Digital Colour Processing

Brian V. Funt

School of Computing Science
Simon Fraser University, Burnaby, B.C.
Canada V5A 1S6
funt@sfu.ca

Abstract

The members of the Computational Vision Laboratory at Simon Fraser University have been studying colour for over a decade. I describe some of the main colour issues, the progress we have made in understanding them and the application of our methods to colour-based object recognition and digital photography.

Introduction

Brockton Point in Stanley Park is one of my favourite spots in Vancouver, the world for that matter, so I've decided to stop there and write this paper on my laptop. It's a beautiful sunny day, blue water, blue sky, fresh white snow on the mountains with a few fluffy clouds clinging to them. I've taken a short walk to the water, inhaled the salt air, and now back in my car, I open my laptop. I can hardly see the characters on the screen; perhaps this wasn't such a good plan after all! My eyes have adapted to the bright sunlight, so the screen which normally seems quite bright to me, is now very dim. I notice also that the white background that I'm typing against looks yellowish, almost a tinge of orange. A few minutes later as my eyes adapt to the relative darkness of my car's interior, the screen now appears a bit brighter, but it's still yellow-orange. "Why?" I ask myself.

It's questions like this that have kept me fascinated with colour since 1984. What is colour? How do we see colour? Why do we see colour—what's its use to us? What's the relationship between the physics of reflection and our perception of colour? How does the context in which we see a colour change its appearance? Why does knowledge not affect our colour perception (e.g., I know my screen is white, so why now does it still look yellow-orange to me despite this knowledge?) Why does the colour clothes look change from the store to home? Why do colours not look even more different than they do given how different the colours of the illuminating lights often are?

I cannot answer all these questions, but I would like to describe some of the results we have obtained in my lab at Simon Fraser University over the past several years. I will be describing work done jointly with my recent students: Graham Finlayson, Janet Dueck, Kobus Barnard, Vlad Cardei and Subho Chatterjee, Louis Brassard and my research assistant, Michael Brockington. This work builds on earlier work with Ian Harder, Brigitte Dorner, Mark Drew and Jian Ho.

I come to the field of colour from an Artificial Intelligence background, so I naturally take the view that colour perception can be explained and modelled as a computational process. I'm interested in understanding how people think and perceive and using that understanding to produce better colour. To me, constructing computational models of colour perception is a part of artificial intelligence because there is no way to measure colour that does not relate to human perception. Colour is a perceptual quantity, not a physical one.

What we speak of when we speak about colour is our experience, not a physical phenomenon. Of course there are underlying physical phenomena creating our experience, but we can experience the same physical phenomenon differently under different circumstances, very much like the way a word can have different meanings in different contexts. Similarly we may experience two different physical phenomena as appearing the same. Hence in general, there is no one-to-one mapping between the spectrum of the light reflected into our eye's (the underlying physical phenomenon) and the colour we will perceive it to be.

The Main Problem of Colour Perception

To explain colour perception, we must explain how it is that we see colours as relatively stable despite changes in the incident illumination. I make the assumption that colour—like the rest of visual perception—is there to give us information about the world, the surface properties of objects in particular, and so the stability and reliability of the information is important. The problem of colour stability arises because the light reaching our eyes from an object is the product of the object's surface reflectance and the spectrum of the light illuminating the object. We do not have direct access to the properties of the incident light, so somehow we must estimate them from the light we receive from the object. To make matters worse, our eyes only measure the spectrum at extremely low resolution.

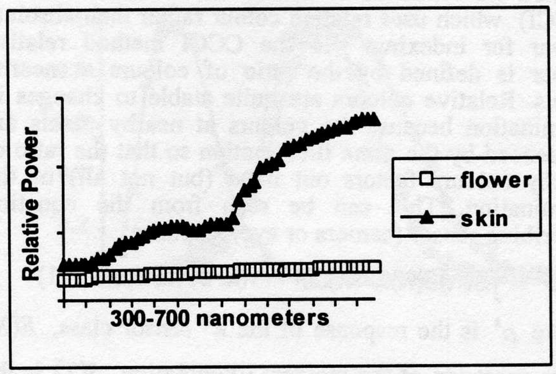
A typical spectrometer's resolution is 1 nanometer, which means that it yields 301 samples in the visible range of 400-700nm. If we consider this to be 301 "pixels" then the eye is, if we picture the spectrum as a 1-dimensional 'image', zooming out and reducing it down to only 3 pixels. Extending the analogy to 2-dimensions, this would be like reducing at 640x480 image to 6x5. What's important about this analogy is that the eye is not simply subsampling the values at 3 narrow points in the spectrum, rather it is sampling 3 weighted averages of the spectrum just as a single image pixel in the zoomed-out, reduced image represents a weighted average of the pixels in the original. It's easy to imagine that a single 6x5 image could have resulted by zooming out from a great many different 640x480 images. Similarly, there can be many spectra that can lead to the same triplet of responses in the eye. Such spectra are called metamers.

Metamers reflect the fact that the relationship between spectra and sensor triplets is many-to-one, as is the relationship between full and reduced images. We cannot perfectly reconstruct the original spectrum or image from its reduced counterpart. Information has been lost in the reduction that cannot be recovered. This is an important observation with respect to colour perception, since it means that there is no hope of recovering the precise surface reflectance properties of surfaces in a scene from a 3-band image of that scene. Nor is there any hope of recovering the exact spectrum of the light illuminating the scene.

The best that we can expect is a 3-parameter specification of the reflectance and illumination properties. This is a conclusion which may seem obvious, but a very common misconception about colour is that there is a one-to-one correspondence between colours and wavelengths—a misconception which is exacerbated by the all-to-common-but-entirely-misleading terminology, 'dominant wavelength'. Yes, to each wavelength there corresponds a unique colour, but to each colour there is no unique wavelength which produces it, once again reflecting the many-to-one relationship between spectra and colours.

With all these obstacles to colour perception, one might easily wonder how colour perception is possible at all. There are several factors that mitigate against the impossible. To begin with, while the spectral resolution of the eye is limited to 3 samples, the spatial resolution is very high. Neighbouring 'pixels' on the retina generally represent neighbouring scene points which usually will have the same or very similar illumination. In a multi-coloured scene, there will often be times when points of different surface reflectance are lit by the same illumination. Although the incident light's intensity may vary quite rapidly from point to point due to changes in surface orientation, generally its spectral composition varies only slowly. Fundamentally, all colour constancy methods rely on the fact that the spectrum of the incident illumination is effectively constant over substantial areas.

A second key factor aiding colour perception is the fact that the functions describing surface reflectances of common objects, such as grass or mud, and the spectra of typical illuminants, such as daylight and tungsten light bulbs, possess some interesting properties: the main one being that there is a high degree of correlation within each function across the different wavelengths. If you look at plots of spectra such as those in Figure 1 you will see that they tend to vary pretty slowly with respect to wavelength. Using standard statistical techniques such as principal components analysis, this regularity can be exploited to reduce the number of parameters required to describe the spectra to somewhere between 4 and 9 from the original 301.[COHEN64] In other words, there is a closer fit between the 3 types of colour-sensitive cone sensors receiving spectral signals we have in our eyes and the actual number of parameters required to describe those spectra than it would at first appear. Unfortunately, 3 parameter models of reflectance and illumination spectra based on principal components analysis provide too crude an approximation, so there remains a mismatch between the data our eyes records and the complexity of the underlying world.



Spectral reflectance functions of a flower petal and skin are quite smooth and slowly varying.
Figure 1

Object Recognition

One of the reasons we might need colour constancy is for object recognition. Colour is an obvious distinguishing feature of an object, but without colour descriptors that remain independent of the colour of the incident illumination, colour will be of limited use. We cannot expect to recognize an object based on its collection of colours if those colours vary dramatically with the illumination.

Although colour can be an important feature of an object, it was not until Swain and Ballard's[SWAIN91] colour indexing work that it was shown how phenomenally useful colour can be in object identification. Their method identifies an object by comparing its colour histogram to the colour histograms of known objects stored in a database of colour histograms. A colour histogram in essence describes an object by the amount of image area each of its colours occupies. An object matches a known model in the database to the extent that its distribution of colour areas resembles those of the model. What I find particularly surprising about Swain and Ballard's method is how well it works even though it ignores all geometric and shape information. Ignoring shape information in many ways is an asset for the method because it means that objects match even when bent out of shape or rotated in the image.

Swain and Ballard's method, however, does have one significant failing, which is that it is quite sensitive to changes in the colour of the incident illumination. When the illumination changes, all the colours in the image change. The colours change enough that the colour histograms of an object and its model in the database no longer match. One way to overcome this problem would be to preprocess the images with a colour constancy algorithm; however, in many ways colour constancy could be considered a harder problem than object recognition. Is it possible to use colour in object recognition without solving the colour constancy problem?

That last question led me and Finlayson[FUNT95] to a method we've called *colour constant colour indexing* (CCCI) which uses relative colour rather than absolute colour for indexing. In the CCCI method relative colour is defined by the ratio of colours at nearby pixels. Relative colours are quite stable to changes in illumination because the colours at nearby pixels are influenced by the same illumination so that the ratio of nearby colours factors out most (but not all) of the illumination. This can be seen from the equation describing sensor (camera or eye) response:

$$\rho^k = \int E(\lambda) S(\lambda) R^k(\lambda) \partial ll \qquad (k=1, ..., 3) \qquad (1)$$

where ρ^k is the response of the kth sensor class, $E(\lambda)$ is the spectrum of the incident illumination, $S(\lambda)$ is the percent spectral reflectance of the surface and $R^k(\lambda)$ is the relative sensitivity function of the kth sensor class.

For two nearby pixels, A and B, the ratio of responses for the kth sensor type at the two locations is

$$\frac{\rho_A^k}{\rho_B^k} = \frac{\int E(\lambda) S_A(\lambda) R^k(\lambda) \partial\lambda}{\int E(\lambda) S_B(\lambda) R^k(\lambda) \partial\lambda} \qquad (2)$$

where $S_A(\lambda)$ and $S_B(\lambda)$ represent the surface reflectance at the scene points corresponding to pixels A and B. While it is tempting to simplify equation (2) by cancelling out $E(\lambda)$ on the top and bottom, obviously we cannot legitimately do so in general since $E(\lambda)$ appears within the integral. It is interesting, however, that experimentally such ratios are found to be quite illumination independent. Why should this be the case?

One way it would become true that $E(\lambda)$ could be cancelled out would be if the sensor sensitivity function $R(\lambda)$ were extremely narrowband (i.e., the Dirac delta) so that its response was only non-zero at a single wavelength, λ_0. By the sifting theorem, (2) then becomes

$$\frac{\rho_A^k}{\rho_B^k} = \frac{\int E(\lambda_0) S_A(\lambda_0) R^k(\lambda_0) \partial\lambda}{\int E(\lambda_0) S_B(\lambda_0) R^k(\lambda_0) \partial\lambda} = \frac{S_A(\lambda_0)}{S_B(\lambda_0)} \qquad (3)$$

in which case the illuminant's effect can be cancelled out and eliminated.

The idea that narrowband sensors result in illumination-independent ratios led Finlayson et. al.[FINLAY94a] to consider the general conditions that lead to stable ratios and I'll discuss those results below; for the moment, however, assume that the sensors are sufficiently narrowband that (3) holds. In that case the ratio of neighbouring pixels will be illumination independent so long as the illumination is not varying spatially between the two pixels. The CCCI algorithm uses these illumination-independent ratios representing relative colour in place of Swain and Ballard's absolute colour.

CCCI histograms the colour ratios in place of the histogramming the absolute colours and then matches ratio histograms instead of colour histograms. In other words, in CCCI the R,G,B colour triplets are replaced by the ratio triplets R_A/R_B, G_A/G_B, B_A/B_B which are then quite stable with respect to changes in illumination colour. These ratio triplets are histogrammed and compared with ratio histograms stored in the database of model objects. CCCI results with ratio histograms were comparable in object-recognition accuracy with those of absolute colour histogram matching. When the illumination was varied, CCCI continued to perform well; whereas, regular colour indexing failed completely.[FUNT95]

For computational efficiency CCCI's ratios are computed by taking the finite-difference approximation to the laplacian applied to the logarithm of the image, which is equivalent to the ratios of neighbouring pixels.

Modelling Illumination Change

Equation (3) incorporates a model of the effects of changing from one scene illumination to another in that it says that the ratio of sensor responses under the two illumination conditions will be identical. Equivalently, we can say that the effect of a changing the illumination can be modelled by a simple scaling of the sensor response. For example, if the sensor response in kth band under illuminant A is ρ_A^k and the response under illuminant B is ρ_B^k and their ratio is $K_{A,B}^k$ then

$$\rho_A^k = K_{A,B}^k \rho_B^k \qquad k=1...3 \qquad (4)$$

The same holds, but with different scale factors, for each image band. In other words, if we know the sensor responses under A we can predict the sensor responses under B simply by multiplying by the appropriate scale factors without reference to the surface reflectance. Each band is scaled independently of the others so there is no interaction between bands. If we arrange the sensor 3-tuple as a vector then this independence amounts to saying that a change from one illumination to another can be modelled as a simple diagonal matrix transformation.

Independently scaling the image bands in this way to accommodate illumination changes is a very common operation in colour, for example, it is at the heart of von Kries adaptation, but the accuracy of the model is often taken for granted when it should not be, since in the absence of constraints on $E(\lambda)$ and $S(\lambda)$, (3) is only guaranteed to hold for narrowband sensors.

Having very narrowband sensors turns out not to be the only way for ratios to be illumination independent, however. Finlayson et. al.[FINLAY94a] form new "sensors" from linear combinations of the original sensors. The linear transformation creating the new sensors is called a 'spectral sharpening' transformation because it comes from the original observation that narrowband, or sharpened, sensitivity functions that approximate Dirac delta functions will have illumination-independent ratios.

There are several different strategies that can be used to sharpen sensors. One is simply to solve for the transformation that maximizes the amount of positive sensor response within a chosen spectral range. Experiments with different choices for the range, for example 530nm to 580nm versus 500nm to 600nm, result in very similar sharpening transformations. This strategy, which we call 'sensor-based sharpening', in essence aims to find the narrowest sensor by squeezing its response into a narrow range.

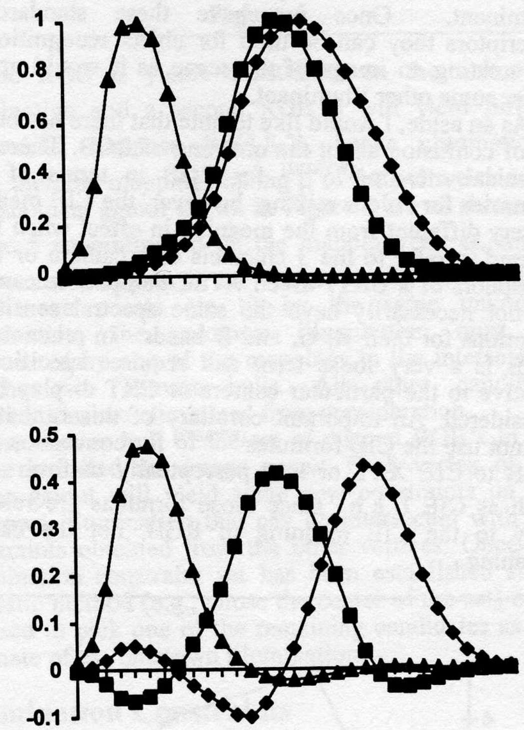
Figure 2 shows the results of sharpening the sensitivity functions of the human cones. Clearly, the resulting sensitivity functions do not look all that narrowband.

Since the goal of spectral sharpening is to find sensors for which the ratios of sensor responses remain independent of the illumination, another sharpening strategy is to solve directly for the sensors transformation leading to the most illumination-independent ratios. This strategy which we call 'database sharpening' uses a database of common illuminant and reflectance spectra over which the stability of the ratios is optimized. Database sharpening leads to sharpened sensors closely resembling those of sensor-based sharpening.

It's quite surprising that it's possible to attain relatively stable ratios since à priori the illumination can only be cancelled from (2) when the sensors are extremely narrowband. The fact that stable ratios do result says something about the illuminations and reflectances that generally arise in the world. If they were completely unconstrained then the ratios would always be unstable for anything but perfectly narrow sensor sensitivities.

The constraints on illuminations and reflectances can be expressed in terms of finite-dimensional models. As mentioned above, principal component analysis reveals that reflectances and illuminations can be described fairly accurately with a smaller number of parameters than one might at first expect. Working with models of a limited number of parameters places constraints on the reflectance and illumination spectra to be considered.

Finlayson et.al[FINLAY94b] show that for the case where illumination spectra span only a 3-dimensional space and reflectances span only a 2-dimensional space then it's possible to find a linear transformation of the original sensor sensitivity functions such that sensor-response ratios are completely independent of the illumination. Similar results hold for the case of 2-dimensional illuminants and 3-dimensional reflectances. While such low-dimensional models do not model illuminations and reflectances very accurately, these results of Finlayson et. al. define the circumstances under which illumination change can be modelled perfectly by independent scaling of the image bands, or equivalently by a diagonal matrix transformation, even though the sensor sensitive functions are broadband.



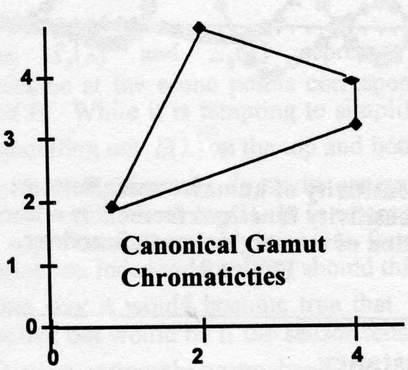
Top: Sensitivity of human cones. Bottom: Sharpened sensitivity functions formed as a linear combination of the original cone functions.
Figure 2

Colour Constancy

A colour constancy method is one which extracts illumination-independent colour descriptors from image data. Psychophysical experiments with human subjects have shown that although human colour constancy is quite good, it is far from perfect.[BRAINARD92] It would be very useful to have a model that accurately predicted the errors in human colour constancy, but in the interim we have been working on machine colour constancy without explicit reference to the psychophysical experiments.

At least with machine colour constancy the problem can be defined well. First we chose some illumination as our standard, or canonical, illumination. The choice of canonical illumination matters little so long as it is not unusual, as for example laser illumination would be. Then consider the 3-band RGB image obtained by any standard colour camera of a scene under some other, unknown illumination. The machine colour constancy problem requires converting the RGB at every pixel to be what it would have been had the same scene been illuminated by the canonical illumination. In this way, all the RGB values in the image of the scene measured under the unknown illumination are converted to a standardized RGB descriptors relative to the canonical illuminant. Once we have these standardized descriptors they can be used for object recognition or for creating an image of the scene as it would appear under some other illuminant.

As an aside, I would like to note that there has been a lot of confusion about the meaning of RGB. There is a technical meaning[WYSZ82] for RGB in terms of CIE primaries for colour mixing; however, the CIE meaning is very different from the meaning in effect when RGB is used to refer to the 3 channels of a camera or the 3 phosphors of a CRT. Even NTSC-compatible cameras do not necessarily have the same spectral sensitivity functions for their R, G, and B bands. In other words, RGB is a very loose term and requires specification relative to the particular camera or CRT display being considered. An important corollary of this is that one cannot use the CIE formulas[WYSZ82] for conversion from RGB to CIE XZY or to a perceptually uniform space such as CIE $L^*a^*b^*$ since those formulas are relevant only to the CIE meaning of RGB, not the camera meaning.



Simple canonical gamut consisting of the interior of the convex hull of the 4 chromaticities: (1,2), (2,5), (4,4), (4,3).
Figure 3

Returning to the colour constancy problem, two of well-known methods, which work under limited circumstances, are the grey-world algorithm and the white-patch algorithm. The grey-world algorithm assumes that the average of all colors in an image is grey, i.e. the red, green and blue components of the average color are equal. The amount the image average departs from grey determines the illuminant RGB. The white-patch algorithm, which is at the heart of many of the various retinex[LAND71] algorithms, presumes that in every image there will be some surface or surfaces such that there will be a point or points of maximal reflectance for each of the R, G, and B bands.

In my lab we have developed two different approaches to machine colour constancy, both of which I believe work better than all other methods to date. One is based on a process of elimination that exploits the constraints that the gamut of RGB's in the image provides about what the scene illumination could not have been. The second method learns the relationship between image RGB's and scene illumination using a neural network.

Gamut-Constraint Method

The gamut-constraint method[FINLAY96, BARNARD95, FINLAY95] builds on initial work by Forsyth.[FORSY90] I will try and provide an intuitive explanation of this method since both Forsyth's and Finlayson's papers are quite technical, which I believe has led to the Forsyth's method and its derivatives not being as widely used as they should be. I will refer to the collection of methods as the *gamut-constraint method* without always trying to distinguish between the contributions of the various authors.

The fundamental observation of the gamut-constraint method is that not all possible RGB values will actually arise in images of real scenes. In advance, one might presuppose that the full range of positive values of R, G and B might be possible; however, if we take a database of hundreds of reflectance spectra from a wide variety of common objects and compute the RGB values that will arise if they are illuminated by the canonical illuminant, we find that the RGB's confine themselves to a subset of the theoretically possible values. The convex hull of this set of RGB's obtained under the canonical illuminant is called the *canonical gamut*. All RGB's inside the convex hull might arise because positive mixtures of the reflectance spectra in the database could be used to create new reflectances.

Now consider an RGB triple α arising in an image of a scene under some unknown illumination. What does α's presence reveal about the illumination? Since the canonical gamut represents the full set of RGB's ever expected to occur, α must have originated from the canonical gamut; however, since α has been obtained under an illuminant different from the canonical one it may no longer lie within the canonical gamut. The transformation required to map it back to the canonical gamut encodes information about the unknown illumination.

Consider the following simple example in which our database has only 4 reflectances so that the canonical gamut is defined by the convex hull of the 4 points as shown in Figure 3.

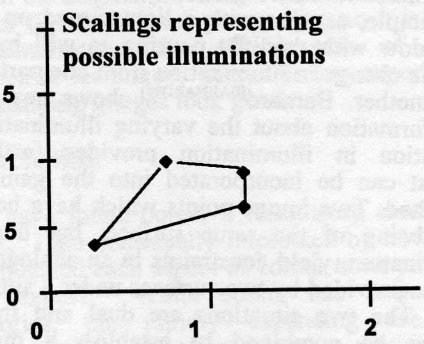
Intensity has been eliminated by moving to two-dimensional chromaticity coordinates (*r,g*) defined by $r = R/B$ and $g = G/B$. Finlayson[FINLAY96] points out that this choice of chromaticity coordinates is crucial since it preserves, in 2-dimensional coordinates, the diagonal model of illumination change that was present in the original 3-dimensional coordinates. In other words, to the extent that changes in illumination can be modelled by independent scaling of the R,G,B channels, then they can similarly be modelled by independent scaling of the resulting 2 chromaticity channels.

Consider an image RGB triple, α=(6,10,2), converted to chromaticity coordinates (3,5), which turns out not to lie within the canonical gamut. What does it take to map it to the canonical gamut? If we suppose that α corresponds to one of the 4 known reflectances, say that represented by (1,2) in the canonical gamut, then to map it there requires a scaling of the first component by 0.33 and the second component by 0.4. On the other hand, it might correspond to the canonical gamut point (2,5) in which case a scaling of 0.67 and 1.0 is needed. The other 2 canonical gamut points yield 2 more scaling pairs. Figure 4 plots the four mappings as the points

Of course it might have been the case that α corresponds to one of the points inside the convex hull of the canonical gamut. However, only linear scalings are involved, so mapping to those interior gamut points would only result in scalings within the interior of the convex hull of the mappings in Figure 4. The convex hull of the set of mappings, therefore, represents the complete set of mappings that could take α into the canonical gamut.

Based on the above discussion about spectral sharpening, the effect of changing from one scene illumination to another is presumed to change all the image RGB's by global scalings of each of the bands independently. Each point within the convex hull in Figure 4 therefore represents a different hypothesis about the unknown illumination since each such point describes just such a scaling. In this case it is the scaling modelling the change in RGB created by moving from the canonical illumination to the unknown illumination.

The convex hull in Figure 4 shows the constraints that finding α in the image imposed on what the unknown illumination might be. The illumination must be represented by one of the points within the convex hull because these are all the illuminations that could possibly have resulted in one of the colours in the canonical gamut appearing as α.



The four scaling pairs (0.33,0.4), (0.67,1.0), (1.33,0.8) and (1.33,0.6) that transform chromaticity (3,5) to the 4 vertices of the canonical gamut.

Figure 4

One RGB α yields constraints on the unknown illumination and a second RGB β will yield further constraints. Suppose β=(4,8,2), hence chromaticity (2,4), then the mappings taking β to the hull vertices of the canonical gamut shown in Figure 3 are as shown in Figure 5 superimposed on the mappings for α. Since both α and β appear in the image, and by assumption, both scene points are lit by the same unknown illumination, the unknown illumination must be represented by one of the mappings in the intersection of the two convex hulls. All other candidate illuminations are eliminated from consideration.

The convex hull of the set of distinct RGB's in the image is called the *image gamut*. Each vertex of the image gamut will yield some new constraints on the unknown illumination that can be intersected with the constraints obtained from the other vertices. Once the illumination constraint set has been established some heuristic method (e.g., chose the center of the set) must be used to pick one of the remaining candidates as the estimate of the unknown illumination.

Illumination Constraints

The gamut-constraint method can be extended by adding information about the set of possible illuminations to that about the set of possible surface reflectances. As with surface reflectance, measurement of lots of different light sources reveals quite a restricted gamut. The constraints implicit in the illumination gamut can then be intersected with the constraints from the image gamut to further eliminate illuminations as candidates for the unknown illumination.

In many scenes there is not a single illumination, but rather the illumination varies gradually throughout the scene. For example, a room with a light turned on as well as a window with daylight pouring in will have quite a dramatic change in illumination from one part of the room to another. Barnard[BARNARD96] shows how to extract the information about the varying illumination field. Variation in illumination provides further constraints that can be incorporated into the gamut-constraint method. Two image points which have been identified as being of the same surface, but under different illumination, yield constraints in an analogous fashion to those provided by two surfaces under a single illumination. The two situations are dual and their constraints can be combined to establish a more accurate estimate of the illumination throughout the scene than would be the case under spatially constant illumination.

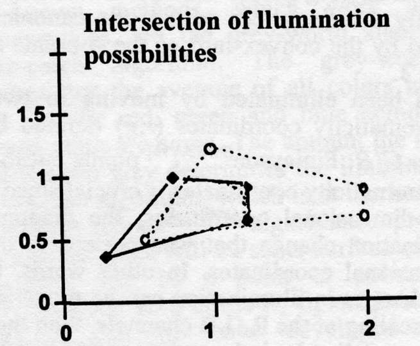
Interreflection between surfaces provides another common source of varying illumination. Since the interreflected light takes on some of the colour of the surface from which it is reflected, colour can be very useful in analyzing interreflections.[FUNT93] Also when Interreflection is known to be the source of an illumination variation, then the colour shift can be exploited in establishing colour constancy.[FUNT91]

Neural Network Colour Constancy

Some of our most recent work on colour constancy has been based on neural networks. Is it possible for a neural network to learn the relationship between an image of a scene and the chromaticity of scene illumination? If this relationship could be extracted by a neural network, then the trained network would be able to determine a scene's illumination from its image, which would then allow correction of the image colors to those relative to a standard illuminant, thereby providing color constancy.

Using a database of surface reflectances and illuminants, along with the spectral sensitivity functions of our camera, we generated thousands of images of randomly selected illuminants lighting 'scenes' of 1 to 60 randomly selected reflectances. During the learning phase the network is provided the image data along with the chromaticity of its illuminant. After training, the network outputs (very quickly) the chromaticity of the illumination given only the image data. We obtained surprisingly good estimates of the ambient illumination lighting from the network even when applied to scenes in our lab that were completely unrelated to the training data.

Previous color constancy algorithms generally employed assumptions either about the surface reflectances that will occur in a scene or about the possible spectral power distributions of scene illuminants. In contrast, the neural network approach involves no built-in constraints. It is an adaptive model that makes no explicit assumptions about the input data. All rules are implicitly learned from the training set. The experimental results (see below) show that the neural network out performs the grey-world and white-patch algorithms, especially in the case of scenes



Dashed lines show convex hull of mappings (0.5,0.5), (1.0,1.25), (2.0,1.0) and (2.0, .75) which take (2,4) to the vertices of the canonical gamut from Figure 3 superimposed on the mappings for (3,5). The mappings describing the unknown illumination are now limited to the shaded intersection of the two convex sets.

Figure 5

containing a small number (1 to 5) of distinct RGB measurements, and performs comparably to the gamut-constraint method. Good performance with only a small number of distinct RGB's means that the network is particularly well suited for processing small, local image regions. This is important because generally a scene contains more than one source of light, so the assumption that the scene illumination is constant will, at best, hold true only locally within an image.

The neural network we used is a Perceptron[HERTZ91] with one hidden layer. The neural network's input layer consists of 1000 to 2000 binary inputs representing the chromaticity of the RGB's present in the scene. The hidden layer has a much smaller size, usually about 16-32 neurons and the output layer is composed of only two neurons Each image RGB from a scene is transformed into an rg-chromaticity space. This space is coarsely, but uniformly, sampled so that all chromaticities within the same sampling square are treated as equivalent. Each sampling square maps to a distinct network input 'neuron'. The input neuron is set either to 0 indicating that an RGB of chromaticity rg is not present in the scene, or 1 indicating that rg is present. Discretizing the image gamut in this way has the disadvantage that it forgoes some of the resolution in chromaticity due to the sampling; however, tests show that this is not too important. What is important is that it provides a permutation-independent representation of the gamut which reduces the size of the input data set even though it is at the cost of a large input layer.

The output layer of the neural network produces the values r and g (in the chromaticity space) of the illuminant. These values are reals ranging from 0 to 1.

Comparison of Colour Constancy Methods

The network was trained using back-propagation algorithm, without momentum[RUMEL] on a training set consisting of a large number of synthesized scenes, each with a random set of from 1 to 60 surface reflectances. The illuminant database contains the spectral power distributions of 89 different illuminants and the reflectance database contains the percent spectral reflectance functions obtained from 368 different surfaces. The training set was composed of 8900 'scenes' (i.e. 100 scenes for each illuminant) and each scene had a random number of colors ranging from 1 to 60.

After training was completed, the network was tested on a different sets of 100 randomly generated scenes. For scenes with 10 distinct surfaces, the root mean square error in correcting the images for the change in illumination was: 1.07 for doing no correction at all; 0.43 for the white-patch algorithm; 0.45 for the grey-world algorithm; 0.52 for the gamut-constraint algorithm [3] without illumination constraints; 0.31 for the gamut-constraint algorithm with illumination constraints; and 0.32 for the neural net.

To compute the RMS error, the chromaticities of all the surfaces in the scene are corrected on the basis of each algorithm's illumination estimate. This yields an image as the algorithm would expect it to be under the standard illuminant. The difference between the true chromaticities under the standard illuminant and those estimated by the algorithm is measured by the RMS error taken over all surfaces in the scene.

Digital Photography

There are now many digital colour CCD cameras on the market and these are being combined with colour printers to provide complete digital photography systems. The low-end cameras have roughly 320x240 pixels with a dynamic range of at most 8 bits per colour channel. The high-end cameras such as Kodak's DCS 460 have 2000x3000 pixels and a dynamic range of 12 bits per channel. In either case, producing the correct colour under different illumination conditions is a problem and one which we can address once we are able to provide reliable estimates of the scene illumination.

Given the chromaticity of the illumination either by the gamut-constraint algorithm or the neural network, the image can be adjusted to have more or less the same RGB values at each pixel as it would have had if it had been taken under the canonical illumination. We have a patent pending on the neural network approach to correcting image colours in this way.

The ratios r_c/r_e and g_c/g_e, where (r_c,g_c) and (r_e,g_e) are the chromaticity coordinates of the canonical and estimated illuminants, define the scale factors by which to transform the chromaticity of each image pixel. Let the new chromaticity of some pixel be (r_p,g_p). In the diagonal chromaticity space the third component $b_p = 1$. This leaves the pixel's intensity to be determined, for which the pixel's intensity in the uncorrected image suffices so the new chromaticity is simply scaled to match the original intensity. Unfortunately there are no colour figures in these proceedings so it is not possible to demonstrate the effectiveness of our colour correction techniques here, but our experience is that the resulting images look good.

Conclusion

I have found colour a fascinating area in which to work. I am continually impressed by how much more complicated each aspect of colour turns out to be than I first expect. While colour has been studied for centuries, the rapidly expanding use of digital representations of colour creates a new urgency for a better understanding of colour. I think it is safe to predict that in the next few years black and white printers will become as uncommon as black and white CRT displays have become in the last few years and that digital photography will replace film.

Digital colour straddles both computer vision and image processing: for computer vision we need stable colour descriptors and it is these same descriptors we also need in processing images to improve colour fidelity in colour reproduction.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada and by a gift from Hewlett-Packard Incorporated

References

[BARNARD95] Barnard, K., Computational Colour Constancy: Taking Theory into Practice, *MSc thesis*, Simon Fraser Univ., School of Computing Science, 1995.

[BARNARD96] Barnard. K., Finlayson, G., Funt, B., Colour Constancy for Scenes with Spectrally Varying Illumination, *ECCV'96 Fourth European Conference on Computer Vision*, Vol. II, pages 3-15, April 1996.

[BRAINARD92] Brainard, D.H. and Wandell, B.A., Asymmetric color matching: how color appearance depends on the illuminant. *J. Opt. Soc. Am. A*, 9, 1433-1448, 1992.

[COHEN64] Cohen, J., Dependency of The Spectral Reflectance Curves of The Munsell Color Chips, *Psychon. Sci.* 1:369-370, 1964.

[FINLAY94a] Finlayson, G., Drew M., and Funt, B., Spectral Sharpening: Sensor Transformations for Improved Color. *J. Opt. Soc. Am. A*, 11(5):1553-1563, 1994

[FINLAY94b] Finlayson, G., Drew, M., and Funt., B., Color Constancy: Generalized Diagonal Transforms Suffice, *J. Opt. Soc. Am. A*, 11(11):3011-3020,1994

[FINLAY95] Finlayson, G., Funt, B. and Barnard, J., Colour Constancy Under a Varying Illumination, *Proc. Fifth Intl. Conf. on Comp. Vis.*, Jun 1995.

[FINLAY96] Finlayson, G. Color in Perspective, PAMI Oct 1996. Vol 18 number 10, p1034-1038

[FORSY90] Forsyth, D., A novel algorithm for color constancy, *Intl. Journal of Computer Vision* 5:5-36, 1990.

[FUNT91] Funt, B.V., Drew, M.S., and Ho, J., Color Constancy from Mutual Reflection. *Intl. Journal of Computer Vision*, 6(1):5-24, April 1991.

[FUNT 93] Funt, B.V. and Drew, M.S., Color space analysis of mutual illumination, *IEEE Trans. Patt. Anal. and Mach. Intell.*, 15(12):1319-1325, 1993.

[FUNT95] Funt, B.V., and Finlayson G.D., Color Constant Color Indexing, *IEEE Trans. Patt. Anal. and Mach. Intell*, 17(5), May 1995.

[HERTZ91] Hertz, J., Krogh, A., and Palmer, R.G. *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, 1991

[LAND71] Land, E.H., and McCann, J.J., Lightness and retinex theory, *J. Opt. Soc. Am*, 61:1-11, 1971.

[RUMEL] Rumelhart, D.E., Hinton, G.E., and. Williams, R.J, Learning Internal Representations by Error Propagation in D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Vol.1: Foundations. MIT

[WYSZ82] Wyszecki, G. and Stiles, W., *Color Science: Concepts and Methods Quantitative Data and Formulas* 2nd ed., Wiley New York, 1982.