First-principles approach to image lightness processing

D. Andrew Rowlands and Graham D. Finlayson;
Colour & Imaging Lab, School of Computing Sciences, University of East Anglia, Norwich, UK

Abstract

There are a variety of computational formulations of retinex but it is the center/surround convolutional variant that is of interest to us here. In convolutional retinex, an image is filtered by a center/surround operator that is to designed to mitigate the effects of shading, which in turn compresses the dynamic range. The parameters that define the shape and extent of these filters are tuned to give the "best" results. In their 1988 paper, Hurlbert & Poggio showed that the problem can be formulated as a regression, where corresponding pairs of images with and without the effects of shading are related by a center/surround convolution filter that is found by solving an optimization.

This paper starts with the observation that finding sufficiently large representative pairs of images with and without shading is difficult. This leads us to reformulate the Hurlbert & Poggio approach so that we analytically integrate over the whole sets of shadings and albedos, which means that no sampling is required. Rather nicely, the derived filters are found in closed form and have a smooth shape, unlike the filters derived by the prior art. Experiments validate our method.

Introduction

The famous retinex theory [1, 2] of color vision pioneered by Land postulates that the human visual system (HVS) does not actually perceive a scene directly in terms of its radiance distribution. Instead, it was proposed that the HVS evolved to discount the illuminant, which means that the perception of a given scene is correlated to its reflectance distribution rather than the actual flux entering the eye.

Land describes color in terms of a triplet of lightness values that correspond to three separate retinex systems, one for each of the long, medium, and short types of cone receptors in the eye [2]. Lightness, measured on a relative scale from dark to light, is the psychophysical interpretation of luminance, a relationship that is often approximated as being logarithmic [3]. In retinex theory, lightness response is correlated with the relative reflectance of a scene object rather than its luminance [1]. White is generated when lightness is placed at the top of the scale in each of the three retinex systems, while color arises from differences in the lightness response between the systems [2].

The original retinex algorithms [1, 2] were path-based computations and several variants of this type of approach were developed [4]. In 1986, Land proposed an alternative technique [5], the idea being to discount the illumination by dividing the scene flux at each small area of interest by a weighted average of the flux from an extended surround. In terms of computation on digital images, this idea can be interpreted as the convolution of an image with a center/surround filter. This version of retinex, which assumed a $1/r^2$ functional form for the surround, was developed further by Jobson *et al.* who used a Gaussian surround [6]. They also developed a multiscale version which, in effect, uses a sum of Gaussians of different spatial extents when defining their center/surround [7, 8, 9]. Other investigations into the shape of the center/surround have also been carried out [10, 11, 12].

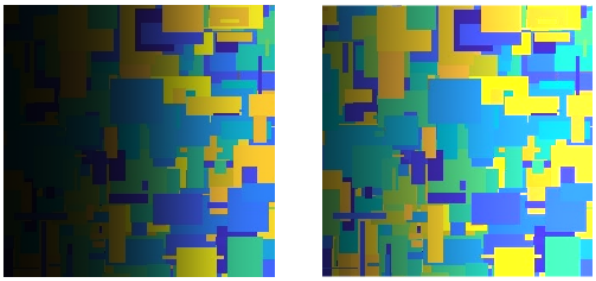

Figure 1. *The left image with a significant shading field is convolved with a center/surround filter, resulting in the image on the right, where the effects of shading are clearly reduced.*

Arguably, and understandably, much of the prior art has been directed towards designing retinex convolution filters that result in a visually pleasing processed output image. However, a more fundamental way of approaching the filter design question was proposed by Hurlbert & Poggio [13], who formulated center/surround retinex as a least-squares optimization problem. Specifically, as training examples they used scan lines along many pairs of randomly-generated Mondrian images, where each pair comprised an albedo image and the same image with a randomly-generated shading (illumination gradient) superimposed. They then used a least squares argument to calculate a filter that could be used to recover the albedo images to good accuracy. However, because of the problem formulation, the computed filter was not smooth. Indeed, it had a shape that would unlikely be implemented in computational or biological systems.

Our aim is to build upon the work of Hurlbert & Poggio by reformulating their approach so that the best filter is solved for analytically in closed form. A key part of our approach is to provide a method for integrating over all possible albedo distributions and all possible shadings of a given functional form. Given this *model* approach, we can solve for the optimal filter in closed form, and so the derived filters are smooth. Our approach is flexible so that it is easy to change the model assumptions, and our method can be tuned to different input classes of images. For example, Fig. 1 illustrates the capability of our method in terms of shading removal using a filter optimised for Mondrian images.

Readers familiar with tone mapping will be aware that the center/surround retinex (linear filtering in the logarithmic domain) is unlikely to deliver shading-free images or, indeed, preferred images. However, most spatially varying tone mappers, including edge-sensitive variants such as those that use bilateral filtering,

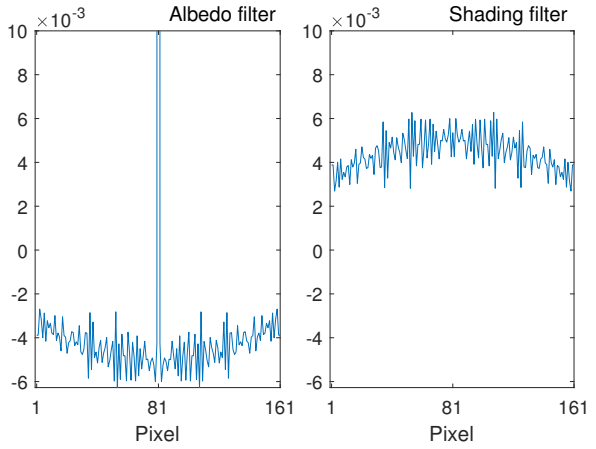

Figure 2. *1D albedo and shading convolution filters f_r and f_e solved for using the Hurlbert & Poggio method [13]. The albedo filter center extends almost to unity. The sum of f_r and f_e is a delta function.*

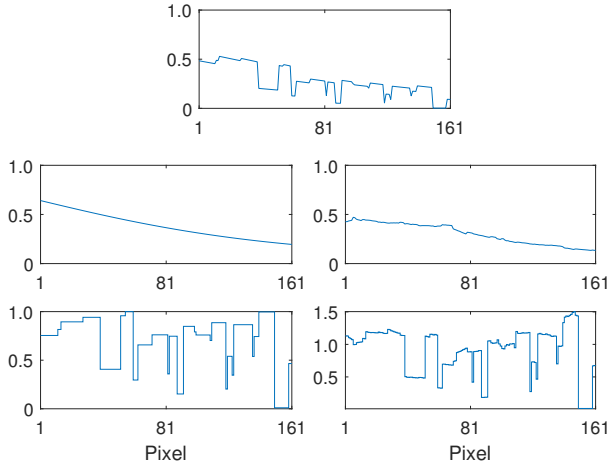

Figure 3. *Shading (middle panel) and albedo (lower panel). Ground truth data is on the left and recovered data is on the right. The color signal (top panel) is necessarily identical in both cases.*

make an assumption about how spatial information is integrated. Thus, the work we present here could also be applied in the context of more advanced algorithms.

Before outlining our methodology, we close this introduction by briefly summarising the novel approach taken by Hurlbert & Poggio [13].

Hurlbert & Poggio's least squares formulation

For each RGB image channel (or a combined luminance channel), let the color signals be defined by

$$c'(x,y) = r'(x,y)\,e'(x,y), \tag{1}$$

where r' and e' are the albedo and shading components, and x,y denote the pixel locations [4]. These color signals are assumed to be the linear image pixel values in the channel. By defining $c(x,y) = \log c'(x,y)$, $r(x,y) = \log r'(x,y)$, and $e(x,y) = \log e'(x,y)$, the logarithm of a color signal can be written as a sum,

$$c(x,y) = r(x,y) + e(x,y). \tag{2}$$

In order to determine a linear operator that is able to extract the albedo image $r(x,y)$ from $c(x,y)$, consider pairs of randomly-generated Mondrian images as training examples. Mondrian images consist of random arrangements of rectangular patches of random sizes [2] and are widely used in visual experiments [14, 15, 16, 17].

Here each Mondrian pair is defined by a randomly-generated albedo image with randomly-generated constant albedo value in each patch, and the same albedo image superimposed with a randomly-generated and smoothly-varying shading. Instead of dealing directly with the 2D images, vertical and horizontal scan lines were taken instead [13], which amounts to using many pairs of 1D vectors $r(x)$ and $c(x)$, each of length p pixels.

A set of N color signal vectors $\{c(x)\}$ can be constructed from N randomly-generated combinations of $r(x)$ and $e(x)$ according to the above prescription. These can be arranged as the rows of a color signal matrix C as follows,

$$C = R + E, \tag{3}$$

where R and E are the corresponding albedo and shading matrices [13]. These matrices all have dimension $N \times p$.

Consider the linear matrix operators L_r and L_e defined as follows,

$$CL_\mathrm{r} \approx R, \quad CL_\mathrm{e} \approx E. \tag{4}$$

(The vectors and matrices defined here are the transposes of those used in Ref. [13] and so matrix rows in Ref. [13] correspond to matrix columns here). By over-constraining the system so that $N \gg p$, the least-squares solutions are obtained as

$$L_\mathrm{r} = \left(C^\top C\right)^{-1} C^\top R, \quad L_\mathrm{e} = \left(C^\top C\right)^{-1} C^\top E, \tag{5}$$

where $\top$ denotes the transpose operator. The matrix $\left(C^\top C\right)^{-1} C^\top$, which allows us to solve least-squares regressions in closed form, is known as the Moore-Penrose pseudoinverse. Furthermore, we see that the driving forces behind least squares are the auto- and cross-correlation terms. For L_r, these are $C^\top C$ and $C^\top R$, respectively. This is a point that we will return to as it is important for the arguments set forth in this paper. Note that $L_\mathrm{r} + L_\mathrm{e} = I$, where I is the $p \times p$ identity matrix, which means that the method always returns the correct color signal.

Although the operators L_r and L_e are full $p \times p$ matrices, a 1D albedo convolution filter f_r of length p, which we generalize to 2D later in this paper, can be extracted from L_r simply by taking the central column. (In principle, the central column can be interpreted as a convolution filter that is optimised for the central pixel of the training vectors). This filter is then used to estimate the albedo vectors. Note that a 1D shading convolution filter f_s can similarly be extracted from L_e, and $f_\mathrm{r}(x) + f_\mathrm{e}(x) = \delta(x - x_\mathrm{o})$ due to the logarithmic separation, where x_o denotes the filter center.

Figure 2 illustrates example Hurlbert & Poggio filters. These were obtained using 2,000,000 training examples and p = 161; each training sample comprised a single 161 component albedo vector (a scan line through a random 2D Mondrian image with random albedo values in the range [0,1]) and a single 161 component shading vector (either a random linear ramp or a random

slowly-varying sinusoid in the range [0,1], with equal probability). Flipped pairs (the same data in reverse order) were also included in the dataset since ultimately we wish to solve for a symmetric operator, and so N = 4,000,000.

Figure 3 illustrates a typical result of applying the filters of Fig. 2. The top panel of Fig. 3 is an example of a random color signal $c(x)$, which is the product of the ground truth randomly generated shading (left, center panel) and albedo (left, lower panel). The recovered shading and albedo are shown on the right center and lower panels, respectively.

Note that Fig. 3 shows data in the primal domain but we actually carry out the computation in the logarithmic domain by taking the log of the albedo (and using log units for the shading) according to Eq. (2), and exponentiating the result. This log-process-exponentiate procedure is also used when we process real images, except that for real images a scalar is added so that the maximum log value is zero and the exponentiated result always returns values in the [0,1] interval.

Method

Here we build upon Hurlbert & Poggio's least squares formulation of retinex [13]. Our contributions:

- We reformulate the base theory to include the set of *all* color signals rather than a sum of randomly chosen examples.
- We show that we can solve for the shading and albedo autocorrelation matrices in closed form by integrating over the sets of all shading and albedo images according to a model assumption. Significantly, the color signal autocorrelation is shown to depend directly on the autocorrelations for the shadings and albedos.
- Our autocorrelation optimisation via integration leads to smooth and hence more biologically plausible filters compared to the prior art.
- We derive closed-form expressions for the autocorrelation matrices for some example scene statistics.
- We extend our method so that it can be applied to 2D images.

The set of all color signals

In Ref. [13] the color signal matrix was constructed as the sum of a randomly generated albedo matrix and a randomly-generated shading matrix, as described by Eq. (3). Here, let us instead begin by constructing sets of n albedo vectors $\{r\}$ and m shading vectors $\{e\}$. We seek to construct a color signal matrix that includes *all* $n \times m$ possible combinations of r and e.

A construction method that can deliver this outcome was proposed in Ref. [18] in a different context, where the aim was to combine spectral reflectance and spectral power distribution products. Here we use a similar approach except that we are dealing with pixel vectors instead. Furthermore, we are in the logarithmic domain and so products become sums.

First consider the $n\times p$ matrix R and the $m\times p$ matrix E formed by arranging the sets $\{r\}$ and $\{e\}$, respectively, as their rows:

$$R = \begin{bmatrix} R_{11} & \cdots & R_{1p} \\ R_{21} & \cdots & R_{2p} \\ & \vdots & \\ {}_{n1} & \cdots & R_{np} \end{bmatrix}, \quad E = \begin{bmatrix} E_{11} & \cdots & E_{1p} \\ E_{21} & \cdots & E_{2p} \\ & \vdots & \\ E_{m1} & \cdots & E_{mp} \end{bmatrix}. \tag{6}$$

Now consider the k^{th} row of E, which is a single illumination gradient vector $e_k(x)$ with $k \in \{1, \cdots m\}$, and construct an $n\times p$ matrix E_k with n identical rows, each defined by the chosen $e_k(x)$. Its matrix elements can be written

$$E_k = \begin{bmatrix} E_{k1} & \cdots & E_{kp} \\ E_{k1} & \cdots & E_{kp} \\ & \vdots & \\ E_{k1} & \cdots & E_{kp} \end{bmatrix}. \tag{7}$$

The total color signal matrix C can now expressed as the sum of two large concatenated albedo and shading matrices R_{c} and E_{c},

$$C = E_{\text{c}} + R_{\text{c}}, \tag{8}$$

where

$$E_{\text{c}} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_m \end{bmatrix}, \quad R_{\text{c}} = \begin{bmatrix} R \\ R \\ \vdots \\ R \end{bmatrix}. \tag{9}$$

In summary, E_{c} is the concatenation of m different shading matrices E_k defined by Eq. (7) with $k = 1, \cdots m$, and R_{c} is the concatenation of m identical albedo matrices R defined by Eq. (6). Each matrix in Eq. (8) has dimension $(m\times n)\times p$.

Using the above, the least squares equations defined by Eq. (5) are generalized as follows,

$$L_{\text{r}} = \left(C^\top C\right)^{-1} C^\top R_{\text{c}}, \quad L_{\text{e}} = \left(C^\top C\right)^{-1} C^\top E_{\text{c}}. \tag{10}$$

These matrix operators again have dimension $p \times p$ and again L_{r} + $L_{\text{e}} = I$.

Closed-form analytic solutions

A general matrix element of the color signal autocorrelation matrix $C^\top C$ can be obtained by multiplying the i^{th} row of $C^\top$ by the j^{th} column of C defined by Eqs. (8) and (9). Collecting terms yields

$$\begin{aligned} \left[C^\top C\right]_{ij} = &\frac{1}{m}\sum_{k=1}^{m} E_{ki}E_{kj} + \frac{1}{nm}\sum_{k=1}^{n} R_{ki}\sum_{k=1}^{m} E_{kj} \\ &+ \frac{1}{n}\sum_{k=1}^{n} R_{ki}R_{kj} + \frac{1}{nm}\sum_{k=1}^{m} E_{ki}\sum_{k=1}^{n} R_{kj} \end{aligned}, \tag{11}$$

where autocorrelation has been defined to include a normalisation by the number of sample points. In matrix form,

$$C^\top C = C^\top E_{\text{c}} + C^\top R_{\text{c}}, \tag{12}$$

where

$$\begin{aligned} C^\top E_{\text{c}} &= E^\top E + \langle R\rangle^\top \langle E\rangle \\ C^\top R_{\text{c}} &= R^\top R + \langle E\rangle^\top \langle R\rangle . \end{aligned} \tag{13}$$

Here:

- $C^\top C$ is the color signal autocorrelation matrix for the set of all $m \times n$ color signals,

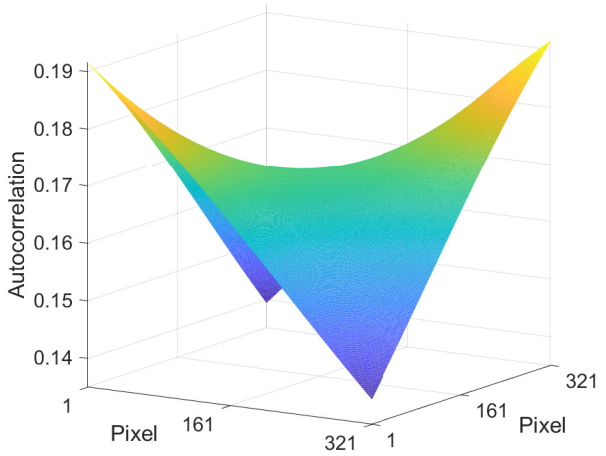

Figure 4. *Shading autocorrelation matrix for a 50:50 mix of linear ramps and sinusoids with p = 321, λ_{min} = 2, and primal-domain shading values restricted to the range [0,1].*

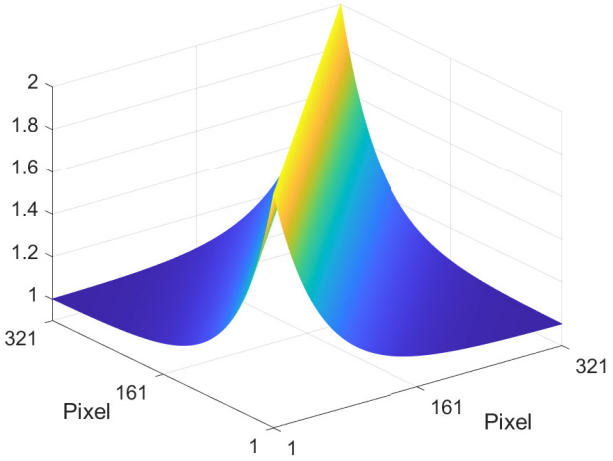

Figure 5. *Albedo autocorrelation matrix in the logarithmic domain for Mondrians with α = 0.981, which corresponds to an expected patch length of 52.6 pixels. The primal-domain albedo values were restricted to the range [0,1].*

- $E^\top E$ is the shading autocorrelation matrix for the starting set of m vectors $\{e\}$,
- $R^\top R$ is the albedo autocorrelation matrix for the starting set of n vectors $\{r\}$,
- $\langle E \rangle$ is a row vector defined by the mean of $\{e\}$,
- $\langle R \rangle$ is a row vector defined by the mean of $\{r\}$.

The practical utility of the decomposition given above is that closed-form analytic solutions can be derived for all of the quantities involved. Moreover, given functional forms for the possible albedo and shading vectors, the entire parameter space can be integrated over analytically by letting the number of training vectors $m, n \to \infty$. In other words, there is no need to sample the parameter space; the training set includes *all* possible instances of the albedo and shading vectors.

Clearly, the nature of the resultant matrix operators L_r and L_e will depend upon the functional form of the albedo and shading vectors $\{r\}$ and $\{e\}$. This means that convolution filters can be extracted from matrix operators that are optimal (in the least squares sense) for specific categories of scenes and shadings.

Shading autocorrelation matrix

Hurlbert & Poggio [13] used an example training set comprised of a 50:50 mix of random linear ramps (with both positive and negative gradients) and sinusoids (with random amplitude, wavelength and phase), all non-negative in the range $[u, v]$. We will show elsewhere that by integrating over the whole of this set analytically, the shading autocorrelation matrix elements (assuming a uniform probability distribution) are given by the following closed-form expression:

$$\left[E^\top E\right]_{ij} = \frac{1}{24}\left(u^2+uv+v^2\right)\left(5+\frac{\mathrm{Sa}\left(k_{\max}(y-x)\right)}{2}\right) + \frac{1}{6}\left(xy-\frac{(x+y)}{2}+\frac{1}{12}\right)(v-u)^2, \tag{14}$$

where Sa is the sampling function or unnormalized sinc function. Here $x = (i-1)/(p-1)$, $i = 1, 2, \cdots p$ and $y = (j-1)/(p-1)$, $j = 1, 2, \cdots p$ denote pixel locations on the $p \times p$ pixel grid. The maximum wavenumber, which corresponds to the minimum allowed wavelength, is defined by $k_{\max} = 2\pi/\lambda_{\min}$. The mean vector required by Eq. (13) is simply a constant defined by $\langle E \rangle_i = 3(u+v)/8$.

The above matrix is illustrated in Fig. 4. However, in our calculations we implemented Eq. (14) using logarithmic units with $[u, v] = [-6, 0]$.

Albedo autocorrelation matrix

We recently proposed that the albedo autocorrelation matrix $R^\top R$ for Mondrian images, which has a Toeplitz structure, can be used as a proxy for that of real scenes [19]. If the albedo values are restricted to the range $[0, 1]$, the matrix elements in the logarithmic domain are given by

$$\left[R^\top R\right]_{ij} = 2\,\alpha^{|j-i|} + \left(1-\alpha^{|j-i|}\right), \tag{15}$$

where the parameter α controls the typical size of the Mondrian patches. This controls the extent at which the autocorrelation falls to its minimum value with distance away from the matrix diagonal (see Fig. 5 for an illustration). Physically, a larger α means that the scene contains larger regions of constant albedo value on the average, while $\alpha = 0$ corresponds to a completely random scene.

The above expression assumes a uniform probability distribution for the albedo values, which in general does not lead to a mean that corresponds with real scenes. However, least squares can be used to fit this analytic expression to the albedo correlation matrix for any given image dataset, which can be determined numerically by averaging over many image scan lines [19].

2D convolution filter

After solving Eq. (10) and extracting a 1D convolution filter such as that illustrated in Fig. 6, a straightforward way to construct a symmetric 2D convolution filter, which we again denote f_r or f_e, is simply to replicate the surround radially, interpolating as necessary. Naturally, the surround subsequently needs to be normalised so that its sum equals that of the 1D surround. The filter center (i.e. the central pixel where the spike is located) should

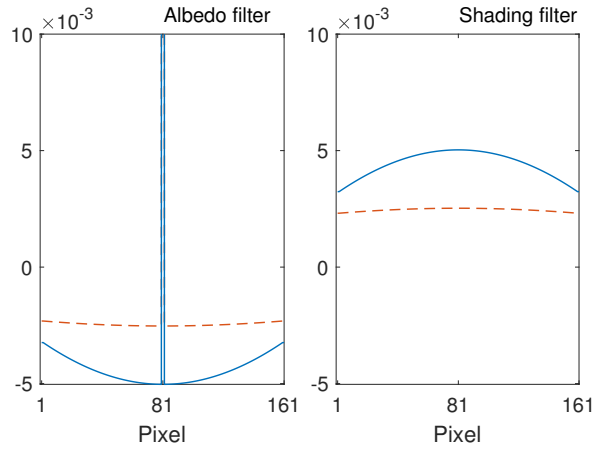

Figure 6. ***Blue solid curves****: Same as Fig. 2 except that the filters have been obtained analytically in closed form. Here the Mondrian albedo patch-length parameter α = 0.85.* ***Orange dash curves****: Filters obtained when α is increased to α = 0.99.*

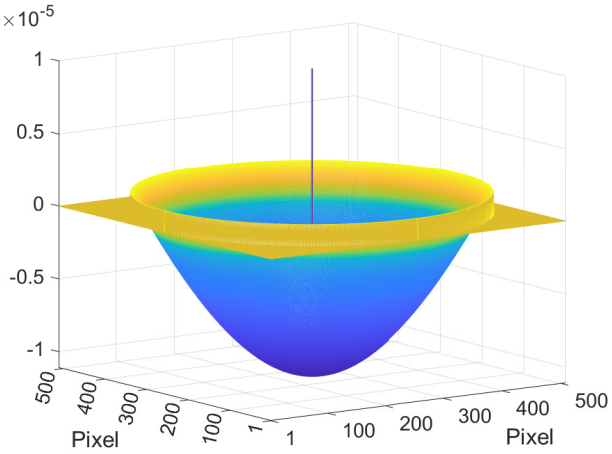

Figure 7. *2D albedo convolution filter optimised for the Kodak dataset.*

be left unaltered so that this pixel is identical in 1D and 2D. We implement the convolution in the Fourier domain using the flip-flip technique discussed in Ref. [9].

Results

Here we describe preliminary investigations into the shape of the filters and the ability of our method to process real images.

Filter shape

Since Eq. (10) can be solved analytically, smooth convolution filters are obtained. For example, by utilizing Eqs. (14) and (15) for the shading and albedo autocorrelation matrices, Fig. 6 shows the 1D filters f_{r} and f_{e}, which estimate the albedo and shading images, obtained using the same parameters as the Hurlbert & Poggio filters illustrated back in Fig. 2. Note that the extent of the filter depends upon p, the length of the training vectors, and this is built into our closed-form solutions.

Investigations show that the shape characteristics of the filters are strongly influenced by the shape of the shading and albedo autocorrelation matrices, $E^{\top}E$ and $R^{\top}R$. The former depends upon the functional form and range assumed for the shadings, and the latter depends upon the correlation between neighbouring pixels in the albedo image, which is controlled by the parameter α. For example, a larger α produces an albedo autocorrelation matrix with a wider shape (i.e. the distance between maximum and minimum autocorrelation is larger) since there will be larger regions of constant albedo value on the average. In turn, this leads to a shallower filter f_{r}, as illustrated in Fig. 6 (orange dot-dash curves).

Lightness processing

Recall that Fig. 1 demonstrated the ability of our method to remove a synthetic shading from a Mondrian image. Here, we test how well we can remove synthetic shadings applied to the 24 images in the Kodak dataset. Figure 8 shows a typical example of a Kodak image with a synthetic shading applied and the recovered shading-free image that was obtained using our method.

An error analysis using multiple examples is given below, but first we summarise how the convolution filter was obtained. Equation (10) was solved using our analytic expressions for the shading and albedo autocorrelation matrices. The shading matrix was as previously described by Eq. (14) and Fig. 4, which corresponds to a 50:50 mix of linear ramps and slowly-varying sinusoids. For the albedo autocorrelation matrix, we used the method of Ref. [19] to 'fit' to the observed path autocorrelation matrix for the Kodak image dataset. That is, we only used the path statistics drawn from the Kodak images as a means to derive the parameters that define an analytic albedo autocorrelation matrix. Furthermore, we only considered the luma channel rather than filtering each channel separately. In this case, the best fit was found for scale and offset parameters given by 0.2051 and 0.8524, and the expected patch-size parameter was found to be $\alpha = 0.981$. (See earlier section on albedo correlation matrix and also Ref. [19] for further details). This enabled us to determine the best 1D convolution filter, which was converted to the 2D filter f_{r} illustrated in Fig. 7.

Error analysis

Here we quantitatively test the ability of our method to remove synthetic shadings applied to the Kodak image dataset. Let us denote the i^{th} image with shading as C_i and the best solved-for convolution filter by f_{r}, so C_i is the product of the i^{th} shading field E_i (randomly selected) and R_i (luma channel of one of the 24 Kodak images, again chosen randomly). Note that the images C_i and R_i are assumed to have values in the interval $(0,1]$, i.e these images are *not* in log units.

Now we use our method to try and estimate R_i from C_i. Let $\hat{R}_i$ denote the albedo image estimated by convolving the logarithm of C_i with f_{r},

$$\hat{R}_i = \exp(\log(C_i) \star f_{\mathrm{r}}), \tag{16}$$

where $\star$ denotes convolution. We would like to measure how close $\hat{R}_i$ is to R_i. Post-convolution we have found it is useful to allow a scaling of the estimated albedo. Indeed, we can arrive at the same colour image C_i given the pairs (R_i, E_i) and $(\alpha R_i, (1-\alpha)E_i)$ as there is an in-built scaling ambiguity. Thus, in considering how close $\hat{R}_i$ is to R_i we will allow a constant scaling term $k_{R,i}$ so that $||k_{R,i}\hat{R}_i - R_i||$ is smallest in a least-squares sense. Here and in the

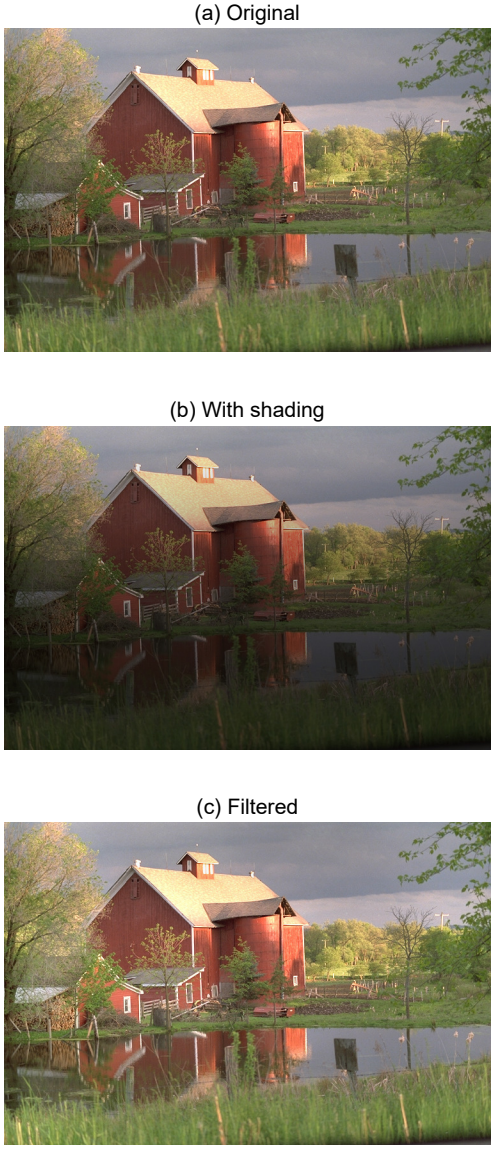

Figure 8. *(a) Original Kodak image. (b) Same image with a randomly generated shading applied. (c) Result of removing the shading by applying an albedo convolution filter optimised for the Kodak dataset.*

next two equations, $||.||$ denotes the Frobenius norm. Our percentage recovery error $errorR()$ is defined as

$$errorR(\hat{R}_i, R_i) = \frac{100 \times \left|\left|k_{R,i}\hat{R}_i - R\right|\right|}{||R_i||}. \tag{17}$$

Of course, we need to compare our method to the error found when the image C_i is not filtered at all (i.e. we do nothing to remove shading). For consistency, we also allow a per image scaling term $k_{C,i}$ that is designed to minimize $||k_{C,i}C_i - R_i||$ in a least-squares sense. Thus the *null* error, *errorN* (the error with no filtering) is calculated as

$$errorN(C_i, R_i) = \frac{100 \times \left|\left|k_{C,i}C_i - R\right|\right|}{||R_i||}. \tag{18}$$

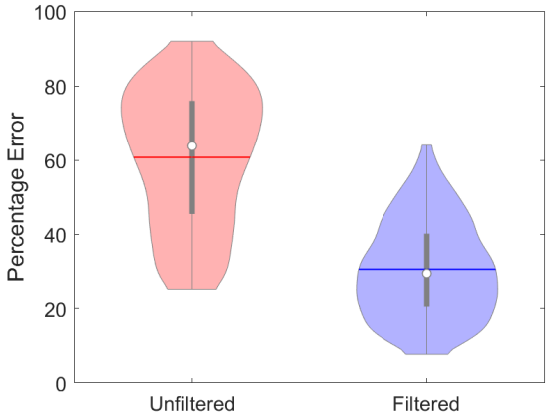

Figure 9. *Left (red): Distribution of percentage error from the ground truth when 10,000 random shadings were applied to randomly selected Kodak images. Right (blue): Same error distribution after filtering. Note: The white circles indicate the medians and the horizontal bars the means.*

For 10,000 randomly generated shaded images C_i ($i = 1, 2, \cdots 10,000$), we calculated the percentage recovery and null errors. These errors are visualised in the violin plot of Fig. 9. Note that we only considered shadings with a mean null error above 25% as below this threshold the visual effect of shading was often not significant but at 25% the shading effect was always clearly evident. The mean of the null error, i.e. the "do nothing" error, on the LHS (red violin) is 0.61 and that of the percentage recovery error on the RHS is 0.31, and so application of the filter halves the error. Furthermore, the largest errors post-filtering are much diminished, as indicated by the top section of the violins.

While it is pleasing that the error is halved, it is not close to zero. But, this is to be expected. First, the Kodak images themselves often have some shading that is mixed with our synthetic shading superimposed on top, and we cannot expect our filter to distinguish these. Furthermore, linear filtering in the log domain can deliver only so much. Removing 50% of the effects of shading with this simple convolutional processing is a good result.

Finally, as a qualitative test we applied the above filter optimised for the Kodak dataset to unseen images, namely the TM-DIED dataset [20], which has images that often lack detail due to the strong shading present. In Fig. 10 we show one of the images before and after the application of the filter. This is a nice qualitative result when we consider that we used a convolution filter that was optimised for a different dataset.

Conclusion

In this paper we reformulated Hurlbert & Poggio's least squares optimisation approach to center/surround retinex [13] and showed that it can be solved analytically. The Hurlbert & Poggio method, based on sampled albedos and shadings, returns filters that are very jagged in shape. Here we showed that their optimisation depends on the autocorrelations of the shading and albedo images, where the actual images we record in the real world are products of shading and albedo. We argue that we should use autocorrelation matrices that account for *all* shadings and albedo images, and we derive analytic expressions for these. When all

Figure 10. *Upper image: Original image from the TM-DIED dataset. Lower image: Output after applying the filter of Fig. 7, which was optimised for the Kodak dataset.*

the likely data is considered, the optimal center/surround filters are smooth. Our method is general in the sense that the shape of the filter can be tuned to the expected statistics of shading and albedo in an image.

Experiments validate our method. Qualitatively, our filter can effectively mitigate the effect of shading from images, and this is also seen quantitatively for a large set of semi-synthetic test images (actual images with shading superimposed).

Acknowledgments

This research was funded in part by EPSRC grant number EP/S028730/1. G. D. Finlayson also gratefully acknowledges the support of York University, Toronto, Canada (where he was a VISTA distinguished visiting scholar, funded by the Canada First Research Excellence Fund and York University).

References

[1] E H Land and J J McCann, Lightness and Retinex Theory, *J. Opt. Soc. Am.* 61, 1-11 (1971).

[2] E H Land, The Retinex Theory of Color Vision, *Scientific American* 237(6), 108-128 (1977).

[3] C Poynton, The rehabilitation of gamma, In B E Rogowitz and T N Pappas, eds., *Human Vision and Electronic Imaging III, Proc. of SPIE/IS&T Conference, Vol. 3299*, pgs. 332-336, 1998.

[4] A C Hurlbert, Formal connections between lightness algorithms, *J. Opt. Soc. Am. A*, 3(10), 1684-1693 (1986).

[5] E H Land, An alternative technique for the computation of the designator in the retinex theory of color vision, *Proc. Natl. Acad. Sci.* 83, 3078-3080 (1986).

[6] D J Jobson, Z Rahman, and G A Woodell, Properties and performance of a center/surround retinex, *IEEE Trans. Imag. Proc.* 6(3), 451-462 (1997).

[7] D J Jobson, Z Rahman, and G A Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, *IEEE Trans. Imag. Proc.* 6(7), 965-976 (1997).

[8] K Barnard and B Funt, Investigations into multi-scale retinex, In *Colour Imaging in Multimedia, Derby, UK*, pgs. 9-17, 1998.

[9] A B Petro, C Sbert, and J-M Morel, Multiscale retinex, *Image Processing On Line* 4, 71-88 (2014).

[10] J-M Morel, A-B Petro, and C Sbert, What is the right center/surround for retinex? In *IEEE International Conference on Image Processing*, pgs. 4552-4556, 2014.

[11] J-L Lisani, J-M Morel, A-B Petro, and C Sbert, Analyzing center/surround retinex, *Information Sciences* 512, 741-759 (2020).

[12] J-L Lisani, A-B Petro, and C Sbert, Center/surround retinex: Analysis and implementation, *Image Processing On Line* 11, 434-450 (2021).

[13] A C Hurlbert and T A Poggio, Synthesizing a Color Algorithm from Examples, *Science, New Series* 239(4839), 482-485 (1988).

[14] J J McCann, S McKee, and T Taylor, Quantitative studies in Retinex theory, a comparison between theoretical predictions and observer responses to Color Mondrian experiments, *Vision Res.* 16, 445-458 (1976).

[15] A Valberg and B Lange-Malecki, "Colour constancy" in Mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast, *Vision Res.* 30(3), 371-380 (1990).

[16] J J McCann, Lessons learned from Mondrians applied to real images and color gamuts, In *Proc. IS&T/SID Seventh Color Imaging Conference*, pgs. 1-8, 1999.

[17] A Hurlbert, Colour vision: Is colour constancy real?, *Current Biology* 9(15), R558-R561 (1999).

[18] J Paul, *Digital Image Colour Correction*, PhD thesis, University of East Anglia, 2006.

[19] D A Rowlands and G D Finlayson, Mondrian representation of real world image statistics, In *Proc. London Imaging Meeting*, pgs. 45-49, 2023.

[20] V Vonikakis, TM-DIED: The Most Difficult Image Enhancement Dataset, downloaded from Flickr repository, 2023.

Author Biography

D. Andrew Rowlands received his PhD in Physics from the University of Warwick, U.K. He has held research positions at the University of Bristol, U.K., Lawrence Livermore National Laboratory, U.S., Tongji University, China, and the University of Cambridge, U.K. He is author of the IOP book "Physics of digital photography" and is currently a researcher in the Colour & Imaging Lab at the University of East Anglia, U.K.

Graham D. Finlayson is currently a Professor with the School of Computing Sciences, University of East Anglia, U.K. His research interests include color, physics-based computer vision, image processing, and the engineering required to embed technology in devices. He is a fellow of the Royal Photographic Society, the Society for Imaging Science and Technology, and the Institution for Engineering Technology.