Fast Multispectral2Gray

Ali Alsam¹ and Mark S. Drew²

¹ Høgskolen i Sør-Trøndelag - HiST, Avdeling for informatikk og e-Iring, Trondheim Norway, ali.alsam@gmail.com

² School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada V5A 1S6, mark@cs.sfu.ca

Abstract

A standard approach to generating a grayscale equivalent to an input multi-spectral image involves calculating the so-called structure tensor at each image pixel. Defining contrast as associated with the maximum-change direction of this matrix, the gray gradient is identified with the first eigenvector direction, with gradient strength given by the square root of its eigenvalue. However, aside from the inherent complexity of such an approach, each pixel's gradient still possesses a sign ambiguity, since an eigenvector is given only up to a sign. This is ostensibly resolved by looking at how one of the color channels behaves, or how the the luminance changes, or how overall integrability is affected by each sign choice. Instead, we would like to circumvent the sign problem in the first place, and also avoid calculating the costly eigenvector decomposition. We suggest replacing the eigenvector approach by generating a grayscale gradient equal to the maximum gradient amongst the color or multi-spectral channels' gradients, in each of x, y. color or But in order not to neglect the tensor approach, we consider the relationship between the complex and the simple approaches. We also note that, at each pixel, we have both forward-facing and backward-facing derivatives, which are different. In a novel approach, we consider a tensor formed from both. Then, over a standard training set, we ask for an optimum set of weights for all the maximum gradients such that the simple maxima scheme generates a grayscale structure tensor to best match the original, multi-spectral, one. If we use only forward-facing derivatives, a fast Fourier-based solution is possible. But instead, we find that a simple scheme that equally weights maxima in the forward-facing and backward-facing directions produces superlative results if a reset step is included, in a spatial-domain solution. Grayscale results are shown to be excellent, and the algorithm is very fast.

1. Introduction

Color images contain information about the intensity, hue and saturation of the physical scenes that they represent. From this perspective, the conversion of color images to black and white has long been defined as: The operation that maps RGB color triplets to a space which represents the luminance in a color-independent spatial direction. As a second step, the hue and saturation information are discarded, resulting in a single channel which contains the luminance information.

In the color science literature there are, however, many standard color spaces that serve to separate luminance information from hue and saturation. Standard examples include: CIELab, HSV, LHS, YIQ etc. We note, however, that the luminance obtained from each of these color spaces is different.

Assuming the existence of a color space that separates luminance information perfectly, we obtain a grayscale image that preserves the luminance information of the scene. Since this information has real physical meaning related to the intensity of the light signals reflected from the various surfaces, we can redefine the task of converting from color to black and white as: An operation that aims at preserving the luminance of the scene.

In recent years, research in image processing has moved away from the idea of preserving the luminance of a single image pixel to methods that include spatial context, thus including simultaneous contrast effects. Including the spatial context means that we need to generate the intensity of an image pixel based on its neighborhood.

The move towards spatial algorithms is motivated by two factors. The first is: For certain applications, preserving the luminance information per se might not result in the desired output. As an example, an equi-luminous image may easily

have pixels with very different hue and saturation. However, equating gray with luminance results in a flat uniform gray. The other factor is the advent of multi-spectral imaging where the captured data is not represented in terms of lightness, hue and saturation.

To proceed, we state that a more encompassing definition of color to grayscale conversion is: An operation that reduces the number of channels from 2, or more, to 1 while preserving certain, user-defined, image attributes. As an example, Bala and Eschbach [1], introduced an algorithm to convert color images to grayscale while preserving color edges. This is achieved by transforming the RGB image to an opponent color space, extracting the luminance image and adding chrominance edges that are not present in the luminance. In this sense, the algorithm aims at preserving chrominance edges. Gooch *et al.* [2], proposed an iterative optimization to map colors to gray maintaining chromatic plus luminance difference as well as possible, i.e. the method aims to preserve color contrast. Rasche *et al.* introduced an algorithm that aims to preserve contrast while maintaining consistent luminance. Similarly, Neumann *et al.* [3] proposed to calculate contrast based on the Coloroid system. The calculation yields a gradient field which is integrated to result in a grayscale image.

Based on the previous examples, we concluded that there is a agreement among researchers in the field that color to gray algorithms should be optimized to preserve color-contrast. We further remark that the calculation of contrast is dependent on the chosen color space and does not generalize to higher dimensional spaces. In other words, none of the previously stated methods can be used, without modification, to convert multi-spectral images to gray.

In this paper, we present an algorithm that converts color and multi-spectral images to grayscale while preserving the *maximum local contrast*. The algorithm builds upon the work of Socolinsky and Wolff [4, 5], who developed a technique for multichannel image fusion with the aim of preserving contrast. In their work, these authors use the Di Zenzo structure-tensor matrix [6] to represent contrast in a multiband image. The interesting idea added in [6] was to suggest re-integrating the gradient produced in Di Zenzo's approach into a single, representative, gray channel encapsulating the notion of contrast.

The Di Zenzo matrix allows us to represent contrast at each image pixel by utilizing a 2×2 symmetric matrix whose elements are calculated based on the spatial derivatives of the color channels in the horizontal and vertical directions. Socolinsky and Wolff defined the maximum absolute color contrast to be the square root of the maximum eigenvalue of the Di Zenzo matrix along the direction of the associated eigenvector. In [4], they noted that the key difference between contrast in the grayscale case and that in a multiband image is that, in the latter, there is no preferred *orientation* along the maximum contrast direction. In other words, contrast is defined along a line, not a vector. To resolve the resulting sign ambiguity, the authors suggest having the orientation of one of the color channels, or alternatively the luminance function, serve as a representation of a smooth function indicating vector sense. In another approach [7], overall integrability is minimized by choosing each sign, in a Markov random field algorithm.

The contribution of this paper is threefold. Firstly, we present an algorithm to calculate contrast in color and multiband images that *does not require the calculation of the eigen-decomposition of the Di Zenzo matrix*. Instead, we make use of simple gradient maxima of each of the color channels, *both in the forward-facing gradient direction and in the backward-* facing one. Using maxima results in a much faster achromatic approximation of *n*-dimensional contrast but also, and most importantly, obviates the sign-ambiguity problem and as a result generally removes unpleasant artifacts that result from getting the sign wrong. An important reset step, similar to that in Retinex [8], is inserted to avoid the influence of large gradients, We found that both the inclusion of forward- and backward-facing derivatives as well as the reset step produced subjectively better results.

The second contribution is that our algorithm incorporates a definition of orientation that is independent of any particular single color channel (such as a distinguished chrominance direction [2]).

Thus we import into the gray regime the same definition of color or multi-spectral contrast, but generated in a simple way involving spectral-channel gradient maxima. The resulting scheme can be implemented as a very fast relaxation algorithm.

2. Background

It is possible to divide the solution domain of color to gray transformations into two groups. In the first, are global projection based methods. In the second, are spatial methods.

Global methods can further be divided into image independent and image dependent algorithms. Image independent algorithms, such as the calculation of luminance, assume that the transformation from color to gray is related to the cone sensitivities of the human eye. Based on that, the luminance approach is defined as a weighted sum of the red, green and blue values of the image without any measure of the image content. The weights assigned to the red, green and blue channels are derived from vision studies where it is known that the eye is more sensitive to green than red and blue.

The luminance transformation is known to reduce the contrast between color regions [1, 5]. A classical example that is used to demonstrate this property is averaging two black and white checkerboard patterns with gray values reversed. In this case, the features of both channels are completely obliterated.

To improve upon the performance of the image-independent averaging methods, we can incorporate statistical information about the image's color, or multi-spectral, information. Principal component analysis (PCA) achieves this by considering the color information as vectors in an *n*-dimensional space. The covariance matrix of all the color values in the image is analyzed using PCA and the principal vector with the largest principal value is used to project the image data onto that vector's one dimensional space [9]. It has been shown, however, that PCA shares a common problem with the global averaging techniques [5]: Contrast between adjacent pixels in an image is always less than in the original. This problem becomes more noticeable when the number of channels increases [5].

Spatial methods are based on the assumption that the transformation from color to grayscale needs to be defined such that differences between pixels are preserved. Bala and Eschbach [1] introduced a two step algorithm. In the first step, the luminance image is calculated based on a global projection. In the second, the chrominance edges that are not present in the luminance are added to the luminance. Similarly, Grundland and Dodgson [10], introduced an algorithm that starts by transforming the image to YIQ color space. The Y-channel is assumed to be the luminance of the image and treated separately from the the chrominance IQ plane. Based on the chrominance information in the IQ plane, they calculate a

single vector: The predominant chromatic change vector [10]. The final grayscale image is defined as a weighted sum of the luminance Y and the projection of the 2-dimensional IQ onto the predominant vector.

Gooch *et al.* [2], proposed an iterative optimization to map colors to gray maintaining chromatic plus luminance difference as well as possible. The proposed optimization is unconstrained, involves multiple local minima and, for the best results, requires user input. With similar reasoning Socolinsky and Wolff [4, 5] proposed that the best grayscale image is defined as an image which, when differentiated, returns gradients that are, in a least square sense, as close as possible to the color image.

Finally, in a recent study, Connah et al. [11] compared six algorithms for converting color images to grayscale. Their findings indicate that the use of spatial algorithms (i.e., using local operations) results in a visually preferred rendering.

3. Multi-spectral, Color and Gray Contrast

Given the a rectangular grid $\Omega = \{1, ..., w\} \times \{1, ..., h\}$, let $\mathbf{P} : \Omega \longrightarrow \mathbb{R}^n$ be a multiband image. Define $g : \Omega = \{1, ..., w\} \times \{1, ..., h\} \longrightarrow \mathbb{R}$ as the unknown grayscale image. For the multiband image, let the two components of the gradient for channel k = 1..n be $\rho_x^k = \delta_x \mathbf{P}(x, y, k)$, $\rho_y^k = \delta_y \mathbf{P}(x, y, k)$ respectively define the partial derivatives. Similarly, let the grayscale derivatives in the horizontal and vertical directions be $g_x = \delta_x g(x, y)$ and $g_y = \delta_y g(x, y)$. For the multiband image, we can form Di Zenzo's structure tensor \mathbf{Z} , at each pixel, as the symmetric 2×2 matrix

$$\boldsymbol{Z}(x,y) = \begin{pmatrix} \sum_{k=1}^{n} \rho_x^k \rho_x^k & \sum_{k=1}^{n} \rho_x^k \rho_y^k \\ & & \\ \sum_{k=1}^{n} \rho_x^k \rho_y^k & \sum_{k=1}^{n} \rho_y^k \rho_y^k \end{pmatrix}$$
(1)

Since ${\pmb Z}$ is symmetric, its eigenvectors form an orthogonal matrix, ${\pmb V}$, with columns ${\pmb v}$ such that

$$\boldsymbol{Z}(x,y) = \boldsymbol{v}_i \lambda_i \boldsymbol{v}_i^T, \quad i = 1..2$$
⁽²⁾

We note that the eigenvector associated with the largest eigenvalue points in the (unsigned) direction of of maximum contrast [6].

For the grayscale image g, the Di Zenzo matrix becomes instead

$$\mathbf{Z}_{g}(x,y) = n \begin{pmatrix} g_{x}g_{x} & g_{x}g_{y} \\ & & \\ g_{x}g_{y} & g_{y}g_{y} \end{pmatrix}$$
(3)

with the n to match the Z from eq. (1) for the multiband image. The eigen decomposition of \mathbf{Z}_g is:

$$\mathbf{Z}_{g}(x,y) = \mathbf{u}_{i}\gamma_{i}\mathbf{u}_{i}^{T}, \quad i = 1..2$$

$$\tag{4}$$

The question we would like to answer is: Given Z, belonging to the original multiband image, what grayscale representation g produces a Di Zenzo matrix \mathbf{Z}_g that matches the matrix Z for the input image best? Before we attempt to answer this question, we note that second eigenvalue of the Di Zenzo matrix \mathbf{Z}_g , for the grayscale image, will always have a value of zero. On the other hand, the matrix \mathbf{Z} will, in general, have two nonzero eigenvalues. Based on this, we state that: the contrast information of a multi-band image cannot, generally, be encoded in a grayscale reproduction. Thus, the contrast of the grayscale representation can only be optimized to be as close as possible to the original.

To resolve the problem of matching the two dimensional contrast in \mathbf{Z} with the one dimensional \mathbf{Z}_g , Socolinsky and Wolff [4, 5] argued that the optimal gradient direction for generating an output grayscale image is in the maximumeigenvalue direction $\boldsymbol{v} = \boldsymbol{v}_1$. Based on this approximation we can define a one dimensional contrast at each image pixel as:

$$C_m(x,y) = \pm \sqrt{\lambda_1} v_1 \tag{5}$$

We define an energy functional:

$$W = \sum_{\Omega} [C_g - C_m]^2 \tag{6}$$

Based on Eq. (6), the best quadratic grayscale approximation g of \mathbf{P} is the one that minimizes the value of W. Clearly, a solution is obtained by setting $C_g = C_m$. We note, however, that the sign of $C_m(x, y)$ in Eq. (5) is not defined. As previously mentioned, Socolinsky and Wolff [4, 5] recommended that the sign is based on the luminance contrast. This choice is, however, problematic for neighboring iso-luminance regions. Thus, to conclude, we state that two approximations are needed to define a one-dimensional contrast representation of a multiband image: The first is that contrast is defined along the eigenvector associated with the maximum eigenvalue and the second is that the sign, the vector-sense direction, can be obtained by considering the contrast in the luminance channel.

We argue that we can solve both these problems by a straightforward scheme of using the *maximum change* over all the image channels as a simple but effective approximation to the eigenvector approach. Moreover, we notice that at each pixel there are two possible choices for the maximum gradient, since we can use either front-facing and backward-facing derivatives (speaking of the horizontal direction; and similarly for the vertical direction). Here we use a combination of *all* the gradient information available in all these color gradients derived from maxima, such that we get a best approximation to Di Zenzo's definition of contrast.

4. Fast Maximum Algorithm

In Di Zenzo's paper [6] the author states that simpler representative gray versions of the color gradient than the eigenvector-based one above could be used. E.g., the root-mean-square, over color, of the color-channel x-gradients could possibly be used as the x-component of the representative gray gradient. Alternatively, "the RMS could be replaced by the sum, or even the maximum, of the absolute values of the differences involved."

Here, our motivation is to make Di Zenzo's method *fast*, by omitting the eigenvalue determination step entirely, and replacing it by a simple calculation of the *maximum*, as suggested. However, at each pixel we note that we have in fact two possible maxima in R,G,B: one forward-facing and (the negative of) another, backward-facing, as in Fig. 1. So far as we are aware, no-one else [12] has made use of a simple maximum scheme over color channels, or of all the information available at each pixel of a multi-valued image. Here, we look at both differences ending at a pixel, West and East for the horizontal maximum, and North and South for the vertical.

Fig. 1 shows the Cartesian direction gradients for each of the R,G,B color channels. Here, in the East direction, we would choose Red as the maximum horizontal component. And for the vertical South direction, we would choose Blue. Cycling through Red, Green, and Blue we would thus take as the correct gray gradient component the maximum over all of R,G,B, separately in each of the two coordinate directions, x and y, and separately for the forward-facing and backward-facing directions. The same argument is used for higher number of channels.

In the simplest incarnation of the idea now, we could simply use the East and South maxima as the gray gradient components, and re-integrating these into a single scalar field, we arrive at a grayscale output. We found that results obtained on color and multi-spectral image tests we tried (section) justify the suitability of using both directions of gradient, in that tests using only forward-facing gradients gave less-excellent results than with using both.

In the next section, we examine just what combination of the front-facing and backward-facing maxima over the color channels gives a matrix Z_g that best matches the matrix Z for the input color image.

5. Algorithm Design

4.1 Learning Maximum-Gradient Weights

In Fig. 1, we show the representative gray gradient as a dashed vector, centered on the current pixel. This is the gradient that would be found using an eigenvalue decomposition of eq. (1). If we use a simple forward-difference definition of derivative, then, for the x direction we consider the difference between the pixel to the right and the current pixel. Usually, we take that forward-facing value to be the x-derivative and leave it at that. However, here we wish to match the (dashed in Fig. 1) gradient, determined by the Di Zenzo construction, as well as possible and under the constraint that we utilize only the maxima in each channel. Since we in fact have available the color-differences in all four directions, we make use of these to obtain a best approximation of the Di Zenzo gradient shown in Fig. 1 from the available maximum differences shown as the longest undashed arrows.

To do so, we randomly select pixels from various images. For each location, we generate a Di Zenzo matrix \mathbf{Z}_{g} associated with the gray image produced by using maxima from color channel gradients. We select weights $\alpha, \beta, \gamma, \delta$ at each pixel location so as to to optimally combine the N,S,E,W color gradients, where in each of the four directions, the gradient in that direction is taken to be the largest change over R or G or B. An optimum is defined as corresponding to that set of weights multiplying these four maximum change vectors that produces the best match to the Di Zenzo matrix \mathbf{Z} for the original, color image composed of both front- and back-facing derivatives.

That is, at a particular x, y, we look for weights $\alpha = \{\alpha, \beta, \gamma, \delta\}$ such that the gradient ∇g satisfies the following



Figure 1. Using maximum horizontal and vertical gradients, over color channels, versus using the maximum-contrast eigenvector (dashed) of the Di Zenzo matrix. Here, Red is max in the N direction, Blue is max for S, Red is max for E, and Green is max for W.

optimization:

with color gradient $\nabla \rho = \{\rho_x^k, \rho_y^k\}, k = 1..3,$

find scalar-field gradients

$$abla g^E_x \;=\; \stackrel{ ext{max_abs}}{k} \left(
ho^{k,E}_x
ight)$$

(7)

where E is the East-facing derivative;

and similarly for W,S,N.

Now combine the candidate gradient terms into a single gradient pair

$$g_x = \alpha g_x^E + \beta g_x^W$$

$$g_y = \gamma g_y^S + \delta g_y^N$$
(8)

where α is chosen by the minimization

$$\overset{\min}{\alpha} \|Z - Z_g\| \tag{9}$$

with \mathbf{Z}_{g} formed via eq. (3) using $\nabla g = \{g_x, g_y\}$.

In a study in [12], we used a standard set of color images ¹ and sampled 10,000 pixel locations from each image.

We found median values given by $\alpha = \{0.530, 0.430, 0.537, 0.443\}$. Since these values are all close to 0.5, we therefore set $\alpha = 0.5$ in subsequent development below.

 $^{^1} We \ used \ the \ KodakCD \ images \ from \ http://www.cipr.rpi.edu/resource/stills/kodak.html.$

4.2 Fourier Approach

We could, alternatively, consider using instead only one-sided gradients, i.e., $\alpha = \{1, 0, 1, 0\}$. In that case, we would fall back on standard techniques for solving for the output gray image. For a rectangular domain, Fourier methods are available for such solutions [13], yielding the same results as a spatial-gradient approach but with increase in speed. However, we found that in fact combining two-sided gradients, along with a reset step (below), produced better results than one-sided gradients; therefore, in this paper we utilize spatial relaxation techniques, as outlined in the next section.

6. Numerical Implementation:

6.1 Relaxation Method

Knowing that our assumption is that the best grayscale image is one having gradients as close as possible to a one dimensional representation of the multiband gradients we can state that we wish to minimize:

$$\min_{g} \int \int_{\Omega} \|\nabla \mathbf{P} - \nabla g\|^2 dx dy \tag{10}$$

Given our forward-facing and backward-facing derivatives, we solve Eq. 10 using Jacobi iteration, with homogeneous Neumann (so-called "natural") boundary conditions to ensure zero derivative at the image boundaries. To handle boundary points, one inserts so-called "ghost" points outside the actual image, with values given by setting the appropriate derivative at each boundary to zero.

The iteration steps are expressed as:

$$g_{x,y}^{t+1} = \frac{1}{4} [g_{x+1,y}^t + g_{x-1,y}^t + g_{x,y+1}^t + g_{x,y-1}^t + p_{x+1,y}^{max} + p_{x-1,y}^{max} + p_{x,y+1}^{max} + p_{x,y-1}^{max}]$$
(11)

This is a generalization of the standard solution [14] of the Poisson equation corresponding to Eq. 10. The difference is that, whereas the usual approach uses only one-sided gradients $\nabla \mathbf{P}$, here we decided to make use of all the information available at a pixel and use instead the maxima for two-sided derivatives. Convergence and uniqueness are similar to that for the standard approach.

Equation 11 can be solved with an arbitrary guess image g. In our implementation we use the luminance image as the starting point g^0 . As iteration proceeds, information is passed to neighboring pixels (relaxation) and the closer we get to an image that satisfies the equality of output gradient to input maximum-gradient.

We could, also, go on to insert another term in Eq. 10, corresponding to allowing the user to specify how closely the solution should actually match the input luminance – we found this to be unnecessary.

However, we did find it useful, to speed up convergence, to weight the generated solution's similarity to the original luminance gradient, decreasing this weighting with the number of iterations, as in simulated annealing.

6.2 Reset Step

Further, to control the dynamic range of the solution image, we introduced a reset step, as in Retinex [8]. Thus, at every iteration level where we impose the constraint that the resultant pixel values, $g_{x,y}^{t+1}$ are between zero and m where m is the maximum intensity level that can be displayed by the output device. Thus if $g_{x,y}^{t+1}$ is less than zero we clip the value to zero; and if it greater than m we clip it to m. Therefore, to Eq. 11 we add the conditions

if
$$(g_{x,y}^{t+1} < 0) \quad g_{x,y}^{t+1} = 0,$$
 (12)

if
$$(g_{x,y}^{t+1} > m)$$
 $g_{x,y}^{t+1} = m$

Without this reset step, we found that image appearance had less contrast. As far as we are aware, this paper is the first use of this constraint inside this kind of iteration.

7. Experiments

To test the performance of the proposed algorithm, we conducted five experiments. In the first, we compared the algorithm's output with that achieved by Socolinsky and Wolff [5]. In the second, we tested the performance compared to the Neumann *et al.* algorithm [3]. In the third, we compared our results with those reported by Bala and Eschbach [1]. In the fourth experiment, we present a comparison with Grundland and Dodgson [10]. Finally, in the fifth experiment we used multispectral, 31-channel, images to compare the performance of our algorithm with that achieved by Socolinsky and Wolff. For the first four experiments we also provide the luminance output. For the multispectral data we compared to the results achieved by PCA.

7.1 Color Image Comparison: Luminance, Socolinsky-Wolff and Present method

Fig. 2 shows four examples of the new, fast method, compared to using either luminance or Socolinsky and Wolff's method [5]. Results are seen to be an improvement over the luminance, and also over the standard approach. In particular, in any situation where we can easily see color edges, as in the painting in the top row of Fig. 2, it is easy to discern integrability errors (bending and warping) associated with Socolinsky and Wolff's method (as discussed in [11]), whereas here the present method in large measure mitigates this problem. As well, since the new method relies on not just the luminance change, but on the change over all three color channels to generate its approximation of contrast, the method tends to change the dynamic range so that dark and bright areas are more visible. Overall, the new method is a substantial improvement.

7.2 Color Image Comparison: Luminance, Neumann et al. and Present method

Fig. 3 shows three examples of the new method compared to using both the luminance and the method of Neumann et al. [3]. (Note that the Neumann et al. results were obtained from the authors website.) While the Neumann et al. method seems to improve upon the performance of luminance, this improvement is not consistent. For example, in the first image, there is better separation between the strawberries and the blueberries; however, the close-up second row image comparison shows the output of Neumann et al. does not produce the best results for all inputs — here, the luminance is clearly better. Comparing to the output for the method proposed here, we notice that the green leaf is wrongly brighter than the white plate and details have clearly been lost. In our output, there is good separation in intensity between the raspberry and the raisin, whereas this is lost in the luminance image.

The flower image in row 3 provides further indications that the Neumann *et al.* method can produce artifacts. Here we note that the separation between the white pot and the green leaves has been lost and details in the petals and leafs are poor. In close-up, row four provides evidence that the Neumann *et al.* method can in some measure outperform luminance, in that the petals and leaves are at least disambiguated. We notice, however, that for all the images, the new method provides a grayscale output that has very high local contrast and results in a reproduction that is visually appealing. For example, for the flower in row four we notice delicate textures on the petals, as well as a correct grayscale difference between petals, leaves, and pot.

In the last row, the geometric pattern is essentially lost in the luminance image. We notice that there is considerably more and indicative gradation in gray, for the present method, which is lost in some measure in the Neumann *et al.* output.

7.3 Color Image Comparison: Luminance, Bala-Eschbach, and Present method

Fig. 4 shows two examples of the present method, compared to using either luminance or Bala and Eschbach's method [1] (the output of the Bala and Eschbach method was kindly provided by the authors). To enable the reader to appreciate the difference between the methods, we cropped out details from both images, shown in the second and fourth row of Fig. 4. We note that for the different objects in the scenes the output achieved by Bala and Eschbach has the same overall effect as that of luminance, assigning near-equal tone values to the green and the pink caps. In contrast, the present method correctly separates these.

For the image in the second and fourth rows, the differentiation between petal and leaf for either of the other methods is not nearly as accurate as for the present method. Indeed, the improvement is quite compelling. We further note that for the other methods the chromatic boundaries between the different objects are enhanced with a sharpening effect which doesn't necessarily appear natural — e.g., the central (vertical) leaf boundary has a distracting bright border on some parts of its edge, in the Bala and Eschbach method. The proposed method, however, achieves separation between different color regions by changing the tonal value of the different image regions. This in turn results in an output that appears more natural than that of Bala and Eschbach.

7.4 Color Image Comparison: Luminance, Grundland-Dodgson, and Present method

Fig. 5 shows example comparisons between luminance, Grundland and Dodgson's approach [10], and the present method. Similar to the previous experiment, we provide the full scale images along with cropped details. For the first image (two women), we note that the present method increases the dynamic range of the scene, especially, in the dark region. Visual details in the top right corner of the image are arguably more apparent than even what is present in the original color version. As previously stated, the increase in the dynamic range is due to the use of contrast over all color channels.

In the detail crop of the two women image in the second row, we note that the skin-tone of the woman's face is darker in the Grundland and Dodgson output than it is in either the luminance or the present method. By definition, Grundland and Dodgson's method adds (or subtracts) chroma to the output of luminance. In the case of this image of two women this approach darkens skin tones, whereas in the present method skin tones look correct.

The same problem is seen in the window image, where the center of the bottom flower appears lighter than the petals in the Grundland and Dodgson output. This is due to subtracting the chroma of the red petals from the luminance, making them darker than the very dark center that has no chroma component. These tests reveal a symptomatic problem that the present method does not share.

For the fourth row, the present method shows more detail. And this is again true for the fifth row; in particular, the present method does not lose the prominent red stripes on the flowers.

7.5 Multispectral Image Comparison: PCA, Socolinsky-Wolff and Present method

For a multispectral experiment we used four of the images captured by Nascimento *et al.* [15]. Fig. 6 shows a comparison between the PCA approach, Socolinsky and Wolff [5] and the present method. In the first column of Fig. 6 we show a false-color representation of the 31 channel data. To obtain the color image we used channels, 25, 15 and 8. These were chosen to roughly correspond with red, green and blue. A PCA image is composed by applying PCA to 3-vector pixel values, and assigning R,G,B values according to pixel weights on the three PCA eigenvectors.

As seen in the case of the color experiments above, we notice the increase in dynamic range achieved with the present method compared to both PCA and Socolinsky and Wolff's method. Note that details, especially in dark regions, are better using Socolinsky and Wolff than in the PCA output. However the contrast is better in PCA. For all the images the output of the present method is a striking improvement over both PCA and the standard approach of Socolinsky and Wolff .

However, a closer examination of the image in the fourth row shows that noise in flat regions is more noticeable in the present method than that in PCA and Socolinsky-Wolff — the white table surfaces, e.g., have noise. This is the case because both PCA and Socolinsky-Wolff include an averaging operator: PCA averages over all the pixels and Socolinsky-Wolff averages over all the gradients. By definition the new approach includes no averaging, thus, for noisy input data, we recommend that noise analysis be included as a preprocessing step. Having said that, at normal resolution the added noise is not noticeable compared to the substantial overall improvment, in the present method, over the other two approaches.

8. Conclusions

We have outlined a new, fast method for converting still color and multispectral image data to grayscale, inspired by but not tied to the Di Zenzo structure tensor approach. Instead of having to calculate eigenvectors at each pixel, we calculate a simple maximum gradient based on the maximum over all color channels. Results are seen to be an improvement over the luminance and also better than the three other methods compared (and, parenthetically, better than simply using histogram-equalization on the luminance). Since the method is not dependent on a distinguished direction deriving from a 3-band image, it works as well for multi-spectral images as for color ones; indeed, the output is convincingly good, compared to either PCA or a standard approach. As well, the algorithm is very fast, typically converging in only 5-7 iterations.

References

- R. Bala and R. Eschbach. Spatial color-to-grayscale transform preserving chrominance edge information. In 14th Color Imaging Conference: Color, Science, Systems and Applications., pages 82–86, 2004.
- [2] A.A. Gooch, S.C. Olsen, J. Tumblin, and B. Gooch. Color2gray: salience-preserving color removal. ACM Trans. Graph., 24(3):634–639, 2005.
- [3] L. Neumann, M. Cadik, and A. Nemcsics. An efficient perception-based adaptive color to gray transformation. In Workshop on Computational Aesthetics in Graphics, Visualization and Imaging, pages 73–80, 2007.
- [4] D.A. Socolinsky and L.B. Wolff. A new visualization paradigm for multispectral imagery and data fusion. In CVPR, pages I:319–324, 1999.
- [5] D.A. Socolinsky and L.B. Wolff. Multispectral image visualization through first-order fusion. *IEEE Trans. Im. Proc.*, 11:923–931, 2002.
- [6] S. Di Zenzo. A note on the gradient of a multi-image. Comp. Vision, Graphics, and Image Proc., 33:116–125, 1986.
- M.S. Drew, D. Connah, G.D. Finlayson, and M. Bloj. Improved colour to greyscale via integrability correction. In B.E. Rogowitz and T.N. Pappas, editors, *Human Vision and Electronic Imaging XIV*, 2009.
- [8] B. Funt, F. Ciurea, and J. McCann. Retinex in matlab. J. of Electronic Imaging, 13(1):48-57, 2004.
- [9] T. M. Lillesand and R. W. Kiefer. Remote Sensing and Image Interpretation. New York: Wiley, second edition, 1994.
- [10] M. Grundland and N. A. Dodgson. Decolorize: Fast, contrast enhancing, color to grayscale conversion. In Pattern Recognition, 40(11), pages 2891–2896, 2007.
- [11] D. Connah, G.D. Finlayson, and M. Bloj. Seeing beyond luminance: A psychophysical comparison of techniques for converting colour images to greyscale. In 15th Color Imaging Conference: Color, Science, Systems and Applications., pages 336–341, 2007.
- [12] Ali Alsam and Mark S. Drew. Fast colour2grey. In 16th Color Imaging Conference: Color, Science, Systems and Applications., pages 342–346. Society for Imaging Science & Technology (IS&T)/Society for Information Display (SID) joint conference, 2008.
- [13] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving Poisson equations in computer vision

problems. IEEE Trans. Patt. Anal. and Mach. Intell., 12:435–445, 1990.

- [14] B. K. P. Horn. Robot Vision. MIT Press, 1986.
- [15] S.M.C. Nascimento, F. Ferreira, and D. Foster. Statistics of spatial cone-excitation ratios in natural scenes. Journal of the Optical Society of America A, 19(3):1484–1490, 2002.



(a) (b) (c) Figure 2. (a): Original color image. (b): Luminance. (c): Socolinsky-Wolff grayscale. (d): Present method.



(a) (b) (c) Figure 3. (a): Original color image. (b): Luminance. (c): Neumann et al. grayscale. (d): Present method.



(a) (b) (c) Figure 4. (a): Original color image. (b): Luminance. (c): Bala and Eschbach grayscale. (d): Present method.







(a) (b) (c) Figure 6. (a): Three band color image. (b): PCA. (c): Socolinsky-Wolff grayscale. (d): Present method.