Perception, 2004, volume 33, pages 1463 – 1473

DOI:10.1068/p5321

# A biologically inspired algorithm for the recovery of shading and reflectance images

Adriana Olmos, Frederick A A Kingdom

McGill Vision Research, Department of Ophthalmology, McGill University, 687 Pine Avenue West, Room H4-14, Montréal, Québec H3A 1A1, Canada; e-mail: adriana.olmos@mcgill.ca

Received 3 February 2004, in revised form 20 April 2004; published online 8 December 2004

**Abstract.** We present an algorithm for separating the shading and reflectance images of photographed natural scenes. The algorithm exploits the constraint that in natural scenes chromatic and luminance variations that are co-aligned mainly arise from changes in surface reflectance, whereas near-pure luminance variations mainly arise from shading and shadows. The novel aspect of the algorithm is the initial separation of the image into luminance and chromatic image planes that correspond to the luminance, red – green, and blue – yellow channels of the primate visual system. The red – green and blue – yellow image planes are analysed to provide a map of the changes in surface reflectance, which is then used to separate the reflectance from shading changes in both the luminance and chromatic image planes. The final reflectance image is obtained by reconstructing the chromatic and luminance-reflectance-change maps, while the shading image is obtained by subtracting the reconstructed luminance-reflectance image from the original luminance image. A number of image examples are included to illustrate the successes and limitations of the algorithm.

## 1 Introduction

The idea of decomposing an image into its reflectance and shading components, a process referred to as 'intrinsic image segmentation', was first introduced by Barrow and Tenenbaum (1978). Figure 1 illustrates the principle. In simple mathematical terms an image, $I(x, y)$, can be considered as the product of a reflectance image, $R(x, y)$, and a shading image, $S(x, y)$:

$$I(x, y) = R(x, y)S(x, y) \; . \tag{1}$$

The term reflectance, as used here, includes both spectral as well as intensive reflectance, corresponding respectively to the perceptual dimensions of lightness and colour. The term shading, as used here, encompasses two forms of spatially non-uniform illumination: the first, shading, which is, strictly speaking, the variation in luminance caused by a change in the angle of a surface with respect to the direction of illumination; the second, shadows (or 'cast' shadows), formed by occlusion of the light source. Although others have employed terms such as 'illumination image' to describe the pattern of non-uniform illumination in an image (Barrow and Tenenbaum 1978), we will use the term 'shading' because it is less ambiguous than the term 'illumination', which is used not only to refer to spatially non-uniform illumination but also to the light source.

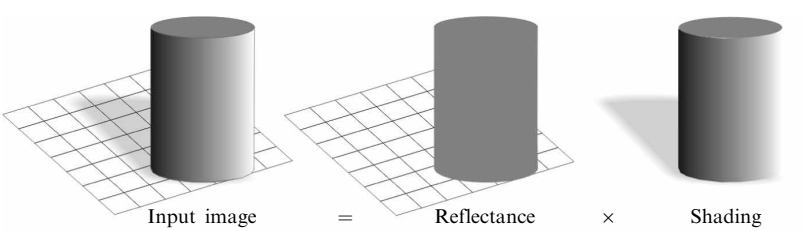


**Figure 1.** Decomposition of an image into its reflectance and shading components.

The separation of an image into its shading and reflectance images is an ongoing research topic in both human and computer vision. In human vision, the ability to discriminate shading from reflectance variation is important for a variety of visual tasks, perhaps the most obvious being the identification of the lightness of surfaces that lie in shadow—ie the achievement of lightness constancy (Gilchrist 1979; Gilchrist et al 1983; Arend 1994). Discriminating reflectance from shading is also necessary for shape-from-shading (Lehky and Sejnowski 1988; Ramachandran 1988; Bergström 1994; Sun and Perona 1997), recognising objects from shadows (Cavanagh and Leclerc 1989), and identifying the direction of object motion from shadow trajectories (Knill et al 1996).

The first computer vision algorithm aimed at separating reflectance from shading was the retinex model of Land and McCann (1971). The goal of the retinex was to recover the reflectance of surfaces in the context of gradual illumination gradients, and to this end it was successful. A problem with the retinex, however, as noted by a number of researchers (eg Arend 1994; Finlayson et al 2002b), is that illumination gradients are not always gradual; shadows often have sharp edges, and the retinex would tend to recover these as reflectance changes. Finlayson et al (2002b) have recently developed a modified version of the retinex that avoids this problem by allowing the user to select the luminance gradients to be removed depending on whether they are shading or reflectance gradients. Other types of prior knowledge, for example of the illumination source and the surface composition of the scene, have also been exploited in reflectance–shading separation algorithms (Gevers and Stockman 2000). Yet another approach has been to use multiple images of a scene photographed under different lighting conditions, with the reflectance map being the component common to all images (Weiss 2001).

Most recently, the computer vision community has exploited colour information for reflectance–shading separation (Finlayson et al 2002a, 2002b; Tappen et al 2003). The basic idea is as follows. Shading, unlike reflectance, is almost exclusively luminance-defined; although sometimes tinged with colour (for example blue when formed in sunlight), shading tends to have minimal colour contrast (Párraga et al 2002). Objects, on the other hand, tend to vary in spectral as well as intensive reflectance, that is they are both luminance- and colour-defined. These relationships are illustrated in figure 3, which shows a shadow falling across a grass/pavement border. The shadow is primarily a change in luminance (bright to dark), whereas the grass/pavement border is a change in both colour (green to grey) and luminance (dark to light). Colour is therefore a potential cue for helping disambiguate shading from reflectance variations through the following rules: luminance variations that are accompanied by colour variations are variations in reflectance; luminance variations that are unaccompanied by colour variations are variations in illumination (Rubin and Richards 1982; Cavanagh 1991; Mullen and Kingdom 1991; Fine et al 2003; Tappen et al 2003). Recently one of us has provided compelling evidence from the study of shape-from-shading that the human visual system has in-built knowledge of these rules (Kingdom 2003).

The principle behind the computer vision algorithms that exploit colour for shading–reflectance separation is to classify image discontinuities based on an analysis of the relationships between corresponding points in the $R, G, B$ planes of an image.[1] Specifically, if the ratio between pixel values on either side of a discontinuity is near-equal in all three $R, G, B$ planes, the discontinuity is classified as shading; and if not, reflectance. Once classified, the discontinuities are preserved as derivatives, and the resulting classified-derivative maps reintegrated to produce the shading and reflectance images. These algorithms are undoubtedly the most successful to date.

[1] The $R, G, B$ planes of an image result from filtering the image through three overlapping spectral sensitivity functions whose peaks roughly correspond to the long, middle, and short wavelength parts of the visible spectrum. These might be the red, green, and blue sensors of a digital camera, or the red, green, and blue phosphors of a CRT monitor.

The algorithm presented here, while adopting a similar approach to some of the algorithms described above in terms of constructing reflectance and shading images from derivative maps, differs from previous colour-based algorithms in the way that the colour information is analysed. Rather than basing the classification process on the relationships between the $R, G, B$ planes of an image, we base them on the relationship between image planes that correspond to the luminance, red – green, and blue – yellow channel responses of the primate visual system (De Valois 1965; Macleod and Boynton 1979; Derrington et al 1984; Sankeralli and Mullen 1997). It has been argued that the luminance, red – green, and blue – yellow channels of the primate visual system are an efficient way of coding the intensive and spectral content of natural images (eg see Wandell 1995). Our algorithm uses the fact that shading discontinuities will be present to a significant degree only in the luminance image plane, whereas reflectance changes appear in all three image planes.

Our algorithm represents an alternative method for the quick estimation of shading and reflectance based on the use of colour information. We stress at the outset that we make no claims regarding the superiority of our algorithm compared to its predecessors. Nor do we maintain that our algorithm represents an accurate model of how the primate visual system separates shading from reflectance variations. In particular, the reconstruction of shading and reflectance images from derivative maps via integration may have no counterpart in any biological visual system.[2] Furthermore, colour is only one of a number of cues available for separating reflectance from shading, as evidenced by our ability to identify shading in black-and-white images. Our aim is to explore the feasibility of basing reflectance – shading separation on the colour-opponent channels of primate vision, as these channels are likely to play an important role in the segmentation of 'intrinsic-images' in primates (Kingdom 2003).

## 2 The algorithm

A flow diagram of the algorithm is provided in figure 2. The individual steps are listed below.

*Step 1. Conversion to L, M, S images*

All images were photographed with a Nikon Coolpix 5700 digital camera, whose gammas and $R, G, B$ sensor spectral sensitivity functions had been calibrated. Each $R, G, B$ digital image was gamma-corrected and then converted to an $L, M, S$ digital image, where $L, M, S$ stands for images filtered through the spectral functions of, respectively, the long-wavelength-sensitive, middle-wavelength-sensitive, and short-wavelength-sensitive cones of the human visual system. Note that $R, G, B$ and $L, M, S$ refer to two-dimensional images, and are thus abbreviations for $R(x, y)$, $G(x, y)$, etc, where $x$ and $y$ are pixel coordinates. The abbreviated form *RGB*, *LMS* etc will be used throughout the text.

A brief exposition of the camera calibration process and method for generating the *LMS* images is provided in Appendix 1, while full details can be found in the McGill colour-calibrated image database (http://tabby.vision.mcgill.ca/), where the images shown in figure 4 can also be viewed and downloaded.

[2] The idea that there exists a stage in human vision analogous to integration is controversial. Integration is a defining characteristic of a number of models of lightness processing, of which the retinex (Land and McCann 1971), Gilchrist et al's edge-integration (Gilchrist 1979; Gilchrist et al 1983) and later anchoring model (Gilchrist et al 1999) are probably the best known (see also Arend 1994, and review by Kingdom and Moulden 1988). Integration is arguably implicit in 'filling-in' models of early vision (eg Grossberg and Todorovíc 1988; Pessoa et al 1998). However, a number of recent models of brightness/lightness perception do not employ an integration stage (Blakeslee and McCourt 1999, 2004; Dakin and Bex 2004). In these models, relatively coarse-scale filters instantiate 'filling-in' at a relatively early stage in visual processing.

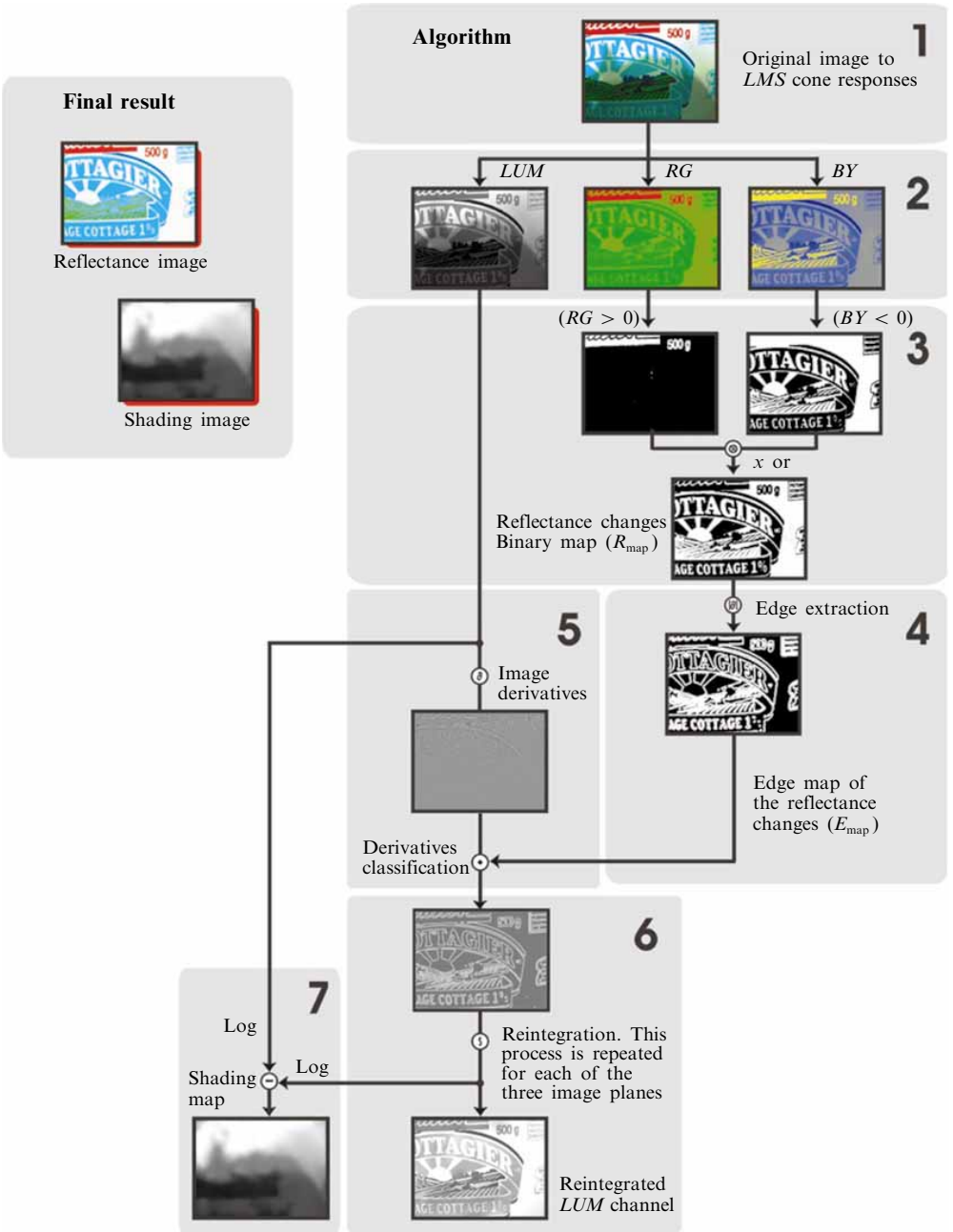


**Figure 2.** Flow diagram of the algorithm. 1—*LMS* image conversion; 2—combination of *LMS* images into *LUM*, *RG*, and *BY* image planes; 3—binarisation of the *RG* and *BY* images; 4—location of reflectance edges; 5—classification of reflectance derivatives; 6—reintegration of the reflectance image; 7—extraction of the shading image. Steps 5 to 6 were repeated for each *LUM*, *RG*, and *BY* image plane.

**Figure 3.** Shadow falling across the grass/pavement border. The shadow is primarily a change in luminance (bright to dark), whereas the grass/pavement border is a change in both colour (green to grey) and luminance (dark to light).


**Figure 4.** Examples from the algorithm: (a) face, (b) shadow of a leaf projected onto the ground, (c) close-up of a plastic container with a shadow projected onto it, (d) two hands projecting a shadow onto the pavement in a parking lot.

*Step 2. Combining the L, M, S images*

Conventionally, the three post-receptoral channel responses of the primate visual system are modelled in terms of combinations of three cone-contrast values: $\Delta L/L$, $\Delta M/M$, and $\Delta S/S$. For example, the red–green channel is typically modelled as $\Delta L/L - \Delta M/M$ (Norlander and Koenderink 1983; Stromeyer et al 1985; Cole et al 1993; Sankeralli and Mullen 1997). The denominator in each cone-contrast term represents the level of cone adaptation, and is typically calculated as the cone response to the DC of the stimulus. While arguably appropriate for simple stimuli such as gratings or noise patterns presented briefly, where one can assume a uniform level of adaptation, cone-contrast definitions with fixed denominators based on the DC level of the image are inappropriate for natural scenes where adaptation levels are likely to vary across the image. In particular, chromatic image planes generated with cone-contrast definitions with fixed denominators will tend to pick up shadows, even if those shadows are devoid of colour contrast in the sense that they produce equal ratio changes in all three cone responses across space. Considerations such as these have led us to use the following 'shadow-invariant', pixel-based definitions of colour-opponent responses as used by Párraga et al (1998):

$$\begin{aligned} LUM &= L + M\,, \\ RG &= \frac{L - M}{LUM}\,, \\ BY &= \frac{S - \frac{1}{2}LUM}{S + \frac{1}{2}LUM}\,, \end{aligned} \tag{2}$$

where $L$, $M$, and $S$ are the input cone-filtered images, and $RG$, $BY$, and $LUM$ are the output red–green chromatic, blue–yellow chromatic, and luminance image planes.

*Step 3. Binarisation of the chromatic images*

In this stage we combined the information from the $RG$ and $BY$ images as a first step towards establishing the location of reflectance changes. We first determined the range of values ($r$) in the $BY$ image. The reason for obtaining $r$ from the $BY$ image plane is because the $BY$ image is the more likely of the two chromatic images to pick up shading (for instance, blue shadows due to sky light). $r$ was calculated as follows:

$$r = \max(BY) - \min(BY)\,, \tag{3}$$

where $\min(BY)$ and $\max(BY)$ stand for the minima and maxima values of the $BY$ image plane. The $RG$ and $BY$ images were then combined to produce a single image termed $R_{\text{map}}$ as follows:

$$\begin{aligned} &\text{if } (a \leqslant r < b) \quad \{R_{\text{map}} = (RG > 0)x \quad \text{or} \quad (BY < 0)\} \\ &\text{otherwise} \qquad\quad \{R_{\text{map}} = (RG > T_{RG}) \quad \text{and} \quad (BY < T_{BY})\}\,, \end{aligned} \tag{4}$$

where the threshold values of $a$ and $b$ were set to 0.8 and 1.6, respectively. These values were found as the min and max average values of 25 $BY$ image planes taken from different images. $T_{RG}$ and $T_{BY}$ are values computed as $T_{RG} = \overline{RG} - \sigma_{RG}$ and $T_{BY} = \overline{BY} - \sigma_{BY}$. $\overline{RG}$ and $\overline{BY}$ are the means of the individual chromatic image planes analysed; $\sigma_{RG}$ and $\sigma_{BY}$ are their standard deviations.

*Step 4. Locating reflectance edges*

The boundaries in the binary $R_{\text{map}}$, which define the locations of the reflectance changes, were then identified with the use of the Canny edge-detector algorithm (Canny 1986). The resulting edge-map is termed $E_{\text{map}}$. Since the actual edge-detector algorithm employed is not critical, the details of implementation are given in Appendix 1.

The choice of the Canny edge-detector is somewhat arbitrary, and other edge-detectors such as the Sobel (Jain et al 1995), would probably work as well.

*Step 5. Classification of reflectance derivatives*
The purpose of this stage was to compute and classify the image derivatives (or local changes) of the *RG*, *BY*, and *LUM* image planes. The simplest discrete approximation to compute the derivatives within an image along the $x$ and/or $y$ coordinates is the difference between two adjacent pixels. This was implemented by convolving the image with a filter mask:

$$D_x = I * f_x \quad \text{and} \quad D_y = I * f_y\,, \tag{5}$$

where $I$ could be any of our image planes, * denotes convolution, $D_x$ and $D_y$ are the resulting image derivatives and the filter masks are defined as $f_x = [1,\ 0,\ -1]$ and $f_y = [1,\ 0,\ -1]'$.

Once the image derivatives of each image plane have been computed, they are classified according to the reflectance edge-map ($E_{\text{map}}$). Specifically, all derivatives in the *RG*, *BY*, and *LUM* image planes not present in $E_{\text{map}}$ are removed by setting them to zero. The effect of this operation will primarily be to remove all the derivatives from the *LUM* image that are due to shading.

*Step 6. Reintegration of reflectance image*
In order to obtain the reflectance image it was necessary to reintegrate each of the classified *RG*, *BY*, and *LUM* derivative planes (Finlayson et al 2002a). We employed the technique described by Weiss (2001), which involves finding the pseudo-inverse of an over-constrained system of derivatives. Briefly, if $f_x$ and $f_y$ are the filters used to compute the derivatives in the $x$ and $y$ directions, and $LUM_x$ and $LUM_y$ are the classified reflectance derivatives of the *LUM* image plane, the reconstructed reflectance image $LUM_r(x, y)$ is given by:

$$LUM_r(x,y) = g * \{[f_x(-x,-y) * LUM_x] + [f_y(-x,-y) * LUM_y]\}\,, \tag{6}$$

where * denotes convolution, $f_x(-x,-y)$ is a reversed copy of $f_x(x,y)$, and $g$ is the solution of

$$g * \{[f_x(-x,-y) * LUM_x] + [f_y(-x,-y) * LUM_y]\} = \delta. \tag{7}$$

The reflectance image, derived from the *RG* and *BY* image planes is found in the same way. An interesting property of this technique is that the computation can be performed with a fast Fourier transform (FFT), which speeds up computation. More details about this technique can be found at http://www.cs.huji.ac.il/∼yweiss/.

*Step 7. Obtaining the shading image*
The shading image was obtained in one of two ways: (a) by subtracting, in the logarithm domain, the original luminance image plane from the reconstructed luminance-reflectance image plane, or (b) by reintegrating the *LUM* image derivatives that were not classified as reflectance changes in step 5.

Finally, the full spectral and intensive reflectance image was obtained by transforming the *LUM*, *RG*, and *BY* reconstructed image planes back into the *LMS* image by using the following equations:

$$\begin{aligned} S &= -\frac{LUM(1+BY)}{2(BY-1)}\,,\\ M &= \frac{LUM(1-RG)}{2}\,,\\ L &= LUM - M\,. \end{aligned} \tag{8}$$

Notice that the *LMS* images shown in the paper were converted to *RGB* images for display.

## 3 Results

Example results are shown in figure 4. In each figure, most of the shading and reflectance patterns have been successfully separated into two images. There are notable failures, however. For example, in figure 4a, part of the frame of the glasses appears incorrectly in the shading image, and some of the shading below the chin is present in the reflectance image. That the algorithm is not robust when dealing with very strong cast shadows is also revealed in figure 4d, where part of the shadow appears in the reflectance image. Nevertheless, the examples demonstrate that the algorithm results in reasonable reflectance – shading separation.

The algorithm described here is based on the postreceptoral channels of the primate visual system, that is through the use of *RG*, *BY*, *LUM* image planes formed from the combination of *L*, *M*, *S* cone-filtered images. We have found, however, that generating the *RG*, *BY*, *LUM* image planes from the *R*, *G*, *B* gamma-corrected camera responses appears to work just as effectively, suggesting that the precise shape of the spectral functions underlying the initial transformation of the image is not critical. It remains to be determined if there are any advantages, or disadvantages, of basing reflectance – shading separation on an analysis of *RG*, *BY*, *LUM* image planes, as opposed to the *R*, *G*, *B* image planes of previous colour-based algorithms (Finlayson et al 2002a, 2002b; Tappen et al 2003).

## 4 Future work

There are a number of avenues that could be explored to improve the performance of the algorithm. For example, we have not optimised the relative weights given to the *RG* and *BY* image planes in the stage that identifies reflectance changes (step 5). Optimum weights could be obtained via an examination of the magnitudes and types of colour contrast observed across the range of shadow/shading boundaries found in natural scenes.

One of the main motivations for the use of colour in our algorithm is that the human visual system appears to exploit colour information for intrinsic-image segmentation. However the evidence for this comes from studies that use relatively simple stimuli, such as sine-wave plaids or Mondrian-like patterns (Kingdom and Kasrai 2002; Kingdom 2003; Kingdom et al 2004). To our knowledge, there is as yet no psychophysical evidence that colour information is used by humans for intrinsic-image segmentation in natural scenes. One possible way to test whether colour is used for this purpose would be to employ algorithms such as the one described here to completely, or partially, remove shading from natural colour images, and test whether observers are better able to discriminate the with-shading from the without-shading images when they are in colour compared to black-and-white. It might also be interesting to compare normal observers with colour-deficient observers on such tasks.

## 5 Conclusion

We have described an algorithm that separates the reflectance and shading components of natural images. The algorithm exploits colour information represented in a form corresponding to the response of the postreceptoral chromatic and luminance channels of the primate visual system. The algorithm is reasonably successful, but in its present form is not robust to strong cast shadows. Reflectance – shading separation algorithms may have a practical use in exploring intrinsic-image segmentation in biological vision systems.

**Acknowledgments.** We would like to thank the following persons: Alejandro Párraga and Tom Trościanko at Bristol University for their help with the camera calibration; Nilima Nigam from McGill University for her advice on the reconstruction of boundaries; and Mark Drew from Simon Fraser University and William Freeman from Massachusetts Institute of Technology for helpful comments. This research is supported by the Canadian Institute of Health Research grant MOP-11554 given to Fred Kingdom.

**References**

Arend L E, 1994 "Surface colors, illumination, and surface geometry: Intrinsic-image models of human color perception", in *Lightness, Brightness and Transparency* Ed. A L Gilchrist (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 159 – 213

Barrow H G, Tenenbaum J, 1978 "Recovering intrinsic scene characteristics from images", in *Computer Vision Systems* Eds A R Hanson, E M Riseman (Orlando, FL: Academic Press) pp 3 – 26

Bergström S S, 1994 "Color constancy: Arguments for a vector model for the perception of illumination, color, and depth", in *Lightness, Brightness and Transparency* Ed. A L Gilchrist (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 257 – 286

Blakeslee B, McCourt M, 1999 "A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction" *Vision Research* **39** 4361 – 4377

Blakeslee B, McCourt M, 2004 "A unified theory of brightness contrast and assimilation incorporating oriented multiscale filtering and contrast normalization" *Vision Research* in press

Canny J, 1986 "A computational approach to edge detection" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** 679 – 689

Cavanagh P, 1991 "Vision at equiluminance", in *Limits of Vision* Eds J J Kulikowski, I J Murray, V Walsh, volume 5 of *Vision and Visual Dysfunction* Ed. J Cronly-Dillon (Boca Raton, FL: CRC Press) pp 234 – 250

Cavanagh P, Leclerc Y, 1989 "Shape from shadows" *Journal of Experimental Psychology: Human Perception and Performance* **15** 3 – 27

Cole G R, Hine T, McIlhagga W, 1993 "Detection mechanisms in L-, M-, and S-cone contrast space" *Journal of the Optical Society of America A* **10** 38 – 51

Dakin S C, Bex P J, 2004 "Natural image statistics explain brightness filling-in" *Proceedings of the Royal Society of London, Series B*, in press

Derrington A M, Krauskopf J, Lennie P, 1984 "Chromatic mechanisms in lateral geniculate nucleus of macaque" *Journal of Physiology* **357** 241 – 265

De Valois R L, 1965 "Analysis and coding of colour vision in the primate visual system" *Cold Spring Harbor Symposia on Quantitative Biology* **30** 567 – 579

Fine I, MacLeod D I A, Boynton G M, 2003 "Surface segmentation based on the luminance and colour statistics of natural scenes" *Journal of the Optical Society of America A* **20** 1283 – 1291

Finlayson G D, Hordley S D, Drew M S, 2002a "Removing shadows from images" *European Conference on Computer Vision, Copenhagen, ECCV'02 Proceedings* part 4, pp 823 – 836

Finlayson G D, Hordley S D, Drew M S, 2002b "Removing shadows from images using retinex" IS&T/SID *Tenth Colour Imagining Conference* pp 73 – 79

Gevers T, Stockman H M G, 2000 "Classifying colour transitions into shadow-geometry, illumination highlight or material edges" *IEEE International Conference on Image Processing (ICIP), Vancouver, Canada, 2000. Proceedings* part I, pp 521 – 525

Gilchrist A L, 1979 "The perception of surface blacks and whites" *Scientific American* **240**(3) 112 – 123

Gilchrist A L, Delman S, Jacobsen A, 1983 "The classification and integration of edges as critical to the perception of reflectance and illumination" *Perception & Psychophysics* **33** 425 – 436

Gilchrist A, Kossyfidis C, Bonato F, Agostini T, Cataliotti J, Li X, Spehar B, Annan V, Economou E, 1999 "An anchoring theory of lightness perception" *Psychological Review* **106** 795 – 834

Grossberg S, Todorovic D, 1988 "Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena" *Perception & Psychophysics* **43** 241 – 277

Jain R, Kasturi R, Schunck B G, 1995 *Machine Vision* Computer Science Series (McGraw-Hill International Editions)

Kingdom F A A, 2003 "Colour brings relief to human vision" *Nature Neuroscience* **6** 641 – 644

Kingdom F A A, Beauce C, Hunter L, 2004 "Colour vision brings clarity to shadows" *Perception* **33** in press

Kingdom F A A, Kasrai R, 2002 "Colour vision facilitates intrinsic-image segmentation" *Perception* **31** Supplement, 12, abstract

Kingdom F, Moulden B, 1988 "Border effects on brightness: A review of findings, models and issues" *Spatial Vision* **3** 225 – 262

Knill D C, Kersten D, Mamassian P, 1996 "Implications of a Bayesian formulation of visual information for processing for psychophysics", in *Perception as Bayesian Inference* Eds D C Knill, W Richards (Cambridge: Cambridge University Press) pp 239 – 286

Land E H, McCann J J, 1971 "Lightness and retinex theory" *Journal of the Optical Society of America* **61** 1 – 11

Lehky S R, Sejnowski T J, 1988 "Network model of shape-from-shading: neural function arises from both receptive and projective fields" *Nature* **333** 452 – 454

MacLeod D I A, Boynton R M, 1979 "Chromaticity diagram showing cone excitation stimuli by stimuli of equal luminance" *Journal of the Optical Society of America* **69** 1183 – 1186

Mullen K T, Kingdom F A A, 1991 "Colour contrast in form perception", in *The Perception of Colour* Ed. P Gouras, volume 6 of *Vision and Visual Dysfunction* Ed. J Cronly-Dillon (London: Macmillan) pp 198 – 217

Norlander C, Koenderink J J, 1983 "Spatial and temporal discrimination ellipsoids in color space" *Journal of the Optical Society of America* **73** 1533 – 1543

Párraga C A, Brelstaf G, Trościanko T, Moorehead I R, 1998 "Color and luminance information in natural scenes" *Journal of the Optical Society of America A* **15** 563 – 569

Párraga C A, Trościanko T, Tolhurst D J, 2002 "Spatiochromatic properties of natural images and human vision" *Current Biology* **12** 483 – 487

Pessoa L, Thompson E, Noë A, 1998 "Finding out about filling in: a guide to perceptual completion for visual science and the philosophy of perception" *Behavioral and Brain Sciences* **21** 723 – 802

Ramachandran V S, 1988 "Perception of shape from shading" *Nature* **331** 163 – 166

Rubin J M, Richards W A, 1982 "Color vision and image intensities: When are changes material?" *Biological Cybernetics* **45** 215 – 226

Sankeralli M J, Mullen K T, 1997 "Postreceptoral chromatic detection mechanisms revealed by noise masking in three-dimensional cone contrast space" *Journal of the Optical Society of America A* **14** 906 – 915

Smith V C, Pokorny J, 1975 "Spectral sensitivity of the foveal cone photopigments between 400 and 700 nm" *Vision Research* **15** 161 – 171

Stromeyer C F III, Cole G R, Kronauer R E, 1985 "Second-site adaptation in the red – green chromatic pathways" *Vision Research* **25** 219 – 237

Sun J, Perona R, 1997 "Shading and stereo in early perception of shape and reflectance" *Perception* **26** 519 – 529

Tappen M F, Freeman W T, Adelson E H, 2003 "Recovering intrinsic images from a single image", in *Advances in Neural Information Processing Systems* **15** [Neural Information Processing Systems, NIPS 2002, Vancouver, BC, Canada] Eds S Becker, S Thrun, K Obermayer (Cambridge, MA: MIT Press) pp 1343 – 1350 [also available online at http://books.nips.cc/papers/files/nips15/VS13.pdf]

Wandell B A, 1995, chapter 9 in *Foundations of Vision* (Sunderland, MA: Sinauer)

Weiss Y, 2001 "Deriving intrinsic images from image sequences" *Proceedings of the 8th ICCV, Vancouver, 9 – 12 July* 15, volume 2, pp 68 – 76, http://www.cs.huji.ac.il/∼yweiss/

## Appendix 1: Camera calibration and generation of *LMS* images

The spectral sensitivities of the three camera sensors were obtained by taking photographs of a white target (made of cyanoacrylate powder) illuminated by a constant-current light source, through a set of narrow-band optical interference filters spanning the range from 400 to 700 nm. The radiance of each filtered image was measured with a Topcon SR-1 spectroradiometer. The set of radiance measures and their associated camera digital values was then used to gamma-correct each image. Finally, each image was converted from an $R, G, B$-sensor to $L, M, S$-cone representation [with the Smith and Pokorny (1975) cone spectral sensitivity functions] by using a conventional linear matrix transform [eg as described in Wandell (1995), Appendix 2].

## Appendix 2: Details of Canny (1986) edge-detector algorithm

The various steps of the Canny algorithm are as follows:

*Image smoothing.* The image was first smoothed by a circularly symmetric Gaussian kernel (of standard deviation $\sigma = 1$ pixel for our case):

$$S(x, y) = G(x, y, \sigma) * I(x, y)\,,$$

where $I(x, y)$ is the image, in this paper $R_{\text{map}}$. $G(x, y, \sigma)$ is the Gaussian smoothing filter, and $S(x, y)$ the Gaussian-smoothed image.

*Differentiation.* The simplest discrete approximation to a derivative along the $x$ and/or $y$ dimension is the difference between two adjacent pixels. The smoothing and differentiation steps can be combined into a single operation by using a mask that is defined as the first derivative of a Gaussian.

*Non-maxima suppression.* After the image has been convolved with the smoothing and differentiation filter, edges can be located at points of local maxima, $M(x, y)$, computed as follows:

$$M(x, y) = (D_x^2 + D_y^2)^{1/2}\,.$$

The resulting image $M(x, y)$, however, is insufficient for identifying edges. A process referred to as 'non-maxima suppression' is employed to thin the ridges in $M(x, y)$ by suppressing all values along the line of the gradient perpendicular to the ridge that are not peak values. Ridge-thinning instantiates the principle of 'good continuation' for edges and contours.

*Edge thresholding.* This is a common procedure for reducing the number of spurious edge fragments. Here a threshold of near-zero was applied to obtain the final binary edge map.

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 33 2004

www.perceptionweb.com