https://doi.org/10.2352/lim.2024.5.1.23

Optimal filter shape for convolution-based image lightness processing

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

Abstract

In the convolutional retinex approach to image lightness processing, a captured image is processed by a centre/surround filter that is designed to mitigate the effects of shading (illumination gradients), which in turn compresses the dynamic range. Recently, an optimisation approach to convolutional retinex has been introduced that outputs a convolution filter that is optimal (in the least squares sense) when the shading and albedo autocorrelation statistics are known or can be estimated. Although the method uses closed-form expressions for the autocorrelation matrices, the optimal filter has so far been calculated numerically. In this paper, we parameterise the filter, and for a simple shading model we show that the optimal filter takes the form of a cosine function. This important finding suggests that, in general, the optimal filter shape directly depends upon the functional form assumed for the shadings.

Introduction

The famous retinex theory [1, 2] of lightness perception postulates that the human visual system (HVS) at least partially attempts to discount the illuminant, which means that scenes are perceived more in terms of their object reflectances rather than the light flux that actually enters the eye. Consequently, our psychophysical interpretation of lightness is approximately correlated with scene reflectance.

Recently, a method for optimising the convolution-based approach to image lightness processing has been introduced [3]. It was shown that the optimal centre/surround filter (in the least-squares sense) depends primarily upon the autocorrelation matrices for the shadings and reflectances (albedos). By using a model approach, the autocorrelation matrices can be obtained in closed form, which results in smooth filters with shape and magnitude that adapt to the scene statistics, which are either known or can be appropriately estimated. This means that filters can be designed for specific datasets or scene categories [3]. For example, Fig. 1 shows an image from the TM-DIED dataset [4], which was processed using a lightness filter optimised for that dataset. The method differs from typical approaches to convolutional retinex where arbitrary parameters that define the shape and extent of the filter are tuned to give visually pleasing results [5, 6, 7, 8, 9, 10, 11].

As the prior-art method uses closed-form expressions for the autocorrelation matrices [3], it follows that, at least in principle, a closed-form expression could be obtained for the optimal filter. However, in the derivation of this prior-art method, the optimal filter calculation involves taking the inverse of a correlation matrix, and this fact alone complicates the search for a closed-form solution. Nevertheless, it would be advantageous to have a simple parameterised model of the optimal lightness filter.

In this paper, we adopt a slowly-varying sinusoidal model of shading. This model accounts for linear shading ramps and some shadings that are substantially more curved. By adopting

Figure 1. *Image (upper) processed by an optimal lightness filter (lower).*

this model, we show that the prior-art numerical results can be effectively parameterised in closed-form by three numbers, each of which in turn is a function of a single parameter that determines the shape of the autocorrelation of albedos.

Summary: Optimal lightness processing

Let the colour signals be defined as

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

where r' and e' are the image albedo and shading components, respectively, and x,y denote the pixel locations. 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 above equation can be converted to a sum [12] in the logarithmic domain,

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

We seek to derive a matrix operator L_{r} that, when applied to a colour signal image, will *best* recover the corresponding albedo image in a least-squares sense. It can be shown [3] that such an operator depends upon the *autocorrelation statistics* of the albedos and shadings.

In order to understand autocorrelation, assume that we have an infinitely large dataset of colour signal images, along with the component albedo and shading images. Now consider one-dimensional (1d) *scan lines* of length p pixels that pass through the centres of the dataset images. The scan lines can be taken in all directions by rotating the dataset through all angles. Let us also assume that we know the functional form for the sets of possible shading scan lines, $\{e(x)\}$, and albedo scan lines, $\{r(x)\}$. Of central interest are the shading and albedo *autocorrelation*

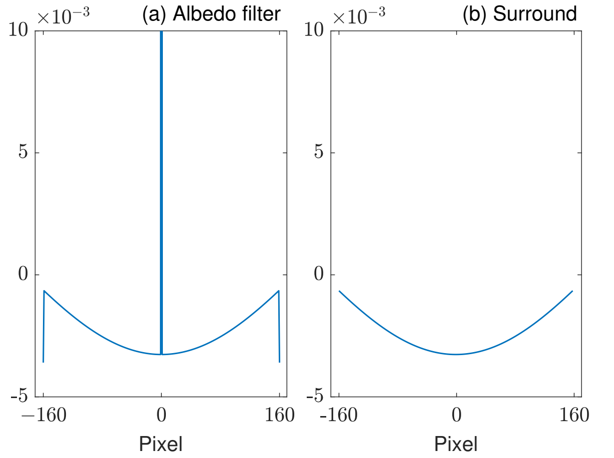

Figure 2. *(a) Example of an optimised 1d centre/surround filter, f_r, with p = 321. The centre extends almost to unity but is shown close to the origin here for clarity. (b) Extracted surround function, $f_{r,s}$, for the same filter.*

matrices C_{ee} and C_{rr}, the matrix elements of which are

$$[C_{ee}]_{ij} = \int_u^v p(e')\, e_i(x)\, e_j(x)\, de', \tag{3a}$$

$$[C_{rr}]_{ij} = \int_a^b p(r')\, r_i(x)\, r_j(x)\, dr'. \tag{3b}$$

Here $e_i(x)$, $r_i(x)$, $e_j(x)$, and $r_j(x)$ are shading and albedo values on a scan line at pixel locations i and j, respectively, and these have been normalised to the ranges $[u,v]$ and $[a,b]$ for shadings and albedos, respectively. The integration is over the infinitely large number of possible scan lines according to the probability density functions, $p(e')$ and $p(r')$. Both C_{ee} and C_{rr} have dimension $p \times p$.

Crucially, it can be shown [3] that the $p \times p$ least-squares matrix operator L_r for recovering the albedo directly depends upon C_{ee} and C_{rr},

$$L_\mathrm{r} = \left(C_{ee} + C_r^\top C_e + C_{rr} + C_e^\top C_r\right)^{-1} \left(C_{rr} + C_e^\top C_r\right). \tag{4}$$

(Here C_r and C_e are vectors representing the mean scan lines, which are typically constants). The significance of Eq. (4) is that matrix operators (and associated convolution filters - see below) that are optimal for specific *image datasets* or *scene categories* can be obtained by using Eqs. (3a) and (3b) to model their autocorrelation statistics. The relevant formulae are summarised below. (See Ref. [3] for further details).

Convolution Filter: Rather than apply the matrix operator L_r to recover albedo scan lines, we can instead use the central column of L_r as a 1d *convolution filter*. This can be convolved with colour signals (in the logarithmic domain) to recover the corresponding albedo scan lines,

$$r(x) \approx f_\mathrm{r}(x) \star c(x). \tag{5}$$

Significantly, f_r turns out to be a *centre/surround filter*. An example is illustrated in Fig. 2(a). The sharp discontinuities at the extremities of the filter will be discussed later in the paper. Since symmetry was built into the construction of the scan lines, a symmetric 2d filter, $f_{\mathrm{r,2d}}$, can be constructed simply by interpolating and renormalising the surround [3], and so

$$r(x,y) \approx f_{\mathrm{r,2d}}(x,y) \star c(x,y). \tag{6}$$

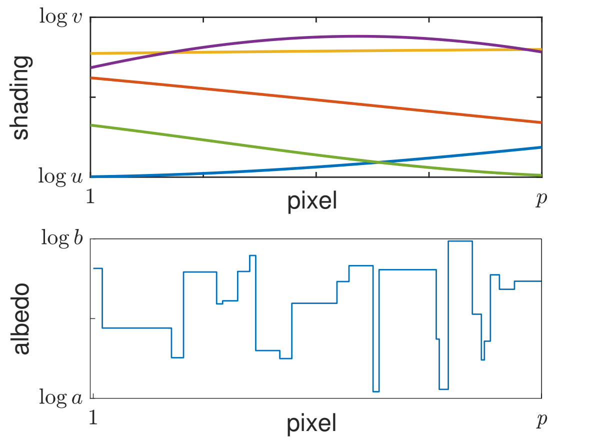

Figure 3. *(upper) Example sinusoidal scan lines. (lower) Example albedo scan line.*

In order to preserve chromaticity, the filter should be applied only to luminance colour signals [7].

Autocorrelation matrix formulae

Shadings: Let us assume that the shading scan lines take the functional form of slowly varying sinusoids,

$$e_i(x) = \frac{A}{2} + \frac{A}{2}\sin(kx + \phi), \tag{7}$$

where k is the wavenumber, ϕ is the phase, and logarithmic units have been used so that A is the amplitude in the range $[\log u, \log v]$. By assuming uniform probability density functions for k, ϕ, and A and integrating over *all* possible scan lines according to Eq. (3a), it can be shown [3] that the autocorrelation matrix elements are given by

$$[C_{ee}]_{ij} = \frac{\log^2 u + \log u \log v + \log^2 v}{12} \times \left(1 + \frac{\sin(k_{\max}(y-x))}{2k_{\max}(y-x)}\right), \tag{8}$$

where $x = (i-1)/(p-1)$ and $y = (j-1)/(p-1)$ for $i = 1,2,\cdots p$ and $j = 1,2,\cdots p$, with p being the length of the scan lines. Here $k_{\max} = 2\pi/\lambda_{\min}$, where $\lambda_{\min}$ is the minimum allowed wavelength for the sinusoids. The mean shading scan line turns out simply to be a constant,

$$[C_e]_i = \frac{\log u + \log v}{4}. \tag{9}$$

Albedos: It has been shown that the albedo autocorrelation matrices of real world image datasets can be matched to those of Mondrian image datasets [13]. If albedo values are normalised to the range $[a,b] = (0,1]$ (before the logarithm is taken), then the Mondrian model is given by

$$[C_{rr}]_{ij} = 1 + \alpha^{|j-i|}. \tag{10}$$

For a given pixel i on a Mondrian scan line such as that illustrated in Fig. 3, the parameter α defines the probability that the adjacent pixel at $i+1$ takes on the same value (a step), or a different value in the range $(a,b]$ (a jump). The average or expected step length, s, is related to α by

$$s = \frac{1}{1-\alpha}. \tag{11}$$

The mean albedo value is given by

$$[C_r]_i = -1. \tag{12}$$

Note that scale and offset parameters can be applied to Eq. (10) to alter the albedo probability density function away from a uniform distribution, however we set these to unity in this paper.

The significance of the expected step length, s, defined by Eq. (11) is that it can be matched to the average or expected object size in real image datasets [13]. In other words, it measures how correlated albedo values of neighbouring pixels are in images. It is a key quantity that affects the shape and magnitude of the optimal filter.

Method

Surround construction

In order to obtain a functional form for the surround, a value for the surround needs to be assigned where the filter centre is located. To proceed, note that Eq. (4) can be modified to recover shadings rather than albedos, in which case the central column of the matrix operator would be a 1d shading filter, $f_\mathrm{e}(x)$, rather than an albedo filter [3]. Furthermore, it can be shown that

$$f_\mathrm{r}(x) + f_\mathrm{e}(x) = \delta(x - x_0), \tag{13}$$

where x_0 denotes the filter centre [3]. (The fact that the albedo and shading filters sum to give a delta function is a consequence of the fact that the optimisation is carried out in the logarithmic domain. This filter property differs from that of other formulations of convolutional retinex, where typically the surround is normalised to spatially integrate to unity [5]). From Eq. (13) it follows that the surround, which we denote as $f_\mathrm{r,s}$, can be constructed as

$$f_\mathrm{r,s}(x) = \begin{cases} f_\mathrm{r}(x), & x \neq x_0 \\ f_\mathrm{r}(x) - 1, & x = x_0 \end{cases} . \tag{14}$$

Here the spatial grid is defined by the set of p integers $\{x\}$ in the range

$$-\mathrm{int}\left(\frac{p}{2}\right) \leq x \leq \mathrm{int}\left(\frac{p}{2}\right), \tag{15}$$

where p is the filter length and int denotes the integer part. The filter centre is located at $x_0 = 0$. Given a functional form for the surround, the centre/surround filter can be constructed by reversing Eq. (14).

Another issue to address is that in Fig. 2(a) it is evident that the example 1d centre/surround filter, f_r, has observable discontinuities at the two edge pixels. This is found generally and is due to a natural property of Toeplitz matrix inverses [14], which appear in Eq. (4). Although these two pixels have negligible effect on the overall influence of the filter in terms of shading removal, one approach to removing them would be to introduce a regularisation term that favours continuity when solving the regression [3]. However, the regularisation can slightly modify the shape of the filter, which is not appropriate in the current context as we are aiming to parameterise the true closed-form analytic result. Therefore, we choose to simply omit the two edge pixels when constructing our surround.

Figure 2(b) shows the type of surround function, which we seek to parameterise, extracted from the example filter of Fig. 2(a).

Parameter normalisation

In order to obtain a filter surround parameterisation that is valid for any chosen filter size, appropriate normalisations are required.

- When using the Mondrian model for the albedo autocorrelation matrix, the manner by which the expected step length for the 1d Mondrian patches or steps, s, affects the filter shape depends upon the filter size. This means that s should be normalised in proportion to the filter length, p. Given that s is defined by Eq. (11), let us introduce a normalised expected step length, s_n, that is defined as a percentage of the filter length,

$$s_\mathrm{n} = \frac{s}{p} \times 100 = \frac{100}{(1-\alpha)p}. \tag{16}$$

  If we choose p to be the same length as the shortest side of the image, then s cannot be longer than p, in which case the maximum value for s_n is 100. The minimum value for s_n is $100/p$.
- The minimum wavelength for the sinusoidal shadings, λ_min, must scale in proportion to the filter length, p. In the present manuscript, we assume that $\lambda_\mathrm{min} = 2p$, i.e. twice the length of the filter.

Parameter fitting

The procedure for parameterising the filter surround, $f_\mathrm{r,s}$, can be outlined as follows:

- Choose a sufficient number of sample points (pixels) for performing the parameterisation. This defines the filter length, p.
- Randomly choose a small number of test values for the normalised Mondrian step length, s_n.
- For each s_n, solve Eq. (4) to obtain the matrix operator, L_r, using the expressions for the shading and albedo autocorrelation matrices defined by Eqs. (8) and (10) and the mean vectors defined by Eqs. (9) and (12).
- For each s_n, extract the 1d centre/surround filter as the central column of L_r, and obtain the surround function using Eq. (14).
- Find a common function such as a polynomial, denoted $\tilde{f}_\mathrm{r,s}(x)$, that fits the surround function $f_\mathrm{r,s}(x)$ for each of the test s_n values to an acceptable error.
- After finding a suitable $\tilde{f}_\mathrm{r,s}(x)$, perform the same fitting procedure using a fine grid of s_n values in the interval $[100/p, 100]$. This will reveal the functional form of the coefficients that define $\tilde{f}_\mathrm{r,s}(x)$.
- Since the coefficients will be functions of the normalised expected step length, s_n, generate look-up tables for the coefficients. Alternatively, if possible, find functions that approximate the coefficients to an acceptable error.

Results

The parameter fitting described above was carried out for sinusoidal shadings (see Fig. 3). Logarithmic units were used with shadings normalised to the range $[\log(u), \log(v)] = [-6, 0]$, which corresponds to $[u, v] = [0.025, 1]$ when exponentiated. The albedo values were normalised to the full range, $(0, 1]$, before taking the logarithm. A spatial grid of 1801 points was used for the surround parameterisation, i.e. $p = 1801$, and the expected albedo step length, s, was also divided into p increments between 1 pixel and 1801 pixels so that $100/1801 \leq s_\mathrm{n} \leq 100$.

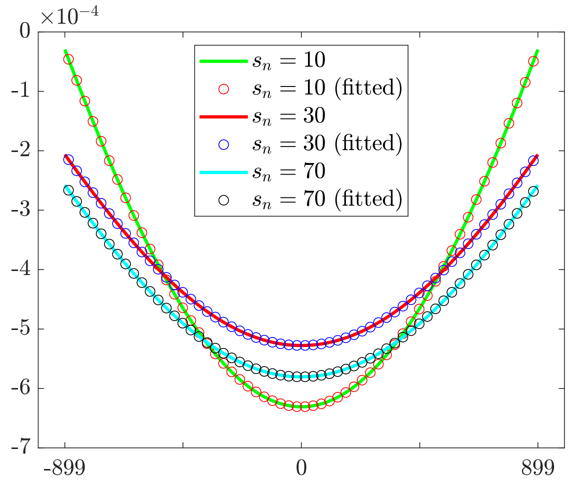

Figure 4. *Filter surround, $f_{r,s}$, (coloured lines) together with the parameterised result, $\tilde{f}_{r,s}$, (coloured circles) for a selection of normalised expected step values, s_n. The filter length was chosen to be p = 1801 pixels.*

Significantly, over most of the range of s_n values, specifically $0.4 \le s_n \le 100$, it can be shown that the filter surround can be parameterised by the following simple function,

$$\tilde{f}_{r,s} = \frac{a_0}{p} + \frac{a_1}{p}\cos\left(\frac{wx}{p}\right), \tag{17}$$

where p is the filter length and the spatial grid $\{x\}$ has been defined by Eq. (15). The coefficients a_0, a_1, a_2 and w are all functions of s_n. Clearly, a_0/p corresponds to an offset, a_1/p is the amplitude of the cosine term, and w/p is the period of the cosine.

For *very* small s_n values in the range $[100/p, 0.4]$, it can be shown that an extra term should be added to Eq. (17),

$$\tilde{f}_{r,s} = \frac{a_0}{p} + \frac{a_1}{p}\cos\left(\frac{wx}{p}\right) + \frac{a_2}{p}\cos\left(\frac{2wx}{p}\right), \tag{18}$$

where the new coefficient a_2 is again a function of s_n.

As illustrated in Fig. 4 for a selection of s_n values, the difference between the filter surround and its parameterised version (described by Eq. (17) or (18)) is imperceptible. Indeed, the RMS error is found to be negligible with an average value of 9.34×10^{-9} over the entire parameter range. The accuracy of the results suggests that these equations would be the functional form of the closed-form result if Eq. (4) were to be solved analytically.

Since the coefficients a_0, a_1, a_2, and w are all functions of s_n, in practice these could be stored in a look-up table and the coefficient values corresponding to any chosen s_n could be generated by interpolation. Nevertheless, it would be very useful if the coefficients themselves could be parameterised. It was found that a reasonable fit could be obtained by using fourth-order rational functions of the form

$$g(s_n) = \frac{b_0 + b_1 s_n + b_2 s_n^2 + b_3 s_n^3 + b_4 s_n^4}{c_0 + c_1 s_n + c_2 s_n^2 + c_3 s_n^3 + s_n^4}, \tag{19}$$

where $g = a_0$, a_1, w, or a_2. Figure 5 shows a comparison between the coefficients obtained numerically and those by fitting using Eq. (19). The RMS errors for the fits were 0.0318, 0.0780, 0.0046, and 0.0250 for a_0, a_1, w, and a_2, respectively. This means that when Eq. (19) was used to generate the coefficients, the average RMS error (over the entire range of s_n values) between the filter surround and its parameterised version increased from 9.34×10^{-9} to 1.45×10^{-5}.

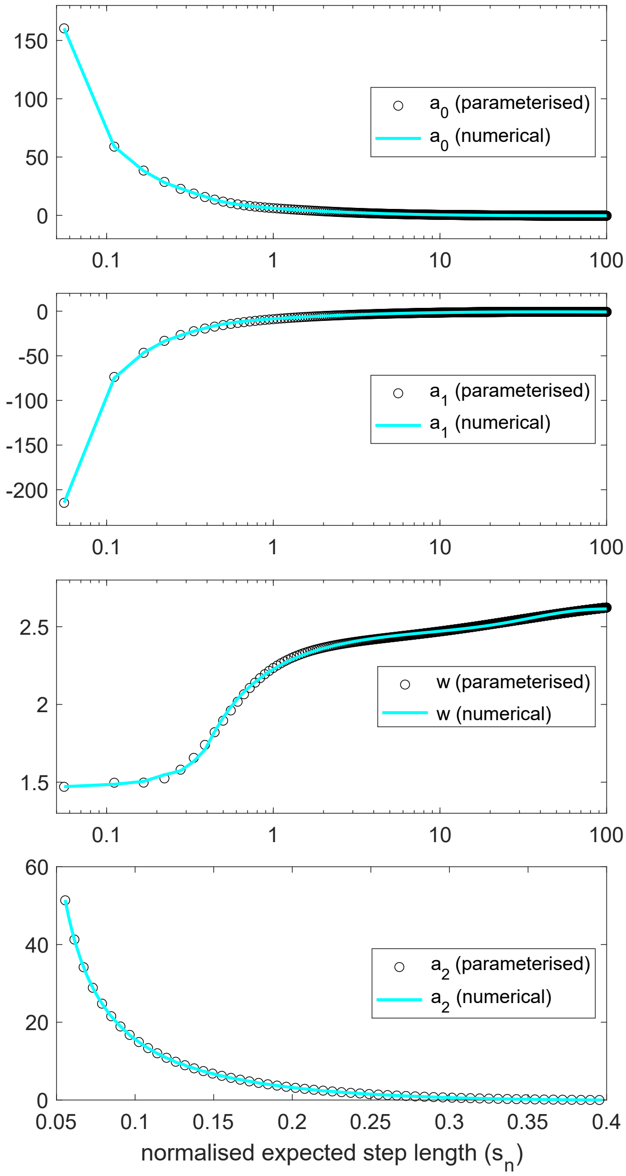

Figure 5. *Numerically obtained coefficients (cyan lines) for Eqs. (17) and (18) as a function of normalised expected Mondrian step length, s_n, together with the parameterised fits (black circles) that were obtained using Eq. (19). A logarithmic scale has been used for the horizontal axes to display the a_0, a_1 and w coefficients. A fine linear grid has been used to parameterise the a_2 coefficient up to s_n = 0.4.*

Conclusion

In this paper we have established that, for slowly varying sinusoidal shadings, the optimal centre/surround convolution filter for mitigating shading from images takes the form of a cosine function. This suggests that, in general, there is a direct link between the functional form assumed for the shadings and the shape of the optimal filter. This finding opens up interesting possibilities for future investigation. For example, it may be possible to find shadings that lead to a Gabor filter [15, 16], which is thought to be the type of filter involved in perception by the HVS.

Acknowledgments

This research is funded by the University of East Anglia and EPSRC grant number EP/S028730/1.

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] D A Rowlands and G D Finlayson, First-principles approach to image lightness processing, in *Proc. 31st Color Imaging Conference*, pages 115–121, 2023.

[4] V Vonikakis, TM-DIED: The most difficult image enhancement dataset (2021). https://sites.google.com/site/vonikakis/datasets/tm-died.

[5] 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).

[6] 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).

[7] K Barnard and B Funt, Investigations into multi-scale retinex, in *Colour Imaging in Multimedia '98, Derby, UK*, pages 9–17, 1998.

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

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

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

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

[12] A C Hurlbert and T A Poggio, Synthesizing a color algorithm from examples, *Science, New Series* 239(4839), 482–485 (1988).

[13] D A Rowlands and G D Finlayson, Mondrian representation of real world image statistics, in *Proc. London Imaging Meeting*, pages 45–49, 2023.

[14] M Dow, Explicit inverses of Toeplitz and associated matrices, *ANZIAM J.* 44(E), E185–E215, (2003).

[15] D Gabor, Theory of Communication, *J. Inst. Electr. Eng.* 93, 429–457 (1946).

[16] G H Granlund, In search of a general picture processing operator, *Computer Graphics and Image Processing* 8(2), 155–173 (1978).

Author Biography

Andy Rowlands received his PhD in Physics from the University of Warwick, UK. He was subsequently awarded an EPSRC fellowship in theoretical physics, which he held at the University of Bristol, UK. He has also held research positions at Lawrence Livermore National Laboratory, USA, at Tongji University, China, and at the University of Cambridge, UK. He authored the IOP book "Physics of digital photography" and is currently a researcher in the Colour & Imaging Lab at the University of East Anglia, UK.

Graham Finlayson is a Professor in the School of Computing Sciences, University of East Anglia, UK and respectively a visiting and adjunct professor at The Norwegian University for Science and Technology and Simon Fraser University. Professor Finlayson's research spans colour image processing, physics-based computer vision and visual perception. He is interested in taking the creative spark of an idea, developing the underlying theory and algorithms and then implementing and commercialising the technology. He has founded and spun out two companies - Imsense (2006) and Spectral Edge (2011) where he, respectively, served as CTO and CSO - which were acquired by industry majors in 2010 and 2019. Professor Finlayson is a fellow of the Society of Imaging Science and Technology, the Institute for Engineering Technology and the Royal Photographic Society.