https://doi.org/10.2352/lim.2023.4.1.11
©2023 Society for Imaging Science and Technology

Mondrian representation of real world scene statistics

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

Abstract

In material appearance we are interested in objectively measuring a physical aspect of a material, such as reflectance, and also in understanding how we see that material. In perceptual experiments we typically display simple stimuli to an observer, record their response, and then try to build a theory of why the observer responded in a given way. Often the latter model is implemented as a computer algorithm, and many of these are, for example, now implemented in camera pipelines for smartphones. However, the stimuli that are shown to observers are necessarily either very simple, such as rectangular patches of colour, or small in number. This raises the question as to whether the recorded responses to simple stimuli actually shed light on how we perceive scenes in the real world.

In this paper, we look at a specific example of perceptual stimuli: Mondrian images, and investigate the extent to which, in the sense of their autocorrelation matrix, they represent the real world. We show that by modelling paths of pixels through an image using a statistical model that captures the statistics of real Mondrians, the autocorrelation matrix of Mondrians is Toeplitz, and moreover this Toeplitz structure is also found in real images. Although Mondrian images do not contain typical visual cues, our path model can be tuned to replicate the statistics of real images in the autocorrelation sense. The practical utility of this method is that paths through images and their autocorrelation statistics are a key tool for developing algorithms to predict the perceptual response to complex scenes. For example, this approach is at the foundation of retinex image processing. Experiments validate our method.

Introduction

The autocorrelation matrix describes the correlation between all pairs of elements of a vector or vectors that represent the same stimuli. In the domain of imaging, it is widely used to describe the spectral statistics of surface reflectance and illumination spectra [1] and is central to colour characterisation and calibration [2, 3, 4]. Autocorrelation matrices are also used in the context of lightness perception algorithms [5], where the vectors that we autocorrelate describe paths of pixel values through images [6]. Famously, path-based computation is at the foundation of the retinex theory of lightness perception [6, 7, 8]. Retinex postulates that the human visual system (HVS) does not perceive a scene directly as an array of radiance (or luminance) patches from which an array of lightness values is interpreted psychophysically. Instead, it is thought that the HVS is able to determine a reflectance map of the scene, from which the array of lightness values is interpreted. For example, this means that areas such as shadow regions, which receive less illumination from the light source, would not be perceived as darkly as the luminance values would suggest. The paths relate pixels in non-proximate regions of an image.

It is the latter application that is of particular interest in this paper. Our goal is to examine the extent that paths in Mondrian-type images can lead to autocorrelation structures that account for the path autocorrelations in real images. In doing so, we provide a possible bridge linking simple (and necessarily austere) psychophysical experiments (where a limited number of simple images are used) and the real world, which has an unbounded number of possible images [9, 10, 11].

We begin by discussing the formal definition of autocorrelation and we then determine the autocorrelation matrix for a large real image dataset. Subsequently, we develop a statistical model for paths through Mondrian images and examine the extent to which our model generates an autocorrelation matrix that matches the statistics of paths in the real world. It turns out that our Mondrian path model by definition *must* result in a Toeplitz autocorrelation structure. We go on to show that we can tune our Mondrian model using a patch-length parameter, which, along with arguments concerning how autocorrelations are used in regression [5], provides excellent correspondence to the autocorrelation matrix for paths drawn from real images.

Autocorrelation matrix

The autocorrelation matrix for a random $1 \times p$ row vector $s = \begin{bmatrix} s_1 & s_2 & \cdots & s_p \end{bmatrix}$ is a $p \times p$ matrix that describes the correlation between all pairs of elements of the vector. It can be written

$$s^\top s = \begin{bmatrix} s_1^2 & s_1 s_2 & \cdots & s_1 s_p \\ s_2 s_1 & s_2^2 & \cdots & s_2 s_p \\ \vdots & \vdots & \ddots & \vdots \\ s_p s_1 & s_p s_2 & \cdots & s_p^2 \end{bmatrix}, \tag{1}$$

where $\top$ denotes the transpose operator. This differs from the auto-covariance matrix in that the expectation values are not defined relative to the mean. (Note we adopt the convention that, given two vectors v and w, the dot product between the vectors is written as vw. The dot-product of v with itself is written as $v^2 \equiv vv$ since the dot-product of a vector with itself is the squared magnitude of the vector).

If there are multiple data values associated with the vector components, a composite $n \times p$ vector S can be constructed, where n is the number of data values [12]. In this case the autocorrelation matrix becomes

$$S^\top S = \frac{1}{n} \begin{bmatrix} \sum_{k=1}^{n} s_{k1}^2 & \sum_{k=1}^{n} s_{k1} s_{k2} & \cdots & \sum_{k=1}^{n} s_{k1} s_{kp} \\ \sum_{k=1}^{n} s_{k2} s_{k1} & \sum_{k=1}^{n} s_{k2}^2 & \cdots & \sum_{k=1}^{n} s_{k2} s_{kp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} s_{kp} s_{k1} & \sum_{k=1}^{n} s_{kp} s_{k2} & \cdots & \sum_{k=1}^{n} s_{kp}^2 \end{bmatrix}. \tag{2}$$

In the context of spectral reflectance and illumination, autocorrelation matrices are commonly encountered where the data values in the vectors represent data from n different spectral reflectance or illumination curves at p specific wavelengths [1], where p is typically 31, representing sampling points from 400 to 700 nm with a 10 nm step. In this paper we are interested in paths of

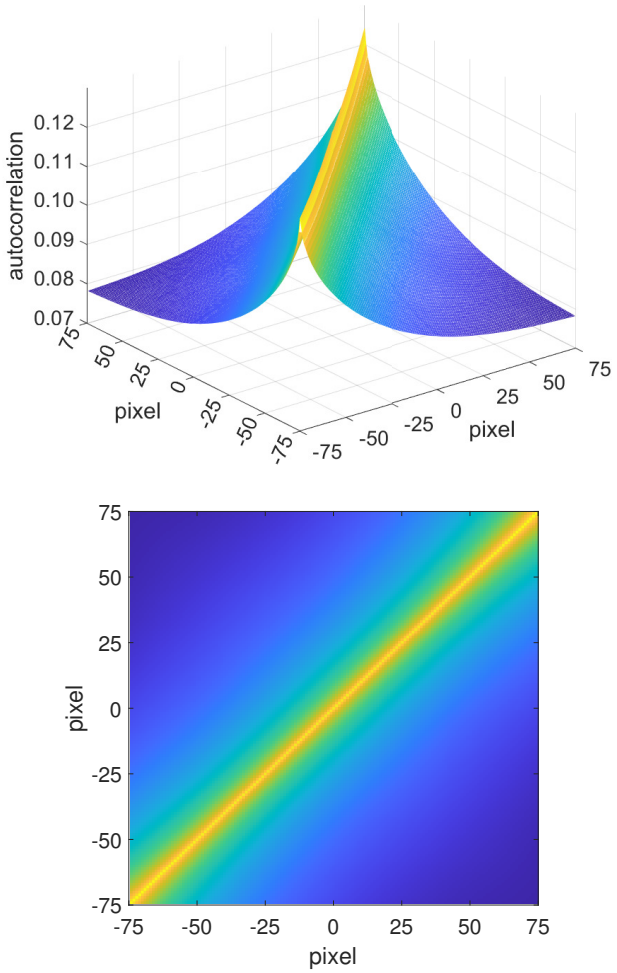

Figure 1. *Autocorrelation matrix for the ImageNet "test" dataset [13] on a 151 × 151 pixel grid plotted in 3d (upper figure) and 2d (lower figure).*

pixel values through images and so each of the n vectors represent strips of pixels of length p [5].

Real world scenes

Fig. 1 shows an autocorrelation matrix calculated for the ImageNet "test" dataset of 100,000 images [13]. The images were processed by first removing the encoding gamma curve before converting to the CIE XYZ colour space with luminance Y normalised to the range $[0, 1]$, which can be approximately correlated with albedo value. A large number of paths were selected through each image and each path was $p = 151$ pixels in length. Both forward and reverse directions were included. The paths chosen were straight lines at uniformly random orientations. These vectors were then used to calculate the autocorrelation matrix via Eq. (2).

The autocorrelation matrix is seen to have an approximately Toeplitz structure. An autocorrelation matrix with a perfect Toeplitz structure is shift invariant [14], which means that all matrix elements remain unchanged as the position is shifted in the direction of the main diagonal. Here this can be understood to indicate that randomness exists on the average. In other words, there is absence of a periodically repeating pattern over long length scales. The off-diagonal elements describe the tendency for like pixel values to cluster together over short length scales. Small datasets or datasets that represent very specific scene categories can show behaviour that deviates away from the perfect Toeplitz structure. This can also happen towards the image borders where illumination patterns may dominate on the average.

An interesting analogy can be made with the high-temperature disordered state in substitutionally-disordered alloys [15], which form a regular lattice structure. At high temperature, there is randomness on the average so that long-range order (LRO) is absent. Nevertheless, short-range order (SRO) will be present, which is the tendency for like or unlike atoms to surround each other over short length scales without affecting the overall symmetry. However, when the temperature is lowered, a phase transition to a LRO state can occur in which the constituent atoms form a pattern that periodically repeats over long length scales.

We now go on to investigate the extent to which autocorrelation matrices for real world scenes can be modelled by paths followed in a Mondrian world.

Mondrian world

Inspired by the abstract grid-based paintings of the Dutch artist Piet Mondrian that first appeared in the early 1920s, Mondrian images consist of random arrangements of rectangular patches of various sizes [6] and are widely used in visual experiments [16, 17, 18, 19, 20].

Figure 2 shows a randomly generated Mondrian image. This was constructed by using a similar approach to that of Ref. [21] where any given patch on the 2d Mondrian grid has a specific width and height, which were randomly drawn from a uniform probability distribution with minimum and maximum width and height limits. The central coordinates of each of the patches were also randomly drawn from a uniform probability distribution and patches generated in this way were superimposed until all pixels on the grid were filled. Parts of patches that fell outside the grid boundary were cropped. Note that only the maximum patch width and height can ultimately be controlled when this construction method is used. Each patch was assigned a random albedo with values drawn from a uniform probability distribution in the range $[0, 1]$.

Figure 2. *Mondrian image randomly generated from a uniform probability distribution with maximum and minimum patch widths and heights taken to be 71 pixels and 11 pixels, respectively. Arbitrary colours have been used to denote the albedo values.*

Figure 3 shows the autocorrelation matrix calculated using Eq. (2) for a large number of paths of length $p = 151$ pixels taken through a very large number of Mondrian images of dimension $p \times p$. The parameters chosen were the same as those used to produce Fig. 2. The autocorrelation matrix is seen to have a perfect Toeplitz structure with minimum autocorrelation reached when the separation from the main diagonal reaches the maximum patch width and height, which were both taken to be 71 pixels in this example.

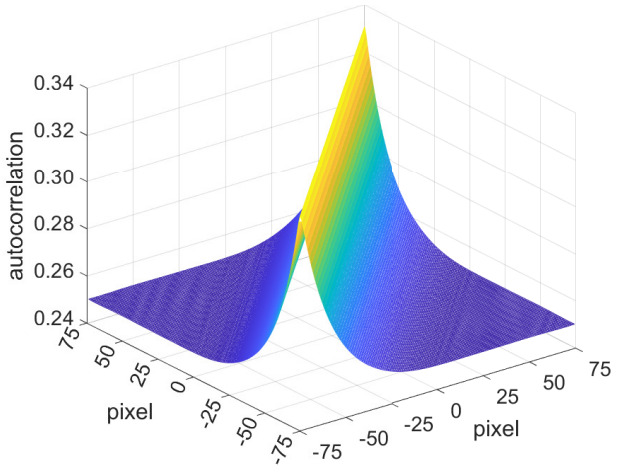

Figure 3. *Autocorrelation matrix for a 151 × 151 pixel grid calculated by taking paths of length p = 151 pixels through a very large number of 2d Mondrian images. A uniform probability distribution was used for the albedo values with maximum and minimum patch widths and heights taken to be 71 pixels and 11 pixels, respectively.*

Mondrian parameterisation

As mentioned above, it is not possible to guarantee a minimum patch width or height due to the construction method used to produce the Mondrian images, which involves the superposition of patches centred at random locations. Nevertheless, after determining the autocorrelation matrix over a range of minimum and maximum patch widths and heights, we found that the shape of the Toeplitz structure and the value at which it falls to its minimum autocorrelation is largely governed by the largest patch diagonal. This leads to the idea that we can effectively parameterise the Mondrian world by using only one dimension.

In 1d, an alternative Mondrian construction method can be used that does not involve the superposition of patches and therefore gives us full control over the probability distribution. The way forward is to introduce direct correlation between adjacent pixels via a "step" parameter α with $0 \leq \alpha \leq 1$ [5]. For a given value s_i located at position i in a path, α describes the probability that the adjacent pixel takes on the same value, $s_{i+1} = s_i$. The probability of a "jump" is then $1-\alpha$, in which case s_{i+1} takes a value in the range $[a,b]$ instead, according to a specified probability distribution. Figure 4 shows the cross-section of such a randomly generated 1d Mondrian image with $\alpha = 0.9$.

For general pixels i, j, the matrix elements of the autocorrelation matrix can now be analytically expressed as follows,

$$\left[S^\top S\right]_{ij} = \int_a^b p(s) s_i \, ds \left(\alpha^{|j-i|} s_i + \left(1-\alpha^{|j-i|}\right) \int_a^b p(s) s_j \, ds \right), \tag{3}$$

where $p(s)$ is the probability density function for the albedo values. Note that $\alpha^{|j-i|} = 1$ if $\alpha = 0$ and $j = i$.

In the case of a uniform probability distribution where the pixel albedo values are continuous random variables that are uniformly assigned in the range $[a,b]$, the probability density function is given by $p(s) = 1/(b-a)$. Substituting into Eq. (3) yields the following expression for the autocorrelation matrix elements,

$$\left[S^\top S\right]_{ij} = \frac{a^2+ab+b^2}{3} \alpha^{|j-i|} + \frac{(a+b)^2}{4} \left(1-\alpha^{|j-i|}\right). \tag{4}$$

Necessarily, if we generate a very large number n of random paths through a Mondrian image according to the prescription described above and then apply Eq. (2), we find that the result converges towards this closed-form solution as n increases. For example, for $\alpha = 0.9$ the mean-squared error falls to 8.1×10^{-7} when n = 100,000. Encouragingly, the Toeplitz structure is approximately unveiled given a small number of paths. The paths available in, say, just 10 randomly generated 2d Mondrians suffices to reveal the closed-form Toeplitz structure.

Importantly, Eq. (4), by definition, is exactly a Toeplitz matrix. That is, our prescription for paths through a Mondrian *must* result in a path autocorrelation that is Toeplitz.

Expected step length

Although the above construction method does not enable a minimum or maximum step length to be specified directly, it is possible to derive an expression for the average or expected step length $\langle n_\mathrm{p} \rangle$ in terms of α.

Consider a pixel i with a randomly generated value. The probability that pixel i is not extended to a longer step is $1-\alpha$, in which case $n_\mathrm{p} = 1$. The probability that the pixel is only extended by one pixel is $\alpha(1-\alpha)$, in which case $n_\mathrm{p} = 2$. Continuing this argument, if pixel j is located n pixels away from i, the probability that pixel i is extended to a step with total length $n_\mathrm{p} = k$ is $\alpha^{k-1}(1-\alpha)$ if $k < n$ and α^{k-1} if $k = n$. Therefore

$$\langle n_\mathrm{p} \rangle = \left((1-\alpha) \sum_{k=1}^{n-1} k\alpha^{k-1} \right) + n\alpha^{n-1} = \frac{1-\alpha^n}{1-\alpha}. \tag{5}$$

Finite step lengths are obtained provided that $0 \leq \alpha < 1$, in which case taking the limit $n \to \infty$ yields the result

$$\langle n_\mathrm{p} \rangle = \frac{1}{1-\alpha}. \tag{6}$$

The example 1d Mondrian plotted in Fig. 4 with $\alpha = 0.9$ corresponds to an average step length $\langle n_\mathrm{p} \rangle = 10$.

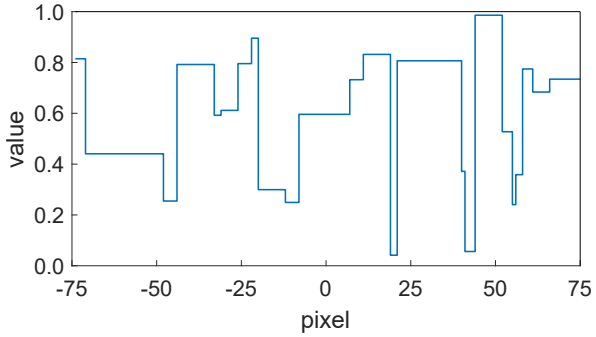

Figure 4. *Cross-section of a 1d Mondrian vector of length p = 151 pixels randomly generated from a uniform probability distribution with albedo pixel values in the range $[0,1]$. Here α = 0.9 and $\langle n_\mathrm{p} \rangle$ = 10.*

Discussion

It is instructive to first analyse the special case where all 1d Mondrian vectors are equally likely to occur. This means that a uniform probability distribution must be used with α set to zero so that correlation between pixels is absent. In this case Eq. (4) reduces to

$$\left[S^\top S\right]_{ij} = \frac{1}{3}\delta_{ij} + \frac{1}{4}\left(1-\delta_{ij}\right), \tag{7}$$

where all pixel values along a path have been normalized to the range $[0,1]$. As shown in Fig. 5, maximum autocorrelation is

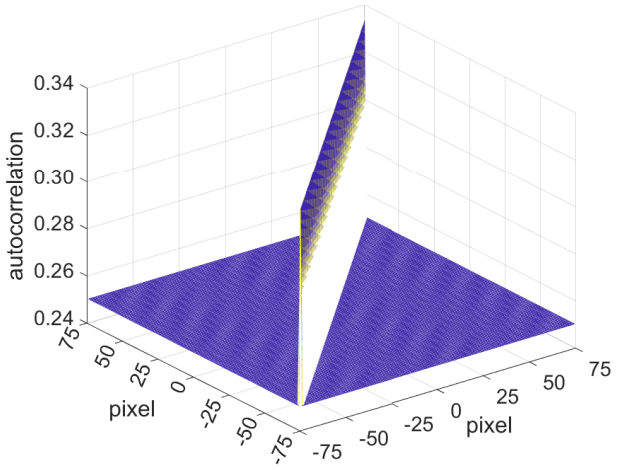

Figure 5. *Autocorrelation matrix assuming Maximum Ignorance with Positivity [2]. Maximum autocorrelation is seen along the main diagonal with a value 1/3 and minimum autocorrelation everywhere else with a value 1/4.*

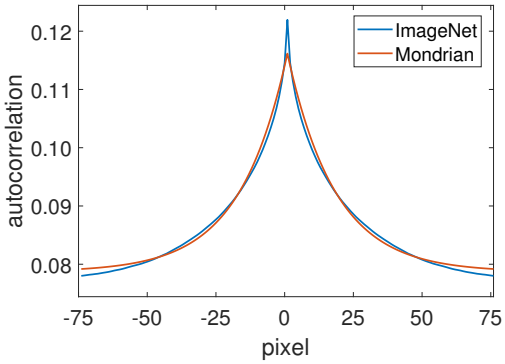

Figure 6. *Cross-section of the autocorrelation matrix for the ImageNet dataset [13] shown in Fig. 1 and the fitted Mondrian autocorrelation matrix. Here the step parameter α = 0.9717. The scale and offset are 0.4497 and -0.0338, respectively.*

seen along the main diagonal and minimum autocorrelation is seen everywhere else. In fact, this corresponds to the so-called Maximum Ignorance with Positivity assumption [2] that was made in connection with camera characterisation (colour correction), where the vectors are all possible spectral colour signals with positive power.

As the Mondrian step parameter α is increased, the autocorrelation matrix gains a Toeplitz structure analogous to that seen in Fig. 3 for the 2d case. Increasing α or the expected step length $\langle n_{\mathrm{p}} \rangle$ increases the separation at which the autocorrelation drops to its minimum value of 1/4 with distance away from the main diagonal. Note that in the context of camera characterisation, this Toeplitz structure was arbitrarily modelled in Ref. [3] by using a Cauchy function.

However, the ImageNet dataset used to produce the real world autocorrelation matrix illustrated in Fig. 1 contains linearised pixel values in the full range $[0,1]$ but has lower minimum and maximum autocorrelation. This can be attributed to the fact that if a uniform probability distribution is used, the mean pixel value is given by $\mu = (a+b)/2$, which is unlikely to correspond to the real world since natural scenes do not necessarily have an average albedo or relative luminance that is midway between the upper and lower limits. In fact, standard exposure strategy in photography assumes that the average luminance for typical scenes is approximately 18% of the maximum [22].

To proceed, let us allow the autocorrelation to be linearly altered in the following manner: $S^{\top}S \rightarrow mS^{\top}S + c$. We wish to take our Mondrian autocorrelation matrix and fit it to the real-world ImageNet dataset autocorrelation matrix shown in Fig. (1). We use linear regression to solve

$$\begin{bmatrix} A_{11} & 1 \\ A_{12} & 1 \\ \vdots & \vdots \\ A_{pp} & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{12} \\ \vdots \\ B_{pp} \end{bmatrix}, \tag{8}$$

where $\{A_{ij}\}$ and $\{B_{ij}\}$ are the columns of the $p \times p$ Mondrian autocorrelation matrix and the $p \times p$ ImageNet autocorrelation matrix, respectively, concatenated as single vectors. Since A depends on the step parameter α via Eq. (4), the α value that yields the minimum mean-squared error (MSE) between A and B provides the optimum fit.

We obtained the following optimised values: step parameter $\alpha = 0.9717$, scale $m = 0.4497$, and offset $c = -0.0338$, which correspond to a minimum mean-squared error MSE = 9.7×10^{-7}. This α value corresponds to an average step length $\langle n_{\mathrm{p}} \rangle = 35.3$ pixels. Fig. 6 shows a cross-section of the main diagonal of the ImageNet and fitted Mondrian autocorrelation matrices. Since the ImageNet matrix varies slightly along its main diagonal as it is not perfectly Toeplitz, the slice along the main anti-diagonal was chosen for illustrative purposes.

Of course, while the fit is pleasingly close, the reader might wonder whether we can scale autocorrelations and add an offset. What does this mean? Well, it does make sense in the context of regression. For example, Ref. [5] attempts to find an *operator* that removes shading variations by regressing shading-confounded image data to shading-free ground-truth data. By appealing to Ref. [23], we state (without formal proof) that if we additionally require paths of uniform pixel values (such as a single-coloured image) to be mapped without error, then the autocorrelations can be scaled in the manner we propose.

An alternate approach that we are investigating is to use a skewed normal distribution truncated to the range $[a,b]$, which would, we believe, lead to us to be able to match the real world image autocorrelation structure and provide the prescription to generate the corresponding Mondrians that have the same autocorrelation properties. Finally, we remark that our model could be applied to replicate the autocorrelation statistics of different classes of scenes, each of which may have a different average albedo or relative luminance and different expected step length.

Conclusion

In this paper we calculated the autocorrelation matrices for paths of pixels in images drawn from a very large real world scene dataset and paths drawn from the Mondrian world and found that they share a common Toeplitz structure. Furthermore, we showed how we can generate Mondrians so that their autocorrelations are in close approximation to those of the real world. The importance of this lies in part in algorithms that are based on signal autocorrelations. In the context of human vision, lightness perception is often explored from an image-path viewpoint, and, in one study [5], the optimal lightness operator is based entirely on the autocorrelations. In conclusion, we have the surprising yet powerful result that optimising the operator for Mondrians will work equally well for images in the real world. More speculatively, we believe that our result helps to strengthen the argument that small numbers of psychophysical image stimuli might deliver results that work in the real world in general.

Acknowledgments

This research is funded in part by EPSRC grant number EP/S028730/1.

References

[1] M S Drew and G D Finlayson, Analytic solution for separating spectra into illumination and surface reflectance components, J. Opt. Soc. Am. A 24(2), 294-303 (2007).

[2] G D Finlayson and M S Drew, The Maximum Ignorance Assumption with Positivity, Proc. IS&T/SID Fourth Color Imaging Conference: Color Science, Systems and Applications, pgs. 202-205 (1996).

[3] J A S Viggiano, Minimal-Knowledge Assumptions in Digital Still Camera Characterization I: Uniform Distribution, Toeplitz Correlation, Proc. IS&T/SID 9th Color Imaging Conference: Color Science and Engineering: Systems, Technologies, Applications, pgs. 332-336 (2001).

[4] G D Finlayson, J Vazquez-Corral and Fufu Fang, The Discrete Cosine Maximum Ignorance Assumption, Proc. IS&T 29th Color and Imaging Conference, pgs. 14-18 (2021).

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

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

[7] E H Land, Recent advances in retinex theory, Vis. Res. 26, 7–21 (1986).

[8] 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).

[9] D Field, Relations between the statistics of natural images and the response properties of cortical cells, J. Opt. Soc. Am. A 4, 2379-2394 (1987).

[10] J Portilla and E P Simoncelli, A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients, International Journal of Computer Vision 40, 49–70 (2000).

[11] J Koenderink, M Valsecchi, A van Doorn, J Wagemans, and K Gegenfurtner, Eidolons: Novel stimuli for vision research, Journal of Vision 17(2), 7 (2017).

[12] J A Gubner, Probability and Random Processes for Electrical and Computer Engineers, Cambridge University Press, Cambridge, UK, 2006, pg. 337.

[13] J Deng, W Dong, R Socher, L-J Li, Kai Li and L Fei-Fei, ImageNet: A large-scale hierarchical image database, 2009 IEEE Conference on Computer Vision and Pattern Recognition, pgs. 248-255 (2009).

[14] B N Mukherjee ans S S Maiti, On some properties of positive definite Toeplitz matrices and their possible applications, Linear Algebra and its Applications 102, 211-240 (1988).

[15] F Ducastelle, Order and Phase Stability in Alloys, North-Holland, 1991.

[16] 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).

[17] 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).

[18] J J McCann, Lessons Learned from Mondrians Applied to Real Images and Color Gamuts, Proc. IS&T/SID Seventh Color Imaging Conference, pgs. 1-8 (1999).

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

[20] S Han, D Alais, and R Blake, Battle of the Mondrians: Investigating the Role of Unpredictability in Continuous Flash Suppression, i-Perception 9(4), 1–21 (2018).

[21] Y-T Lin and G D Finlayson, An investigation on worst-case spectral reconstruction from RGB images via Radiance Mondrian World assumption, Color Res Appl. 48(2), 230-242 (2023).

[22] D Connelly, Calibration Levels of Films and Exposure Devices, The Journal of Photographic Science 16(5), 185-193 (1968).

[23] G D Finlayson and M S Drew, White-Point Preservation Enforces Positivity, Proc. IS&T/SID Sixth Color Imaging Conference: Color Science, Systems and Applications, pgs. 47-52 (1998).

Author Biography

D. Andrew Rowlands received his PhD in Physics from the University of Warwick, U.K. He has held postdoctoral positions at University of Bristol, U.K., Lawrence Livermore National Laboratory, U.S., Tongji University, China, and University of Cambridge, U.K. 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, 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.