,[2–5] to its last form as a center/surround spatially opponent operation[3] related to the neurophysiological functions of individual neurons in the primate retina, lateral geniculate nucleus, and cerebral cortex. Hurlbert[6,7] looked at the problem of color constancy and showed that there is no mathematical solution to the problem of removing lighting variations. Moore *et al.*[8,9] implemented a version of the Retinex in analog VLSI for real-time dynamic range compression, but encountered scene context-dependent limitations and hence failed to achieve a generalized implementation.

In our research,[10–19] we do not use the Retinex as a model for human vision color constancy. Rather, we use it as a platform for digital image enhancement by synthesizing local contrast improvement, color constancy, and lightness/color rendition. The intent is to transform the visual characteristics of the recorded digital image so that the rendition of the transformed image approaches that of the direct observation of scenes. Special emphasis is placed on increasing the local contrast in the dark zones of images of wide dynamic range scenes—scenes that contain brightly lit and dark regions—so that it matches our perception of those dark zones. Basic study of the properties of the center/surround Retinex led us in the direction of using a Gaussian surround used by Hurlbert[6,7,20] as opposed to the $1/r^2$ surround originally proposed by Land,[2,3] or the exponential surround used by Moore[8,9] for analog VLSI resistive networks. Since the width of the surround affects the rendition of the processed image, multiple scale surrounds were found to be necessary to provide a visually acceptable balance between dynamic range compression and graceful tonal rendition. This is discussed in more detail in Sec. 2.

The final visual defect in performance was the color ''graying'' due to global and regional violations of the gray-world assumption intrinsic to Retinex theory. A color restoration was essential for correcting this and took the form of a log *spectral* operation similar to the log *spatial* operation of the center/surround. This produces an interaction between spatial and spectral processing and results in a trade-off between strength of color constancy and color rendition. The color restoration yields a modest relaxation in color constancy, perhaps comparable to human color vision's perceptual performance (see Sec. 3). Barnard and Funt[21,22] developed a neural network to provide color constancy and rendition. They were ''uncomfortable with [our]

Paper RTX-10 received Nov. 2, 2002; revised manuscript received Aug. 6, 2003; accepted for publication Aug. 27, 2003.
1017-9909/2004/$15.00

procedure as the effect [was] hard to characterize.''[21] However, their network requires a calibration of the algorithm against known illuminants, a procedure we do not require.

In the scientific and signal processing community, the linear representation of a scene's radiometric characteristics is a widely accepted standard. Most image reconstruction and restoration algorithms rely on a linear image representation and use linear end-to-end image metrics to reproduce an image that is as radiometrically close to the original scene as is technically possible.[23–25] In the computer graphics and imaging world, most display devices are linearized to achieve correct reproduction of intensity.[26,27] In the course of our experiments, we have noted that this commonly accepted linear representation often fails to produce a realistic rendition of the observed scene. The images either have saturated bright regions to compensate for the dark regions, or clipped dark regions to compensate for the bright regions. Even when the dynamic range of the scene is narrow enough to be completely captured by the dynamic range of the imaging device, the resultant image is a poor representation of the observed scene, being too dark and too low in overall contrast. A nonlinear representation, such as the multiscale Retinex with color restoration (MSRCR), provides the necessary dynamic range compression that encompasses the full dynamic range of the scene that is needed to produce images that approach the direct perception of natural scenes. Section 5 lays out these ideas in more detail. A comparison of the Retinex with the traditionally used image enhancement techniques for enhancing images of wide dynamic range scenes propels one toward the acceptance of the nonlinear representation as the appropriate one. Section 4 covers this issue in more detail.

2 Multiscale Retinex

The basic form of the multiscale retinex (MSR) is given by

$$R_i(x_1,x_2)=\sum_{k=1}^{K} W_k\{\log I_i(x_1,x_2) -\log[F_k(x_1,x_2)*I_i(x_1,x_2)]\} \quad i=1,\cdots,N, \tag{1}$$

where index i references the i'th spectral band, (x_1,x_2) is the pixel location in Cartesian coordinates, and $*$ is the convolution operator. N is the number of spectral bands—$N=1$ for grayscale images, and $N=3$, $i\in R, G, B$ for typical color images. I is the input image and R is the output of the MSR process. F_k is the k'th (Gaussian) surround function, W_k is the weight associated with F_k, and K is the number of surround functions, or scales. The F_k are given as:

$$F_k(x_1,x_2)=\kappa \exp[-(x_1^2+x_2^2)/\sigma_k^2],$$

where σ_k are the standard deviations of the Gaussian surrounds. The magnitude of σ_k controls the extent of the surround: smaller values of σ_k result in narrower surrounds. The MSR output is normalized by $\kappa=1/[\Sigma_{x_1}\Sigma_{x_2}F(x_1,x_2)]$. The MSR reduces to the single scale retinex (SSR) when $K=1$, with the additional constraint that $W_1=1$.

As mentioned in Sec. 1, we found that multiple surrounds were necessary to achieve a graceful balance between dynamic range compression and tonal rendition. The number of scales used for the MSR is, of course, application dependent. We have found empirically, however, that a combination of three scales representing narrow, medium, and wide surrounds is sufficient to provide both dynamic range compression and tonal rendition. Figure 1 shows the input image, the output of the MSR, and the outputs when the different surround functions are applied to the original image. These are obtained by setting $k=1$ and $W_k=1.0$ in Eq. (1). As is evident from Fig. 1, single scale Retinexes cannot attain the goal that we are striving for: visual realism. The failure of the narrow and medium surround cases is self-evident; the wide-surround case, however, deserves additional discussion because it produces a ''nice'' output image. Where it fails is in encompassing the wide dynamic range of the scene, with the result that features that were visible to the observer are obscured in the Retinexed image. The MSR image contains features from all three scales simultaneously, providing dynamic range and tonal rendition. However, for the example shown in Fig. 1, it is also obvious that the MSR does not provide very good tonal rendition. The image was chosen specifically because it has large monochrome areas, and the Retinex computation forces them toward middle gray, resulting in color desaturation. A method to deal with this problem is discussed in Sec. 3. It should be noted, however, that the MSR computation provides a completely color constant result, similar to that produced by the SSR computation.[13] Figure 2 shows an example of the color constancy that the MSR provides.

3 MSR with Color Restoration

The general effect of MSR processing on images with regional or global gray-world violations is a graying out of the image either in specific regions or globally. This desaturation of color can, in some cases, be severe. We can, therefore, consider the computation that is needed to mitigate this desaturation as a color restoration (CR). The CR process should produce good color renditions for any degree of graying. In addition, the CR should preserve a reasonable degree of color constancy, since that is one of the basic motivations for the MSR. However, 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 color dependency is tolerable to the extent that the visual defect is not visually too strong.

We consider the foundations of colorimetry,[28] even though it is often considered to be in direct opposition to color constancy models and is 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.[29] The reason for this choice is simply that it serves as a foundation for creating a relative color space, and in doing so uses ratios that are less dependent on illuminant spectral distributions than raw spectrophotometry. We compute a

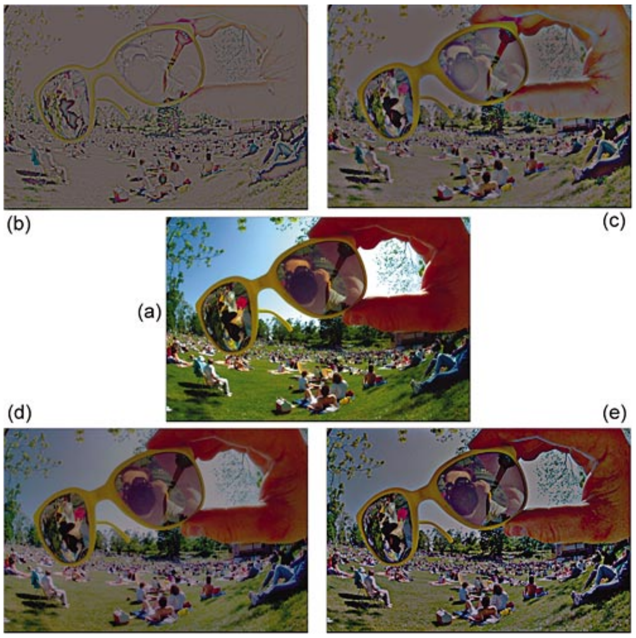


Fig. 1 (a) The original input, (b) narrow surround $\sigma=5$, (c) medium surround $\sigma=20$, (d) wide surround $\sigma=240$, and (e) MSR output with $W_k=1/3$, $k=1,2,3$. The narrow-surround acts as a high-pass filter, capturing the fine detail in the image, but at a severe loss of tonal information. The wide-surround captures the fine tonal information but at the cost of fine detail. The medium surround captures both dynamic range and tonal information. The MSR is the average of the three renditions.

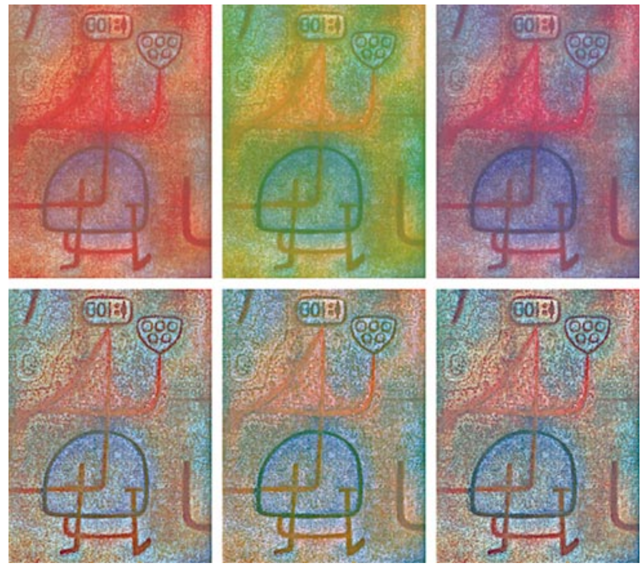


Fig. 2 The image shows a painting by Paul Klee. The effect of changing the illuminant was simulated by red, blue, and green shifting the original image (top row). The bottom row shows the MSR output for each case. Note that the MSR outputs are *almost* color constant, much like the human visual system.

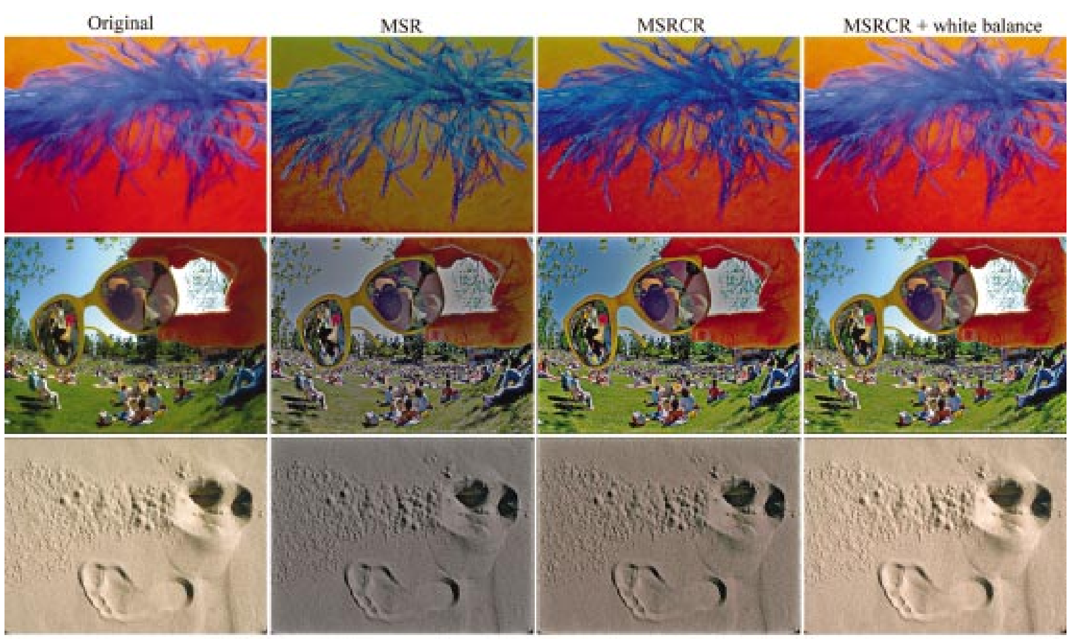


Fig. 3 Scenes that violate the gray-world assumption, and the MSR, MSRCR, and MSRCR with white-balance[31] outputs. Note that while all the processed outputs are sharper than the originals, the MSR output is considerably more desaturated than the MSRCR output, which still shows some color loss. The MSRCR with white balance corrects the latter problem.

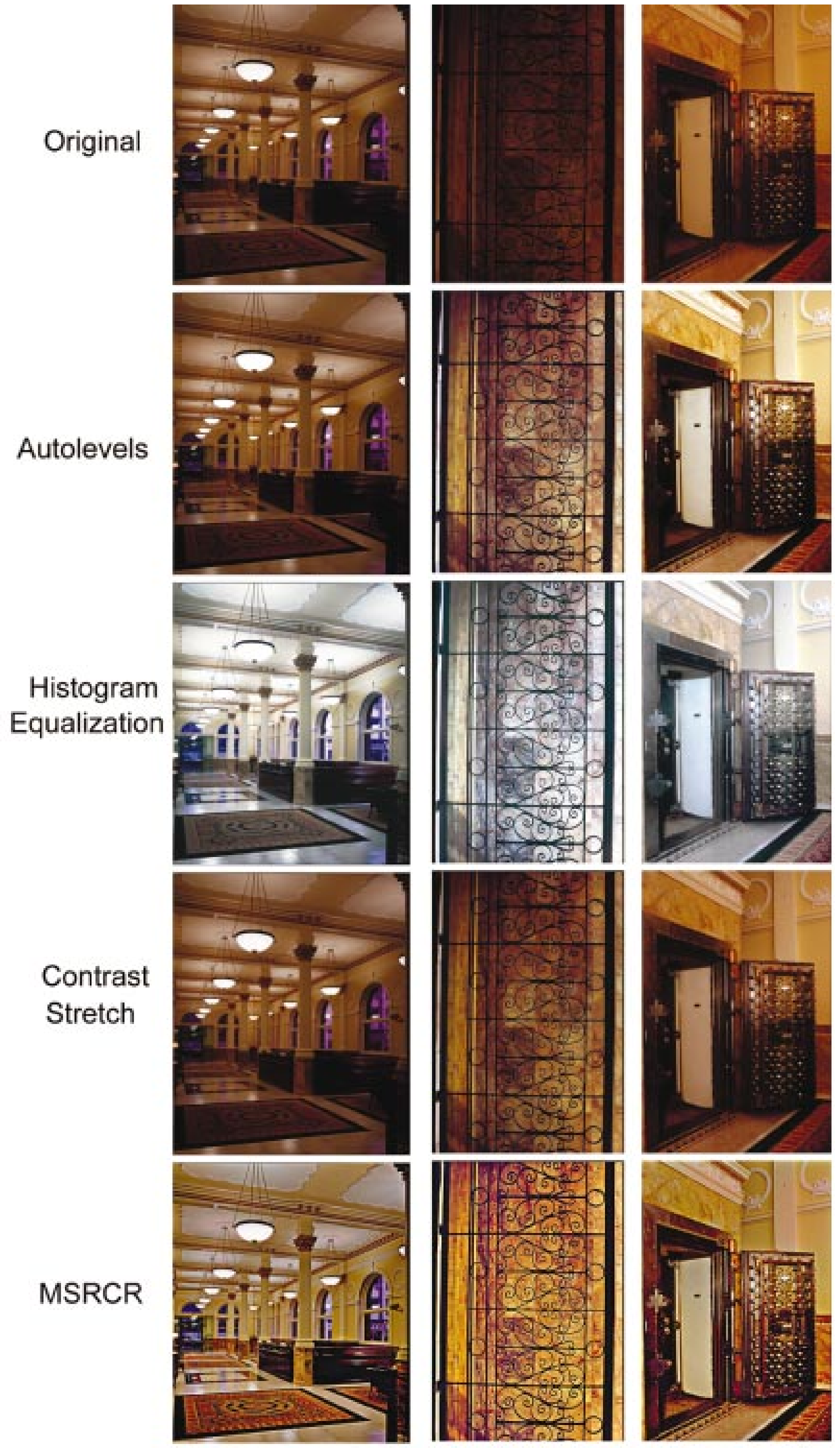


Fig. 4 Comparison of the MSRCR with commonly used automatic image enhancement techniques. It is evident that if the original scene has a wide dynamic range, then some of these techniques do not affect the original image at all. Aside from the MSRCR, histogram equalization provides the strongest dynamic range compression, but suffers considerably from color distortion.

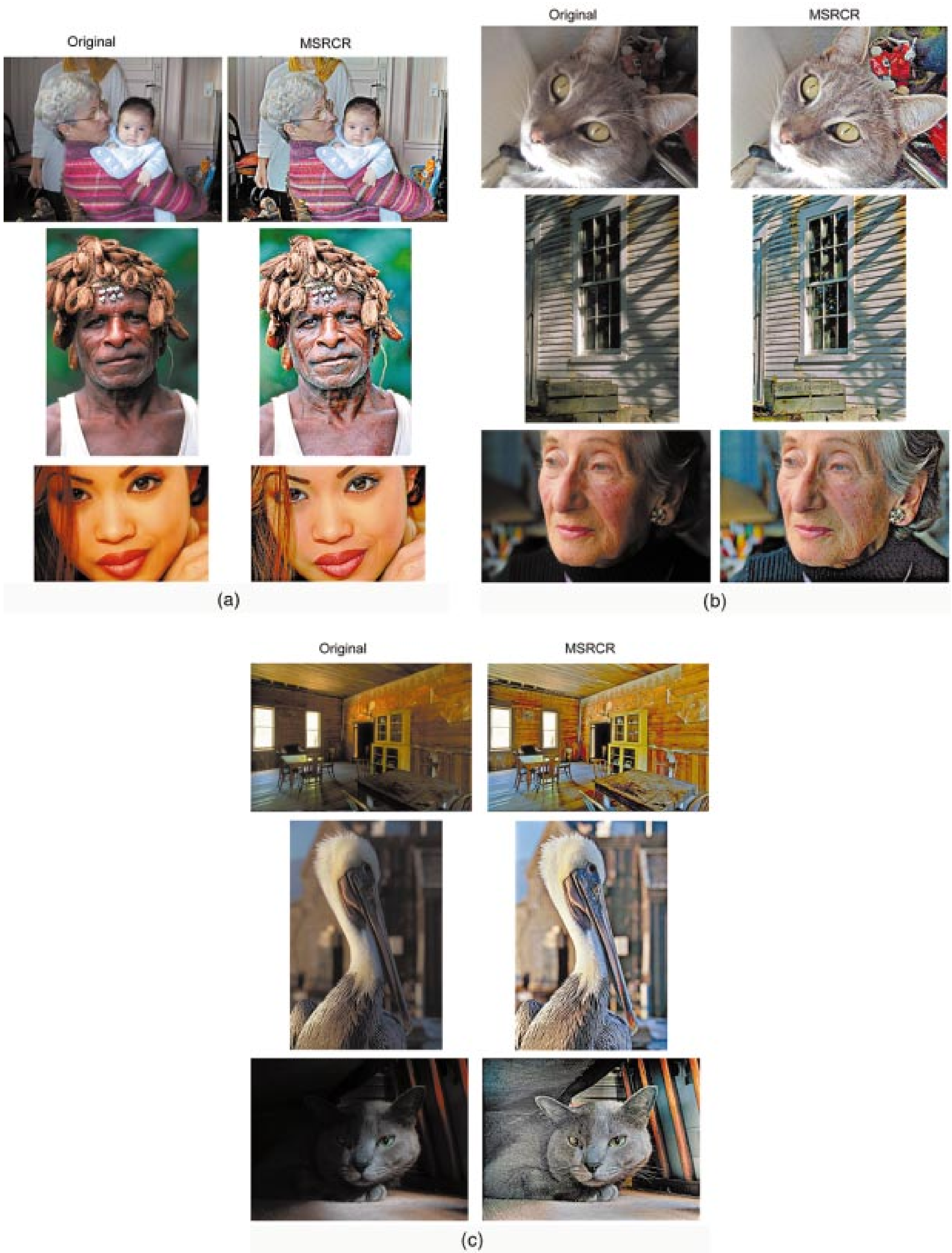


Fig. 5 Retinex examples to illustrate that the strength of the enhancement matches the degree of visual deficit in the original image. (a) Subtle enhancements: the original images are of a type that is generally acceptable to the viewer. However, the MSRCR processed results are slightly sharper than the originals, and tend to be more realistic. (b) Moderate enhancements: the original images in this case are moderately underexposed or have slight shadows. The MSRCR process removes the effects of the shadows, making the processed image closer to the observed image. (c) Strong enhancements: the original image in this case have strong shadows or regions of brightness that lead to underexposure. By enhancing the details in the darker regions, the MSRCR processed results reduce the impact of underexposure, and also (almost) completely eliminate the effects of dark shadows.

Fig. 6 Visual inadequacy of the linear representation. All of the original images were acquired under bright, sunlit conditions using a Nikon D1 set in linear mode. However, the presence of bright reflectance sources in each image—the lighthouse in the top and bottom rows, and the rubber sides on the sneaker—causes the camera to compensate in a manner that leads to poor image renditions.

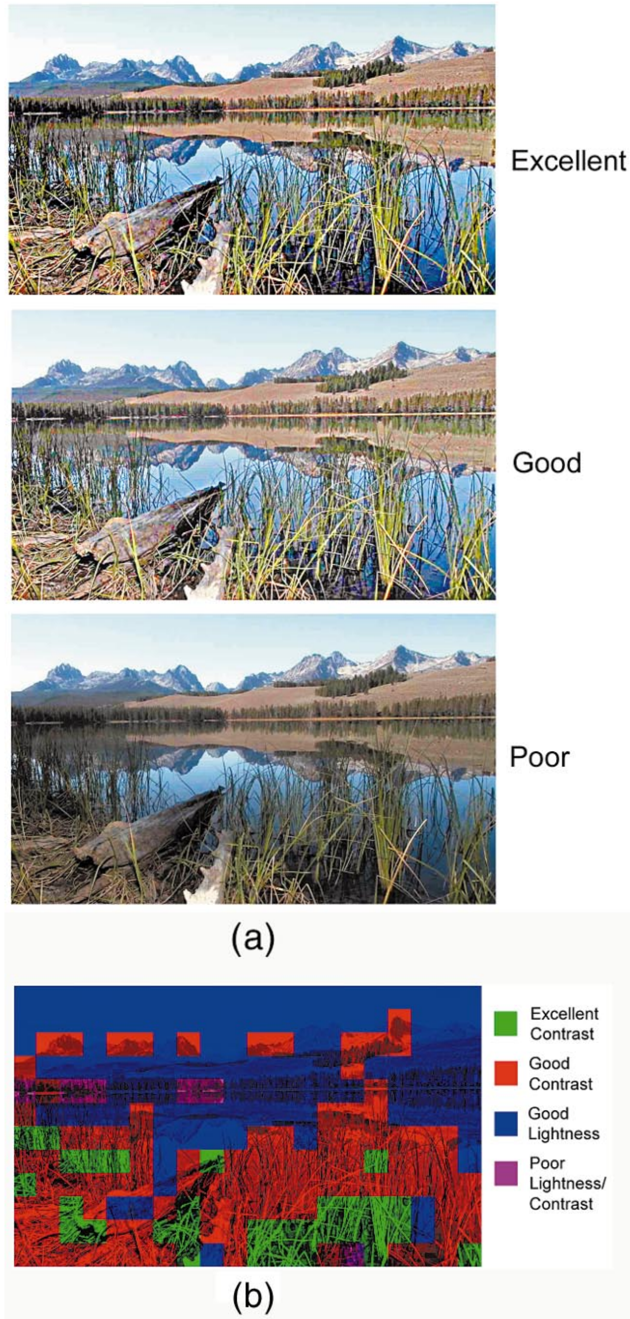


Fig. 7 (a) Visual measures for automating visual assessment: images are assigned to one of three global classes, excellent, good, and poor. The classes are bases on global and regional brightness and contrast measures. (b) Visual map showing regional classes: the combination of the regional classes is used to derive the global classification.

color restoration factor α based on the following transform:

$$\alpha_i(x_1,x_2)=f\left[I_i(x_1,x_2)\bigg/\sum_{n=1}^{N} I_n(x_1,x_2)\right], \quad (2)$$

where $\alpha_i(x_1,x_2)$ is the color restoration coefficient in the i'th spectral band, N is the number of spectral bands, I_i is the i'th spectral band in the input image, and $f(\ )$ is some color space mapping function. Combining the color restoration term in Eq. (2) with the MSR given in Eq. (1), gives the MSRCR[30]

$$R_i(x_1,x_2)=\alpha_i(x_1,x_2)\sum_{k=1}^{K} W_k\{\log I_i(x_1,x_2) - \log[F_k(x_1,x_2)*I_i(x_1,x_2)]\}. \quad (3)$$

MSRCR has been implemented in a commercial software package, PhotoFlair, available from TruView Imaging Company. The results of applying this transformation to images with significant monochrome areas are shown in Fig. 3. It is noticeable that the color restoration term does not completely restore the bright colors that are in the original image (see middle row in Fig. 3). This effect can be ameliorated by using a white balance process that is the subject of a current patent application.[31] In essence, the white balance process ensures that bright areas in the original image do not get desaturated to middle gray. Barnard and Funt[21] hypothesized that MSR processing suffers from ‘‘... color bleeding at certain color edges due to the local contrast enhancement.’’ Though we see this effect in computer-rendered images with very sharp edge transitions, we have not observed it to be a major source of concern in the many thousands of images we have processed. Barnard and Funt also point out in Ref. 21 that this ‘‘... is normally not noticeable in images of typical natural scenes.’’

While we have called this additional computation a color restoration, depending on the form of $f(\ )$, this can be considered as a spectral analog to the spatial Retinex computation. If $f(\ )=\log(\ )$, then Eq. (2) becomes

$$\alpha_i(x_1,x_2)=\log\left[NI_i(x_1,x_2)\bigg/\sum_{n=1}^{N} I_n(x_1,x_2)\right],$$

and the internal form of the Retinex computation and the color restoration computation is essentially the same. This mathematical and philosophical symmetry is intriguing, since it suggests some underlying unifying principle between the two computations: both computations are contextual, highly relative, and nonlinear. We speculate that the visual representation of wide ranging scenes is a compressed mesh of contextual relationships, even at the stage of lightness and color representation. This sort of information representation would certainly be expected at more abstract levels of visual processing, such as shape information composed of edges, links, and the like, but is surprising for a representation so closely related to the raw image.

4 Comparisons

Before delving into some of the philosophical and developing scientific aspects of the MSRCR, we feel that it would be of interest to the readers to see a comparison of the MSRCR with several traditionally used image enhancement algorithms. Since the MSRCR is an automatic processing algorithm, we have confined our comparison to other automatic image enhancement techniques.

Figure 4 shows a comparison of the MSRCR with three other automatic image enhancement methods: autolevels, histogram equalization, and automatic image contrast stretch. All three methods belong to the class of histogram modification techniques. In addition, they are global because the relationship between the input and the processed image can be described with a single lookup table (LUT).

Autolevels. Autolevels is a commonly used image enhancement function, which derives its popularity from the facts that it is fast, automatic, and provides (fairly) good processed results for input images that have low dynamic range in at least one color channel. (Popular image enhancement software such as Adobe Photoshop, JASC Paintshop Pro, GNU GIMP, TruView PhotoFlair, and many others implement some version or another of autolevels). It is similar in implementation to contrast stretch, except a predefined parameter is used to clip the tails of the histogram as a percentage of the total number of pixels in the image. The new endpoints of the histogram are then mapped to the full representation dynamic range by applying a gain.

Histogram equalization. Histogram equalization is a well known technique that is used to maximize the entropy of an image.[32–34] Entropy of an image is maximized when all gray levels occur with equal probability. When an image has regions that are very dark (bright), it tends to have a histogram that is skewed toward the lower (higher) values, and often with a large peak at the lower (upper) end. Histogram equalization uses the cumulative histogram distribution to map the original histogram to a histogram that has a uniform distribution, i.e., maximum entropy. This results in moving pixels from the lowest values to higher values, making the darker regions brighter.

Contrast stretch. The only difference between automatic contrast stretch and autolevels is that contrast stretching techniques typically do not clip the histogram. Rather, they use the minimum and maximum values of the histogram to compute a gain, which they then use to stretch the histogram to the full dynamic range. The obvious problem with this method is that if the image contains even one pixel at the highest and lowest levels, then this method provides no enhancement. This is the reason why clipping was introduced in the autolevels approach, making the enhancements more robust and useful overall.

Autolevels and contrast-stretch techniques are quite good if the images have a narrow dynamic range, i.e., do not contain significant numbers of bright and dark pixels simultaneously. If this condition is violated, then the autolevels and contrast stretch methods leave the image virtually unchanged, as can be seen in the examples in the second and third columns in Fig. 4. Histogram equalization techniques tend to work quite well in compressing the wide dynamic range in the image and bringing out details in the darker regions. However, they suffer from quite severe color distortion, leading to visually unacceptable images. In a previous work, we also compared the performance of the MSRCR to the so-called homomorphic filtering method.[15] However, the results showed that the homomorphic filter was not capable of compressing the same dynamic range as the Retinex, and suffered from weak contrast.

The MSRCR computation differs from the other automatic image enhancement techniques in two major ways. First, the relationship between an image and its MSRCR enhanced output *cannot* be described by a single LUT. MSRCR is a nonlinear contextual operation. This means that the output representation of the same input value will be different, depending on what surrounds the original pixel. Second, the image processing techniques described earlier, though automatic, make adjustments based on the image content. In other words, though the histogram equalization, for example, uses the same *procedure* each time, the LUT is very different from image to image. The MSRCR, however, does not make adjustments on a per scene basis: the scales, weights, and gains and offsets are canonical, i.e., the same set of values is used for every image. In some sense this is also true of the autolevels process, where the predefined parameter is a percentage of total pixels, but which has different values for each image.

5 Direct Viewing of Scenes and the MSRCR

Our work with the Retinex[13,14] has led us away from the world of color and into the world of contrast and lightness perception of visually complex natural scenes.[35,36] While the MSRCR synthesizes color constancy, dynamic range compression, and the enhancement of contrast and lightness, the emphasis is on the latter: the MSRCR brings the perception of dark zones in recorded images up in local lightness and contrast to the degree needed to mimic direct scene viewing. In the world of natural images, only those images with very modest dynamic ranges do not need enhancement comparable to what the MSRCR provides, and for these the exposure must be very accurate to achieve a good visual representation. Wide ranging reflectance values in a scene, and certainly, strong lighting variations, demand a strong enhancement to produce a representation that is anything like the visual realism of direct observation. The dynamic range compression of the Retinex computation is the basis for the contrast and lightness enhancement, and its generic character forms the basis for using it as an automatic enhancement. Some examples of MSRCR enhancements will serve to convey the degree to which images need to be improved, and to provide a demonstration that the MSRCR does, in fact, perform this task with considerable agility (Fig. 5) and without human intervention. These examples highlight a major facet of MSRCR performance: the degree to which the image is (automatically) enhanced is commensurate with the degree of visual deficit in the original acquired image.

During the course of developing and experimenting with the MSRCR, we repeatedly observed certain (puzzling) features of the imaging process that led us to reexamine some of the most basic ideas about the imaging process. If the goal of imaging is to produce a good visual representation—image—of the observed scene, then the idea that *imaging is a replication process that produces minimal distortion of measured signals or radiometry* is clearly untenable. Instead, the idea that *imaging is a process of (profound) transformation that intrinsically involves nonlinear spatial processing* does produce visually acceptable images. Hence, the traditional wisdom of linear radiometric representation with minimum distortion is clearly inadequate in representing the full dynamic range and, hence, the direct visual perception of most natural scenes. Figure 6 shows a set of examples where the scene has been shown in its linear and nonlinear (MSRCR) representations. All of these images were taken on a very bright, clear January day. There are negligible lighting variations, so virtually all signal variations are due to reflectance and topographical changes. Even for this restricted dynamic range case, the linear representation does not convey the direct perception of these scenes. In general, the linear representation is not a good visual representation. This observation is consistent with the conclusion of a study of the data handling and processing for color negative film scanning.[37] Tuijn describes the correction for all transfer functions, so that the image data is linear, and then explains that this is often visually inadequate—weak in contrast and color. To explore this further, we displayed known linear data taken with a Nikon D1 camera in linear mode on a linearized color computer monitor (gamma correction of 1.6). For a wide array of images, the displayed image is too dark (Fig. 6), and the MSRCR enhancement (also shown for comparison) is required to produce a good visual representation. The linear representation can approach a good visual rendering for a very restricted class of scenes—those with diffuse illumination and restricted ranges of reflectance, or those where white surfaces do not contain significant detail so can be saturated. Even for this cooperative class, a substantial degree of (nonlinear) processing is required to achieve a good visual representation.

While image data can be quite arbitrary in a statistical sense (the histograms of images vary widely) we observe that the MSRCR processed data are not as arbitrary. As noted in a previous work,[14] histograms of MSRCR processed images tend toward a characteristic Gaussian-like shape. More recently, we have studied regional means (visual lightness) and standard deviations (visual contrast), and found that they tend to converge on consistent global aggregates.[35] This implies that a good visual representation can be associated with well-defined statistical measures for visual quality. In scientific terms, this implies the existence of a canonical visual image as a statistical practical ideal. Such a defined ideal can then serve as the basis for the automatic assessment of visual quality. By following the general idea that the MSRCR brings regional means and standard deviations up to higher values, and that these approach an ideal goal, we have constructed a set of *visual measures*.

The general idea behind visual measures is that good visual representations seem to have some combination of high regional visual lightness and contrast.[35] To compute the regional parameters, we divide the image into nonoverlapping 50×50 pixel blocks. For each block, a mean I_b and a standard deviation σ_b are computed. This regional scale is sufficiently granular to capture the visual sense of regional brightness and contrast. Both the global contrast and lightness can then be measured in terms of the regional parameters. The overall lightness is measured by the image mean $\mu = \overline{I_b}$, which is also the ensemble measure for regional lightness. The overall contrast $\overline{\sigma_b}$ is measured by taking the mean of regional standard deviations σ_b, and it provides a gross measure of the regional contrast variations. A classification of excellent, good, or poor is then based on how many of these regional blocks exceed a given contrast and brightness threshold. The global standard deviation of the image did not relate, except very weakly, to the overall visual sense of contrast.[38]

The measures were set empirically on a small diverse test image set, and then were applied to a broad array of images of all sorts. Figure 7(a) shows a sample of the automatic visual quality assessment by classification into one of three classes: poor, good, and excellent. The classification scheme is based on the map shown in Fig. 7(b). While more study and development is underway, these early results do hold the promise that the idea of a canonical visual image with well-defined statistical properties can lead to a new statistical measure of visual quality.

The Retinex experience provides new avenues for the study for statistical image processing; it also suggests deterministic pathways. The generic character of the Retinex computation suggests that some new quantitative definition of visual information may be possible. A deterministic definition would contrast with previous statistical ones based on information theory.[23,39] Specifically, the MSRCR is (approximately) performing a log operation on the ratio of each pixel in each spectral band to both spatial and spectral averages. The suppression of spatial and spectral lighting variations is achieved at the expense of accepting a significant degree of context dependency. Simply put, the MSRCR appears to mimic human perception in producing color and lightness that are influenced by the visual setting in which they occur. The exchange of spatial and spectral lighting dependencies for spatio-spectral context effects appears to be a very basic element of human vision and the MSRCR computation. While we do not have a clear definition of information in a semantic sense, or visual information as some subset of all information, the idea that information is derived from contextual relationships is appealing. The additional factor of a log function suggests a compactness that may be leading in the direction of symbolic representation, the symbol being the ultimate conciseness and carrier of meaning.

The establishment of context relationships is central to at least the senses of vision and hearing. Music seems to be based on pitch relationships, with certain ratios producing consonance or dissonance in varying degrees. Speech recognition must contend with the difficulties of speaker variations, the interdependencies of phonemes, and all manner of extraneous variations in loudness, temporal rates, degrees of clarity, and the like. For vision, the awesome task of transforming the signals of vision into the sense of vision must succeed in extracting information in the presence of all manner of extraneous variations, as well as find some very concise ultimately symbolic representation.[40] Context must be a critical element of vision information as it is in speech and music, where isolated acoustical events become perceived as a fluid temporal mesh of meaningful words or melody, harmony, and rhythm. Signals are not meaningful in isolation. For vision, contextual relationships such as edge connectedness, textural uniformity, and color reflectance differences seem fundamental to deriving visual information. Perhaps the Retinex transformation moves one step in this direction by reducing extraneous variations, increasing spatial and spectral differences, and providing a foundation for a structure of relatedness, which with subsequent processing can become symbolic.

6 Conclusions

The visual image remains an enigma full of surprises, some of which we have encountered in our experiences with Retinex image processing. Though we do not understand the intricacies that allow the human vision system to encompass very wide dynamic ranges and provide color constancy, we have developed an approach that seems to mimic these behaviors. Because of this, our thinking about the imaging process has changed in basic ways

1. Imaging should be considered as a process of transformation rather than replication with minimal distortion. This is evident in comparing direct observation of a scene with its captured image representation, and by comparing the type and degree of enhancement that is needed to make the captured image *look* like the observed scene.
2. The statistical convergence of MSRCR enhancements to a histogram that closely matches Gaussian distributions leads us to postulate the existence of a canonical visual image with consistent statistical aggregate characteristics. Further, these can be used to construct entirely new visual measures, which can be the basis for the automatic assessment of visual quality of arbitrary images by the computer.
3. A new deterministic definition of visual information emerges from the computational form of the Retinex, namely, that visual information is in some sense the log of spatial and spectral context relationships within the image.

A computation like the MSRCR appears to have two very useful properties simultaneously: a diminishment in the dependence of the appearance of the image on extraneous variables, such as spatial and spectral lighting, and the construction of compact context relationships. The former is inherently useful because it can lead to better image classifications, and the latter because it shows very clearly that the appearance of a color is dependent not only on the spectral characteristics of a pixel, but also its surround. Together, these properties may be able to provide a basis for bringing more advanced levels of visual intelligence into computing.

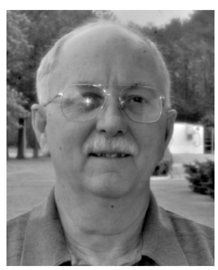
Acknowledgments

Z. Rahman's work was supported with the NASA cooperative agreement NCC-1-01030 under the aegis of the Synthetic Vision Sensors element of the Aviation Safety Program.

References

1. E. Land, ''The retinex,'' *Am. Sci.* **52**, 247–264 (1964).
2. E. Land, ''Recent advances in retinex theory and some implications for cortical computations,'' *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5163–5169 (1983).
3. E. Land, ''Recent advances in retinex theory,'' *Vision Res.* **26**(1), 7–21 (1986).
4. E. Land and J. J. McCann, ''Lightness and retinex theory,'' *Vision Res.* **61**(1), 1–11 (1971).
5. E. Land and J. J. McCann, ''Method and system for reproduction based on significant visual boundaries of original subject,'' U.S. Patent No. 3,553,360 (1971).
6. A. C. Hurlbert, ''Formal connections between lightness algorithms,'' *J. Opt. Soc. Am. A* **3**, 1684–1693 (1986).
7. A. C. Hurlbert, ''The computation of color,'' PhD Thesis, Massachusetts Institute of Technology (Sep. 1989).
8. A. Moore, J. Allman, and R. M. Goodman, ''A real-time neural system for color constancy,'' *IEEE Trans. Neural Netw.* **2**, 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., pp. 370–376, Morgan Kaufmann, San Mateo, CA (1991).
10. Z. Rahman, D. Jobson, and G. A. Woodell, ''Multiscale retinex for color image enhancement,'' *Proc. IEEE Intl. Conf. Image Process.* (1996).
11. Z. Rahman, D. Jobson, and G. A. Woodell, ''Multiscale retinex for color rendition and dynamic range compression,'' *Proc. SPIE* **2847**, 183–191 (1996).
12. D. Jobson, Z. Rahman, and G. A. Woodell, ''Retinex image processing: Improved fidelity for direct visual observation,'' *Proc. IS&T 4th Color Imaging Conf. Color Sci. Syst. Appl.*, pp. 124–126 (1996).
13. D. J. Jobson, Z. Rahman, and G. A. Woodell, ''Properties and performance of a center/surround retinex,'' *IEEE Trans. Image Process.* **6**, 451–462 (Mar. 1997).
14. D. J. Jobson, Z. Rahman, and G. A. Woodell, ''A multi-scale Retinex for bridging the gap between color images and the human observation of scenes,'' *IEEE Trans. Image Process. Special Issue Color Process.* **6**, 965–976 (Jul. 1997).
15. Z. Rahman, G. A. Woodell, and D. Jobson, ''A comparison of the multiscale retinex with other image enhancement techniques,'' *Proc. IS&T 50th Anniversary Conf.*, pp. 426–431 (1997).
16. Z. Rahman, D. Jobson, and G. A. Woodell, ''Resliency of the multiscale retinex image enhancement algorithm,'' *Proc. IS&T 6th Color Imaging Conf. Color Sci. Syst. Appl.*, pp. 129–134 (1998).
17. D. Jobson, Z. Rahman, and G. A. Woodell, ''Spatial aspect of color and scientific implications of retinex image processing,'' *Proc. SPIE* **4388**, 117–128 (2001).
18. D. Jobson, Z. Rahman, and G. A. Woodell, ''Feature visibility limit in the enhancement of turbid images,'' *Proc. SPIE* **5108**, 24–30 (2003).
19. Z. Rahman, D. Jobson, and G. A. Woodell, ''A method for digital image enhancement,'' U.S. Patent No. 5,991,456 (1996).
20. A. C. Hurlbert and T. Poggio, ''Synthesizing a color algorithm from examples,'' *Science* **239**, 482–485 (1988).
21. K. Barnard and B. Funt, ''Investigations into multi-scale retinex,'' in *Colour Imaging: Vision and Technology*, pp. 9–17, John Wiley and Sons, New York (1999).
22. K. Barnard and B. Funt, ''Analysis and improvement of multi-scale retinex,'' *Proc. IS&T 5th Color Imaging Conf. Color Sci. Syst. Appl.*, pp. 221–226 (1997).
23. F. O. Huck, C. L. Fales, and Z. Rahman, *Visual Communication: An Information Theory Approach*, Kluwer Academic Publishers, Norwell, MA (1997).
24. J. A. McCormick, R. Alter-Gartenberg, and F. O. Huck, ''Image gathering and restoration: Information and visual quality,'' *J. Opt. Soc. Am. A* **6**(7), 987–1005 (1989).
25. C. L. Fales and F. O. Huck, ''An information theory of image gathering,'' *Inf. Sci. (N.Y.)* **57,58**, 245–285 (1991).
26. C. Poynton, *A Technical Introduction to Digital Video*, John Wiley and Sons, New York (1996).
27. C. Poynton, ''The rehabilitation of gamma,'' *Proc. SPIE* **3299**, 232–250 (1998).
28. W. D. Wright, *The Measurement of Colour*, 2nd Ed., Hilger and Watts, London (1958).
29. P. Lennie and M. D. D'Zmura, ''Mechanisms of color vision,'' *CRC Critical Rev. Neurobiol.* **3**, 333–400 (1988).
30. D. J. Jobson, Z. Rahman, and G. A. Woodell, ''A multi-scale Retinex for bridging the gap between color images and the human observation of scenes,'' *IEEE Trans. Image Process. Special Issue Color Process.* **6**, 965–976 (Jul. 1997).
31. G. A. Woodell, D. Jobson, and Z. Rahman, ''Method of improving a digital image having white zones,'' U.S. Patent Application No. 20030026494 (2001).
32. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA (1993).
33. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, 2nd ed., Academic Press, Orlando, FL (1982).
34. J. C. Russ, *The Image Processing Handbook*, 3rd ed., CRC Press, Boca Raton, FL (1999).
35. D. Jobson, Z. Rahman, and G. A. Woodell, ''The statistics of visual representation,'' *Proc. SPIE* **4736**, 25–35 (2002).
36. Z. Rahman, D. Jobson, and G. A. Woodell, ''Retinex processing for automatic image enhancement,'' *Proc. SPIE* **4662**, 390–401 (2002).
37. C. Tuijn, ''Scanning color negatives,'' *Proc. IS&T 4th Color Imaging Conf. Color Sci. Syst. Appl.*, pp. 33–38 (1996).
38. D. Jobson, Z. Rahman, and G. A. Woodell, ''The statistics of visual representation,'' *Proc. SPIE* **4736**, 25–35 (2002).
39. F. O. Huck, C. L. Fales, and Z. Rahman, ''Information theory of visual communication,'' *Philosophical Trans. Royal Soc. London A* **354**, 2193–2248 (Oct. 1996).
40. S. Zeki, *Inner Vision: An Exploration of Art and the Brain*, Oxford University Press, New York (1999).

Zia-ur Rahman is a research associate professor of applied science and adjunct assistant professor of computer science at the College of William & Mary, where he is currently an investigator on a NASA grant working on topics concerning nonlinear image fusion and enhancement. Before joining W&M, he worked for 7 years as a research scientist with Science and Technology Corporation, working under contract to NASA Langley Research Center on advanced concepts in image processing. He jointly holds a US patent with NASA researchers on the Retinex nonlinear image enhancement technique that has wide-ranging applications. He is also vice president for research and development of TruView Imaging Company, which develops Retinex-based products.



Daniel J. Jobson received his BS degree in physics from the University of Alabama, Tuscaloosa, in 1969. He is a senior research scientist at the NASA Langley Research Center, Hampton, Virginia. His research has spanned topics including the design and calibration of the Viking Mars lander cameras, colorimetric and spectrometric characterization of the Mars surface, 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, and nonlinear image enhancement methods that mimic the dynamic range compression and lightness/color constancy of human vision.

Glenn A. Woodell is an engineering technician with the Sensors Research Branch, supporting NASA Langley's Aviation Safety program. He lead the NASA Langley image enhancement support for the shuttle Columbia investigation and is a co-inventor of the internationally patented Retinex image enhancement technology that received a 1999 Space Act Award and the 2003 Holloway Non-Aerospace Technology Transfer Award. He is a co-inventor of the recent Visual Servo technology and has published several papers on this and related topics.