A System for Finding Changes in Colour

D.A. Forsyth *
Robotics Research Group
Department of Engineering Science
Oxford University
England

July 23, 1987

Abstract

We acquire a great deal of information about the world about us by perceiving the colour of objects. We can manage without colour, but certain tasks can become rather difficult - telling ripe from unripe fruits, or finding green tennis balls lost in grass, for example. Yet colour information is seldom used by the machine vision community, although, given the present state of image understanding technology, it is undesirable to waste cues as to the nature of the world observed.

The colour signal received by an imaging device does not itself tell us much about the world. We need to compute from this signal the surface reflectances of the objects that were imaged. To do so, we need to be able to find changes in the colour signal. This requires an understanding of what a noteworthy difference in colour is. We discuss metrics on the colour space, colour constancy, and some models of colour vision. We describe a working colour edge finder, based on an opponent coding scheme, which uses Fleck's edge finder to generate a colour edge map.

1 INTRODUCTION

There is a fundamental theme of matching edges or features in a large number of vision techniques: in stereo, in motion stereo, in motion, and in grouping. I propose that a grey-level edge is an insufficiently rich description of what is happening in an image to allow effective, unambiguous matching. We cannot waste information: we should use all cues in the image that allow us to assert something about the world. Recording the colour at each side of a change in colour, and then attempting to match such changes in colour will produce better results than matching edge-points, simply because the richer description can reduce the number of spurious matches.

Being able to perceive the colour of objects is a useful skill. For example:

- Ripe and unripe tomatoes have no great difference in brightness.
- Consider matching (for stereo, or motion, or whatever) grey-level images of a bookshelf: there is likely to be a considerable picket-fence effect. The matches may be disambiguated by recording the colour present on each side of an edge in each image.
- A robot has to deal in different ways with three kinds of object: one red, one yellow, and one green. We can easily construct these objects to be hard to tell apart using brightness or shape properties alone; the robot need only be a dichromat to tell the objects apart.
- We have a grey level image of a red ball on a green lawn; if the ball in a grey level image has the same intensity as the lawn, an edge detector can find no edges that would provide us with clues to its presence. Yet its hue makes it easily distinguishable in a colour image.
- Consider the problem of deciding whether two collections of pixels represent the projection into the image of the same surface. We know that the grey level information at pixels depends both on the reflectance and the orientation of a surface. But the colour information at a pixel depends as well on the shape of the surface reflectance function, which, by making assumptions about the spatial variation of the spectral composition of the illuminant, we are able to recover. Hence we are in a good position to group the points.
- Look at figure 1. In it, you should be able to find four impala ram and one lamb. If you cannot, try looking at the colour version of this picture. (This is less compelling than it should be, as I have not been able to present the colour pictures here. However, the components should themselves prove helpful.)

We can use colour information to aid us in solving grouping problems associated with calculating the shape of surfaces. The techniques of Brady [5] and Fleck [10] require one to compute a representation based on pairings of edge points. Many spurious pairings of edges can arise, and the computation is $O(n^2)$, where n is the number of edge-points. We can avoid this problem by grouping the edges, so that we form representations only for pairs that are in some sense interesting. Colour information can be used to advantage here, as it can be used to propose "good" groupings, using the fact that the colour of a surface does not change with its orientation. We may in this way obtain a large simplification in the complexity of the computation, if not in its order.

The most useful information is concentrated at changes, and we need to be able to find changes in colour. We must first discuss how we aquire colour information, and what it represents. Thereafter we discuss a metric on the colour space, an idea fundamental to finding changes in colour. We discuss the vexing issue of colour constancy, and we present a working colour edge finder.

2 EARLY COLOUR VISION

We have a scene consisting of surfaces with surface reflectance $S(x, y, \lambda)$, lit by an illuminant, $E(x, y, \lambda)$. The signal that is available is then the product of S and E, multiplied with $K(x, y)$, which is constant with respect to wavelength, and which expresses the viewing geometry. We

*The author acknowledges the support of the Rhodes Trust

AVC 1987 doi:10.5244/C.1.39

will refer to this signal as the colour signal, and denote it C(x, y, λ) following Wandell [33], and Maloney [22]. This signal is sampled by photoreceptors distributed both spatially and spectrally.

The retina of the eye consists of tiling of at least four types of receptor: the rod, and at least three types of cone, which we shall refer to as R, G, and B cones, distinguished by the spectral sensitivity of their respective pigments. Notice that, although each type of cone possesses a distinct peak in its spectral sensitivity in a different place, we do not see red solely through the action of the R cone: rather, we perceive all colour as a result of interaction between the cones. The belief is current in the colour literature that there is only one type of pigment in any given receptor, and the receptors are labelled according to their pigment. Notice that at each location on the retina, there is but one type of receptor. A receptor of class i, receiving light reflected from a point (x, y) responds with

$$\int_{-\infty}^{\infty} C(x,y,\lambda)\rho_i(\lambda)d\lambda$$

where $\rho_i(\lambda)$ is the spectral sensitivity of that receptor.

For rods, $\rho(\lambda)$ has a broad area of support, suggesting that they are primarily of value in estimating the luminance of any signal reaching the extra-foveal region of the eye, which is populated with rods, whereas the region of support of the cones is narrower, although in no sense narrow. Although there are no foveal rods [29], there is reason to believe that the rods contribute to colour vision. [23]. We will not discuss in this paper the role of the rods, although one may plausibly suggest that they play a role in colour constancy.

The suggestion that each receptor receives light reflected from only one point in space is unfortunate. Furthermore, we have given no account of the considerable chromatic aberration in the eye, which effectively acts as a low pass filter for spatial frequencies, with the filter passband narrowing as the wavelength decreases. There are a number of interesting issues to do with the foveal population of blue cones, indicating that the blue system and the red-green system are in some ways substantially different [26] [25] [34] [35].

To digitize scenes, we use a CCD camera and a number of different gelatine filters; a red component, a green component and a blue component were digitised for the images shown. A filter to cut near-visible infrared is used to prevent the camera responding to these wavelengths, as gelatine filters are infrared translucent.

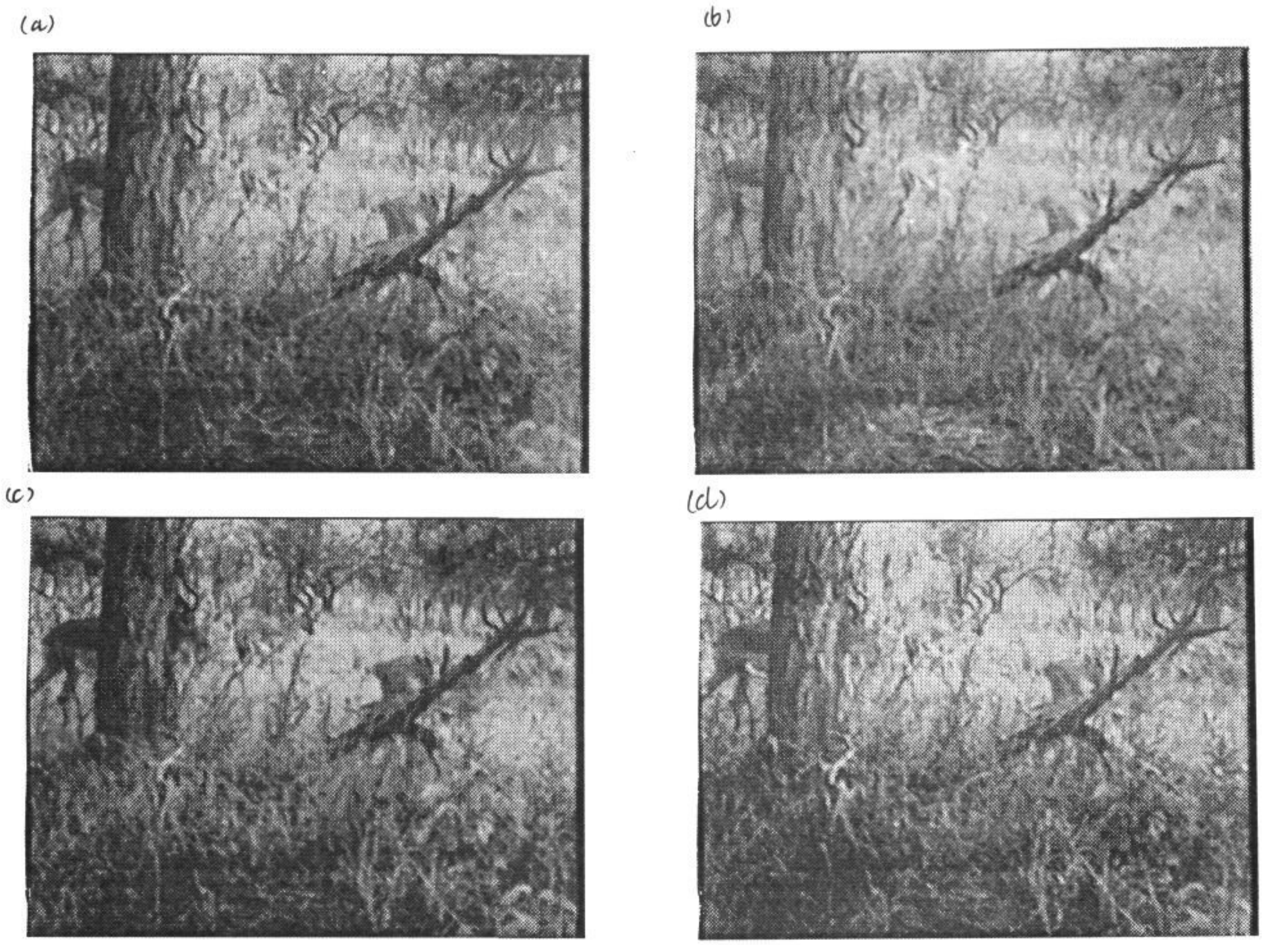


Figure 1: **Impala Rams:** there are four impala rams and a lamb in this image. (a) shows the intensity image of the impala. (b) shows the red component of the colour version of this picture. (c) shows the green component of the colour version. (d) shows the blue component of this picture.

The following differences between our camera setup and the eye should be noticed:

- The chromatic aberration of the lens involved is so small at the frequencies of interest as to be of no importance. Each colour channel has then the same spatial resolution.
- There are as many kinds of receptors at each point as we care to put there, within the limitations of the physics of available filters.

These differences are considerable, but the kernel of the problem still remains: given the receptor responses, and knowing the receptor sensitivities, we wish to recover the surface reflectance function, up to a constant (I refer to this rather loosely as the shape of the surface reflectance function.) It is clear that this is not possible without constraining assumptions. It is also clear that the human visual system is fairly good at solving this problem, to such an extent that most people do not realise that a considerable component of the structure of the signal reaching their eye is due to the lighting - this tends to come as a shock to beginners in photography [17] [24].

The problem is referred to in the literature as that of colour constancy, and is at present only really understood at a phenomenological level. Fundamental to all computational work in the subject is the assumption that illumination changes slowly over space. This is equivalent to the assumption that a sharp change in the colour signal is due to a sharp change in the world of reflecting objects, rather than to a sharp change in illumination.

That the human vision system makes this assumption is best illustrated by considering the illusions involved while watching a film: one perceives surfaces moving across the screen, rather than a blank screen under coloured lighting.

We can present an account of early colour vision as follows:

- Form a best approximation to the colour signal, and find the changes in it.
- From those changes, construct an estimate of the spatial structure of the illuminant.
- From this, construct a map of the surface reflectance functions presumed to have caused the signal.

There are some gross simplifications here, as we have ignored the many modes of colour perception discussed, for example, in Beck [2], and are claiming that the world, at least in the very early stages of a colour vision system, should be modeled as composed of surfaces. For example, an image taken with the lens cap on, will be considered to be an image of a surface that does not reflect any light at all in the absence of other cues.

The earliest attempt to present an algorithm for a colour constancy calculation is Land's Retinex calculation [18], which operates by calculating over a number of paths, the ratios of the colour signal for a each wavelength sample at each point in the path. This is then used to calculate the surface reflectance at each point. The Retinex computation has not, however, been shown to calculate a descriptor of colour that is not influenced by the lighting conditions.

McCann et al. [24], as a result of an experiment where observers attempted to match colour chips to a Mondrian image made of multiple squares of coloured paper, claim that there is a strong correlation between what is perceived and the shape of the surface reflectance function. The experiment is flawed by the fact that they do not attempt to present the colour chips to which the Mondrian patches must be matched under a different illuminant than that illuminating the Mondrian itself, although the colour names that they recover are consistent with a computation being performed that does recover the surface reflectance. This, and Land's Retinex [18] work, suffers from the problem that the structure of the lighting under which a scene is presented appears to simplify the computation. McCann et al. used as illuminants three essentially monochromatic lights; write the lighting as

$$\sum_{i=0}^{i=2} a_i \delta(\lambda - \lambda_i)$$

Then the receptor response for the k'th receptor, which has receptor sensitivity function ρ_k, is p_k, where

$$p_k = \sum_{i=0}^{i=2} a_i \rho_k(\lambda_i) S(x, y, \lambda)$$

by the sampling property. The wavelengths chosen, 450nm, 530nm and 630nm are such that for $i \neq k$, $\rho_k(\lambda_i)$ will in fact be small. So we get

$$p_k = a_k b_k S(x, y, \lambda_k)$$

(without any summing over k: the b_k are constants).

Now the calculation allows us to form the ratio of the surface reflectance at any one point to that at any other. Then we will recover for our point a reflectance of

$$k(1, S(x, y, \lambda_1)/S(x, y, \lambda_0), S(x, y, \lambda_2)/S(x, y, \lambda_0))$$

The problem is evident. If we change the ratios of the original monochromatic sources, then the direction of this vector will not change. But if we change the wavelength of one of the sources, then we recover a vector in a different direction for the same patch of surface. If we relax the assumption about monochromatic light sources the computation becomes even less well behaved.

Horn has proposed a generalisation of the retinex calculation suited to a parallel machine [14], and Blake [3] corrects Horn's proposal such that it can produce a unique solution. However, neither appeal for two reasons:

- Firstly, Blake's proposal requires all colour patches to be unchanging, and hence is only suited to Mondrian world. On the other hand, Horn's proposal will not necessarily produce a unique solution.
- Secondly, it seems from the above that the retinex computation will not in fact recover the surface reflectance at a point (as always, we want it only up to a constant).

Another approach to solving this problem is presented by Maloney and Wandell [21], and detailed by Maloney [22]. They model lighting and surface reflectance functions by finite dimensional linear models. There are indications that such models will be successful for surface reflectance functions, but it is not at all clear that they can be a good model of lighting. Certainly it is hard to account for McCann, McKee, and Taylor's [24] results with a lighting model of low dimension. Barlow discussed trichromacy in the light of such a model, considering the smoothing produced by the shape of the photoreceptor sensitivity curves, but concluded that this could not be used to explain colour constancy effects [1].

In my discussion of finding changes in colour, I avoid the issue of colour constancy by assuming that the spectral content of the illuminant varies slowly over space. Thus, all perceivable spatial changes in colour in images of a given scene are due to changes in the shape of the surface reflectance function. This assumption underlies all colour constancy work, and is not the same as assuming isochromatic illuminants from scene to scene.

The second assumption underlying this work is the assumption that the colour labels attached to a point do not vary with most small changes in viewpoint, when the illumination remains unchanged. For this assumption to be unreasonable requires an unusual setup of lights and surfaces. The work presented computes an internally consistent description of the changes in the colour signal: this must precede any colour constancy calculation, whose goal is to force these labels to be consistent with properties of the surfaces composing the world. For the purposes for which we propose this colour edge finder (stereo, motion and grouping work), this is not unreasonable behaviour, although it is desirable to have a system that can recognise an object's colour under a variety of lighting conditions. This latter goal requires a colour constancy system.

Another issue I have so far avoided is that of highlights, points on surfaces where reflection is essentially specular. Highlights can lead to groups of edges in an image which are essentially dependent on the viewing geometry. There exists some work on the subject, which models the colour signal produced by a highlight as the sum of the colour signal produced by the diffuse reflection at the hightlight, and a (usually large) component which is equivalent to the illuminant reflected off a white surface. I do not attempt to explicitly find highlights at this point, although Shafer [32] has reported some success with his technique, and Gershon [12] with his. At this stage, both techniques suffer from the considerable disadvantage that they assume that the light reflected specularly is due to the illuminant, rather than to mutual illumination of objects.

3 FINDING CHANGES IN THE COLOUR SIGNAL

Colour has been popular as a variable by which images may be segmented, and a number of transformations from RGB space to some form of Hue-Brightness-Saturation space have been proposed to improve this segmentation. However, little work has been done on describing images using colour: images segmented using colour do not tell us much about the world if we cannot tell whether the colour is due to the illuminant or the surface. Ohlander used colour information very successfully in his region segmentation program. He produced a segmenter that recursively segmented colour images using histogram splitting for three different representations of colour in the image. He did not present a way of using the results of the segmenter to describe a scene, nor did he use a very sophisticated representation of colour [28].

Nevatia constructed a colour edge detector by using the Hueckel operator on each of the red, green and blue images with the intention of using the additional information to improve the original edge map, but concluded that most edge information was in the intensity image [27]. It turns out that this is in a sense, correct, but is an unfortunate conclusion: the available colour information, rather than being a source of new edges, is a rich source of descriptors to attach to existing edges.

Machuca and Philips, at the end of their 1983 paper on edge detection using vector field techniques, proposed a colour edge detector operating on the phase of the IQ vector (effectively a representation of hue), but presented few results. Their proposed detector, $\sum_\gamma abs(\Delta\theta)$, where γ is a closed curve around a pixel and $\Delta\theta$ are the changes in the phase of the IQ vector as on tracks along γ, was imaginative, but has obvious problems with spatial resolution [20].

Gershon proposed a colour edge detector based on the presence in the cortex of double opponent cells, but did not present many experimental results [11]. Faugeras has described a system for image processing based on a perceptual model of the colour vision system, but does not address the problem of detecting edges [8].

Two kinds of approach to colour edge finding appear in the literature:

- A metric that reflects psychophysical performance is highly desirable for many reasons (designing road signs, for one). This can be used for edge finding.
- The other approach taken is to attempt to find colour changes, as Nevatia himself did, by finding changes separately in some transformed version of the red, green and blue component images.

Many metrics on colour space, using a line element approach, have been proposed: Helmholtz first suggested the technique [13]. All are intended to mimic the human perception of just noticeable differences. All suffer from the considerable problem that, by virtue of their construction, they allow one to ask the (to my mind meaningless),question, "Is green further from red than blue is from green?".

These metrics are constructed as follows. We have the space R^3, which we regard as a Riemannian manifold, and equip with a metric by constructing the (for our puposes positive definite) metric tensor with components g_{ij}. Then the distance from point $\mathbf{u}$ to point $\mathbf{v}$ is given by:

$$min \int_\gamma g_{ij} dx^i dx^j$$

where γ is a curve from $\mathbf{u}$ to $\mathbf{v}$. In fact, the minimum implies that γ is the geodesic from $\mathbf{u}$ to $\mathbf{v}$. Thus this formulation also requires the numerical solution of non-trivial differential equations merely to tell the distance between two points, as a geodesic must be constructed. This makes it less than desirable as a computational technique. Furthermore, the Riemannian manifold formulation is not easily constructed, and can be difficult to use. It must be allowed, however, that this representation can be used to produce predictions of psychophysical thresholds that are very accurate - see, for example, [16]. It does not capture in an obvious way the following complexities of colour perception:

- As a colour desaturates, it is finally perceived as white; but white itself is not perceived as having a hue. Thus, one may follow different paths of constant (different) hue, to suddenly arrive at a point with no defined hue.

- A similar remark obtains about the hue and saturation of a colour as it loses brightness and becomes black [15].

We would also like to be able to represent the following two effects, which can certainly be done in the manifold formulation, but only in a way that is non-intutive, as few people can directly interpret metric tensors.

- The Abney effect: the percieved hue of a colour changes as an achromatic stimulus is added.

- The Bezold-Brucke effect: the percieved hue of a colour changes as its luminance alone changes.

(for detailed discussion of both effects, see [36]

One may try to represent the colour space using spherical polar coordinates, with R to represent the brightness, θ to represent the hue, and ϕ to represent the saturation of the percieved colour. Then we can deform this space so that the local metric properties of the colour space, and the Bezold-Brucke and Abney effects, are captured. (To do this we can use the MacAdam ellipses, and psychophysical data on the effects mentioned.) This approach to finding changes in colour without using a global metric on the colour space has this problem, that one is not able to meaningfully define an analogue of a second difference, and hence it is extremely difficult to use extant edge finding machinery. I have found this problem with this approach insurmountable.

Alternatively one can accept the claim that almost all changes in colour are found with changes in brightness; and simply track brightness edges, looking at colour values at either side of them.

It is however necessary to consider just to what extent we wish to retain psychophysical plausibility; to do so, we should need to restrict the spatial bandwidth of our machine's blue system, and restrict ourselves to just one receptor at each point over the sensor. This would appear to be unwise, unless the intention is to construct purely a computational model of human colour vision.

If we wish to use the spatial bandwidth of our system, then this option must be rejected. Our sensor has the ability to spatially localise isoluminant changes in colour, a task at which humans are very poor. We should use this ability.

As we have said above, we are attempting to find a property of surfaces from an image: thus, more than anything else, we are interested in changes in the shape of the colour signal, which, by the assumptions above, are assumed to correspond to changes in the shape of the spectral reflectance of the world. It is clear that an opponent coding of a colour signal considers the shape of that colour signal, if we normalise it by the intensity of the two components involved. A (B-Y)/(B+Y) signal will consider the relative size of the humps at the long wavelength and short wavelength ends of the spectral energy density, and a (R-G)/(R+G) signal will similarly consider the finer scale shape at the long wavelength end of the spectrum.

This is an important point. The opponents represent a sensible decomposition of that signal the sensors can reconstruct. Recall that the only signal available to the system is the projection of the colour signal onto three functions. We can then represent the colour signal as an infinite sum. As the first three basis functions, we choose any set of orthogonal functions which span the space spanned by the receptor sensitivity functions, and as the others, any functions, orthogonal to the sensitivity functions, which, as a set, span the rest of the space. We can always choose such functions. Then, to reconstruct the signal, we must simply compute the coefficients of the receptor sensitivity functions in the expansion: we are blind to the rest of the signal, by construction. This result discourages such work as that of Barlow, for example, [1], who considered the response of the receptors to a Fourier decomposition of the spectrum presented. The approach of decomposing the signal onto orthogonal bases is a useful one: however there is a natural basis, which is the one mentioned above. It is possible using this approach combined with a metric on the space of reconstructable colour signals, to predict the human threshold for a just noticeable difference in colour associated with a shift in wavelength. The preliminary results I have had with this approach are encouraging, but not outstanding.

The double opponent cell here begins to make sense. Look at the receptor sensitivity functions in figure 2. Notice that there is a clear separation between the hump at the blue end of the spectrum, whereas the receptors at the longwave end of the spectrum are rather close together. Thus, a signal reconstructed as a linear combination of these functions can qualitatively be described by considering two questions:

- (i)is the hump at the shortwave end of the signal larger or smaller than the composite hump at the longwave end?

- (ii)considering the finer scale structure of the longwave end of the signal, which hump is larger?

Opponent cells can be seen to consider these questions.

The double opponent cell was detected in the goldfish by Daw [7], and has been shown to be common in the cortex of the macaque [19]. Rubin and Richards suggested that this cell detected what they called "spectral crosspoints", points where the spectral energy density functions due to neighbouring surfaces crossed over [30]. It can be seen that the cells in fact detect both these points, and points where

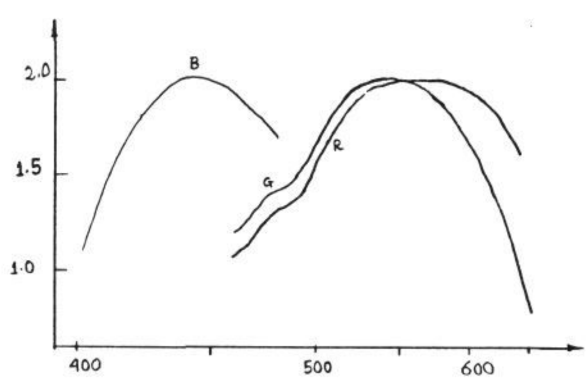


Figure 2:**Receptor sensitivity functions:** The graph shows human cone sensitivity functions plotted as log relative sensitivity against wavelength. The graphs were simplified from Bowmaker and Dartnall, 1980.

the change across wavelength of the colour signal of neighbouring points is of different sign, a fact that they point out in a second memorandum [31]. Look at figure 3. The responses of a set of these cells laid along the x axis will be zero, increase to a maximum, decrease through zero to a minimum, and increase again to zero (with increasing x). The zero crossing will mark the colour change in the stimulus. This can be seen as describing the changes that occur in the shape of the colour signal - we know that the colour signal must consist of linear combinations of the receptor sensitivity functions, and we are considering when we apply what is recognised as an edge finder (a Mexican hat cell) to that encoding, changes in the shape of the colour signal. Given that such changes are crucial to the present theories of colour constancy, which implicitly assume that such changes have been found (unless one wishes to perform a retinex computation along all possible paths), it is tempting to suggest that colour constancy computations occur no earlier in the visual pathway than the cortex. By these arguments we claim that there is great value to the computation described here, which essentially finds and describes changes in the colour signal. We describe those changes in terms of the colour signal; the domain of colour constancy is then to describe such changes in terms of surface properties. Finding the changes first is crucial.

Mollon has expressed a belief that the two chromatic opponent systems are qualitatively different (see, for example [25]). The B-Y system appears to be an older system; it is present in primates that do not have the benefit of an R-G system. The R and G receptors are structurally extremely similar, and the genes for the receptors are carried next to one another on the X chromosome. The psychophysics and physiology of the blue system are different to those for the R-G system. We then see an appealing story of increasing sophistication of representation for a signal: with a B-Y system one can consider only the humps at either end of the colour signal, whereas with an R-G system as well, one may consider also the fine scale structure at the longwave end of the signal.

All indications point to the suggestion that colour edges may be detected by applying a conventional edge detector to an opponent encoding of the image.

4 A WORKING COLOUR EDGE FINDER

Our goal is to generate an enhanced description of an image by finding changes in colour. To do so, we are using an opponent coding of the image and Fleck's state of the art edge-finder [9]. This edgefinder is noteworthy for a number of reasons: its very effective noise suppression, its ability to locate spatially extremely fine changes in grey level (a one pixel wide "thin bar" is reliably detected as two step edges) and its use of considerably more sophisticated labelling for edges than is current.

The way in which this edge finder works is described elsewhere [9]. A quick summary follows:

- The image is decomposed into a pyramid of scaled and resampled versions. The finest scale version is then the image itself. At each scale, at each point in the image, the maximum of four second directional derivatives is taken as the response. These responses, in combination with a sophisticated noise model, are used to label the image. The labelling

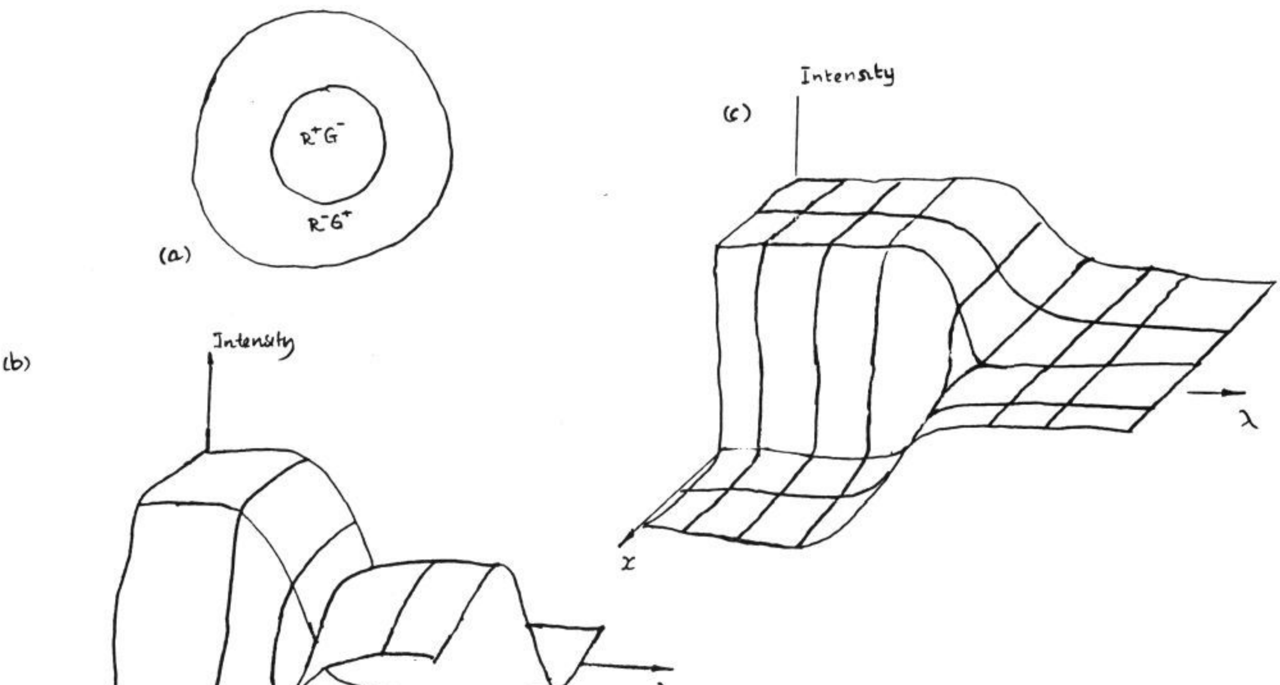


Figure 3: **Double opponent cells:** (a) shows an indication of the spatial response of a double opponent cell. Stimulus (b) is a spectral crosspoint; in (c) the changes across wavelength of the stimulus are of opposite sign.

refers to whether points are in a region which responds as being light or dark with respect to neighbouring regions, rather than on an edge itself. Those pixels labelled as being in a dark response region, which have as neighbours pixels labelled as being in a light response region, are then marked as being on the dark side of a zero crossing; the same (*mutatis mutandis*) technique is used to mark pixels on the light side of a zero crossing. The response at each scale is combined with the responses at the coarser scales to produce the finest scale labelling.

I draw attention to the labelling, however, for this is precisely what we require for a colour edge finder, and the present edge finder in fact works by expanding the set of available labels (we ignore labels to do with crack edges and the like at present, and deal only with the light side - dark side labelling), increasing our ability to describe the image, for our labels describe changes in the colour signal which correspond strongly to changes in surface material in the world.

The images were all digitized using Kodak Wratten gelatine filters, no. 's 25, 47B and 58. Neutral density filters (Kodak Wratten filters, no. 96) were used to compensate for the camera's reduced gain at short wavelengths. The near-infrared cut filter was obtained from Balzer's High Vacuum, and has a response which is essentially neutral in the visible spectrum.

The colour edge finder works like this. The opponent signals are computed, and a pyramid of scaled and resampled versions is constructed for these signals, the intensity signal, and the opponent intensity signal. For those regions in the image where the opponent intensity signal is below some threshold, the opponent signals are adjusted to be in balance, as reliable information about the nature of the colour signal and its changes does not exist at such points.

Notice that the opponent signals that are constructed are not an estimation of the human opponent signals under the circumstances, but, although there are differences in the shape of the responses of the photoreceptors and the filters, we may legitimately apply the same encoding trick as the human vision system. I construct the "composite hump" at the yellow end by considering the average of the red and green images at the point in question.

The response regions for each opponent at a scale are now calculated. Notice that at this stage we do not use the normalised opponent intensity signals, for the simple reason that we wish to force the edges in the opponent signals to line up with those in the intensity signal. I have found that using the opponent signals leads to an average error of less than a pixel (or rather, an occasional error of a pixel, and no greater errors), whereas using the normalised opponents leads to a larger error which is difficult to deal with in a principled way.

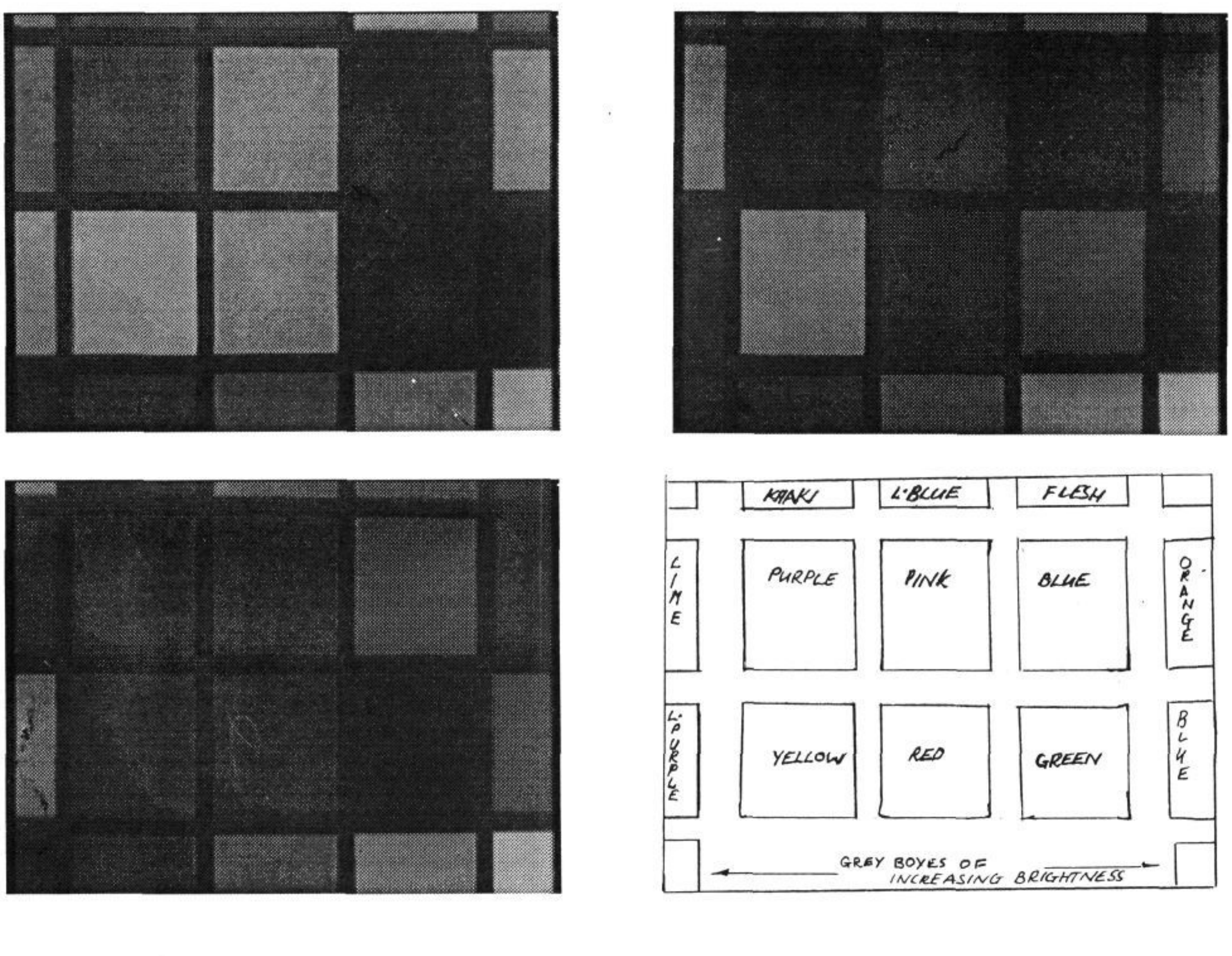


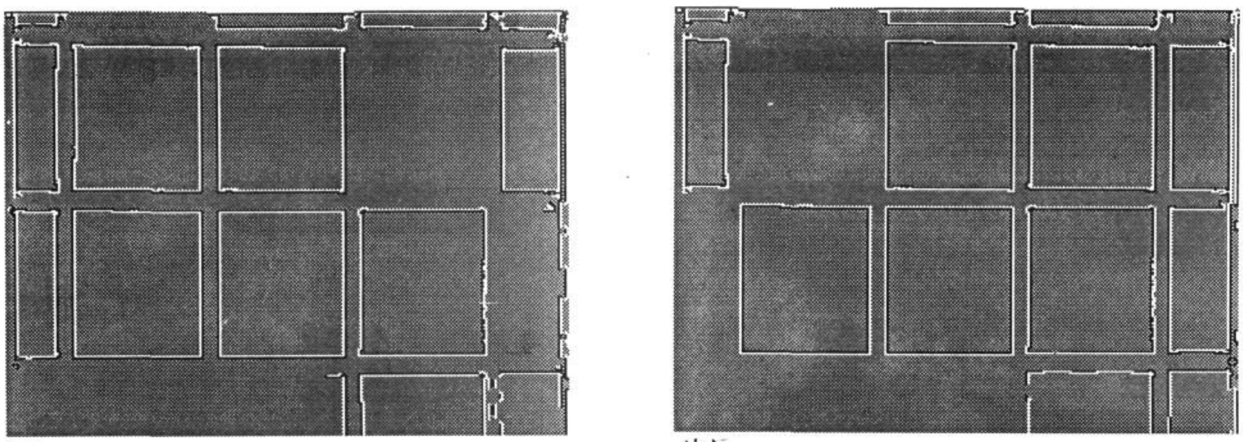
Figure 4: **The Macbeth ColorChecker:** (a) shows the red component of the color image; (b) the green component; (c) the blue component. (d) shows a drawing with color names of that section of the colorchecker that is visible.

These response regions label the image with areas where the opponents have swung one way or the other. Thus we can isolate the position of isoluminant colour changes. When we construct zero crossing maps of the image, the zero crossings are labelled with the colour in the original image, and with the nature of the change in colour across the zero crossing. Keeping the images in register is relatively simple: the fact that the edge finder is so accurate at localising changes, means that the response regions coincide to within at most a single aberrant pixel. Locally, we then require edges in the opponents either to be in register with edges in the brightness image, or to be far from them.

This can easily be enforced. For each pixel marked as being on the light side of a zero crossing, we inspect locally its neighbours which are similarly marked. The neighbours of these pixels which are not marked as being on either side of a zero crossing, are sorted by opponent response, and the majority vote is accepted as a labelling for all pixels visited. The same must be done for all pixels marked as being on the dark side of a zero crossing, mutatis mutandis. This technique will deal with errors in the register of the response images which are a pixel wide; I have found no worse errors. It does however require of a response region that it be three pixels wide, for the technique to work. This is unfortunate.

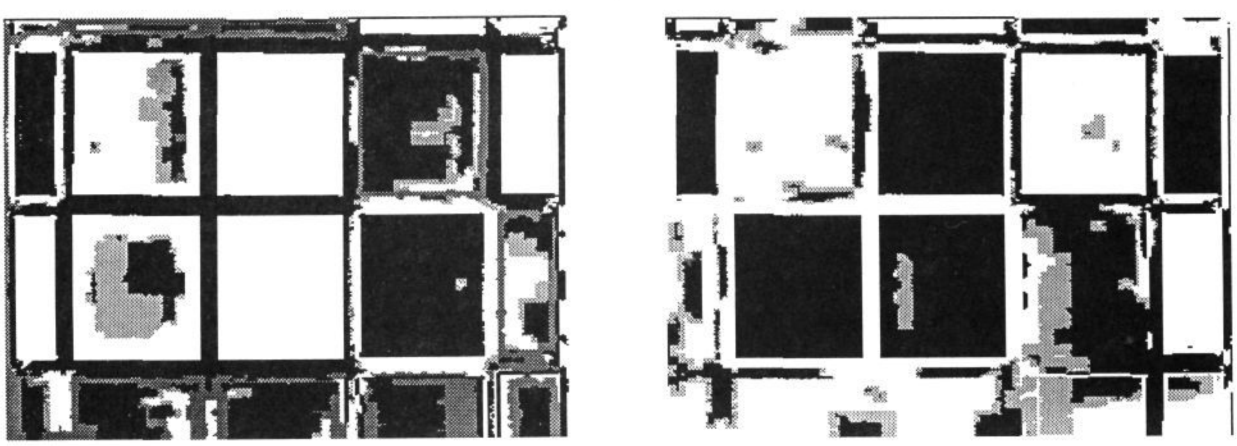
Phantom edges are zero crossings generated from an image associated with inflexions, rather than extrema, of its first derivative. These appear when, for example, a red region is separated from a green region by a white region: the white region is then redder than the green region and greener than the red region. Clark [6] has suggested that these may be rejected when one marks zero crossings by considering the actual change over the puported zero crossing: the program described uses an idea due to Fleck, which works by considering the way these zero crossings tend to arise when one combines a map of responses obtained at a coarse scale with a set obtained at a finer scale. The technique is described fully in [9], in this volume. A version of Clark's technique is also used, as these spurious edges, although not terribly common in intensity images, are for some reason a real problem in opponent encoded images. Opponent zero-crossings are subjected to this technique as well.

Results obtained using this edge finding scheme are rather good. We show below some examples; the images are presented as a set of three component images, and the response regions and zero crossing maps are shown for a number of examples.



(a) (b)

Figure 5: **ColorChecker Edges:** (a) shows the zero crossings of the R-G opponent. (b) shows the zero crossings of the B-Y opponent. The red (blue, resp.) sides of edges are marked white, the green (yellow, resp.) black.



(a) (b)

Figure 6: **ColorChecker Edges:** (a) shows the response regions of the R-G opponent. (b) shows the response regions of the B-Y opponent. The red (blue, resp.) responses are marked white, the green (yellow, resp.) black.

5 ACKNOWLEDGEMENTS

These ideas were improved by discussions with my supervisor, Mike Brady. Andrew Parker helped me with the human vision literature. Margaret Fleck has discussed them and others with me, and generously provided me with her code. Alison Noble discussed vision with me.

References

[1] Barlow, H.B. "What causes trichromacy?: a theoretical analysis using comb filtered spectra" *Vis. Res.* **22**, 635-643, 1982.

[2] Beck, J. *Surface color perception* Cornell University press, 1972.

[3] Blake, A. "Boundary conditions for Lightness computation in Mondrian world" CVGIP **32**, 314-327, 1985

[4] Bowmaker, J.K. and Dartnall, H.J.A. "Visual pigments of rods and cones in a human retina" *J.Physiol.* **298**, 501-511, 1980.

[5] Brady, J.M and Asada, H. "Smoothed local symmetries and their implementation" MIT AI Memo 757, MIT AI Lab, 1984.

[6] Clark, J.J. "Authenticating edges produced by zero crossing algorithms" unpublished pamphlet.

[7] Daw, N. "Color-coded cells in goldfish, cat and rhesus monkey *Investigative Opthalmology* **11**, 411-417, 1972.

[8] Faugeras, O.D. "Digital color image processing within the framework of a human visual model" *IEEE Trans. ASSP*, **ASSP-27** 4, 380-393, 1979.

[9] Fleck, M.M. "The Phantom edge finder *Proc. ALvey Vision Conf.*, 1987.

[10] Fleck, M.M. "Local Rotational symmetries" M.Sc. thesis, MIT, 1985.

[11] Gershon, R. "Empirical results with a model of color vision" *Proc. CVPR 1985.*

[12] Gershon, R., Jepson, A.D., and Tsotsos, J.K. "Highlight identification using chromatic information" *Proc. ICCV 1987.*

[13] Helmholtz, H. van *Handbuch der Physiologischen Optik* 2 ed. Hamburg: Voss, 1896.

[14] Horn, B.K.P.H. "On Lightness" MIT AI Memo 295, MIT AI Lab, 1973.

[15] Kender, J.R. "Saturation, Hue and Normalised color: calculation, digitization effects, and use" Department of Computer Science, Carnegie Mellon University, 1976.

[16] Koenderink, J.J., van de Grind, W.A. and Bouman, M.A. "Opponent color coding: A mechanistic model and a new metric for color space *Kybernetik*, **10** 78-98, 1972

[17] Land, E.H. "Color vision and the natural image: part 1" *Proc. Nat. Acad. Sci. USA*, **45** 1, 115-129, 1959.

[18] Land, E.H. and McCann, J.J. "Lightness and Retinex theory" *JOSA* **61** 1, 1-11, 1971.

[19] Livingstone, M.S. and Hubel, D.H. "Anatomy and Physiology of a color system in the primate visual cortex. *J. Neuroscience*, **4** 1, 309-356, 1984.

[20] Machuca, R. and Philips, K. "Applications of vector fields to image processing" *IEEE PAMI* **PAMI-5** 3, 1983.

[21] Maloney, L.T. and Wandell, B.A. "A computational model of color constancy" *JOSA* **1**, 29-33, 1986.

[22] Maloney, L.T. "Computational approaches to color constancy" PhD. dissertation, Stanford University, Stanford, CA 1984.

[23] McKee, S.P., McCann, J.J. and Benton, J.L. "Color vision from rod and long wave cone interactions: conditions in which rods contribute to multicolored images." *Vis. Res.* **17** 175-185, 1977.

[24] McCann, McKee, and Taylor, 1977 "Quantitative studies in Retinex theory" *Vis. Res.* **16** 445-458, 1976.

[25] Mollon, J.D. "The oddity of Blue" *Nature* **268**, 587-588, 1977.

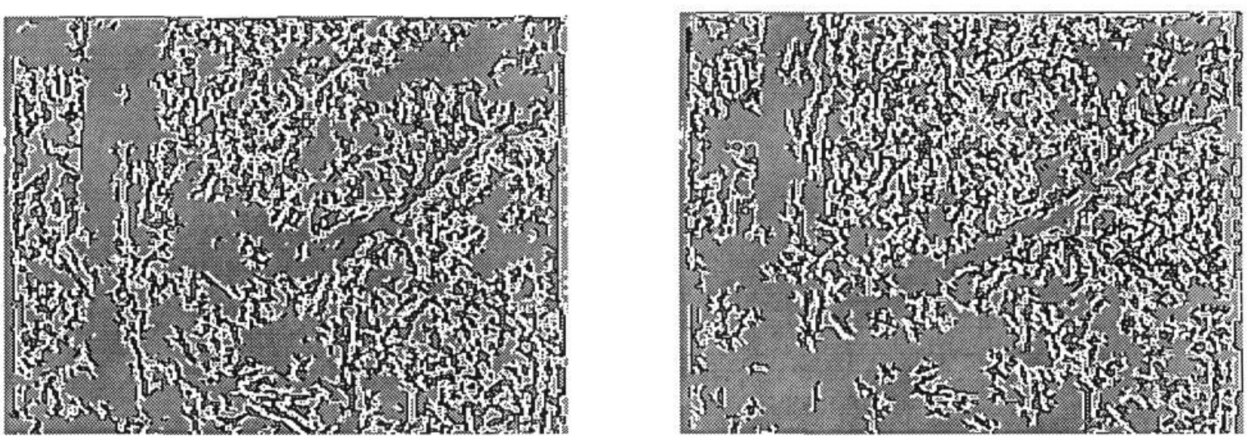


Figure 7: **Impala Edges:** (a) R-G opponent. (b) B-Y opponent. Notice that we now have much more information about the composition of this image.

[26] Mollon, J.D. "A Taxonomy of Tritanopia" *Docum. Ophthal. Proc. Series*, **33** 87-101, 1982.

[27] Nevatia, R. "A Color edge detector and its use in scene segmentation" *IEEE Trans. SMC*, **SMC-7** 11, 820-826, 1977.

[28] Ohlander, R. "Analysis of natural scenes" PhD. thesis, Carnegie Mellon university, 1975.

[29] Osterberg, G. "Topography of the layer of rods and cones in the human retina" *Acta Ophthal. Suppl.* **6** 1-102, 1935.

[30] Rubin, J.M. and Richards, W.A. "Color vision and Image Intensities: When are changes material?" MIT AI memo 631, MIT AI Lab, 1981.

[31] Rubin, J.M. and Richard, W.A. "Color vision: representing material categories" AI memo 764, MIT AI Lab, 1984.

[32] Shafer, S.A. "Using color to separate reflection components" University of Rochester TR-136, 1984.

[33] Wandell, B.A. "The synthesis and analysis of color images" *IEEE PAMI* **PAMI-9** 1, 2-13, 1987.

[34] Williams, D.R., MacLeod, D.I. and Hayhoe, M.M. "Punctate sensitivity of the blue sensitive mechanism" *Vis. Res.* **21**, 1357-1375, 1981.

[35] Williams, D.R., MacLeod, D.I. and Hayhoe, M.M. "Foveal Tritanopia" *Vis. Res.* **21**, 1341-1356, 1981.

[36] Wysiecki, G. and Stiles, W.S. *Color Science: Concepts, and Methods, Qualitative Data and Formulae* John Wiley and Sons, 1982.