Investigations into multi-scale retinex

Kobus Barnard and Brian Funt

kobus@cs.berkeley.edu, funt@cs.sfu.ca

This is a preprint of an article published in Colour Imaging: Vision and Technology, pp. 9-17 John Wiley and Sons (1999).

1. Introduction

Recent work [1, 2, 3, 4] advocates MSR as a method to bridge the gap between what a camera sees and what a human sees, with the goal being to provide image reproductions which are very similar to what the human viewer would have seen, were they present when the picture was taken. For example, humans can see details both in deep shadows and in nearby highly illuminated areas. On the other hand, a photograph of the same scene will either show the shadow as too dark, or the bright area as overexposed. The idea is to process the captured image, possibly in conjunction with the capture itself, so that the reproduction does not have these problems. For example, in Figure 1(a), the detail in the foreground is not visible with a conventional reproduction. Figure 1(b) shows the result using MSR processing before reproduction.

(a) (b)

Figure 1. Image (a) has too much dynamic range for reproduction; the area in shadow is too dark. Image (b) is the result of MSR processing.

In this paper we will focus on colour problems which occur if we utilize MSR as an image processing technique to solve these problems. If MSR is used for this purpose, it is difficult to know whether the colour of the reproduction will be accurate. We believe that the chief conceptual problem using MSR in this way is that a number of image-processing tasks are performed simultaneously without sufficient regard to the interactions occurring between them. The main practical consequence of this is that MSR is not appropriate for applications which are sensitive to colour.

In the image processing/image enhancement context, MSR serves a subset of the following five image processing goals, depending on the circumstances:

1) Compensating for uncalibrated devices (gamma correction)
2) Colour constancy processing
3) Local dynamic range compression
4) Global dynamic range compression
5) Colour enhancement

In the original MSR method all the processing steps are intertwined, and as a result, the colours are changed in image-dependent and unpredictable ways. We will disentangle these tasks, and develop a sound theoretical basis for them. In addition, when the fifth task is required, it is our view that this processing should proceed relative to a well defined colour baseline, and such a baseline is provided by our approach. To do this we use an effective, neural-network-based, colour constancy algorithm [5] to correct for mismatches between the illuminant and the camera balance. Then we use a modified version of MSR which strives to preserve the colour. This results in a well defined baseline for subsequent colour enhancement if required by the application.

2. Possible Goals of MSR Style Processing

The analysis of MSR style processing is dependent on the context of the overall objectives of the processing. Thus we first need to analyze these objectives. One possible objective of MSR processing is “scene replication”. This goal forms a large part of the stated motivation of the original MSR work, where it is described as the production of images which present to the viewer what they would have seen, if they had been present when the original was taken. Accomplishing this may or may not coincide with an image we aesthetically prefer. Similarly, even if an MSR result was not more suggestive of the original scene than a standard reproduction, we may still prefer the MSR result for aesthetic or other reasons. For example, this relationship between what we see when viewing the original scene and what we want the reproduction to be like can be analyzed with respect to any of a number of image/scene characteristics both locally and over the entire images. These characteristics include contrast related ones and colour appearance variables such as hue, chroma, lightness, brightness, and colourfulness (see [6] for discussion pertaining to this terminology). Our experience so far suggests that MSR processing (including our modified version) has an advantage from an information-content perspective. Simply put, with MSR processing of certain images, the viewer can see things that he or she would not see with a conventional reproduction, and thus this processing should be very useful for some applications. However, we have also found that more often than not, viewers prefer the original from an aesthetic point of view. We find that our modified version mitigates this problem, but not nearly enough to be considered a solution. Additional experiments suggest that we may be able to get the best of both by combining MSR results with the original, but such an approach is less satisfying than modifications based on a better understanding of the effects of the processing on a human viewer.

We will now discuss the scene replication goal further. Progress in this area is of potential theoretical interest because it would give us some insight into the function of human vision. In fact, some of the MSR literature is suggestive of another possible goal: modeling human vision. However, applying the same approach to scene replication and to modeling human vision requires care. If we assume that a method for scene replication can double as a model for human vision, then we must be confident that when we view a reproduction, such processing is inhibited. Otherwise, we must advocate that it makes sense to do the processing twice—once to make the reproduction, and once to view it. For example, when viewing a printed photograph, we assume that human colour constancy processing largely applies to the viewing illumination, and does not, therefore, correct for poor camera balance in the reproduction. A model for human vision, including one which doubles as a method of scene replication, must predict this distinction to the extent that it occurs. In the case of colour constancy and dynamic range compression, an argument can be made in terms of the typical angular extents when viewing the reproduction (small) and viewing the scene (large). This is touched upon in [4]. Of course, more research into this is needed.

In the case of certain properties of the human vision system such as simultaneous contrast and contrast perception, we have even less data to guide us with respect to this double processing issue. We know that simultaneous contrast effects are accessible to the viewer looking at standard reproductions, as this is normally how they are brought to our attention. So it is unclear as to what degree applying a simultaneous contrast model as part of image reproduction, where it will be re-processed by the viewer, is warranted, and we leave this analysis for future work.

In this paper we will focus on the goal of producing images which are, at least in some sense, preferable to standard reproductions. We will therefore deal mostly with the current concrete result of MSR research; namely, an algorithm for improving images with respect to dynamic range and illuminant based colour problems (colour constancy). In analyzing MSR research in this context, we make two assumptions worth specifying. First, we assume that if a picture had been taken with the appropriate colour balance (digital photography) or filter/film combination (conventional photography) for some sufficiently sophisticated reproduction system, then the viewer would find the colour appropriate. In other words, we will view colour problems as occurring due to a mismatch between the scene illuminant and that for which the reproduction system is calibrated, as well as limitations in reproduction technology. If these problems do not exist, or have been adequately corrected for, we will try to change the colour as little as possible, except to serve a specific enhancement purpose. If, on the other hand, there are colour problems, then we will first attempt to deal with them, and then perform additional processing changing the corrected colours as little as possible.

Second, we assume that for some applications, darker parts on an image should be brightened so that the details of what is in the shadow are available to human viewers. Again, this may well be at odds with the intentions of the photographer, and quite often will lead to an image that is less aesthetic. Of course, actually brightening dark areas for reproduction, assumes that there is more dynamic range available in accessible input devices as compared to that of typical output devices. We have found that this is marginally the case. Quite often, the brightened dark area is quite noisy. Thus, to some extent, MSR is a method awaiting better input devices.

3. Overview of previous MSR methods

The first MSR method for image processing [7] is a multi-scale embodiment of the well known original Retinex algorithm [8, 9] which was developed to explain human colour vision. The Retinex algorithm estimates the lightness of a surface by comparing its quantum catch in each channel to what amounts to an estimate of the maximum in that channel, found by exploring a number of random paths from the pixel in question, discounting gradual illumination changes. By varying the length and number of the paths, the normalization can be made to have a more local effect, and in [7] a method for efficiently computing the estimates at various scales and combining the results is presented. We mention that one problem with adapting maximum based scaling from the domain of human vision models to the domain image processing is that such methods are sensitive to clipped input pixels [10].

In later versions of Retinex, normalization based on maximum estimate was dropped in favour of normalizing by a Gaussian weighted average of a relatively large surround [11] (the relative values of these two models is still being investigated [12]). The standard MSR method analyzed here builds upon this second form of Retinex, which leads to single-scale Retinex (SSR) [2, 3, 4]. For SSR we have:

$$R_i(x,y,c) = \log\{I_i(x,y)\} - \log\{F(x,y,c) \otimes I_i(x,y)\} \tag{1}$$

where $R_i(x,y,c)$ is the output for channel "i", $I_i(x,y)$ is the image value for channel "i", $\otimes$ denotes convolution, and $F(x,y,c)$ is a Gaussian surround function explicitly given by:

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

with K selected so that:

$$\iint F(x,y,c)\,dx\,dy = 1 \tag{3}$$

In the above, the constant "c" is the scale. The result of applying SSR to image in Figure 2(a) for scales or 15, 80, and 250 is shown in Figure 2(b), 2(c), and 2(d), respectively. The MSR output is simply the weighted sum of several SSR's with different scales:

$$R_{M_i}(x,y,\mathbf{w},\mathbf{c}) = \sum_{n=1}^{N} w_n R_i(x,y,c_n) \tag{4}$$

where $R_{M_i}(x,y)$ is the MSR result for channel "i", $\mathbf{w} = (w_1, w_2, \ldots, w_N)$ where w_n is the weight of the n'th SSR, $\mathbf{c} = (c_1, c_2, \ldots, c_N)$, where c_n is the scale of the n'th SSR, and we insist that $\sum_{n=1}^{N} w_n = 1$. In [2] the authors state that the choice of scales is application dependent, but that for most applications at least three scales are required, and that equal weighting is usually adequate. The MSR result of combing these three SSR results for Figure 2(a) is shown in Figure 2(e). Additional examples are available in the MSR literature.

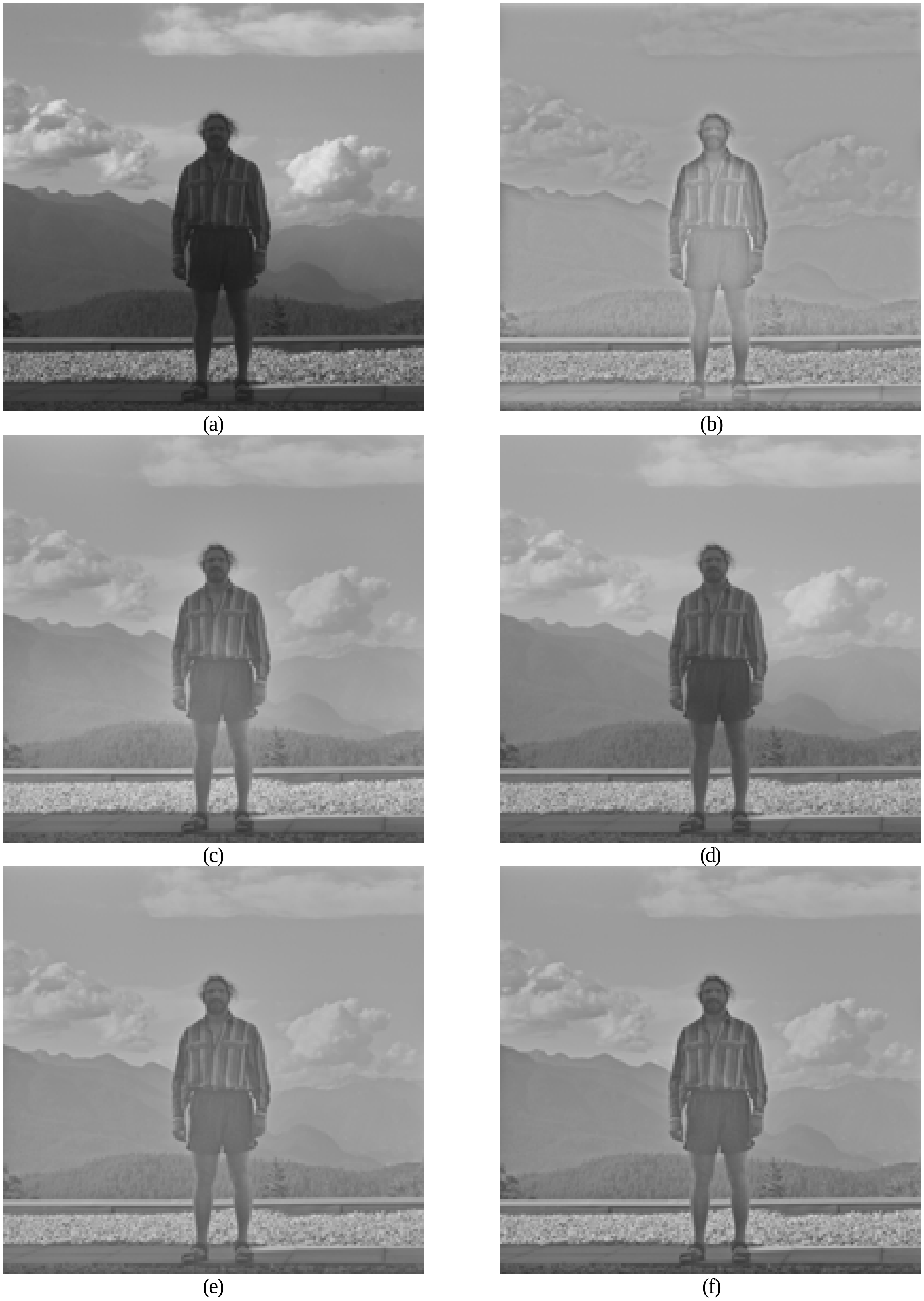


Figure 2. Image (a) is the input. Image (b) is the SSR at scale 15, (c) is for scale 80, and (d) is for scale 250. Image (e) is the result combining the results at these three scales, and (f) is the result with the colour restoration step added. Images (b), (c), (d), and (e) have been adjusted to have a brightness range similar to that of (f), as the MSR gain-offset constants are meant to apply to the complete process, as applied to get image (f).

The result of the above processing will have both negative and positive RGB values, and the histogram will typically have large tails. Thus a final canonical gain-offset is applied as mentioned in [3] and discussed in more detail below.

This processing can cause image colours to go towards grey, and thus an additional processing step was proposed in [1]:

$$R'_{M_i}(x,y,\mathbf{w},\mathbf{c},C) = R_{M_i}(x,y,\mathbf{w},\mathbf{c}) * I'_i(x,y,C) \tag{5}$$

where $I'_i(x,y,C)$ is given by:

$$I'_i(x,y,C) = \log\left(1 + C\frac{I_i(x,y)}{\sum_{i=1}^{3} I_i(x,y)}\right) \tag{6}$$

where we have taken the liberty to use log(1+x) in place of log(x) to ensure a positive result. In [4] a value of 125 is suggested for C; for [13] we empirically settled on a value of 100 for a specific test image. The difference between using these two values is small. In [4] (formula 5) a second constant is used which is simply a multiplier of the result: $I''_i(x,y,C,b) = \beta I'_i(x,y,C)$. However, in our implementation this constant is absorbed in the final gain-offset step.

A few more words about the final gain-offset step are warranted. Figure 8 of [3] shows how clipping is needed to have good contrast, as the resultant image histogram has quite large tails. We assume that the goal of consistently removing these tails led to the pair of gain-offset constants recently published in [4][1]. Although these constants are not optimal for most images, they yield serviceable results for many images, and thus we agree that widespread applicability of the constants is a strength of MSR. Nonetheless, we have found that the appearance of the result is substantially affected by the gain-offset procedure, and that part of the contrast enhancement achieved by MSR can be traced to this procedure, and we thus feel that more research into the relationship of the gain-offset procedure and the rest of MSR processing is warranted. As a start, we have experimented with some input dependent methods for the gain offset adjustment, where heuristics are applied to the image histogram to find gain offset parameters. We have also tried setting the gain based on the offset such that areas which are in the middle of the range (as defined by MSR) are mapped to slightly less than the middle range of the output (35% gives reasonable results). This makes the algorithm dependent on only one parameter and thus is a good method when it is anticipated that a final user adjustment to the gain-offset will be desirable.

---

[1]. We had difficulty obtaining reasonable results using the formula and constants in [4] and we acknowledge the subsequent help of Don Jobson with implementing standard MSR.

4. MSR Style Algorithms and Colour fidelity

Colour fidelity in image reproduction is a complex and active research topic. As a starting point, we use CIE colorimetry which applies to isolated regions. Furthermore, we assume that changing the intensity does not change the perceived non-intensity aspects of colour. This assumption is, of course, only an approximation. Even for isolated colours, colour appearance does change with intensity (see [6], chapter 6, or [14], chapter 5) with the Bezold-Brücke effect perhaps being the most important deviation. Nonetheless, we take this assumption as a serviceable approximation, noting that much colorimetry, such as that embodied by the CIELAB space, uses this assumption as a starting point. We also note that the original MSR also does not address these problems, and that our modified version of MSR is more suited to adding corrections for deviations from the above assumption. We thus acknowledge that an important area for future work is to include more complex colour appearance models, including ones for simultaneous contrast as already discussed.

For the purposes of this work, then, we assume that a first approximation of faithful colour reproduction is to preserve CIE chromaticity, as defined by $x=X/(X+Y+Z)$ and $y=Y/(X+Y+Z)$. If we assume that the region's reproduced chromaticity $x'y'$ matches the region's scene chromaticity xy, then it follows immediately that $X' = kX$, $Y'=kY$, and $Z'=kZ$, where $k = (X'+Y'+Z')/(X+Y+Z)$. Similarly, scaling the reproduction $X'Y'Z'$ by a constant, k, preserves chromaticity. Thus to preserve chromaticity we can manipulate k on a region-by-region basis, but not otherwise.

Normally we deal with intermediate variables such as camera RGB. In order to be confident that a good approximation of scene chromaticity is being reproduced, we need to know the relationship between the input scene XYZ and RGB, and RGB and the reproduction X'Y'Z'. If we know these relationships, then we can map intermediate variables into ones which are linearly related to input XYZ and output X'Y'Z'. Having done so, we can manipulate the reproduced contrast while preserving chromaticity by scaling pixel RGB. Linearity from RGB to X'Y'Z' ensures that the chromaticity reproduced by (kR, kG, kB) is the same as that reproduced by (R, G, B). Similarly, if we work in linear RGB, then we can investigate the extent that other strategies for contrast enhancement alter chromaticity, although an exact characterization will be output device dependent.

Thus to preserve image chromaticity while doing dynamic range compression, we begin with calibrated input and output devices. For mathematical convenience we use the calibration to map real coordinates into ones which are linear (if necessary). If we gauge results on a CRT monitor, this means that the CRT's non-linearity must be taken into account. It is standard to approximate the output of a CRT monitor channel to input X (scaled to be in the range [0,1]) by

x^{γ} where γ is usually in the range of [2.2 to 2.5]. For this reason, CRT non-linearity is often referred to as "gamma". For linear output one must then use a reverse gamma transformation given by: $x^{1/\gamma}$. We note that on some systems intermediate software applies a reverse gamma transformation, so that the apparent gamma is different than the range given above [15]. Due to a variety of reasons, this simple power law is only an approximation [16, 17], and furthermore, the best value for gamma is a function of monitor, monitor settings, and channel. Thus we have found it worthwhile to calibrate our monitor with a spectraradiometer.

There is an interesting relationship with standard MSR and gamma correction. That method uses a channel-independent logarithm, which normally would have the side effect of changing the image colours substantially. However, the operation approximates monitor gamma correction, and thus the colour shift is far less of a problem when displaying the result without gamma correction than would be expected. In fact, applying gamma correction to the result of MSR processing normally gives poor results. Specifically the images look washed out and over gamma corrected. The problem with just accepting and using this coincidence as a conveniently provided gamma correction is that device calibration (gamma correction) is meant to compensate for devices, but now one is committed to a single method, and thus the result is device dependent. Regardless, since MSR can, to some extent, play the role of gamma correction, it is important to ensure proper gamma correction is being applied to the original image when being compared to MSR results on a monitor.

Since device non-linearity affects colour, ideally an algorithm like MSR should be specified with respect to some output device, with the obvious candidate being a linear output device. If the algorithm is so specified, then the user can map the output to any other calibrated device. As it stands, it appears that MSR has some built in gamma, but it is not clear what it is. Based on visual inspection, and the chromaticity results presented below, we suggest that MSR effectively applies a gamma correction in the range of 2 to 2.5, which we will refer to as the MSR equivalent gamma.[2]

Device calibration is also an issue on input. Most capture technology such as CCDs are internally linear, but it is common to supply the data with some gamma correction. This may be partly to relieve the user of the burden of thinking about gamma, but more importantly, encoding the gamma in the signal is a more effective use of the limited bandwidth (usually 8 bits per channel), given that the image is for a human viewer. Thus we often have a calibration concern on input.

Standard MSR uses channel independent logarithms of contrast ratios, and this suggests that the effect of input gamma would roughly be to change the gain-offset adjustment because with

[2]. We have recently been informed that the original MSR method was developed on a monitor with a gamma of about 1.3, and therefore, the equivalent gamma should also be considered 1.3. However, we cannot get reasonable results with this figure.

a simple channel independent logarithm, the gamma would become a scale factor. To investigate this further, we verified that when histogram based, image dependent gain offset adjustment is used, standard MSR is robust with respect to input gamma.

The initial formulation of standard MSR (as defined by equations 1 through 4 above) had a serious colour problem, which was that the image colours tend to be desaturated and greyish. This is due to the manner in which grey-world-based colour constancy processing is applied to relatively small image neighbourhoods. Each pixel's colour is compared to the average of the colours in a surrounding neighbourhood. For regions of constant colour this means that the MSR result will tend towards grey regardless of the colour of the region.

To address this problem, the standard MSR now includes a processing step which puts back some of the colour that was removed (equation 5 above). The intermediate image colours are modified by a non-linear function of the original image colours. In the context of our work, as outlined in the introduction, we are uncomfortable with such a procedure as the effect is hard to characterize. This is because the colour change of the first part (the greying out) is hard to predict, being non-linear and image dependent, and the correction factor is also non-linear.

A second problem with the colour correction step is that it seems to defeat the colour constancy benefit of MSR. A grey wall under blue light, as seen by a camera balanced for a redder light, will be too blue. MSR without colour correction will move the colour of the wall towards grey, and thus achieve some degree of colour constancy. However, if the colour correction step is now used, the colour of the wall will be moved back towards blue!

Another possible colour problem with standard MSR processing is complement colour bleeding at certain colour edges due to the local contrast enhancement. Consider a white card mounted on a yellow background. For simplicity, consider that the red and the green channels of the yellow are similar to that of white, and the blue is substantially smaller. Then only the blue channel will change due to the boundary, and the blue channel of the white near the boundary will be enhanced relative to the others which represent neutral. Hence the white card will have a blue halo near the boundary. We hasten to add that this problem, while evident on test images designed to investigate it, is normally not noticeable in images of typical natural scenes.

In the case of the luminance based processing proposed below, this colour bleeding becomes luminance bleeding. In either case, it is due to the local dynamic range compression method inherent in MSR, and will occur whenever there is a clean edge comparable in length to the smallest mask size (the effect is stronger the longer the edge). The fact that it is barely noticeable in standard MSR results is a consequence of two factors. The first is that part of the dynamic range compression in MSR is due to the move to log space. The second factor is that MSR results are influenced by the colour correction step, which adds a significant amount of the original image colour content into the result. Modified versions of MSR which do not use this step may need to mitigate this effect by using larger mask sizes, but this problem does not occur with typical natural scenes with standard mask sizes.

5. Chromaticity preserving MSR

We now outline an alternative approach to MSR. As mentioned earlier, the main idea is to separate the processing goals/effects of MSR so that each one can be done more optimally. First we ensure that the input is linear. Then we optionally apply colour constancy processing to correct for mismatches between the imaging system and the illumination. This is followed by MSR style processing on an appropriately defined image luminance. The processing here can take many forms, of which two are discussed in detail below. The RGB of the output image pixels are then set so that their chromaticity is the same as in the original linear image, but their luminances are the result of the previous processing step. At this point colour enhancement, such as increasing the colour saturation, can be applied. Finally, the image is mapped into the appropriate space to give linear output on the target device. In the case of a CRT monitor, this can be approximated by a gamma correction. We now provide some additional details.

5.1. Linearization of Input

As discussed above, dynamic range adjustment while preserving colour is most easily achieved if our intermediate colour coordinates such as camera RGB are proportional to scene radiance. Thus we attempt to linearize the input if this is not the case. We have experimented with input from a Sony DXC930 CCD camera, a Kodak DCS-460 camera, as well as Kodak photo CD images. In the case of the Sony camera, we have verified that it is linear over most of its range, and we map the RGB into a space which is proportional to scene radiance using a look up table developed as part of camera calibration. In the case of Kodak DCS-460, we very roughly estimated the input gamma as 1.6. Finally, we linearize photo CD images by inverting the algebra described in [18]. We have not verified how well this actually corresponds to scene radiance, but for the purposes of experimentation we assume it is linear as implied in [18].

5.2. Colour Constancy Processing:

If colour constancy is an issue for the application, it is dealt with next. For the purpose of this paper, we define colour constancy processing as a correction for a mismatch between the illuminant for which the imaging system is calibrated and the actual illumination of the scene. Colour correction so defined is different than simply determining an illuminant independent description of the scene. Most methods available to do this correction implicitly assume that the input is linear, and thus a good result is dependent on the linearity considerations discussed above. In fact, using the above definition for colour constancy processing almost demands reference to a linear space.

Standard MSR has its roots in the latest colour constancy work by Land [11], and colour constancy processing is one of the purported goals of MSR processing. However, the colour constancy processing inherent in standard MSR processing, even without the difficulties due to the colour correction step discussed above, has several weaknesses. First, it attempts to do colour correction in a non-linear space. Second, it is essentially based on the grey world assumption, which is not a major problem, except that there are better approaches available (see, for example, [5, 19, 20, 21, 22]). A more serious problem is that the implementation of the grey world algorithm is not optimal. Colour constancy algorithms generally make some assumption about how the illuminant chromaticity varies spatially (the most common assumption being that it is uniform), and then exploit that assumption. In the case of MSR, the use of a large scale implies some confidence that the illumination uniformity is wide, but the use of smaller scales yields poor colour constancy results due to local violations of the grey world assumption, and leads to a greyed out image. Averaging the results mitigates the errors, but also reduces the chances for good performance, and thus is unsatisfactory. We posit that if illumination uniformity is an issue, it should be dealt with explicitly in the algorithm (as is done in [23]). Otherwise, the illumination chromaticity should be assumed constant, as this gives the most effective colour constancy processing.

The colour constancy algorithm used for our experiments is a neural network trained to predict the chromaticity of the scene illuminant [5]. This is then used to compute an estimate of what the scene would look like, had it been illuminated by an appropriate illuminant for the imaging system. The performance of this algorithm is significantly better than grey world based methods.

Since the neural net approach to colour correction is relatively new, we will discuss it further. The neural net is a multi-layer Perceptron with two hidden layers. As is common, the general structure is pyramidal. The input layer consists of 2500 nodes, the first hidden layer has 400 nodes, the second hidden layer has 30 nodes, and the output layer has 2 nodes. The input to the network is based on the chromaticity space (r,g) given by $r=R/(R+G+B)$ and $g=G/(R+G+B)$. We divide that space into discrete bins, with each input neuron corresponding to one of the discrete bins. The input to each neuron is a binary value representing the presence or absence of a scene chromaticity falling in the corresponding (r,g) bin. Thus we form a (r,g) histogram of the image, and then binarize that histogram.

The output signal from the two output neurons are real valued, and correspond to an estimate of the chromaticity of the scene illuminant. The network is trained to compute this estimate by being presented with many synthesized images together with the chromaticity of the illuminant used to generate each image. The illuminants used for training are 90 sources and complex illuminations measured around our university campus using a spectraradiometer. These include a variety of both indoor and outdoor illuminations and combinations thereof, such as a mix of indoor lighting with that coming through a window. Generating the scenes further requires a database of surface reflectances and a camera model to compute RGB from spectra. Thus this method of colour constancy is specific to a given camera system.

The training of the neural net occurs by re-adjustment of neuron weights using back-propagation without momentum [24] based on the discrepancy between predicted and actual scene illuminant chromaticity. Since we know in advance that some of the input (r,g) space will not be used, we use an adaptive network whereby links to neurons dormant for an entire training epoch are deleted and replaced by ones connected to randomly selected input nodes. This modification substantially reduces training time.

Once the neural net has been trained, it is used to estimate the chromaticity of the scene illuminant from an input image. Here the computation is very efficient, with the most time consuming step being the generation of the chromaticity histogram from the image. Finally the illuminant chromaticity is used to correct the image chromaticity assuming a diagonal model for the effect of illumination change. The diagonal model maps the RGB obtained under one illuminant to those obtained under a second by scaling each channel independently. Thus the input (R, G, B) becomes $(d_R R,\ d_G G,\ d_B B)$ where the constants d_R, d_G, and d_B are equally used for all pixels. The term diagonal model is used because this corresponds to post multiplying (R, G, B) by a diagonal matrix, emphasizing the difference between this model and a full matrix model. The accuracy of the diagonal model is a function of the camera sensors [25]. Furthermore, if the diagonal model does not hold very well, it can sometimes be improved by using a linear transformation the RGB corresponding to “sharper” sensors [26, 27, 28]. In the case of the Sony DXC930 camera used for the colour constancy experiments, we have verified that the diagonal model is a good approximation.

Given a diagonal model of illuminant change, and the chromaticity of the unknown test illuminant (r_T, g_T) the diagonal elements d_R, d_G, and d_B are easily determined as follows. Let (r_C, g_C) be the chromaticity of the known, “canonical” illuminant for which the camera is balanced. Then (d_R, d_G, d_B) is given by:

$$(d_R, d_G, d_B) = \left(\frac{r_C}{r_T}, \frac{g_C}{g_T}, \frac{1 - r_C - g_C}{1 - r_T - g_T}\right) \tag{7}$$

5.3. Chromaticity Preserving Dynamic Range Adjustment

The next step is to apply MSR style processing on an expression of the image luminance. We offer two methods to do this. The first method is simpler and changes the image less, and may be preferable for images from sources known to have small dynamic range. The second method is designed to approximate the dynamic range compression of the original MSR method. The significance of the second method is that it is more appropriate on images with high dynamic range. In order to investigate the relationship of the various methods and input dynamic range we created some images with extended dynamic range by either combining a number of images taken at different apertures, or averaging a large number of images.

For the first method we apply MSR style processing *without* taking logarithms on the image luminance defined by $I_I = \sum_i I_i$ (in the case of three-channels $I_I = I_{red} + I_{green} + I_{blue}$) as follows. For each scale we map the input intensity to the output intensity, $R_I = \sum_i R_i$, using formula (1) which becomes:

$$R_I(x,y,c) = I_I(x,y)/F(x,y,c) \otimes I_I(x,y) \tag{8}$$

with $F(x,y,c)$ given by (2). To get a luminance version of MSR, we simply use formula (4) with the arbitrary channel "i" being replaced by the single intensity result. This method has the appeal that the luminance is in a space which is locally approximately linear, and thus it could be argued that some aspects of the processed image should look more natural than methods using logarithms. We find, for example, that image textures typically look more natural.

With an appropriate choice of scales, the above method can give an arbitrary amount of dynamic range compression. This is the case because a very small scale will remove all intensity differences, and reduce the image to a chromaticity image. Nonetheless, applying the above method to images with large dynamic range often gives a poor result at sharp shadow edges as mentioned above in the discussion of edge effects. The region in shadow is typically brightened significantly, but the edge itself becomes a dark area between two light areas, and thus looks unnatural.

Standard MSR typically does not brighten the shadow as much, but has much less of this edge effect, and the shadow simply looks like a less dark shadow. Again, as mentioned above, one of the reasons for this difference is that part of the dynamic range compression of standard MSR is due to the logarithm operation. This can be verified by applying the processing without any ratios. The observation that the logarithm operation has a definite benefit leads us to the second method for luminance based MSR style processing, first introduced in [29].

This method is designed to provide the same dynamic range compression as original MSR. Here we define the image luminance by the geometric mean of the channels: $I_I = \left(\prod_i^N I_i \right)^{1/N}$.

Although it is possible to use the arithmetic mean (as was done in [13]), the geometric mean is intuitively superior, as it gives a cleaner correspondence between the luminance of standard MSR and the luminance based alternative. Having computed the luminance, standard MSR processing is now applied to it, this time including the logarithm operation. In order to obtain an output luminance comparable to standard MSR, an additional step is needed. This is due to the observation above that standard MSR has some built in gamma. Thus we need to apply a corresponding reverse gamma correction to the MSR luminance result. Again the correspondence between the effect and the desired result is better served by the use of the geometric mean in place of the arithmetic mean. It should be noted that since we are only dealing with luminance, the reverse gamma correction need not be exact, and is adequately implemented with a power function. Specifically we raise the luminance to the 2.2 power, with any power in the range 2 to 2.5 being reasonable. Again, the motivation of this step is to mimic standard MSR global dynamic range compression. If a different global dynamic range compression is required, a different function can easily be substituted.

5.4. Gain-offset Adjustment, Colour Enhancement, and Output

The next step is to apply the offset part of the gain-offset procedure discussed above. Here the range of the luminance result is offset so that some of the dark pixels are clipped at zero. For the instantiation of our method designed to match the standard MSR method with respect to dynamic range compression, an offset of 4 is used**.** Having thus determined the desired relative intensity, we set each channel to the same chromaticity as in the input by:

$$R_i(x,y) = R_I(x,y) \frac{I_i(x,y)}{I_I(x,y)} \tag{9}$$

The processing so far has been designed to preserve chromaticity. However, this is not the same as producing the most pleasing colour. If colour enhancement is desired, then it is best added at this stage. For example, for some applications, increasing colour saturation may be desired.

Next we map the pixels into [0, 255], recalling that the zero point is already set by the bottom clipping of the intensity. One possible solution is to simply scale the range to fit. However, often a better result is obtained by allowing some clipping of the upper range. The chromaticities of the pixels that are clipped will be a slightly incorrect, but this is not normally noticeable. It is not recommended, however, to do the same with the bottom of the range, as this can affect the chromaticities of all the pixels. Instead it is generally better to increase the amount of clipping on the bottom by doing so when the luminance range is adjusted, as described above. Again, for completeness, we add that for the instantiation of our method designed to match the standard MSR method with respect to dynamic range compression, a gain of 2.8 is used.

The final step of the algorithm is to map the output into a space which produces linear output on the target device. In the case of a CRT monitor, this may be approximated by gamma correction.

In summary, we have an algorithm which maintains the dynamic range compression benefits of standard MSR, but is more precise with respect to colour. In addition, the algorithm requires less processing because we only need to perform convolutions on the luminance. Even if convolutions are performed using Fourier transforms, this is a considerable saving. This saving is only slightly diminished if a fast colour constancy method such as the neural net technique described above is used in conjunction with the algorithm

6. Results

We have experimented with images taken with a three chip, 8 bit, Sony DXC930 CCD camera, a Kodak DCS-460 camera, as well as Kodak photo CD images. We have viewed the results using calibrated monitors and a Hewlett-Packard PhotoSmart printer. We have not calibrated the printer, but assuming that it roughly prints sRGB gives results consistent with those using monitors. We have found that the Kodak DCS-460 camera has enough dynamic range beyond reproduction technology to make MSR processing viable. In the case of the Sony camera, we lose dynamic range by using the linear setting which simplifies calibration, but we can simulate high dynamic range image capture by averaging multiple images. Finally, in the case of Kodak photo CD images, our findings are mixed. For one set of a 100 mountaineering images, which typically have extremely large difference in brightness between sunlit and shadow, we have found that the dark areas typically do not contain enough information to be brightened. Likely the original transparencies do not have much more information, but since the photo CD images record a uniform dark result as opposed to a noisy one, we suspect that some information is lost by the photo CD process. On the other hand, the photo CD images available from the Kodak via WWW used for the second and third image in Figure 6 of [4] do give good MSR results.

For all three input devices we have verified that the second form of luminance based dynamic range compression matches the contrast appearance of standard MSR. As mentioned above, we found that the first form of luminance based dynamic range compression is too susceptible to edge effects. However, since these images often look more natural when there are no such effects, we feel that the results are of theoretical importance, and are perhaps suggestive of future research directions. The need for alternative methods is suggested by one interesting overall result: It is rare for a viewer to prefer the result of MSR style processing to the original from an aesthetic point of view, even when there is clearly more information in the processed image.

We now present the results with two images in more detail. The first image (Colour Plate 1(a)) was taken with a Kodak DCS-460 camera. Other than a rough estimate of input gamma, we have not yet calibrated this camera, and therefore we exclude colour constancy processing for this image. We put the subject in late afternoon sun in front of a mosaic which is in the shade. Part of the subject was also in shade. Plate 1(b) shows the input brightened by a factor of 20, showing more detail in the shaded area at the expense of over-exposing the foreground. To illustrate the significance of the MSR equivalent gamma, Plate 1(c) shows the result of standard MSR processing assuming that it is specified for linear input and output. Plate 1(d) illustrates the results assuming that standard MSR has an equivalent gamma of 1.5, and Plate 1(e) is for an equivalent gamma of 2.2.

In Figure 3 we plot each component of the (r,g) chromaticities for 10,000 randomly selected image points from these results versus the corresponding values in the input image. The corresponding points for the luminance based methods will lie on the line y=x (up to quantization error), as this is forced by these methods. Figure 3 shows output "r" versus input "r" for MSR equivalent gammas of 1, 1.5, and 2.2. These plots show that the choice of MSR equivalent gamma is important for colour reproduction accuracy, but we hasten to add that the choice cannot be made on the basis of these plots alone, as contrast is not represented. Visual inspection of a variety of images suggests that an MSR equivalent gamma in the range of 2 to 2.5 is reasonable.

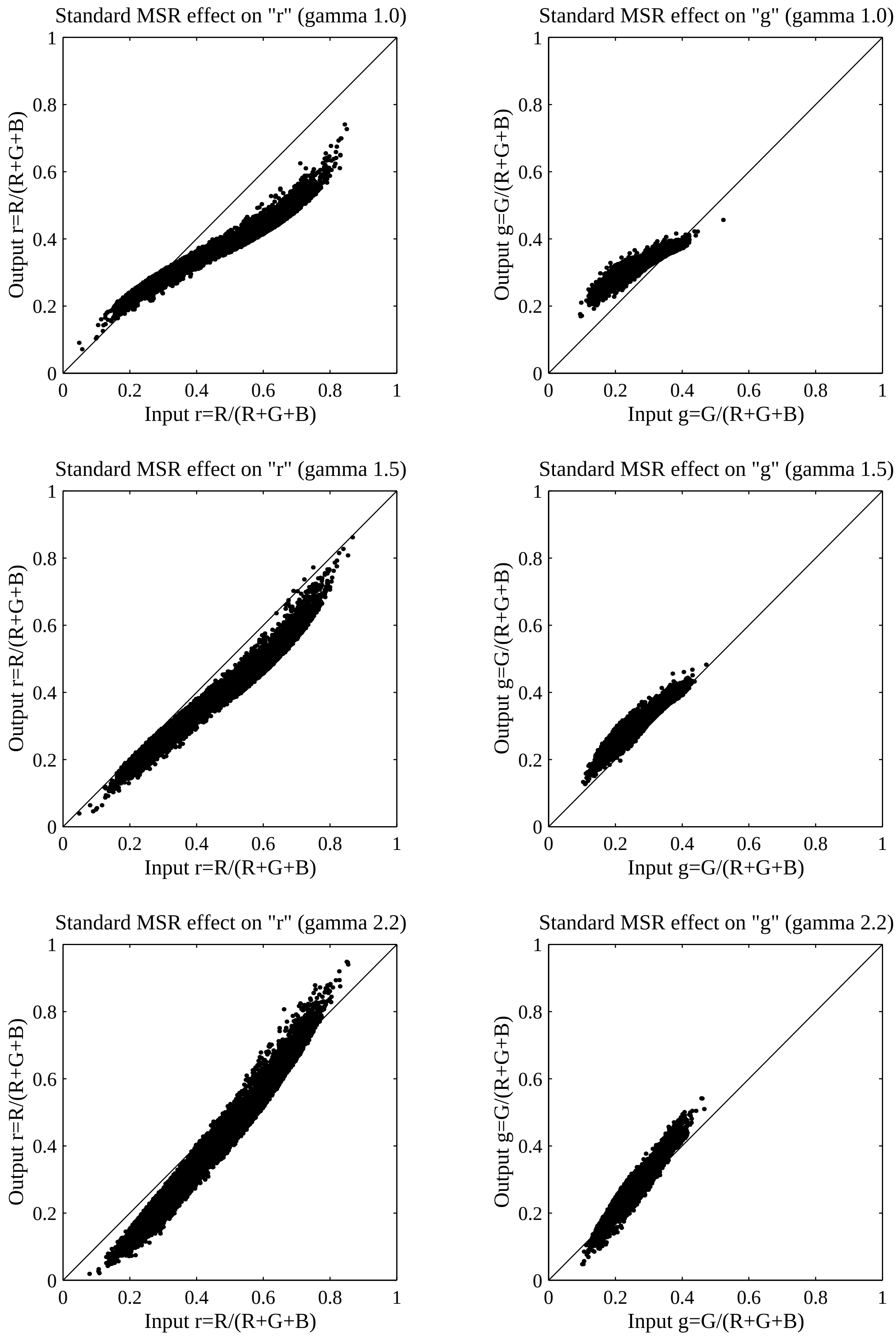


Figure 3. The effect of standard MSR processing on chromaticity for 3 different assumptions as to the appropriate output gamma. In each case, linear input is assumed.

Returning to the image results, Plate 1(f) shows the result of the version of the luminance based approach designed to match standard MSR processing with respect to dynamic range compression. Comparison with Plates 1(a) and 1(b) suggest that preserving the chromaticity yields more accurate colour reproduction, and furthermore, if the reader could see either the real mosaic or the real shirt, this claim would be even better supported. Finally, Plate 1(g) shows the result using the first luminance based method. The edge effects are very noticeable around the subject's left arm. However, we feel this method reproduces the texture of the mosaic better. This indicates to us that we have still far to go to a complete understanding of this kind of processing.

Next we explored the inter-play of the various methods and colour constancy. We took images of the same scene with a shadow of varying strengths using two very differently coloured lights. The first was a regular incandescent bulb which is a good illuminant for the indoor setting on our Sony CCD camera. The second illuminant was a cool white fluorescent together with a blue filter which creates an illuminant similar in chromaticity to that of deep blue sky. The same camera colour temperature setting was always used, creating a colour constancy problem. In one image the incandescent light source was near the camera resulting in an image which was both well colour balanced and devoid of shadows. This was used as a reference. Then shadows of increasing strengths were put across the images. In order to explore the method fully, for each illuminant an image with a extraordinarily dark shadow was taken by combining several images taken at different apertures.

Plate 2 shows some of the results. The input image taken with the blue filter is shown in (a). It is clearly too blue, and the dynamic range is too great for printing. In (b) we have corrected the input chromaticity using the neural net colour constancy algorithm, but the dynamic range problem still remains. In (c) chromaticity preserving MSR has been used to bring out the details in the shadows. The result of using chromaticity preserving MSR without the colour constancy step is shown in (d). Since chromaticity preserving MSR is designed to preserve colour, the image is overly blue just like the input. In contrast, standard MSR without the colour correction step leaves the result too grey (e). Finally, (f) shows the problem with using the colour correction step when colour constancy is considered part of the standard MSR—some of the unwanted blue has been put back into the image.

7. Conclusion

Standard Multi-scale retinex processing works quite well as a method of compressing an image's dynamic range so that the information content can be presented using a reproduction method with less dynamic range. Standard MSR performs a mixture of local (via ratios) and global (via logarithms) contrast adjustment. From our point of view, standard MSR has the drawback that it perturbs the image colours in quite unpredictable ways. We have analyzed the fundamental steps of MSR and disentangled the various operations so that their effects can be handled separately, which also makes it possible to add true colour constancy processing as one of the steps. The resulting algorithm provides better colour fidelity, is more flexible, and is less computationally expensive.

8. Acknowledgments

We are grateful for the support of Hewlett-Packard Corporation and the Natural Sciences and Engineering Council of Canada.

9. References

1. D. J. Jobson, Z. Rahman, and G. A. Woodell, “Retinex Image Processing: Improved Fidelity To Direct Visual Observation,” Proceedings of the IS&T/SID Fourth Colour Imaging Conference: Colour Science, Systems and Applications, Scottsdale, Arizona, November, pp. 124-126, 1996.

2. Z. Rahman, D. J. Jobson, and G. A. Woodell, “A Multiscale Retinex for Colour Rendition and Dynamic Range Compression,” *SPIE International Symposium on Optical Science, Engineering and Instrumentation, Applications of Digital Image Processing XIX*, Proceedings SPIE 2825, Andrew G. Tescher, ed., 1996.

3. D. J. Jobson, Z. Rahman, and G. A. Woodell, “Properties and Performance of a Center/Surround Retinex,” *IEEE Transactions on Image Processing*, March 1997.

4. D. J. Jobson, Z. Rahman, and G. A. Woodell, “A Multi-Scale Retinex For Bridging the Gap Between Colour Images and the Human Observation of Scenes,” *IEEE Transactions on Image Processing: Special Issue on Colour Processing*, July 1997.

5. B. Funt, V. Cardei, and K. Barnard, “Learning Colour Constancy,” *Proc. Fourth IS&T/SID Colour Imaging Conf.*, pp. 58-60, Scottsdale, Nov. 19-22, 1996.

6. Mark D. Fairchild, *Color Appearance Models*, Addison-Wesley, 1998.

7. Jonathan Frankle and John McCann, “Method and Apparatus for Lightness Imaging,” United States Patent No. 4,384,336, May 17, 1983.

8. E. H. Land, "The Retinex theory of Color Vision," *Scientific American*, 108-129, (1977).

9. John J. McCann, Suzanne P. McKee, and Thomas H. Taylor, “Quantitative Studies in Retinex Theory,” *Vision Research*, 16, pp. 445–458, (1976).

10. Brian V. Funt, and Kobus Barnard, “Is Colour Constancy Good Enough,” Proceedings*: 5th European Conference on Computer Vision*,, 1998.

11. E. H. Land, “An alternative technique for the computation of the designator in the Retinex theory of colour vision,” *Proc. Natl. Acad. Sci.* USA, Vol. 83, pp. 3078-3080, 1986.

12. John J. McCann, “Magnitude of color shifts from average quanta catch adaptation,” *Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, pp. 215-220, 1997.

13. B. V. Funt, K. Barnard, M. Brockington, and V. Cardei, “Luminance-based multi-scale Retinex,” *Proceedings AIC Colour 97*, Kyoto, Japan, May 25-30 (1997).

14. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas*, 2nd edition, (Wiley, New York, 1982).

15. M. Stone, “Color in Web-based 3D Graphics,” *Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, pp. 27-32, 1997.

16. R. S. Berns, R. J. Motta, and M. E. Gorzynski, “CRT Colorimetry. Part I: Theory and Practice,” COLOR *research and applications*, Vol. 18, No. 5, pp. 299-314, 1993.

17. N. Katoh and T. Deguchi, “Reconsideration of CRT Monitor Characteristics,” *Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, pp. 33-40, 1997.

18. Eastman Kodak Company, “Fully Utilizing Photo CD Images: Article No. 4—PhotoYCC Colour Encoding and Compression Schemes,” 1994, available from 15://15.Kodak.com/pub/photo-cd/general/pcd045.txt.

19. D. Forsyth, “A novel algorithm for colour constancy,” *International Journal of Computer Vision*, 5, pp. 5-36, 1990.

20. G. D. Finlayson, “Colour Constancy in Diagonal Chromaticity Space,” In *Proceedings: Fifth International Conference on Computer Vision*, pp 218-223, (IEEE Computer Society Press, 1995).

21. W. Freeman and D. Brainard, “Bayesian Decision Theory, the Maximum Local Mass Estimate, and Colour Constancy,” in *Proceedings: Fifth International Conference on Computer Vision*, pp 210-217, (IEEE Computer Society Press, 1995).

22. G. D. Finlayson, P. H. Hubel, and S. Hordley, “Color by Correlation,” *Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, pp. 6-11, 1997

23. K. Barnard, G. Finlayson, and B. Funt, “Colour constancy for scenes with varying illumination,” In *Proceedings of the 4th European Conference on Computer Vision*, pp. II:1-15, Bernard Buxton and Roberto Cipolla, eds., Springer, 1996.

24. J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, 1991.

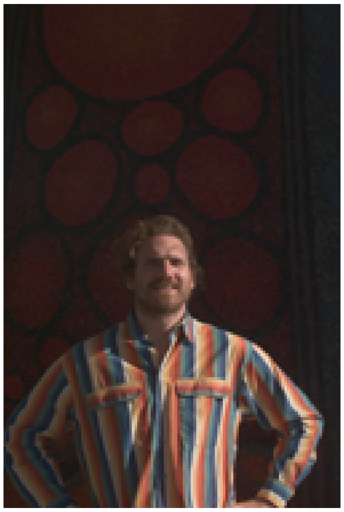
25. G. D. Finlayson, “Coefficient Color Constancy,” Ph.D. thesis, Simon Fraser University, School of Computing (1995).

26. G. D. Finlayson and M. S. Drew and B. V. Funt, “Spectral Sharpening: Sensor Transformations for Improved Color Constancy,” *Journal of the Optical Society of America A*, 11, 5, pp. 1553-1563, 1994.

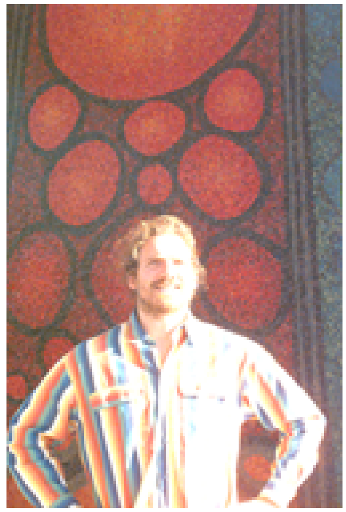
27. G. D. Finlayson and M. S. Drew and B. V. Funt, “Color Constancy: Generalized Diagonal Transforms Suffice,” *Journal of the Optical Society of America A*, 11, 11, pp. 3011-3020, 1994.

28. K. Barnard and B. Funt, “Experiments in Sensor Sharpening for Color Constancy,” *Proceedings of the IS&T/SID Sixth Color Imaging Conference: Color Science, Systems and Applications*,, 1998 (to appear).

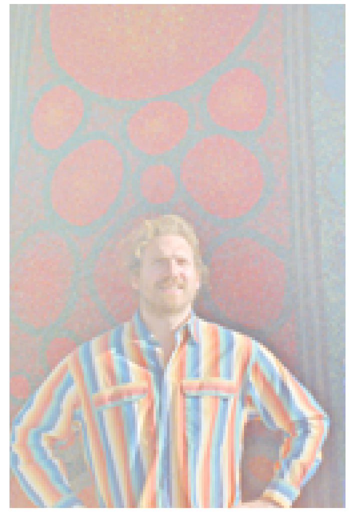
29. Kobus Barnard and Brian Funt, “Analysis and Improvement of Multi-Scale Retinex,” In *Proceedings of the IS&T/SID Fifth Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, pp. 221-226, November 17-20, 1997.



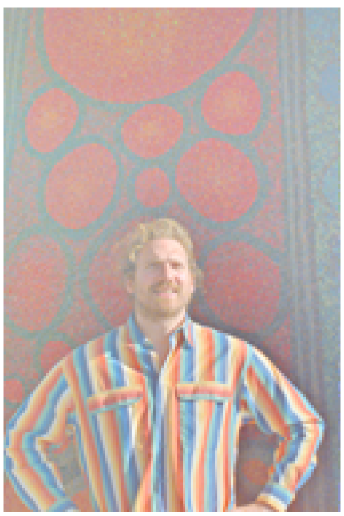
(a)



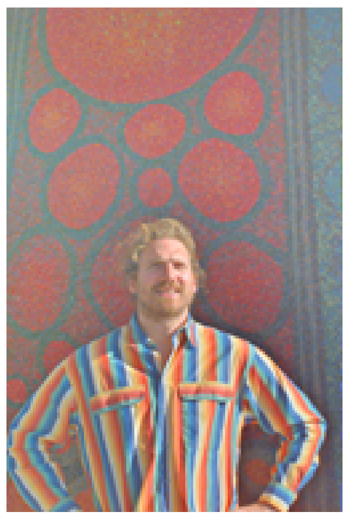
(b)



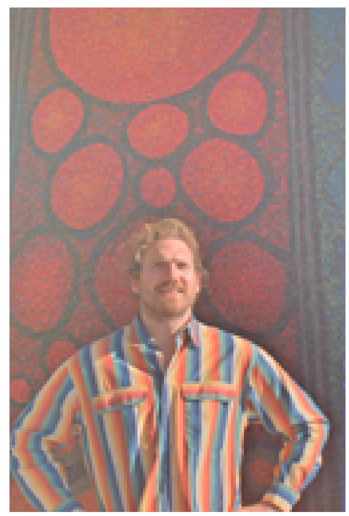
(c)



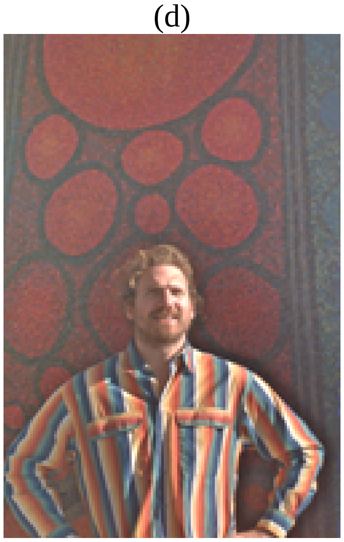
(d)



(e)



(f)



(g)

Plate 1. Image (a) has too much dynamic range to print the foreground and background simultaneously. Image (b) shows the background. Image (c) is the result of standard MSR processing assuming that the input and output are linear. Image (d) is the result of standard MSR processing assuming that the output has gamma 1.5, and (e) is the result assuming gamma 2.2. Image (f) is the result of the modified algorithm using the geometric mean for luminance based MSR processing using log space, with constants designed to match the dynamic range compression of standard MSR. Image (g) the result of the linear form of luminance based MSR processing using the arithmetic mean. This method is more prone to edge halo effects (as evident here around the subject's left arm), but offers some insight into the nature of MSR style processing.

(a) (b) (c) (d) (e) (f)

Plate 2. Image (a) was taken under a light which is far too blue for the camera. The dynamic range is also too high for printing. In (b), the colour balance has been improved with neural net colour constancy processing, but the shadowed area remains very dark. In (c), chromaticity preserving MSR has been used to improve the dynamic range. Image (d) shows the result of applying chromaticity preserving MSR without colour constancy. Image (e) shows the result using standard MSR without the colour correction step, which removes some of the unwanted blue at the expense of greying out the colours. Image (f) shows the result of standard MSR with the colour correction step. Here much of the unwanted blue has been added back.