NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents

Chromatic Aberration: A New Tool for Colour Constancy

by

Jian Ho

B.Sc, Simon Fraser University, 1986

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Master of Science
in the School
of
Computing Science

SIMON FRASER UNIVERSITY

April 1988

ISBN 0-315-48753-4

Approval

Name: Jian Ho

Degree: Master of Science

Title of Thesis: Chromatic Aberration: A New Tool for Color Constancy

Chairman: B.K. Bhattacharya

Brian V. Funt,
Senior Supervisor

Thomas W. Calvert,
Supervisory Committee Member

Leigh H. Palmer,
Supervisory Committee Member

Ze-Nian Li,
External Committee Member

April 14, 1988
Date of Approval

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

CHROMATIC ABERRATION: A NEW TOOL FOR COLOUR CONSTANCY

Author: _____
(signature)

JIAN HO
(name)

27 Apr. 88
(date)

Abstract

Maloney proposed solving the colour constancy problem by using a finite-dimensional linear model. His method depends on the number of sensor classes being larger than the number of basis functions needed to model the surface spectral reflectance. Since the surface spectral reflectances of most natural objects require at least three basis functions for accurate modeling, according to Maloney's paper we must have four sensor classes in order to achieve colour constancy. However, most experiments suggest that the human visual system is tri-chromatic. That is, there are only three classes of sensors in the human visual system. As a result, we must look for other spectral information to help us achieve colour constancy. We claim that chromatic aberration can help here.

From the chromatic aberration which occurs at the edge between two coloured regions under unknown illumination, we derive the difference of the spectral power distributions of the lights reflected from these regions. Using finite-dimensional linear models of illumination and surface spectral reflectances as our basic assumption, we obtain a set of equations for the coefficients that describe the illumination and the surface spectral reflectance. In combination with the equations obtained from the sensors inside each coloured region, we determine the surface spectral reflectances, hence colours, of the regions and the spectral power distribution of the illumination, if the number of sensor classes used is not smaller than that of the dimensions of the finite-dimensional linear model for surface spectral reflectance.

Since using degree three in the linear model approximates most of surface spectral reflectances very well, we only need to use three classes of sensors and chromatic aberration to recover the surface spectral reflectances. Without using chromatic aberration, Maloney had to use at least four sensors. Thus the information provided by chromatic aberration is very valuable.

Acknowledgements

I would like to extend my greatest appreciation to Dr. Brian V. Funt for originally suggesting the topic of chromatic aberration, and for his guidiance and encouragement throughout my thesis work. Many thanks also to Dr. Leigh H. Palmer for his consultation and expertise in optics. In addition, I would like to thank Dr. Thomas W. Calvert, Dr. Ze-Nian Li, Mark Drew and Frank Tong for their helpful discussions and advice. This work is supported by C.D. Nelson Memorial Graduate Scholarship.

Table of Contents

List of Figures

Chapter 1

Introduction

Colour is becoming more and more useful in computational vision today. Just as black and white television sets are gradually abandoned, we are moving from grey level computational vision to computational colour vision.

Colour plays an important role in computational vision. It helps us to identify objects, distinguish regions and recognize patterns. For example, when we see a traffic light, we can use colour to distinguish a red light from a green light, whereas in a black and white image, we can only distinguish the red light from the green light by their relative positions.

We do not notice much of a change in the colours of objects regardless of whether we see them by daylight or artificial light, nor do we notice much change resulting from the variations in natural light in the course of a day. The perceptual ability of assigning stable colours to objects despite lighting variation is called colour constancy. In order fully to use the information which colour provides us, we must achieve colour constancy. By colour constancy, we mean the ability to recover the colour of an object even if it is viewed under different illumination. Human beings have this ability, at least in most circumstances. Before we can answer the question how a similar colour constancy can be achieved in computational vision, we must define the colour of an object. Many definitions for the term colour exist; however, from the point of view of colour constancy, the colour of an object should be defined as an intrinsic property of the object, not in terms of the illumination or the surroundings.

Like other researchers in vision, we take the surface spectral reflectance of an object as the intrinsic chromatic property of that object. The surface spectral reflectance of an object specifies the fraction of incident light reflected by this object at each wavelength. Consequently, the problem of colour constancy is reduced to the problem of recovering surface spectral reflectance. It is not known how the human visual system obtains the surface reflectance, since the receptors in the cortex only receive the product of the illuminance and the surface spectral reflectance. However, we can at least try to find out what machine visual sensing can do.

E.H. Land [Land 71] proposed a theory called "retinex theory" to determine the colour of an object. Land's theory has been shown to be valid under a limited set of physical conditions [Brainard 85]. In 1985, Laurence T. Maloney proved that we could determine the colours of objects with over 90 percent confidence provided we have four sensors and a sufficient number of distinct colours in the image. His result relies on the finite-dimensional model. The finite-dimensional model states that most spectral power distributions of ambient light can be modeled well by a linear combination of a finite number of basis functions; this is also true for surface spectral reflectance [Maloney 85]. We will give more detailed explanations in a later section.

In this paper we will show that some information about the spectral power distribution can be obtained from the chromatic aberration effect that occurs at the boundary between two coloured regions[1]. This information is very useful since the knowledge of the spectral power distribution is similar to getting extra sensors. Thus, by using the finite-dimensional model we can determine the surface reflectance of an

---

[1]Credit is given to Brian Funt for suggesting this idea to me

object, and hence the colour of the object. Chromatic aberration is often considered an undesirable effect in image formation. It causes images to be blurred, because light of different wavelengths travels at different velocities within a medium so that light of shorter wavelength has a higher refractive index than longer ones and hence is focused in front of light of longer wavelength. However, in our work, we turn the negative features of chromatic aberration into positive ones. Although it is known that a variety of chromatic aberration exists in the human visual system, whether chromatic aberration is actually used in colour perception by the human visual system is still a puzzle and needs to be investigated. We will here show that it can be used to help colour constancy perception in machine vision.

Chapter 2

Previous Work

Computer vision scientists have put much effort into solving the colour constancy problem. An early success was achieved by Land in his retinex theory. Land assumed that the illumination gradually changes over the scene. Based on this assumption, he used the ratio of the intensities near the edge of two regions in the image to eliminate the effect of the illumination variation. The algorithm Land used is first to find all the edges of a coloured region. Then sequential products are computed along many arbitrary pathways that wander through the two dimensional array of the intensity image taken through a filter. The sequential products of the pixels within the same coloured region are so close as to be virtually the same. Afterward, the sequential products are normalized to range from zero to one. When a scene is imaged through different filters, e.g., red, green and blue, a 3-tuple descriptor of the sequential products is obtained at each pixel. Land used this tuple to decide the colour of that point [Land 71], [Land 74], and [Land 83]. In this way, Land assumed that the colour of a point having a descriptor (1,1,1) is white.

Land's retinex theory has received much criticism. In fact, the retinex theory does not provide an exact colour constancy result exactly. It simply computes constant colour descriptors in a fixed scene, because all of the colour descriptors in a scene depend on the true colour of the guessed colour patch.

Laurence T. Maloney proposed computational approaches to colour constancy

[Maloney 85]. He used finite-dimensional models of the spectral power distributions of ambient light, $E(\lambda)$, and for surface spectral reflectances, $S(\lambda)$. These models will be presented in detail in the later sections. Maloney's mechanism for achievement of colour constancy relies jointly on the validity of these models and on the presence of a sufficiently large number of different colours in the scene being observed. Maloney's mechanism does not depend on a guess of the colour of a region as does Land's retinex theory. Maloney further assumes that ambient light is constant over the whole scene.

Briefly, Maloney assumes that a linear combination of a finite set of basis illumination functions $E_i(\lambda)$ is sufficient to describe the ambient light, that is

$$E(\lambda) = \sum_{i=1}^{m} \epsilon_i E_i(\lambda).$$

Similarly he assumes that any spectral reflectance is sufficiently well specified by a linear combination of a finite set of basis reflectance functions $S_j(\lambda)$:

$$S(\lambda) = \sum_{j=1}^{n} \sigma_j S_j(\lambda).$$

So the number of parameters required to specify the ambient spectrum is m and the number required to specify the reflectance is n.

Consider now the problem of achieving colour constancy in a scene consisting of regions with k distint colours. The spectral composition of light reflected from a single colour element in the scene is the product of $E(\lambda)$ and $S(\lambda)$. We now analyze the image of the scene with sensors having s different spectral response functions (such as the red-, green-, and blue-filtered images produced within a colour television camera). This analysis produces sk independent relations in the parameters ϵ_i and σ_j. A necessary condition for solution of these equations is that $sk \geq m + kn$, or

$s \geq n + m/k$. Since both s and n are integers and $m > 0$, the condition becomes $s \geq n+1$. If a sufficiently large number of distinct colours, $k \geq$ m, is present in the scene, then a minimum of $s = n + 1$ kinds of sensors with different spectral response functions is required to achieve colour constancy.

Maloney [Maloney 85] shows that if the number of sensor classes is greater than the number of basis functions required to model the surface reflectances of a scene, we can recover the surface reflectance of each region in the scene, provided that there is a "sufficient" number of distinct colour regions in the image. The number of distinct colour regions required depends on the number of sensor classes. In his paper, Maloney also discusses models of light and surface reflectance in detail. He sets out conditions for the feasibility of the set of equations. He also demonstrates the import of his work for biological and machine vision. The main contribution of his paper is to provide a computational approach to colour constancy.

The major problem with Maloney's approach is that it requires the number of classes of sensors to be at least one greater than the number of basis functions in modeling the surface spectral reflectances. It has been shown that at least three basis functions are needed to model most of the surface spectral reflectances [Cohen 64] and [Maloney 85]. This means that there should be four classes of sensors. But it is generally acknowledged that only three sensor classes (the cones) are active in the human colour vision system. The colour constancy problem according to Maloney is therefore reduced to finding a fourth sensor class.

Other researchers have tried to use finite-dimensional models and other information to discover the colour of an object. To recover the descriptors of surface spectral reflectance, M. D'Zmura and P. Lennie [D'Zmura 65] use the mechanisms of light

adaptation with eye movements. They give a procedure for assigning these descriptors to corresponding hues, which are the representations of surface colour independent of object shape and viewing geometry.

Gershon, Jepson and Tsotsos [Gershon 87] also use finite-dimensional linear models for the spectral power distribution of light and the spectral reflectance of surfaces to achieve colour constancy. Their mechanism requires that the average of the spectral reflectances of the surfaces in a scene be known. And they use this average as additional information to accomplish surface spectral reflectance recovery. However, such a requirement is very restricted.

In the present paper, we propose a different approach to the problem of recovering the surface spectral reflectances. We use the mechanism of *chromatic aberration* to find extra information on the spectral power distribution of incoming light in order to recover the spectral power distribution of the illumination as well as the surface spectral reflectances and thus to achieve colour constancy.

Chapter 3

Motivation

Since we only have the cone responses of the product of the illumination and surface reflectance, we need to know more about the spectral power distribution of this product in order to achieve colour constancy for most objects. There may be more than one ways to achieve this goal; one of them is chromatic aberration. A rainbow is a naturally occurring example of the effect of chromatic aberration. When white light passes through a prism, we can see a range of colours from violet to red. We can determine the light composition of a rainbow from its colours. Similarly, we can reconstruct the light since we know the spectrum of colours which results when the light passes through a prism. In this way we obtain more information about the incoming light. Later we will demonstrate that surface spectral reflectance can be recovered with this additional information.

It has been demonstrated that there is much chromatic aberration inside the human eye. Early studies [Wald 47] and [Bedford 57] showed that the average chromatic aberration is approximately 1.75 D between 420 and 660 nm. Here D is an optical unit, standing for Diopter, which is the reciprocal meter (m^{-1}). It is used for the measurement of "curvature." We can imagine the wavefront of a point light source. With different distances from the point source (the point of convergence), the curvature of the wavefront changes. In other words, the *curvature* of any surface is inversely proportional to the *radius of the curvature*. When we use the term Diopter for lenses, Diopter refers to the reciprocal of the focal length of a lens expressed in

meters. B. Gilmartin and R. E. Hogan (1984) found a mean chromatic range of 1.87 D between 488 and 633 nm. When they repeated the experiment under different conditions, the range increased slightly to 1.91 D. Gilmartin and Hogan also found that the range is 2.65 D for the wavelengths 458 and 476 nm. There are many other consistent studies, including those described in [Gilmartin 85]. Peter Alan Howarth and Arthur Bradley (1985) studied 20 subjects and discovered that the average aberration is 1.82 D with a standard deviation of 0.15 D [Howarth 86]. These results prove the existence of chromatic aberration inside the human eye. Malacara showed further that the pupil of the eye affects the amount of chromatic aberration in the eye. Since the extent of the opening of the iris depends upon the ambient light, the chromatic aberration in the eye depends indirectly on the ambient light. However, [HOWARTH84] and certain other researchers think that the chromatic aberration in the eye is a negative feature and therefore try to correct it. We have not found any evidence to show that the human visual system uses the chromatic aberration effect to achieve colour constancy. Since it can be shown however, that chromatic aberration is helpful in computational vision, it will prove to be of positive use.

Chromatic aberration has already been utilized by Juetz, Sincerbox and Yung, who have done spectral range sensing using dynamic chromatic aberration [Juetz 85].

Chapter 4

Chromatic Aberration

Chromatic aberration is an optical phenomenon. When light passes from one medium into another, it is refracted. The angle of refraction is not the same as the incident angle but depends on the refractive indices of the two media. Experiments show that the index of refraction varies with wavelength. For glass, for example, the refractive index decreases as the wavelength increases. In this chapter, we will investigate the spectral information we can obtain from chromatic aberration.

4.1. Assumptions

The first assumption is that all the parameters, such as the focal length for each wavelength, the aperture of the lens, the image plane distance, of the camera are known. This is a very reasonable assumption. The second assumption is that we can always focus on either end of the visible wavelength. This assumption may not be reasonable since consistent focusing on either end is hard to achieve; we will discuss a method for solving this problem later. In later calculations, we assume that the image plane is placed at the image point of the shortest visible wavelength. Although Malacara demonstrated that to a normal eye, distant blue light should always appear fuzzier than distant red light [Malacara 75], this may be due to the fact that the number of red cones inside our eyes is larger than the number of blue cones. It is a common observation when we look at projector that it is harder to read red pen writting than blue pen. Another reason for such assumption is to make the amount of

distortion from chromatic aberration increases with the amount of distortion from diffraction.

4.2. Background

Our later calculations are based on the following background in optics, which must be understood before any results can be derived from chromatic aberration.

4.2.1. Light Refraction

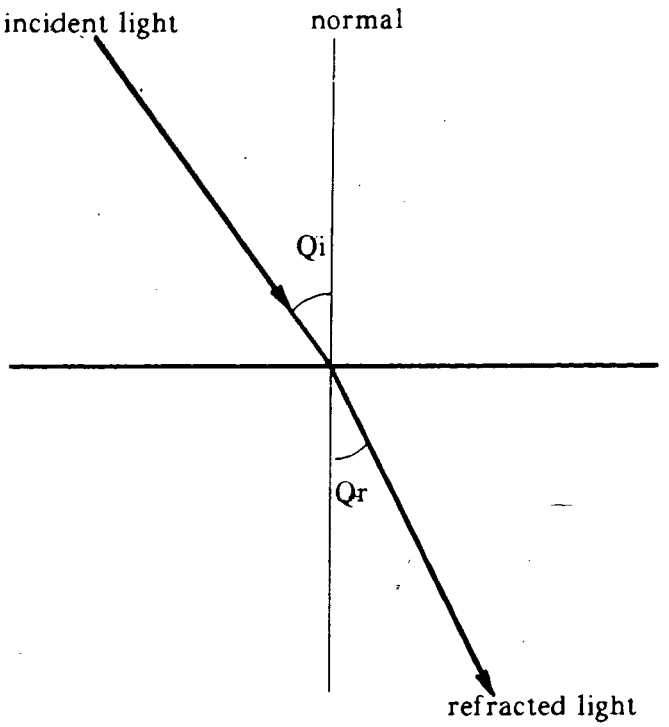


Figure 4-1: Refractive effect on a light ray passing from one medium into another

When a ray of light enters a medium from another, it is refracted. The refracted ray lies in the plane of the incident light and the normal to the surface of the dielectric media at the intersection point of the incident ray and the surface [See *fig.* 4-1] . The angle between the refracted ray and the normal θ_r obeys Snell's law [See *fig.* 4-1]:

$$\frac{\sin\theta_i}{\sin\theta_r} = \frac{n_2}{n_1} = \frac{v_1}{v_2}$$

where θ_i is the incident angle formed between the incident ray and the normal, θ_r is the refracted angle, $n_i = (c/v_i)$ is the index of refraction of medium i, c is the speed of light in a vacuum, and v_i is the wave velocity of light in medium i. Chromatic aberration arises from the fact that the refractive index of a medium varies with the wavelength of the incoming ray. When a light ray, consisting of lights of different wavelengths, enters another medium, the lights with different wavelengths will be refracted with different angles.

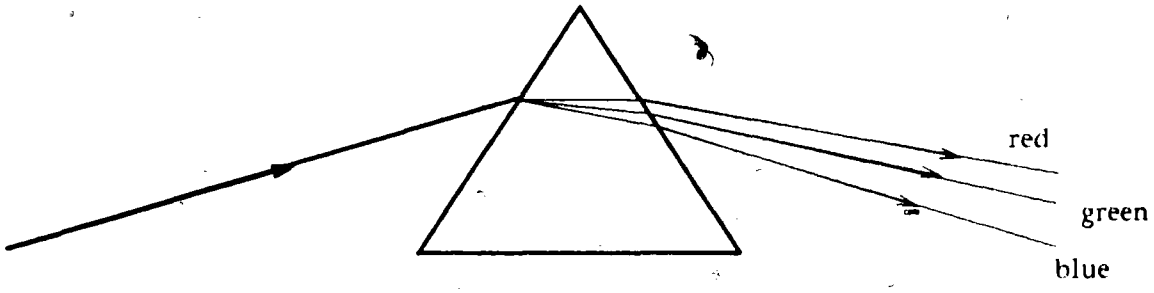


Figure 4-2: Chromatic aberration effect of white light passing through a prism

Since visible white light is composed of lights with all wavelengths from about 400 nm to 700 nm, we can see a spectrum of lights when white light passes through a prism [*fig.* 4-2]. This effect is known as *chromatic dispersion*. See [Klein 70] and [Meyer-Arendt 72] for references.

4.2.2. Index of Refraction

The index of refraction varies continuously with the wavelength. See *fig.* 4-3 for an approximate outline of the refractive index as a function of wavelength. The index function $n(\lambda)$ is normally monotonic, as it is for glass. Therefore, we can prove that the index function of glass has an inverse.

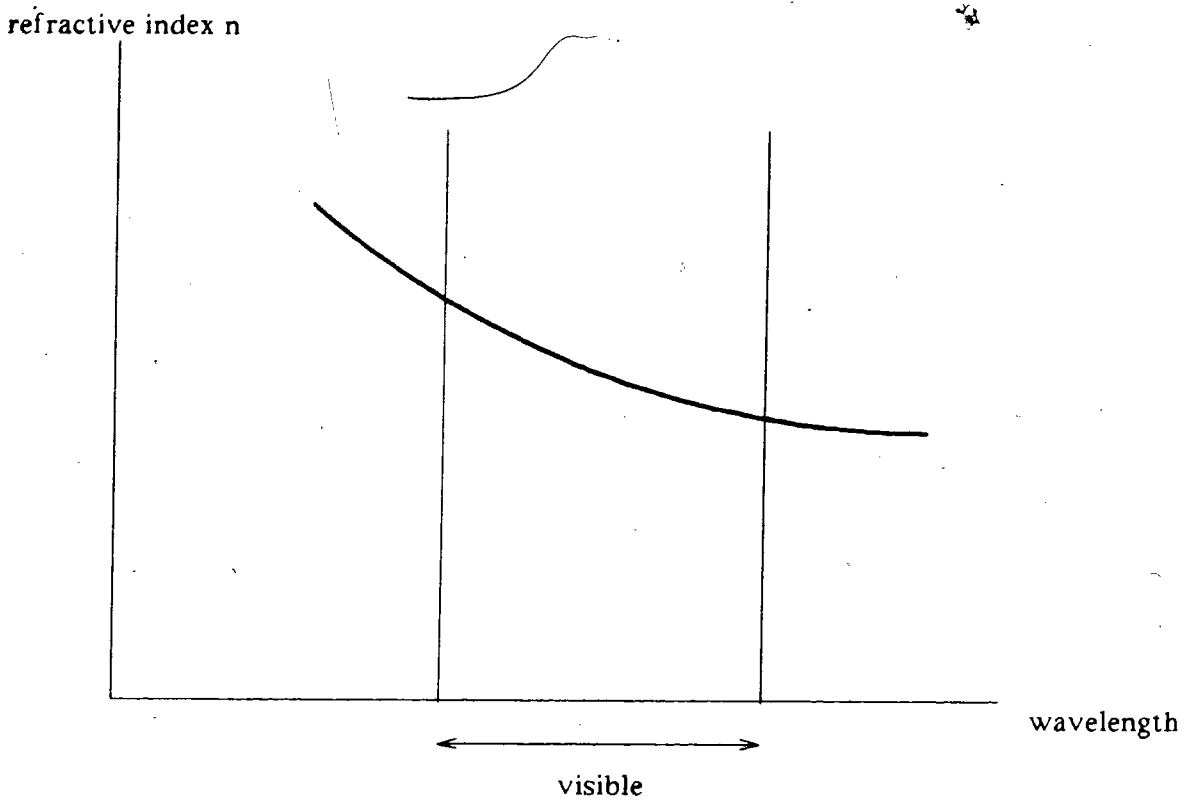


Figure 4-3: Approximation of the refractive index of glass vs. wavelength

4.2.3. Thin Lens

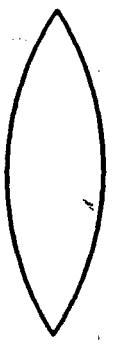


Figure 4-4: Cross-section of a circular double convex thin lens

A camera usually consists of a set of thin convex lenses. However, since the effect of this set of thin lenses is the same as a single thin lens, we must now understand how a thin lens works. (In this paper, the term "thin lens" shall be understood to

mean "convex circular thin lens" [See *fig.* 4-4].) A thin lens consists of two refracting surfaces close enough for their separation to be negligible with respect to the object and image distances and the focal length. A well known law for thin lenses is the **Gaussian thin-lens equation**:

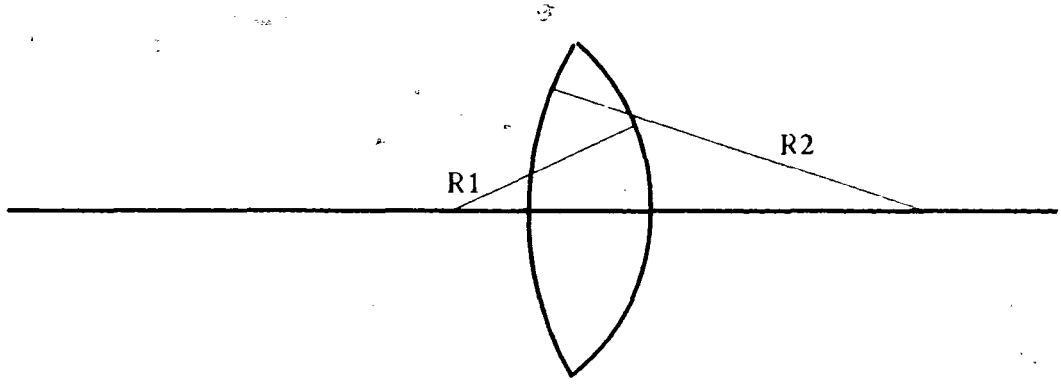
$$\frac{1}{S} + \frac{1}{S'} = \frac{1}{f}$$

where S is the distance between the object and the lens, S' is the distance between the lens and the image of the object and f is the focal length of the lens. If the two surfaces of the lens are fractions of two spheres, then the focal length of the lens is determined by the **Lens Maker's Equation**:

$$\frac{1}{f} = (\frac{n'}{n} - 1)(\frac{1}{R_1} - \frac{1}{R_2}).$$

Here the radii are signed with directions. For the double convex lens, the above equation becomes

$$\frac{1}{f} = (\frac{n'}{n} - 1)(\frac{1}{|R_1|} + \frac{1}{|R_2|}).$$



Convex Lens

Figure 4-5: Circular double convex thin lens with surface radii R_1 and R_2

where R_1 is the radius of one surface while R_2 is the radius of the other [See *fig.* 4-5]. n is the index of refraction of the outside medium, assumed to be vacuum or air here, while n' is the refractive index of glass. In our paper, we suppose the outside medium to be vacuum (*i.e.* $n = 1$ and does not vary with wavelength). Thus, the focal length of a lens depends on the refractive index, and hence on the

wavelength of the incoming light ray. Consequently if the incoming ray is white light, only a fixed wavelength can be in focus while all other different wavelengths will not be in focus.

For a lens whose thickness d is not negligible, the previous equation becomes

$$\frac{1}{f} = (\frac{n'}{n} - 1)(\frac{1}{R_1} - \frac{1}{R_2} + \frac{(n' - n)d}{n'R_1R_2}).$$

4.2.4. Relationship between Wavelength and Focus

As a result of the above discussion, we know that

$$\frac{1}{f} = (\frac{n'}{n} - 1)(\frac{1}{R_1} - \frac{1}{R_2}).$$

If $n = n_0$ and $n' = n(\lambda)$, we can express $n(\lambda)$ as

$$n(\lambda) = 1 + \frac{f(\lambda)n_0}{(1/R_1 - 1/R_2)}.$$

Then we can use the inverse function to find wavelength λ in terms of focal length f.

$$\lambda = \Lambda(f) = N(1 + \frac{fn_0}{(1/R_1 - 1/R_2)}),$$

where $N(y)$ is the inverse function of $n(\lambda)$.

Even with the thick lens equation

$$\frac{1}{f} = (\frac{n'}{n} - 1)(\frac{1}{R_1} - \frac{1}{R_2} + \frac{(n' - n)d}{n'R_1R_2}).$$

It is easy to see that the focal length f strictly increases with the index of refraction n', and hence we can still find the inverse function $N(f)$.

4.2.5. Diffraction

Diffraction can be defined as any "distortion" not accounted for by geometric optics [Meyer-Arendt 72]. For a circular aperture, the result of diffraction will be a series of concentric circles. The energy at any point beyond the first minimum in the diffraction function is small enough to be ignored. We can therefore truncate the spread function at that point. The first minimum occurs at an angular departure of about $0.61\lambda/r$, where r is the radius of the aperture. If the image plane is L from the lens, the first minimum in the image plane will differ by

$m(\lambda) = 0.61\lambda L/r$

from the point predicted by geometric optics.

4.2.6. Spread Function

Without diffraction, we know that a light with wavelength λ will form a circular disk of radius $d(\lambda)$ according to the Gaussian thin-lens Equation.

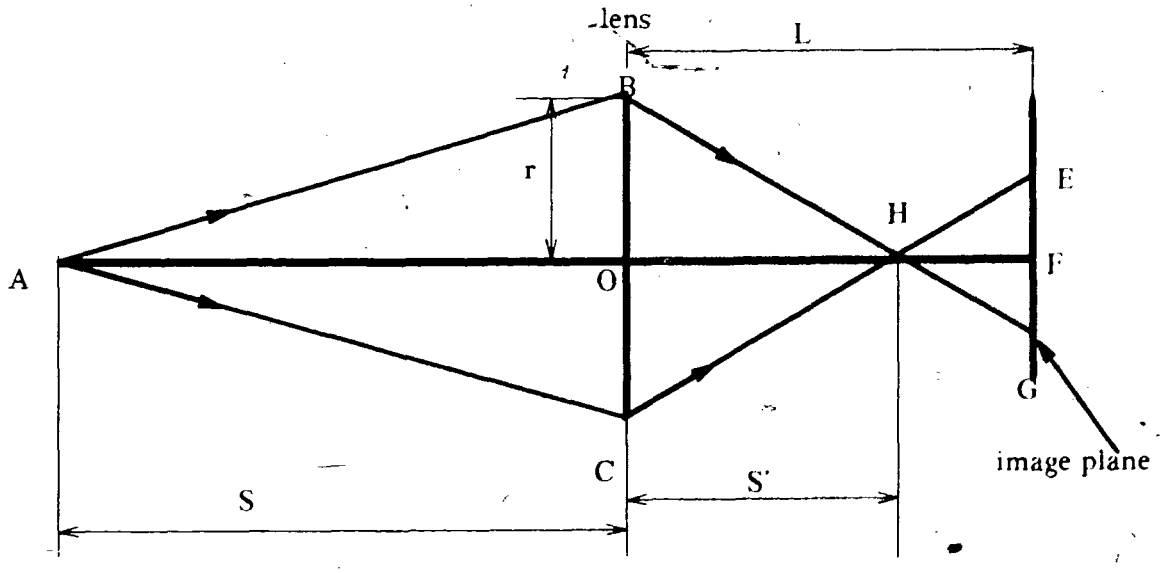


Figure 4-6: This figure shows the case where the image of a point is in front of the image plane

In *fig.* 4-6, S is the distance between the object (or incoming point source) and the

lens. f is the focal length of the lens to the light with λ wavelength. S' is the image distance of the point source. L is the distance between the lens and the image plane. Then we can express the radius of the circular disk in terms of S' and r, where r is the aperture of the lens. We know

$$S' = \frac{1}{1/f - 1/S}$$

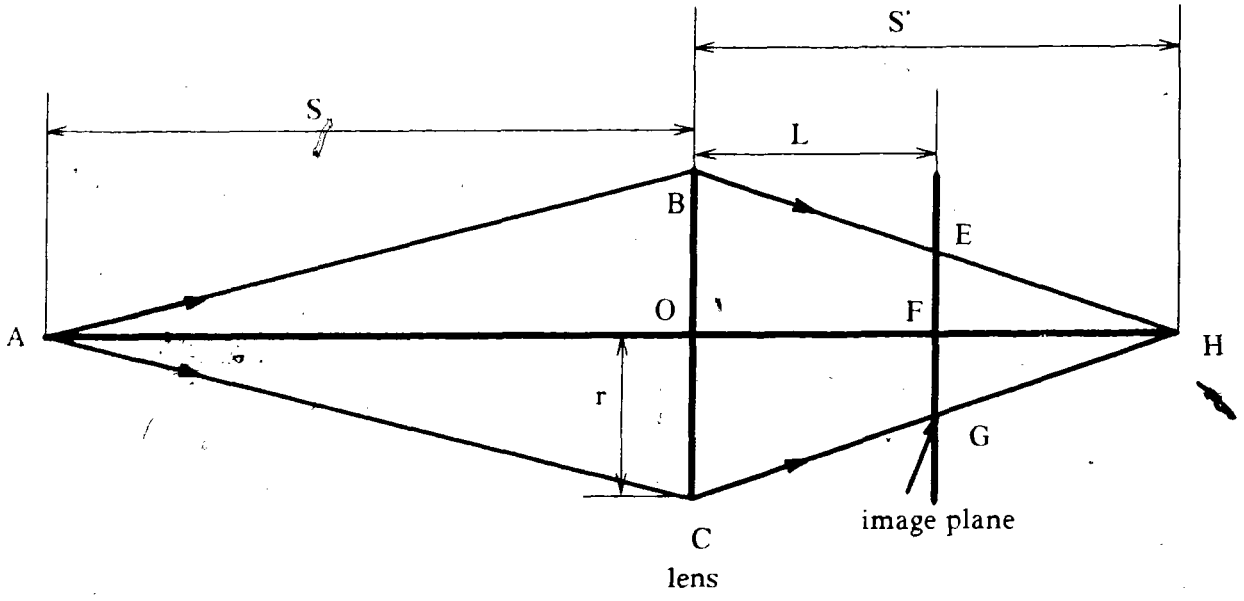


Figure 4-7: This figure shows the case where the image of a point is behind the image plane

There are two cases to be considered. As shown in *fig.* 4-6 and *fig.* 4-7, we can either have $S' \leq L$ or $S' \geq L$. In the case where $S' \leq L$ [*fig.* 4-6], the triangle BHO is similar to the triangle GHF. Thus FG/BO = FH/HO or

$$d(\lambda)/r = (L - S')/S'.$$

That is,

$$d(\lambda) = r(L - S')/S',$$

or

$$d(\lambda) = \frac{r(L(S - f) - fS)}{fS}.$$

In the case where $S' > L$ [*fig.* 4-7] the triangle BHO is similar to the triangle EHF. Thus EF/BO = FH/OH, or $d(\lambda)/r = (S' - L)/S'$. That is

$$d(\lambda) = \frac{-r(L(S - f) - fS)}{fS}$$

where f is a function of λ.

For a distant point source, i.e., for any object distance $S >> r$, the spread function without diffraction effect can be thought of as uniform, because every point receives light of the same intensity.

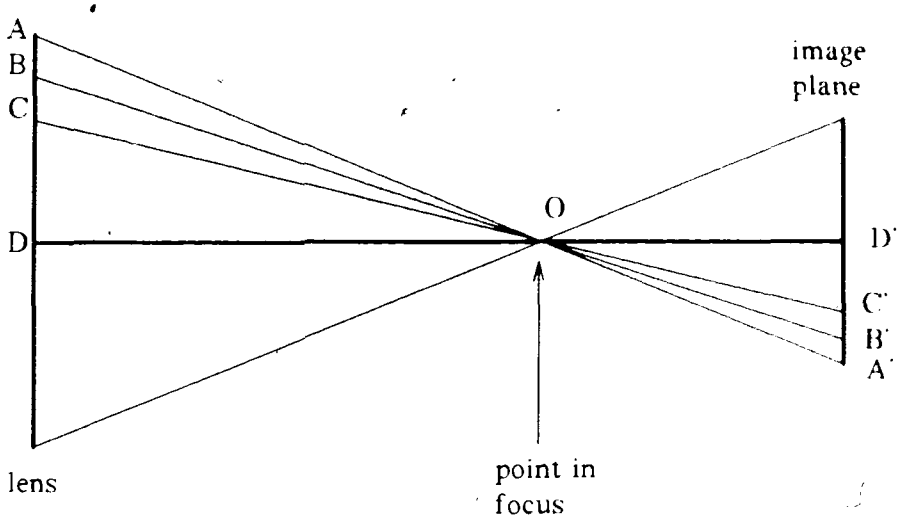


Figure 4-8: For uniform incoming flux, the flux arriving at the image plane will also be uniform

fig 4-8 shows that the spread function is uniform, let us look at *fig* 4-8. This is the one-dimensional case. Assuming that the incoming light energy arrives at the lens uniformly, the density of the energy flux is e. Let points A, B and C in figure 4-8 be the consecutive points in the lens at which the light flux arrives. Since the image plane is parallel to the lens, the triangle ABO is similar to the triangle A'B'O and the triangle BDO is similar to the triangle B'D'O. Thus we have AB/A'B' = BO/B'O and BO/B'O = DO/D'O. That is, AB/A'B' = DO/D'O. Likewise we have BC/B'C' = DO/D'O. Since points A, B and C are consecutive points, we should have AB = BC = $1/e$. Hence A'B' = B'C'. Therefore, the flux density at the image plane should be uniform. This result is independent of the image point being in front of or behind the image plane. The proof of the uniformity of the spread function in the two-dimensional case is similar. We will not prove it here. Therefore, the spread function of wavelength λ without diffraction is:

$$O_\lambda(x) = \begin{cases} 1/[\pi d^2(\lambda)], & x \leq d(\lambda) \\ 0, & x > d(\lambda) \end{cases}$$

When the diffraction effect is considered, the spread function will not be uniform. The spread function due to the diffraction effect is also wavelength dependent. For monochromatic light, the point spread function of wavelength λ when the image is in focus is $O^0_\lambda(x) = [\pi r^2/(\lambda^2 L^2)][2J_1(t)/t]^2$, where r is the radius of the aperture, L is distance between the lens and the image plane, $t = 2\pi rx/(\lambda L)$, d is the distance of the point from the center of the image and $J_1(t)$ is the first-order Bessel function, [Klein 70]. The point spread function $R(x,\lambda)$ for an out-of-focus optical system will be the convolution of $O^0_\lambda(x)$ and the uniform disk with radius $d(\lambda)$.

4.2.7. Cut-off Point for the Point Spread Function

Due to diffraction, the cut-off point will not be the geometric one $d(\lambda)$ as calculated above. Instead, a quantity of $m(\lambda) = 0.61L\lambda/r$ should be added to $d(\lambda)$. Thus, the cut-off point for the point spread function should be

$$D(\lambda) = d(\lambda) + m(\lambda)$$
$$= \frac{r|L[S - f(\lambda)] - f(\lambda)S|}{f(\lambda)S} + 0.61L\lambda/r.$$

If we assume that the image plane is placed at the image point of the shortest visible wavelength λ_{min}, we have

$$L = S(\lambda_{min})$$
$$= Sf(\lambda_{min})/|S - f(\lambda_{min})|.$$

It is obvious that the function $m(\lambda)$ is strictly increasing with wavelength. If the image plane is placed at or in front of the image point of the shortest visible wavelength, the function $d(\lambda)$ will also strictly increase with wavelength. In this case, the cut-off point function $D(\lambda)$ will increase with wavelength. This is one of the reasons why we assume that the image plane is placed at the image point of the shortest visible wavelength.

4.3. Spectrum Information from Chromatic Aberration

In this section, we will use chromatic aberration to find the spectral power distribution of the incoming light. The result of this section is independent of the specific form of the spread functions derived above. The method is valid for any spread function as long as it has the properties described below.

Let us assume that the picture plane is fixed at the image point of either end of the visible spectrum. In order to find the spectral power distribution of the incoming light using chromatic aberration, the spread function must have a finite cut-off.

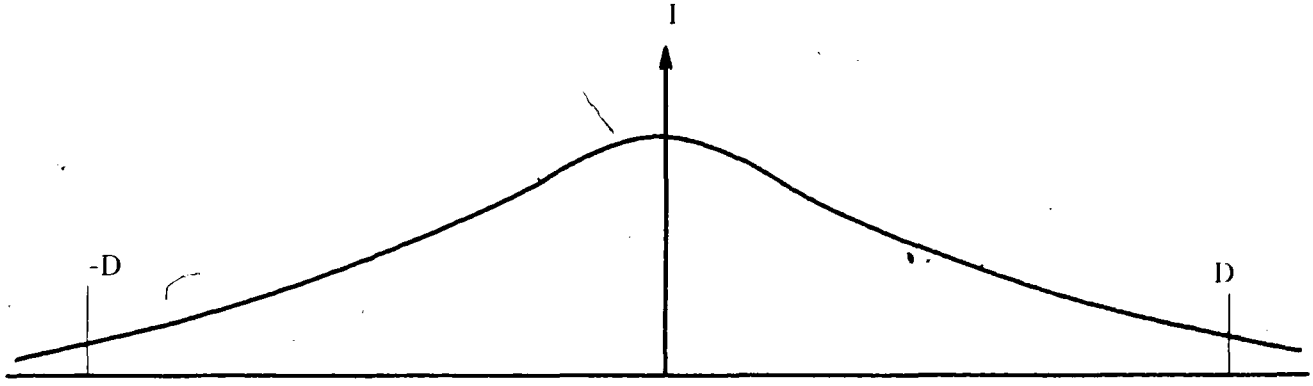


Figure 4-9: This figure shows that at a long distance a spread function is too small to be measured

That is, it must be approximately zero at some sufficiently long distance [See *fig.* 4-9]. This is a reasonable assumption since the intensity at a long distance is small enough not to be measured by a receptor. In addition, the cut-off point must vary monotonically with wavelength. The $D(\lambda)$ is either increasing with wavelength if the picture plane is at the image point of the shortest visible wavelength, or decreasing with wavelength if the picture plane is at the image point of the longest visible wavelength [See *fig.* 4-10]. For convenience, we assume the image plane is located at the image point of the shortest visible wavelength. Thus, the cut-off point function will strictly increase with wavelength.

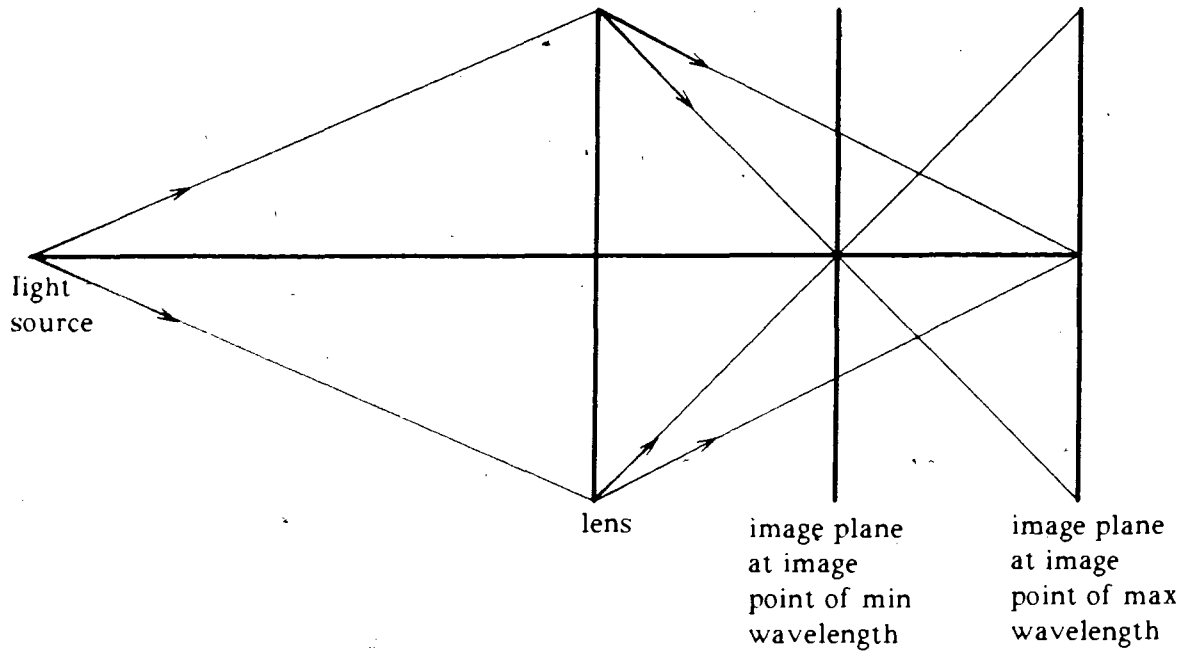


Figure 4-10: Enlarged chromatic aberration effect of a thin lens

We will present the analytical results of the simplest to the most complicated case. We start with the one-dimensional case with a uniform spread function and point light source, and we finally go to the chromatic aberration of an edge with any spread function in two dimensions. Since in some cases it is difficult or even impossible to get a closed-form solution, we propose numerical solutions for those equations (integration equations). In the following discussion, let the visible wavelenghth range be $[\lambda_{min}, \lambda_{max}]$.

4.3.1. Point Source in One Dimension

As in *fig.* 4-11, light from a point source of spectral power distribution $I(\lambda)$ at point O comes through the lens and forms an image in the picture plane, located at the image point of the shortest wavelength λ_{min} of the visible spectrum. So the cut-off point for wavelength λ is $D(\lambda)$ as calculated in the previous section.

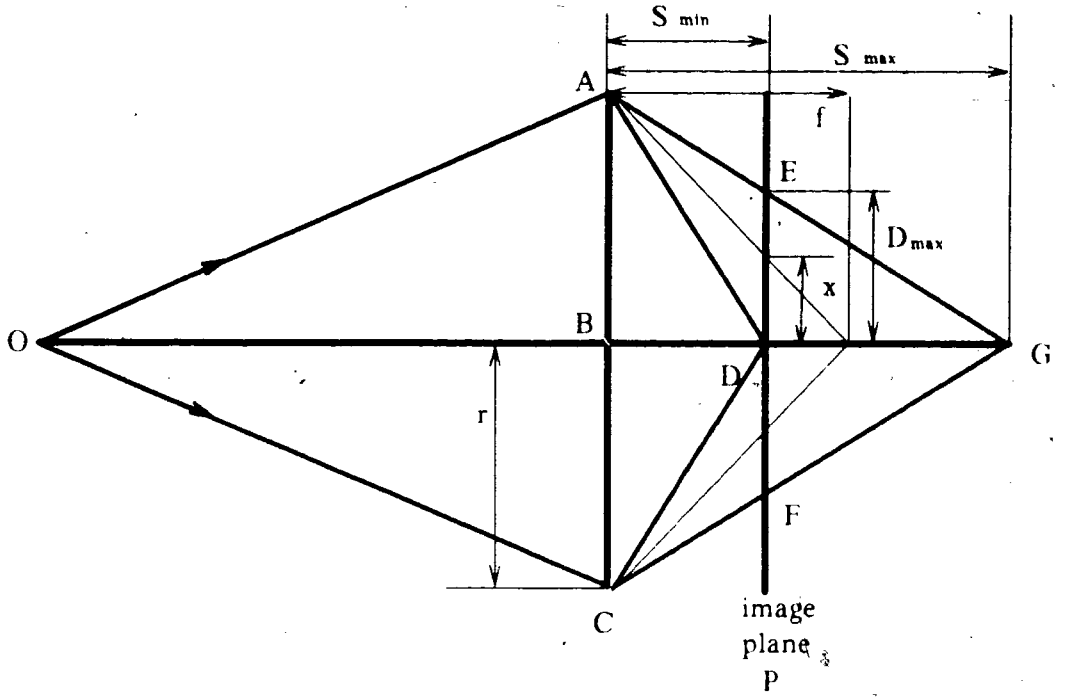


Figure 4-11: Pure chromatic aberration effect of a point source

Let x be a point at the image plane that is located at a distance x from the optical axis. All light with wavelength λ such that $D(\lambda) \geq x$ will affect the intensity of this point. In fact, the intensity at point x consists of all the light $I(\lambda)R(x,\lambda)$ where the function $R(x,\lambda)$ is the defocus spread function for light of wavelength λ at a point with distance x from the center of the image, and $I(\lambda)$ is the spectral power distribution of the incoming light. Because we assume that the cut-off point of the spread function of wavelength λ is $D(\lambda)$, $R(x,\lambda) = 0$ for all the λ such that $D(\lambda) < x$. Since the summation is over the whole visible spectrum, it becomes an integral. And we get

$$p(x) = \int_{-\infty}^{\infty} I(\lambda)R(x,\lambda)d\lambda.$$

where $p(x)$ is the intensity at the point x. Only the visible spectrum $[\lambda_{min}, \lambda_{max}]$ will be in effect for the human visual system, so the integral becomes

$$p(x) = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda)R(x,\lambda)d\lambda.$$

This is an integral equation in which $p(x)$ and $R(x,\lambda)$ are known. Moreover, we want to find $I(\lambda)$.

Because the cut-off function $D(\lambda)$ is strictly increasing in our case, it must have an inverse function. Let $\lambda(D)$ be the inverse function of $D(\lambda)$, then the function $\lambda(D)$ is strictly increasing. Since the $D(\lambda)$ specifies the length of the spread function from the center to one end, the whole length of the spread function is thus $2D(\lambda)$. If the spread function is uniform, i.e., if the function $R(x,\lambda)$ is such that

$$R(x,\lambda) = \begin{cases} 1/[2D(\lambda)], & \lambda \geq \lambda(x) \\ 0, & \lambda < \lambda(x) \end{cases}$$

then we can get a closed-form solution to the integral equation, which becomes

$$p(x) = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda)R(x,\lambda)d\lambda$$

$$= \int_{\lambda(x)}^{\lambda_{max}} I(\lambda)/[2D(\lambda)]d\lambda.$$

Hence, differentiating $p(x)$ we get

$p'(x) = -I(\lambda)\lambda'(x)/[2D(\lambda)]$.

or

$I(\lambda) = -2p'(x)D(\lambda)/\lambda'(x)$.

Substituting $D(\lambda)$ for x, we get

$I(\lambda) = -2p'(D(\lambda))D(\lambda)/\lambda'(D(\lambda))$

where $D(\lambda)$ is as given above. Now the function $I(\lambda)$ is expressed in terms of $D(\lambda)$,

$\lambda(D)$ and $\lambda'(D)$, hence the index of refraction function $n(\lambda)$, its inverse and their derivatives. The solution for the function $I(\lambda)$ is the spectral power distribution of the incoming light, which is what is being sought. Since the function $p(x)$ is decreasing, $p'(D(\lambda))$ is always negative. Hence the negative sign in front ensures that $I(\lambda)$ will be positive.

It is not generally possible to find a solution in closed form for $I(\lambda)$ given an arbitrary spread function $R(x,\lambda)$. The equation then is

$$p(x) = \int_{\lambda_{min}}^{\lambda_{max}} I(\lambda)R(x,\lambda)d\lambda.$$

Using the fact that $\lambda(x)$ is strictly increasing or decreasing, the equation can be solved numerically. If the function $\lambda(x)$ is strictly increasing, that is, the image plane is located at the image point of the shortest visible wavelength, and if the visible wavelength range is divided into n+1 intervals from λ_{min} to λ_{max}, then $\lambda_{min} = \lambda_0 < \lambda_1 < \ldots < \lambda_n = \lambda_{max}$. Due to the monotonically increasing nature of $D(\lambda)$, we have $D(\lambda_0) < D(\lambda_1) < \ldots < D(\lambda_n)$ Let us assume we have the mean measurement of intensity at each interval $[\lambda_{i-1}, \lambda_i]$ as $p(D(\lambda_i))$. In such a case the integration becomes a summation. The equation can be numerically rewritten as follows:

$$p(D(\lambda_n))=I(\lambda_n)R(D(\lambda_n),\lambda_n)(\lambda_n-\lambda_{n-1})$$

$$p(D(\lambda_k)) = \sum_{i=1}^{n} I(\lambda_i)R(D(\lambda_k),\lambda_i)(\lambda_i-\lambda_{i-1})$$

where $1 < k < n$.

The first equation holds because $R(D(\lambda_n),\lambda_i) = 0$ for all $i < n$. The second equation is obtained by changing the integral into summation. Because the spread function cut-off point is $D(\lambda)$, the spread function $R(D(\lambda_k),\lambda_i) = 0$ for all $i < k$. Hence The second equation can be written as:

$$p(D(\lambda_k)) = \sum_{i=k}^{n} I(\lambda_i)R(D(\lambda_k),\lambda_i)(\lambda_i-\lambda_{i-1}).$$

Writing the set of equations in terms of $I(\lambda)$, we get:

$$I(\lambda_n) = p(D(\lambda_n))/\{R(D(\lambda_n),\lambda_n)[\lambda_n - \lambda_{n-1}]\}$$

$$I(\lambda_k) = \{p(D(\lambda_k))- \sum_{i=k+1}^{n} I(\lambda_i)R(D(\lambda_k),\lambda_i)(\lambda_i-\lambda_{i-1})\}/[R(D(\lambda_k),\lambda_k)\{\lambda_k-\lambda_{k-1}\}].$$

We obtain a numerical solution from this recursive expression of $I(\lambda)$. Since the function $p(x)$ usually has discrete values, this set of equations is more useful because of its easy application.

4.3.2. Edges in One Dimension

It is harder to derive the incoming spectral power distribution from an edge than from a point source. So let us assume that the off-center spread function in this case is the same as the spread function at the center.

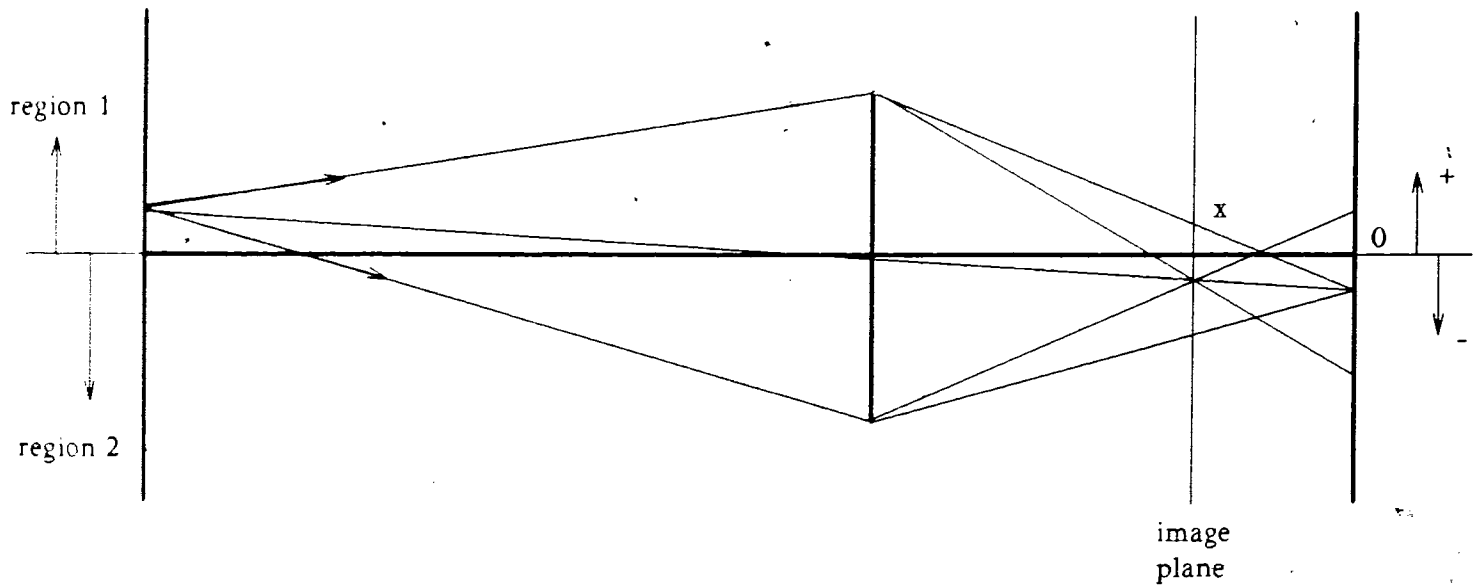


Figure 4-12: The image forming of an off-optical axis point.

(See *fig.* 4-12 for a better understanding.) Let the points in the picture plane above the center be positive and points below negative. At a point x, the intensity depends

on the light emitted from a range of points on the object. The points in region 1 can at most affect a point x in the picture with $x \geq -D(\lambda_{max})$. And the points in region 2 can at most affect a point x in the picture with $x \leq D(\lambda_{max})$. Let $D_{max} = D(\lambda_{max})$. Since all points in the same region will have the same colour, let the spectral power distribution of the reflected light from region 1 be $I_1(\lambda)$ and that from region 2 be $I_2(\lambda)$. Similar to the point source case, we have the following:

For $x > D_{max}$,

$$p(x) = \int_{\lambda_{min}}^{\lambda_{max}} \int_{-D_{max}}^{D_{max}} I_2(\lambda)R(x,\lambda)dxd\lambda.$$

For $x < -D_{max}$,

$$p(x) = \int_{\lambda_{min}}^{\lambda_{max}} \int_{-D_{max}}^{D_{max}} I_1(\lambda)R(x,\lambda)dxd\lambda.$$

For $0 < y \leq D_{max}$,

$$p(y) = \int_{\lambda_{min}}^{\lambda_{max}} \int_{-y}^{D_{max}} I_2(\lambda)R(x,\lambda)dxd\lambda + \int_{\lambda_{min}}^{\lambda_{max}} \int_{-D_{max}}^{-y} I_1(\lambda)R(x,\lambda)dxd\lambda.$$

For $0 > y \geq -D_{max}$,

$$p(y) = \int_{\lambda_{min}}^{\lambda_{max}} \int_{-y}^{D_{max}} I_1(\lambda)R(x,\lambda)dxd\lambda + \int_{\lambda_{min}}^{\lambda_{max}} \int_{-D_{max}}^{-y} I_2(\lambda)R(x,\lambda)dxd\lambda.$$

These are the equations which can be derived from chromatic aberration. The first two equations are quite obvious. For a point $x > 0$ with a distance from the center greater than D_{max}, the intensity at this point is solely due to colour 2. Similarly for a point $x < -D_{max}$, the intensity at that point is solely due to colour 1. When a

point is within the range $(0, D_{max})$, *e.g.*, x, the intensity at this point consists of both colour 1 and colour 2. Therefore, the integral is broken into two parts. The first integral represents the light coming from the region of colour 2 and the second integral represents the light coming from the colour 1 region. We obtain the last equation in the same manner.

Let us now consider the simplest case, in which the spread function is uniform. Since the cut-off point of $R(x,\lambda)$ is $D(\lambda)$ from the center, the whole length of the spread function for λ is $2D(\lambda)$. So

$$R(x,\lambda) = \begin{cases} 1/(2D(\lambda)), & \lambda \geq \lambda(x) \\ 0, & \lambda < \lambda(x) \end{cases}$$

Let $U(\lambda) = 1/(2D(\lambda))$. Assume that the functions $I(\lambda)$ and $R(x,\lambda)$ are so smooth that $I(\lambda)R(x,\lambda)$ has continuous first and second derivatives. In fact, the condition can be relaxed. All that is required is to be able to exchange the places of the integrals in the double integration. So the integral equation can be written as follows:

For $0 < y \leq D_{max}$,

$$p(y) = \int_{-y}^{D_{max}} \int_{\lambda_{min}}^{\lambda_{max}} I_2(\lambda)R(x,\lambda)d\lambda dx + \int_{-D_{max}}^{-y} \int_{\lambda_{min}}^{\lambda_{max}} I_1(\lambda)R(x,\lambda)d\lambda dx$$

$$= \int_{-y}^{D_{max}} \int_{\lambda(x)}^{\lambda_{max}} I_2(\lambda)U(\lambda)d\lambda dx + \int_{-D_{max}}^{-y} \int_{\lambda(x)}^{\lambda_{max}} I_1(\lambda)U(\lambda)d\lambda dx.$$

Taking the first derivative we get

$$p(y) = \int_{\lambda(y)}^{\lambda_{max}} I_2(\lambda)U(\lambda)d\lambda - \int_{\lambda(y)}^{\lambda_{max}} I_1(\lambda)U(\lambda)d\lambda.$$

Taking the second derivative we have

$$p''(y) = -I_2(\lambda(y))U(\lambda(y))\lambda'(y) + I_1(\lambda(y))U(\lambda(y))\lambda'(y).$$

or

$$-I_2(\lambda(y)) + I_1(\lambda(y)) = p''(y)/[U(\lambda(y))\lambda'(y)].$$

That is

$$I_1(\lambda) - I_2(\lambda) = p''(D(\lambda))/[U(\lambda)\lambda'(D(\lambda))]$$

$$= 2p''(D(\lambda))D(\lambda)/\lambda'(D(\lambda))$$

We will get the same result from the last integral equation when a point is in $(-D_{max}, 0)$. Therefore, the chromatic aberration occurring at the edge of two distinct adjacent regions gives us the difference between the spectral power distributions of the incoming lights from the two regions.

We will also obtain the difference of the spectral power distributions of the incoming lights from adjacent regions when the spread function is not uniform. However, the equations have to be solved numerically. We will discuss these equations in the two-dimensional case.

There are two interesting facts concerning the derivatives of the intensity function. The first is $p'(x) = p'(-x)$ for x in $[-D_{max}, D_{max}]$. The second is $p''(x) = -p''(-x)$ for x in $[-D_{max}, D_{max}]$. These equations will also hold in the two-dimensional case. It is easy to see that $p''(0) = 0$. This is consistent with zero-crossing edge finding technique.

4.3.3. Edges In Two Dimensions

Two-dimensional chromatic aberration is a simple extension of the one-dimensional case. Let $D_{max} = D(\lambda_{max})$ as above, and the widths of the colour regions in the picture are assumed to be much larger than D_{max}. (See *fig.* 4-13 for a better understanding of the problem. Region 1 and region 2 are of different colour, $I_1(\lambda)$ and $I_2(\lambda)$ respectively.) As in the one-dimensional case, there are different equations

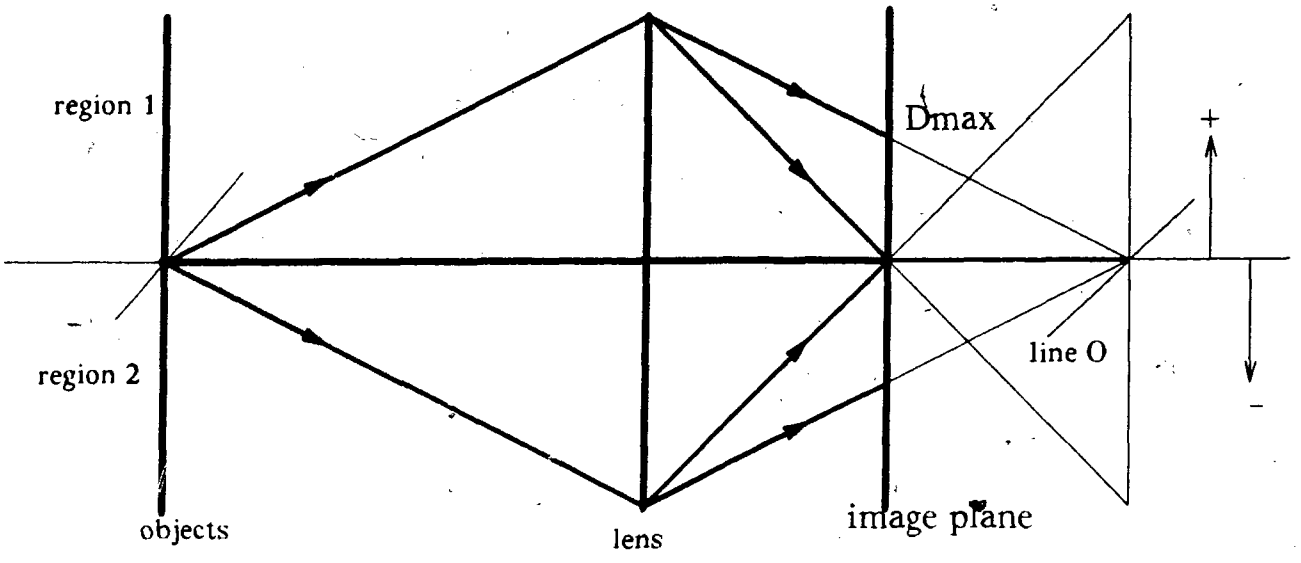


Figure 4-13: Formation of the image in the two-dimensional case

for different points in the picture, depending on the distance from the point to the minimum visible wavelength image of the edge, which is the line O in *fig.* 4-13. Let the λ_{min} wavelength edge image, i.e., line O, be the zero line. Since there is no difference in the horizontal direction, only the distance of a point to this line is significant. However, the intensity of the light at a point is now not just due to a set of points in a line as in the one-dimensional case. It is instead due to a circular disk of points on the object that will affect the point. Thus, we will get a triple integral in the integral equations rather than a double integral as in the one dimensional case. Let s be the distance of the point s from the line O. The equations are:

For $s > D_{max}$,

$$p(s) = \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \leq D^2_{max}} I_2(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda = c_2.$$

For $s < -D_{max}$,

$$p(s) = \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \le D^2_{max}} I_1(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda = c_1.$$

For $0 < s < D_{max}$,

$$p(s) = \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \le D^2_{max} \& x > -s} I_2(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda$$
$$+ \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \le D^2_{max} \& x < -s} I_1(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda.$$

For $-D_{max} < s < 0$,

$$p(s) = \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \le D^2_{max} \& x > -s} I_1(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda$$
$$+ \int_{\lambda_{min}}^{\lambda_{max}} \iint_{x^2+y^2 \le D^2_{max} \& x < -s} I_2(\lambda) R(\sqrt{x^2+y^2}, \lambda) dx dy d\lambda.$$

Note that

$p(s) + p(-s) = c_1 + c_2 = constant.$

Assuming the function $p(s)$ has second derivatives, we have

$p'(s) - p'(-s) = 0,$

or

$p'(s) = p'(-s)$

and

$p''(s) = -p''(-s).$

It is hard to get a closed-form solution to the above set of integral equations even if the spread function is uniform. Hence, we will look for a numerical solution. Let the picture plane be located at the image point of the shortest visible wavelength.

Then the function $D(\lambda)$ is monotonically increasing. Suppose the visible wavelength is divided into $n+1$ values from λ_{min} to λ_{max}. $\lambda_{max} = \lambda_0 > \lambda_1 > \ldots > \lambda_n = \lambda_{min}$. Thus, we have $D(\lambda_0) > D(\lambda_1) > \ldots > D(\lambda_n)$. Here the interval $[\lambda_i, \lambda_{i-1}]$ is assumed to be so small that the spectral power distribution of the incoming light is approximately constant over the interval $[\lambda_i, \lambda_{i-1})$ for $i = 1$ to n. That is, $I_1(\lambda) \approx I_1(\lambda_i)$ and $I_2(\lambda) \approx I_2(\lambda_i)$ for all $\lambda \in [\lambda_i, \lambda_{i-1})$. Here [a, b] means closed interval and (a, b) means open interval. [a, b) means closed at the left end and open at the right end.

After the division of the visible wavelength into a finite number of segments, one can see from *fig.* 4-14 that at the first outer ring only the light of wavelength λ_0 will be present. At the second outer ring, only the light of wavelengths λ_0 and λ_1 will be present, and so on. Let

$\Omega_{00} = \{x^2 + y^2 \leq D^2_{max} \ \& \ y > D(\lambda_1)\}$.

Thus, we have the following solution:

$$p(D(\lambda_1)) = c_2 + [I_1(\lambda_0) - I_2(\lambda_0)]\int_{\lambda_{min}}^{\lambda_{max}} \int\int_{\Omega_{00}} R(\sqrt{x^2+y^2},\lambda)dxdyd\lambda,$$

or

$$I_1(\lambda_0) - I_2(\lambda_0) = [p(D(\lambda_1)) - c_2]/A'_0,$$

where

$$A'_0 = \int_{\lambda_{min}}^{\lambda_{max}} \int\int_{\Omega_{00}} R(\sqrt{x^2+y^2},\lambda)dxdyd\lambda.$$

This equation holds because at point $D(\lambda_1)$, $R(D(\lambda_1),\lambda) = 0$ for all $\lambda < \lambda_1$. Hence the spectral power distribution $I_1(\lambda)$ with $\lambda \leq \lambda_1$ will not affect the intensity of the point $D(\lambda_1)$. $(I_1(\lambda) - I_2(\lambda))$ can be taken out of the integration because $I_j(\lambda)$ is assumed to be approximately constant over the previously divided intervals. The

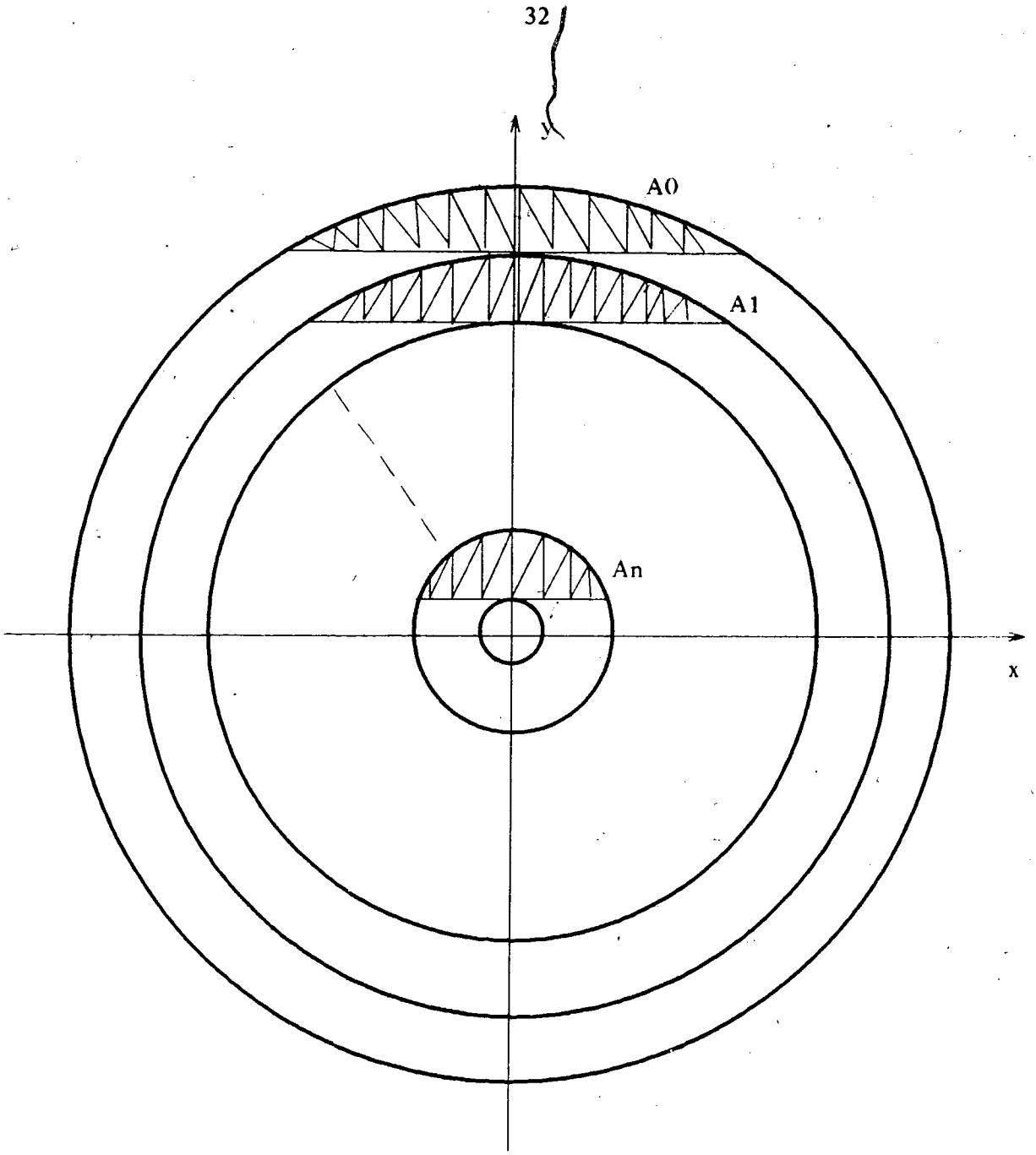


Figure 4-14: Chromatic aberration in the two-dimensional case. The shaded areas are the areas of the A_n's.

spread function $R(r,\lambda)$ is also assumed to be approximately constant with respect to λ over the region $[\lambda_0, \lambda_1)$. This assumption is only made for the convenience of

presentation. Even if it does not hold, the integration equation is still solvable numerically. Given this assumption, $A'_0 = (\lambda_0 - \lambda_1)A_0$, where

$$A_0 = \int\int_{\Omega_{00}} R(\sqrt{x^2+y^2},\lambda_0)dxdy.$$

And the equation for $I_1(\lambda_0) - I_2(\lambda_0)$ becomes

$$I_1(\lambda_0) - I_2(\lambda_0) = [p(D(\lambda_1)) - c_2]/[(\lambda_0 - \lambda_1)A_0].$$

In general, we have

$$I_1(\lambda_k) - I_2(\lambda_k) =$$

$$\{p(D(\lambda_{k+1}))-c_2-\Sigma_{i=0}^{k-1}[I_1(\lambda_i)-I_2(\lambda_i)](\lambda_i-\lambda_{i+1})\int\int_{\Omega_{ki}} R((x^2+y^2)^{1/2},\lambda_i)dxdy\}/[(\lambda_k-\lambda_{k+1})A_k],$$

where

$$\Omega_{ki} \equiv \{x^2+y^2 \leqslant D^2(\lambda_i) \;\&\; x^2+y^2 \geqslant D^2(\lambda_{i+1}) \;\&\; y > D(\lambda_{k+1})\}$$

and

$$A_k = \int\int_{\Omega_{kk}} R(\sqrt{x^2+y^2},\lambda_k)dxdy,$$

with

$$\Omega_{kk} \equiv \{x^2+y^2 \leqslant D^2(\lambda_k) \;\&\; x^2+y^2 \geqslant D^2(\lambda_{k+1}) \;\&\; y > D(\lambda_{k+1})\}$$

and k range from 0 to $n-1$. A_k is the volume of the spread function of wavelength λ_k over the k^{th} darkened area as shown in *fig.* 4-14. The above set of recursive equations provides a way to solve the integral equation numerically. Thus, the difference of the spectral power distributions of the two distinct colour regions can be obtained in this way.

4.4. Discussion

The result of the previous section confirms that the difference of the spectral power distributions of two different incoming lights can be obtained from the chromatic aberration occurring at the edge between these two lights. The algorithm to find the difference of spectral power distributions, given in the previous section, does not depend on the defocus spread function of a given wavelength. For our algorithm to work, certain pausible assumptions have to be made, such as the finite cut-off and the monotonic property of the cut-off points in terms of wavelength.

The main difficulty of using chromatic aberration is that the image must be focused on the shortest visible wavelength, *i.e.* the image plane must be located at the image point of this wavelength. There may be some question as to how such focusing can be achieved. Since all the parameters of the camera are known, if the distance to the object is known, the *Gaussian thin lens* formula can be used to compute the image distance using the focal length of the lens for the shortest visible wavelength; then the image plane is situated at this distance from the lens. If the distance between the object and the lens is not given, infinite distance is assumed. This assumption is usually legitimate, because the distance of the object from the camera will be far greater than the focal length of the camera. Other means can be used to calculate the depth of an object. The literature available on finding depth is quite extensive. Stereo vision, shape from shading, texture, motion, and various other means can be used to find the depth of an object.

4.5. Depth from Chromatic Aberration

Another technique to find the depth of an edge is to use the chromatic aberration information in the image. It is assumed that the image plane is in front of the image point of the shortest visible wavelength, *i.e.*, the image plane distance from the lens is smaller than the distance of the shortest visible wavelength image point from the lens (*fig* 4-15).

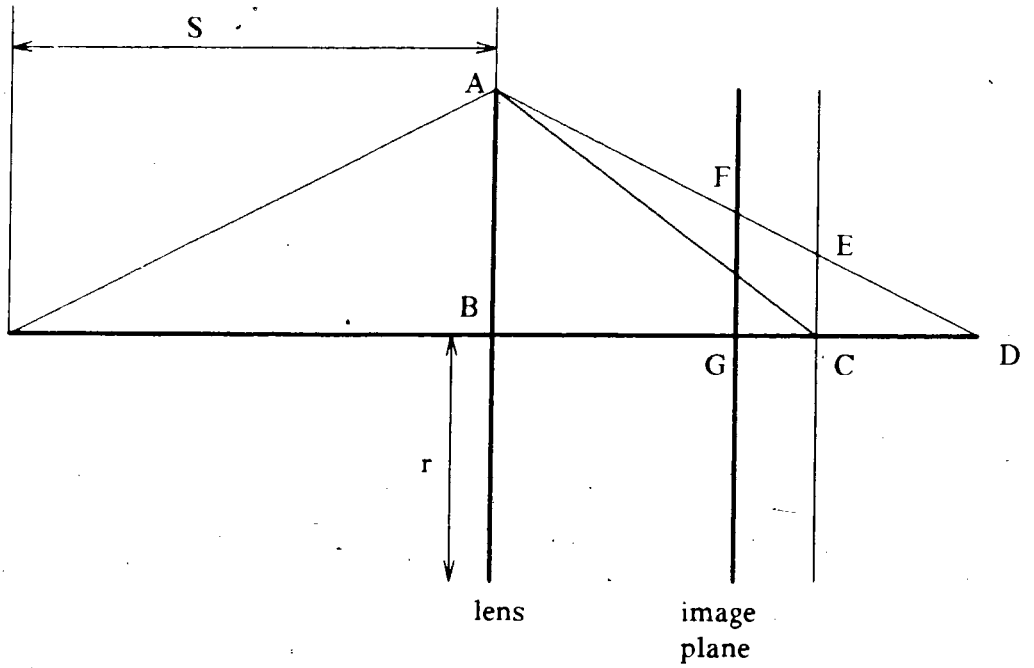


Figure 4-15: Figure for finding the depth of a coloured point or edge from its image of chromatic aberration

Let the distance of a point from the lens be S. Since the triangle ABD is similar to the triangle FGD, we have AB/FG = BD/GD. AB = r = the aperture which is given. GD = BD - BG. Because the image plane distance (BG = L) is known, we have

$$BD = \frac{L}{1 - FG/r}.$$

In the image plane, the intensity beyond the cut-off point of the longest visible wavelength λ_{max} of the incoming spectrum is constant, and between this point and the optical axis is not constant. Hence we can find out the cut-off point of λ_{max}, by

using the first directional derivative of the intensity function along the normal of the edge. Let the distance between the cut-off point and the center of the geometrical image of the point or edge be *a*. The point F in *fig* 4-15 differs from the actual cut-off point by the amount of diffraction $m(\lambda_{max})$ of the wavelength λ_{max}. Thus, we have $FG = a - m(\lambda_{max})$, where $m(\lambda_{max}) = 0.61L\lambda_{max}/r$. So

$$BD = \frac{Lr^2}{r^2 - ra + 0.61L\lambda_{max}}.$$

Since BD is the geometric image distance for the point, we have

$$S = \frac{1}{1/f(\lambda_{max}) - 1/BD}.$$

Because r, L and *a* are known and f is a given function, knowing λ_{max} will provide the object distance S of the point. However, it is not necessary that the incoming light contain the whole visible spectrum. Therefore, different known filters should be used to take the image. We assume that the incoming spectral power distribution contains at least one of the λ_{max}'s of the filters. Since BD is the image distance of the light of wavelength λ_{max}, BD is constant. It is easy to see that $1/f(\lambda)$ decreases with wavelength. Thus, $1/f(\lambda) - 1/BD$, hence 1/S decreases with wavelength if $\lambda_{max} \geq \lambda$, which is certainly true from our assumption. In another words, the calculated object distance is greater than the true object distance if $f(\lambda)$ where $\lambda < \lambda_{max}$ is used in the calculation. As a result, the smallest S calculated by using different filters will provide the object distance of the point. If all of the λ_{max}'s of the filters are within the visible spectrum, using largest of the λ_{max}'s will bring the most reliable results. Therefore, if the incoming spectral power distribution contains at least one of the wavelengths λ_{max} of the filters used, the object distance can be determined with these filters by means of the chromatic aberration effect. In fact, the image plane can also be placed behind the image point of the maximum visible wavelength; then the calculation is almost identical with the one above, except that the λ_{min}'s of the filters are used. This may provide a more accurate result than the

former method because the width of the chromatic aberration effect is much larger and easier to measure.

Pentland's method of finding depth of field using defocusing [Pentland 87] is similar to the technique described here. However, Pentland assumes the spread function for the spectral power distribution of the incoming light to be Gaussian, while our calculation shows that the spread function of a particular wavelength is not Gaussian. Theoretically, the sum of these spread function over the visible wavelength is also not Gaussian. Therefore, our method may provide a more accurate result.

The human visual system selects a particular wavelength to be in focus. Polack demonstrates that the wavelength chosen for focusing by the human eye varies from one end of the visible spectrum to the other, as the fixation varies progressively [See *fig.* 4-16] [Polack 22]. G. Wald, D. R. Griffin, H. H. Emsley, C. Lapicque and I. Borish report that the wavelength in focus for the eye varies from 555nm to 589 nm [Wald 47], [Emsley 52], [Lapicque 73] and [Borish 70]. J. G. Sivak and C. W. Bobier find a variation in the wavelength in focus depending on fixation distance. Their experiments also suggest that the wavelength in focus in the human visual system is determined by a learning process.

In summary, a depth-finding method can be used to calculate the distance of an object to the lens; and this distance can then be used to calculate the distance where the image plane is required to be put. In fact, chromatic aberration can be used as a depth cue.

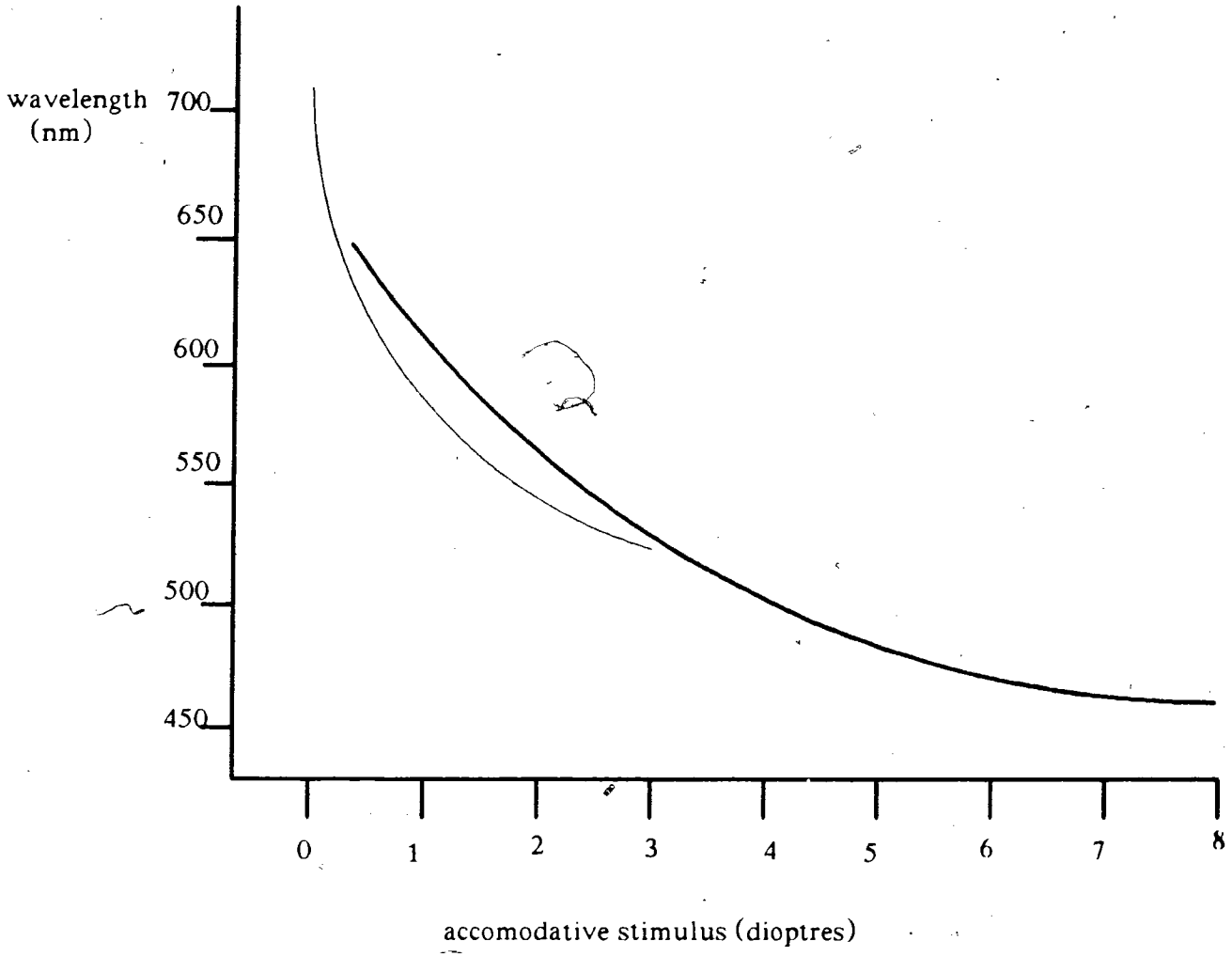


Figure 4-16: Variation of the wavelength in focus on the retina as a function of accommodative stimulus. The thin line represents the mean results obtained by Ivanoff (1953), while the bold line represents the results of Millodot and Sivak (1973). *This figure is reproduced from [Sivak 78].*

4.6. Implementation and Results

The numerical method of solving the integral equation for chromatic aberration has been programmed in C. The formula for thick lenses is used instead of that for thin lenses because the thickness of the lens here used is not negligible. The program has been tested using the real image of an edge. Since there are various technical problems, such as distance measurement, in taking an accurate picture, the picture is a

little distorted from the ideal one. (The photographic profile of the edge is displayed in *fig.* 4-17.)

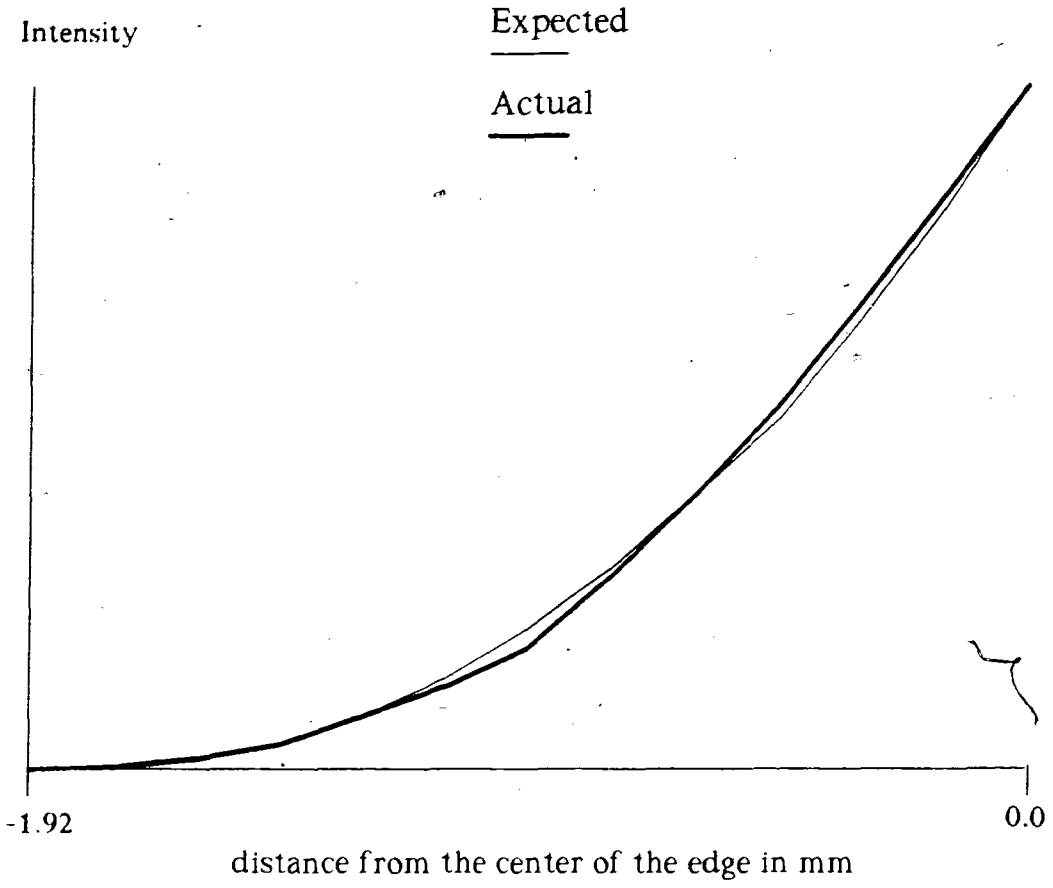


Figure 4-17: Intensity profile of the chromatic aberration effect of an edge

The picture is taken using ILFORD PAN F black and white film and a Nikon FM camera. The lens is a double convex lens made of higher dispersion crown glass. The radius of the lens is 19mm. and the radii of the surfaces of the lens are both 52.9mm. and its thickness is 8.5mm. The illumination we used is tungsten light of 2805°K.

Chapter 5

Colour Constancy

In this chapter, we will show how the chromatic aberration result found in the previous sections helps achieve colour constancy. The approach we use is to model the light and surface spectral reflectance using a finite-dimensional model. The spectral difference provided by chromatic aberration will give us sets of equations and will help us determine the surface spectral reflectance and the illumination.

Here a point in the picture is denoted by a single variable x. At each location, we assume that there are s sensor classes $R_1, \ldots, R_s$. Let $E(\lambda)$ denote the spectral power distribution of illumination, and let $S(\lambda)$ denote the surface spectral reflectance. Then the measurement of the colour signal ρ_i^x at location x with sensor i can be expressed as

$$\rho_i^x = \int E^x(\lambda) S^x(\lambda) R_i(\lambda) d\lambda.$$

where $E^x(\lambda)$ is the spectral power distribution of the ambient light on point x and $S^x(\lambda)$ is the surface spectral reflectance at x.

5.1. Assumptions

In this paper, the surface spectral reflectances and the spectral power distribution of ambient light are assumed to be well described by the finite-dimensional model. In addition, we assume that the illumination (also referred to as "ambient light") remains constant over the whole scene. We also require that the objects lie in a plane

perpendicular to the optical axis of the lens, and that the objects are coloured regions. The edges between these regions are assumed to be sharp.

5.2. Finite-dimensional Models for Illumination and Surface Reflectance

Stiles, Wyszecki and Ohta(1977), Buchsbaum and Gottschalk (1984) show that many surface spectral reflectances are well-approximated by a specific finite-dimensional linear model comprising frequency-limited functions. In 1973, Sallstrom used the finite-dimensional model in colour vision to approximate surface spectral reflectances. Brill, Buchsbaum and Maloney also used this model. Since the publication of Maloney's paper, many computer colour vision researchers have used this model. In his Ph.D. thesis Maloney gives a good justification for the use of finite-dimensional models [Maloney 85].

The finite-dimensional linear model states the following: first, a set of basis functions called $S_i(\lambda)$, with i ranging from 1 to n, are selected. Then, any surface spectral reflectance is represented as the linear combination of these basis functions. Such modeling of the surface spectral reflectance is called the finite-dimensional model for surface spectral reflectance. The constraint we put on the basis functions, as in linear algebra, is that all the basis functions must be linearly independent. That is, if $\psi_1 S_1(\lambda) + \ldots + \psi_n S_n(\lambda) = 0$, all ψ_i's must be zero. The number of basis functions is called the degree of freedom of the linear model. The finite-dimensional linear model for illumination can be similarly defined, except that the basis functions for illumination may differ from those for surface spectral reflectances.

The number of parameters in the finite-dimensional model for surface spectral reflectances is the same as the number of basis functions used, namely, n. If we use m basis functions for the lighting, the number of parameters for the light is m.

The finite-dimensional model of the surface spectral reflectance may not be able to express the surface spectral reflectance of an object exactly. Therefore, we define the error in the finite-dimensional model for the spectral reflectance of an object as the integral of the square of the difference of the real surface spectral reflectance and the approximated function of the model. The parameters of the model should be taken in such a way that the error value is minimal.

5.3. Models of Illumination and Surface Spectral Reflectance

In the following discussion, we take the visible spectrum to be approximately 400nm to 700nm in wavelength. Over the visible spectrum, the surface spectral reflectance curves of natural objects usually are reasonably smooth and continuous. Many experiments on empirical surface spectral reflectances show that the surface reflectances of natural objects are very limited and most of them can be modeled using a few basis functions. Cohen(1964) analyzed the surface spectral reflectances of 150 Munsell chips and computed the characteristic vectors for these chips. The result shows that using the first three of Cohen's characteristic vectors accounts for 99.2 percent of the variance in the fit. Adding another characteristic vector only yields a slight improvement (about 0.5 percent) [Cohen 64]. Maloney(1985) showed that most surface spectral reflectances of natural objects can be modeled adequately by means of three basis functions. (Detailed information can be found in Maloney's Ph.D. thesis [Maloney 85].) Similarly, three to five basis functions suffice in modeling most of the natural daylight.

Let surface spectral reflectance have n basis functions $S_1(\lambda)$, ..., $S_n(\lambda)$, and the spectral power distribution of illumination have m basis functions $E_1(\lambda)$, ..., $E_m(\lambda)$. Now the surface spectral reflectance at point x can be written as

$$S^x(\lambda) = \sum_{k=1}^{n} \sigma_k{}^x S_k(\lambda).$$

Likewise the spectral power distribution of ambient light can be expressed in terms of the basis functions,

$$E^x(\lambda) = \sum_{l=1}^{m} \epsilon_l{}^x E_l(\lambda).$$

Let us have s classes of sensors at each point and let the classes be denoted as $R_k(\lambda)$, k = 1 to s. Hence, within each region we have

$$\rho_k{}^x = \int E^x(\lambda) S^x(\lambda) R_l(\lambda) d\lambda = \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i{}^x \sigma_j{}^x g_{ijk},$$

where $g_{ijk} = \int E_i{}^x(\lambda) S_j{}^x(\lambda) R_k(\lambda) d\lambda$.

5.4. Finding the Colour of a Surface

The measurement of the colour signal at a point taken through a filter is termed the *colour catch*. If there are s classes of sensors at each point, the colour catches ρ_k, k = 1 to s, at each point can be measured. We also assume that the ambient light remains constant over the whole scene. Because $\rho_k{}^x$ is given at each point, there are s equations for the $n + m$ variables of the ϵ_i's and σ_j's within each region. They are

$$\rho_k{}^x = \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i \sigma_j{}^x g_{ijk},$$

where $g_{ijk} = \int E_i(\lambda) S_j(\lambda) R_k(\lambda) d\lambda$ and k = 1 to s. If two regions with different colours are adjacent, there will be $n + n$ variables for the surface spectral reflectances for the two regions and m variables for the spectral power distribution of ambient light. Let the two regions be labeled 1 and 2. Then there are $s + s = 2s$ equations

$$\rho_k{}^x = \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i{}^x \sigma_j{}^x g_{ijk}. \quad \textit{for } k = 1 \textit{ to } s \textit{ and } x = 1 \textit{ or } 2.$$

Given the chromatic aberration result derived earlier, the difference of the two incoming lights $C(\lambda)$ can be determined from the chromatic aberration occurring at the edge. Thus,

$$E(\lambda)S^1(\lambda) - E(\lambda)S^2(\lambda) = C(\lambda).$$

i.e.,

$$\{\sum_{i=1}^{m} \epsilon_i E_i(\lambda)\}[\Sigma_{j=1}^{n} (\sigma_j^{\,1} - \sigma_j^{\,2})S_j(\lambda)] = C(\lambda).$$

Concrete values of the variable λ can be substituted into the function $C(\lambda)$ to get as many equations as one wants, but only independent equations are actually needed. Depending on the dimension of the set of the functions $E_i(\lambda)S_j(\lambda)$'s, a number of independent equations for the variables ϵ_i's, $\sigma_j^{\,1}$'s and $\sigma_j^{\,2}$'s are obtained. The number of independent equations is the dimension of the set of the functions $E_i(\lambda)S_j(\lambda)$'s. Other methods, such as using least squares fitting, can also be used to derive independent equations for the unknowns ϵ_i's and σ_j's. We will address the bound of the dimension of this set below.

5.4.1. Best Case

At best, all the functions $E_i(\lambda)S_j(\lambda)$ are linearly independent. Then we can solve for all $\epsilon_i(\sigma_j^{\,1} - \sigma_j^{\,2})$. Since the unknowns appear as products, it is impossible to solve for absolute values of the ϵ_i's and σ_j's. However, some constraint such as $\Sigma_{i=1}^{m} (\epsilon_i)^2 = 1$ or $\|\Sigma_{i=1}^{m} \epsilon_i E_i(\lambda)\| = 1$ can be used. Here $\|f\|$ stands for the norm of f. Hence, we can solve for ϵ_i's and $\sigma_j^{\,1} - \sigma_j^{\,2}$. In this case, the unknowns ϵ_i's for the spectral power distribution of the ambient light and all pairs of $\sigma_j^{\,1} - \sigma_j^{\,2}$ for the surface spectral reflectances can be recovered. Therefore, there are n equations for the variables $\sigma_j^{\,1}$'s and $\sigma_j^{\,2}$'s from the chromatic aberration. Then using the $2s$ equations from the colour catches within the two regions, s of the $(\sigma_j^{\,1} - \sigma_j^{\,2})$'s can be expressed in

terms of the ϵ_i's and the rest of the $(\sigma_j^1 - \sigma_j^2)$'s. That is, s of the equations for σ_j^1 and σ_j^2's are dependent on the other equations. Therefore, there is a total of $n + 2s - s = n - s$ independent equations for the $2n$ variables σ_j^1's and σ_j^2's. The condition for obtaining a unique solution for the σ's is $2n \leqslant n + s$, or $n \leqslant s$. If, as in the human visual system, there are 3 sensors at each point (i.e. $s = 3$), the degree of freedom of the linear for the surface spectral reflectances can be as large as 3.

5.4.2. Worst Case

At worst, there are still $max\{m,n\}$ number of independent equations based on the information provided by the chromatic aberration at an edge. Since all E_i's are independent, $E_1(\lambda) \not\equiv 0$. Also, because S_j's are independent, all $E_1(\lambda)S_j(\lambda)$ are independent provided $E_1(\lambda)$ has only a finite number of zero points. Thus, the maximum number of independent functions $\geqslant n$. Likewise, the maximum number of independent functions $\geqslant m$. Overall, there are at least $max\{n,m\} + 2s \geqslant m + 2s$ independent equations and $m + 2n$ number of unknowns. The requirement for the feasibility of this set of equations will be $m + 2n \leqslant m + 2s$ or simply $n \leqslant s$. So, with $s = 3$, even in the worst case the degree of freedom of the linear model can be as large as 3 in modeling the surface spectral reflectances.

5.4.3. Algorithm

Usually, the colours of the adjacent regions are so different that the $2s$ equations of

$$\rho_k^x = \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i \sigma_j^x g_{ijk},$$

are independent, where $g_{ijk} = \int E_i(\lambda)S_j(\lambda)R_k(\lambda)d\lambda$. Hence, the solvability of these equations is not an issue. In fact, it does not matter whether these equations are

independent. The functional equation from the chromatic aberration will provide the rest of the information necessary for solving the ϵ_i and σ_j's. The algorithm to solve them is simple. Assume that the functions $E_i(\lambda)S_j(\lambda)$'s are linearly independent. Then use the equation

$$E(\lambda)S^1(\lambda) - E(\lambda)S^2(\lambda) = C(\lambda)$$

or

$$\{\sum_{i=1}^{m} \epsilon_i E_i(\lambda)\}[\sum_{j=1}^{n} (\sigma_j^1 - \sigma_j^2)S_j(\lambda)] = C(\lambda),$$

where the function $C(\lambda)$ is given by the chromatic aberration effect. Substituting a number of different values for the variable λ, we obtain a set of linear equations for the variables $\epsilon_i(\sigma_j^1 - \sigma_j^2)$'s. In the case of $n = 3$ and $m = 3$, nine different values of λ can be used. Solving this set of linear equations will provide all $\epsilon_i(\sigma_j^1 - \sigma_j^2)$'s. However, the finite-dimensional model may not describe the spectral power distribution of ambient light and the surface spectral reflectances with precision, in which case we will not get a satisfactory solution for the variables $\epsilon_i(\sigma_j^1 - \sigma_j^2)$'s. Alternatively, the least squares method or other statistical methods can be used to determine the best fitting solution. Let $\epsilon_i(\sigma_j^1 - \sigma_j^2) = c_{ij}$. Then we will have $\epsilon_i = (c_{ij}/c_{1j})\epsilon_1$, for $i > 1$. In combination with $\Sigma_{i=1}^{m} (\epsilon_i)^2 = 1$ we can solve for the ϵ_i's. After this, all remaining equations for the σ_j^1's and σ_j^2's are linear equations, and standard algorithms can then be applied to solve them.

However, when we solve for ϵ_i's, we have higher order equations instead of linear equations if the condition for ϵ_i's is non-linear. Thus, there will be multiple sets of solutions for the ϵ_i's. But some sets of solutions are not realistic, and thus can be ignored. For instance, using the conditions proposed above, there will be two sets of solutions. However, in one set all ϵ_i's are negative, which is not realistic. Therefore,

we can ignore this solution. In practice, the condition for ϵ_i's should be of as low a degree as possible. Wandell [Wandell 87] proposes a condition for the ϵ vector that $\epsilon_1 = 1$. If ϵ_1 can never be zero, this condition will eliminate the multiple solution problem. In the next section, the dependency of the basis functions for the spectral power distribution of ambient light and for the surface spectral reflectances as well as the problems of multiple solutions arising from the dependency of these functions will be discussed in more detail.

5.4.4. Feasibility of the Solutions

Using the algorithm in the previous section, the spectral power distribution of the ambient light can be determined from chromatic aberration provided all of the $E_i(\lambda)S_j(\lambda)$'s are linearly independent. However, it may not be realistic to assume that $E_i(\lambda)S_j(\lambda)$'s are linearly independent. A better assumption is that the basis functions for the spectral power distribution of ambient light and surface spectral reflectances are the same. In this case, the surface spectral reflectances can still be found, but there are multiple solutions. The multiple solutions have the property that if there is a set of solutions for ϵ_i's and the $(\sigma_j^1 - \sigma_j^2)$'s, then there will also be another set of solutions in which the ϵ_i's correspond to the $(\sigma_j^1 - \sigma_j^2)$'s in the first solution and the $(\sigma_j^1 - \sigma_j^2)$'s of the second solution correspond to the ϵ_i's of the first one.

When the surface spectral reflectance of one region is a multiple of that of the other, absolutely certain solutions for the colour constancy problem cannot be obtained. A simple example is a beam of red light illuminating on a white and a grey patch. This situation is indistinguishable from the situation in which a white light falls on a light red and a dark red patch. Although these situations cannot be distinguished, our previous equations still bring results since the solvability does not depend on the colours of the regions. What one will get in these cases is a finite

number of solutions. The multiple solutions cause confusing circumstances. The same also happens to the human visual system. But the reason for getting a finite instead of an infinite number of solutions is that we use finite-dimensional models which only model a subset of the spectral power distribution of ambient light and of the surface spectral reflectances. For instance, we can either get $x = 1$ and $y = 0$ or $x = 0$ and $y = 1$ as solution to the equation $x + y = 1$, if non-negative integer solutions are desired. However, there is an infinite number of real solutions. A finite-dimensional model models an infinite number of spectral functions, which are only a subset of all possible spectral functions.

If the basis functions for both the spectral power distribution of ambient light and the surface spectral reflectance are the same and the functions $f_i f_j$ are linearly independent for i and j in the range from 1 to n and $i < j$, the number of equations obtained from the functional equation provided by the chromatic aberration effect at the edge is $n + n(n - 1)/2$, where we assume $m = n$. If $m = n = 3$, then f_1f_1, f_2f_2, f_3f_3, f_1f_2, f_1f_3 and f_2f_3 are all linearly independent. Thus, there are six equations, namely the equations for $\epsilon_i x_i$'s, $\epsilon_1 x_2 + \epsilon_2 x_1$, $\epsilon_1 x_3 + \epsilon_3 x_1$ and $\epsilon_2 x_3 - \epsilon_3 x_2$, where $x_i = (\sigma_i^1 - \sigma_i^2)$. Since there are six variables ϵ_i's and $(\epsilon_2 x_3 - \epsilon_3 x_2)$'s and six equations, the set of equations is solvable, but with multiple solutions due to the second order terms in these equations. In general, the illumination and the surface spectral reflectances can still be determined if $n \geq 3$, because the number of equations $n + n(n - 1)/2$ is greater than the number of unknowns $2n$.

Here we will solve the case of $n = 3$ as an example. We first solve for the coefficients of the independent functions $E_i(\lambda)S_j(\lambda)$'s. So we have $\epsilon_i x_i = c_{ii}$ for $i = 1$ to 3, $\epsilon_1 x_2 + \epsilon_2 x_1 = c_{12}$, $\epsilon_1 x_3 + \epsilon_3 x_1 = c_{13}$ and $\epsilon_2 x_3 - \epsilon_3 x_2 = c_{23}$, where $x_i = (\sigma_i^1 - \sigma_i^2)$. Thus, $c_{12}/c_{11} = x_2 x_1 + \epsilon_2/\epsilon_1$, and $c_{11}/c_{22} = \epsilon_1 x_1/(\epsilon_2 x_2)$. These two

equations give $(c_{22}^{*}/c_{11})(\epsilon_1/\epsilon_2) + (\epsilon_2/\epsilon_1) = c_{12}/c_{11}$. We can now solve for ϵ_2/ϵ_1. Since the equation is in the second order, there will be two solutions. From the symmetry of the equation we can derive that if $\epsilon_i = d_i$ and $x_j = e_j$, the other solution is $\epsilon_i = e_i$ and $x_j = d_j$. Once we solve for the ϵ_i's, the rest will be the same as that in the Algorithm section.

After the multiple solutions for the spectral power distribution of ambient light are obtained, they can be verified for consistency with other solutions in the same scene. During the verification process, some of the solutions are eliminated to arrive at a unique solution. Verification can only take place if there are more than two regions of sufficiently different colours in the scene.

5.4.5. Implementation

The algorithm has been implemented using the first three of Cohen's vectors [Cohen 64] for surface spectral reflectance and the first three of Judd's vectors [Judd 64] for spectral power distribution of ambient light. If these particular basis functions are used, all $E_i(\lambda)S_j(\lambda)$ are independent. Therefore, the algorithm for solving the ϵ_i and σ_j's are programmed. Given the difference of the spectral power distribution of the incoming lights and the existence of sensor catches with three sensor classes at each region, the ϵ_i and σ_j's (that is, the colours of the objects) can be recovered.

In practice, finite-dimensional models can not model the spectral power distribution and the surface spectral reflectances with precision. Therefore, satisfactory results cannot be obtained from solving the system of equations which is obtained by substituting different values of λ. Instead, let the visible spectrum of wavelengths be divided into k intervals $\lambda_0 < \lambda_1 < \ldots < \lambda_k$. Assume that the spectral power distribution of the incoming light $I(\lambda)$ is given. Then the least squares method is used to minimize the equation

$$\sum_{l=0}^{k} \{[\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l)][\sum_{j=1}^{m} \sigma_j S_j(\lambda_l)) - I(\lambda_l)\}^2.$$

The set of equations for the minimization are the following:

for t=1 to 3,

$$\sum_{l=0}^{k} E_t(\lambda_l)\{\sum_{j=1}^{m} \sigma_j S_j(\lambda_l)\}[(\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l))(\sum_{j=1}^{m} \sigma S_j(\lambda_l)) - I(\lambda_l)]=0;$$

for t=1 to 3,

$$\sum_{l=0}^{k} S_t(\lambda_l)\{\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l)\}[(\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l))(\sum_{j=1}^{m} \sigma S_j(\lambda_l)) - I(\lambda_l)]=0.$$

A numerical algorithm is used to solve such systems of nonlinear equations. The actual package we use is *minpack*. Since the daylight spectral power distribution is very limited, we can obtain a close approximation for the ϵ_i's, and hence for the σ_j's, which can be taken as the initial values for input into the *minpack* programs. Once we obtain the solution, we also have the approximate spectral power distribution of the illumination and the approximate difference of the surface spectral reflectances. In the next chapter, we will show that the estimated surface spectral reflectance is really a satisfactory approximation if the finite-dimensional models can satisfactorily model the spectral power distribution of the illumination and the surface spectral reflectances.

5.5. Implications

Contrary to normal expectation, no special sensors are needed in analyzing the chromatic aberration effect. All that is required is a light intensity sensor, i.e., a black and white receptor. The interesting part of the finding is that chromatic aberration provides the whole spectral power distribution instead of colour catches, i.e., intensity values of different sensors. Thus, this information is much richer than that obtained by merely using three chromatic classes of sensors. Once the spectral power distribution of the incoming light of a region is known, those of the others in the same scene are easily found.

When there is no chromatic aberration existing in the image, for example, in an image taken through a pinhole camera, other types of distortion, such as diffraction, at the common edge of two regions, can also be used to find the difference of the spectra of the regions. If the aperture decreases, the amount of chromatic aberration decreases, but the amount of diffraction increases. Thus, the technique is not restricted to the presence of a chromatic aberration effect.

By means of the information derived from chromatic aberration, colour constancy can be achieved from only two distinct colour regions with sharp edges. The prerequisites of this achievement are that the illumination is constant over the scene and that the finite-dimensional models are valid for the spectral power distribution of the illumination and for the surface reflectances. Maloney [MALONEY85] requires that there should be a sufficient number of distinct colour regions in the scene. Furthermore, the condition for colour constancy in Maloney's paper is that the degree of freedom of the linear model in modeling the surface spectral reflectances must be less than the number of sensor classes. With three classes of sensors, as in the human

visual system, only two basis functions can be used, which imposes severe limitations. However, if chromatic aberration is used, two distinct colour regions present in the scene will be sufficient. In addition, the number of possible basis functions used to model the surface spectral reflectances can be as large as the number of sensor classes.

Once the surface spectral reflectance of an object is recovered, a colour descriptor which is independent of the illumination is obtained for that object. Therefore, colour constancy is achieved.

Chapter 6

Separation of Illumination and Surface Reflectance from Incoming Light Spectrum

In this chapter, we will show that if the basis functions for the spectral power distribution and surface spectral reflectance have certain properties, it is possible to recover both of the illumination and reflectance from the spectral power distribution of the incoming light. In the following sections, we will prove a number of theorems which justify our claim.

6.1. Exact models

When the spectral power distribution of illumination and surface spectral reflectance are modeled precisely by a finite dimensional model, both the spectral power distribution of the illumination and the surface spectral reflectance can be derived easily and exactly, except for a multiplicative scale factor.

Let the basis functions for the spectral power distribution of the illumination be $E_1(\lambda)$, $E_2(\lambda)$, ... $E_m(\lambda)$, and that for the surface spectral reflectance be $S_1(\lambda)$, $S_2(\lambda)$, ... $S_n(\lambda)$. Let the spectral power distribution of the incoming light be $I(\lambda)$ and that of the illumination be

$$E(\lambda) = \sum_{i=1}^{m} \epsilon_i E_i(\lambda),$$

and let the surface spectral reflectance be

$$S(\lambda) = \sum_{j=1}^{n} \sigma_j S_j(\lambda).$$

Since the spectral power distribution of the illumination and the surface spectral reflectance are precisely modeled, we have $E(\lambda)S(\lambda) = I(\lambda)$. That is,

$$I(\lambda) = \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i \sigma_j E_i(\lambda) S_j(\lambda).$$

If all the product functions $E_i(\lambda)S_j(\lambda)$ are linearly independent, then we can choose *nm* different λ values $\lambda_1, \lambda_2, \ldots, \lambda_{nm}$ so that the matrix M, whose columns consist of $[E_i(\lambda_1)S_j(\lambda_1), E_i(\lambda_2)S_j(\lambda_2), \ldots, E_i(\lambda_{nm})S_j(\lambda_{nm})]^T$, is nonsingular. Hence, when these *nm* values of λ are substituted, *nm* linearly independent equations in terms of $\epsilon_i\sigma_j$'s are obtained. Because of the linearly independent property, the solution for the coefficients $\epsilon_i\sigma_j$'s is unique. If the illumination function $E(\lambda)$ is normalized so that $\epsilon_1 = 1$, then we have unique solutions for the ϵ_i's and σ_j's. Thus, it is possible to derive the exact illumination and the surface spectral reflectance except for a scaling applied to normalize ϵ_1 to 1.

6.2. Approximated Models

In real cases, the finite-dimensional models will not describe the spectral power distribution of illumination and the surface spectral reflectance exactly, but will only approximate the illumination and the surface sprectral reflectance. In mathematical terms, this can be expressed as

$$E(\lambda) \approx \sum_{i=1}^{m} \epsilon_i E_i(\lambda)$$

and

$$S(\lambda) \approx \sum_{j=1}^{n} \sigma_j S_j(\lambda).$$

Thus

$$I(\lambda) \approx \sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i \sigma_j E_i(\lambda) S_j(\lambda).$$

Therefore, the above method can not be applied. We propose to use the least squares fitting of the product of the models to extract the surface reflectance and the illumination. We select the ϵ_i's and σ_j's so that

$$\sum_{\lambda} [\sum_{i=1}^{m} \sum_{j=1}^{n} \epsilon_i \sigma_j E_i(\lambda) S_j(\lambda) - I(\lambda)]^2$$

is minimum, with the condition of $\epsilon_1 = 1$.

In general, the result of the least squares fitting will not be satisfactory. For example, if $f(\lambda)g(\lambda)$ is the result of the least squares fitting of the function $I(\lambda)$, then $[f(\lambda)/c(\lambda)][g(\lambda)c(\lambda)]$ will also be the solution for any non-zero function $c(\lambda)$. In addition, there may be an infinite number of solutions to the least squares fitting. However, if the set of possible illumination distributions and the set of possible surface spectral reflectances satisfy the assumptions stated in the preceding chapter, such fitting will provide satisfactory results. The proof is presented in the next section.

6.3. Theorems

Theorem 1:

Let x_i and y_i, i = 1 to n, be real numbers, then

$$\sum_{i=1}^{n} (x_i y_i) \leq \{\sum_{i=1}^{n} x^2_i\}^{1/2} \{\sum_{i=1}^{n} y^2_i\}^{1/2}.$$

This is a well known theorem. It simply states that the cosine of the angle between the two vectors is less than or equal to 1, and it is known as Cauchy-Schwarz inequality.

Theorem 2:

Let x_i and y_i, i = 1 to n, be real numbers, then

$$[\sum_{i=1}^{n} (x_i+y_i)^2]^{1/2} \leq [\sum_{i=1}^{n} x^2_i]^{1/2} + [\sum_{i=1}^{n} y^2_i]^{1/2}.$$

This theorem is the well known triangle inequality.

Corollary 1:

Define the norm of a function f to be

$$\|f\| = [\sum_{i=1}^{k} f^2(\lambda_i)]^{1/2},$$

where all λ_i's are given distinct values in the domain of the functions. Then

$\|f + g\| \leq \|f\| + \|g\|$.

Corollary 2:

$\|f\| - \|g\| \leq \|f - g\|$

Corollary 3:

$\|f - g\| \leq \|f - h\| + \|h - g\|$.

Fact:

$\|f\| = \|-f\|$.

Given the definition of $\|f\|$, it immediately follows that $\|f\| = \|-f\|$.

Theorem 1, 2 and **Corollary 1, 2** and **3** are well known so we will not give the proofs here. The proofs can be found in most linear algebra textbooks (for example, [Fraleigh 87]).

Theorem 3:

Let x_i and y_i, i = 1 to n, be real numbers, then

$$[\sum_{t=1}^{n} (x^2_t y^2_t)]^{1/2} \leqslant [\sum_{t=1}^{n} x^2_t]^{1/2} [\sum_{t=1}^{n} y^2_t]^{1/2}$$

Proof:

Since $x^2_i \geqslant 0$ and $y^2_i \geqslant 0$,

$$[\sum_{t=1}^{n} x^2_t]^{1/2} [\sum_{t=1}^{n} y^2_t]^{1/2}$$

$$= \{[\sum_{t=1}^{n} x^2_t][\sum_{t=1}^{n} y^2_t]\}^{1/2}$$

$$= \{\sum_{t=1}^{n}\sum_{j=1}^{n} x^2_t y^2_j\}^{1/2}$$

$$\geqslant \{\sum_{t=1}^{n} x^2_t y^2_t\}^{1/2}$$

Corollary 4:

$\|fg\| \leqslant \|f\| \times \|g\|$.

Proof:

The result can be obtained by letting $x_i = f(\lambda_i)$ and $y_j = g(\lambda_j)$ in **Theorem 3.**

Theorem 4:

Let F and G be sets of bounded functions and let $E(\lambda)$ and $S(\lambda)$ be two bounded given functions, and let all functions have the same domain. Let

$$e_F = \mathbf{Min}_{f \in F}\{\|f - E\|\}$$

and let

$$e_G = \mathbf{Min}_{g \in G}\{\|g - S\|\}.$$

If c is the bound for all the funtions in sets F and G and functions $E(\lambda)$ and $S(\lambda)$, then

$$\mathbf{Min}_{f \in F \,\&\, g \in G}\|fg - ES\| \leqslant c(e_F + e_G).$$

Proof:

Since c is the bound for all functions in sets F and G and functions $E(\lambda)$ and $S(\lambda)$, we have $\|f(\lambda)\| \leqslant c$, $\|g(\lambda)\| \leqslant c$, $\|E(\lambda)\| \leqslant c$ and $\|S(\lambda)\| \leqslant c$ for all $f \in F$ and $g \in G$. Let f' in F and g' in G be such functions that $e_F = \|f' - E\|$ and $e_G = \|g' - S\|$ hold. Then

$$\begin{aligned}
&\mathbf{Min}_{f \in F \,\&\, g \in G}\|fg - ES\| \\
&\quad \leqslant \|f'g' - ES\| \\
&\quad = \|f'g' - f'S + f'S - ES\| \\
&\quad = \|f'(g' - S) + (f' - E)S\| \\
&\quad \leqslant \|f'(g' - S)\| + \|(f' - E)S\|
\end{aligned}$$

$$\leq \|f'\|\times\|g' - S\| + \|f - E\|\times\|S\|$$

$$\leq ce_F + ce_G$$

$$= c(e_F + e_G).$$

Corollary 5:

Let $E_i(\lambda)$, i = 1 to m, $S_j(\lambda)$, j = 1 to n, $E(\lambda)$ and $S(\lambda)$ be given bounded functions, and let all functions have the same domain. Let I = [a,b] be a bounded closed interval. Let

$$e_F = \mathbf{Min}_{\epsilon_i \in I}\{\|\sum_{i=1}^{m} \epsilon_i E_i - E\|\}$$

and let

$$e_G = \mathbf{Min}_{\sigma_j \in I}\{\|\sum_{j=1}^{n} \sigma_j S_j - S\|\}.$$

Then there exists some constant c so that

$$\mathbf{Min}_{\epsilon_i,\ \sigma_j \in I}\{\|\sum_{i=1}^{m} \epsilon_i E_i \sum_{j=1}^{n} \sigma_j S_j - ES\|\} \leq c(e_F + e_G).$$

This corollary states that the error in the least squares fitting of the incoming light is proportional to the errors in the individual fitting for the illumination and surface spectral reflectance. If the finite-dimensional linear models closely approximate the illumination and surface spectral reflectance, the error in the least squares fitting of the product of the illumination and the surface spectral reflectance will be very small.

Proof:

Since $E_i(\lambda)$, i = 1 to m, are bounded functions,

$$\sum_{i=1}^{m} \epsilon_i E_i(\lambda)$$

is also a bounded function, for $\epsilon_i \in I$, i = 1 to m. Let

$$F = \{\sum_{i=1}^{m} \epsilon_i E_i(\lambda),\ \epsilon_i \in I,\ i = 1\ to\ m\}.$$

Then F is a set of bounded functions. Similarly

$$G = \{\sum_{j=1}^{n} \sigma_j S_j(\lambda),\ \sigma_j \in I,\ j = 1\ to\ n\}$$

is also a set of bounded functions. The result follows directly from **Theorem 4**.

Theorem 5:

Let $E_i(\lambda)$, i = 1 to m, $S_j(\lambda)$, j = 1 to n, $E(\lambda)$ and $S(\lambda)$ be given bounded functions, and let all functions have the same domain. Let I = [a,b] be a bounded close interval. Let

$$e_F = \mathbf{Min}_{\epsilon_i \in I}\{\|\sum_{i=1}^{m} \epsilon_i E_i - E\|\},$$

where $\epsilon_1 \equiv 1$, and let

$$e_G = \mathbf{Min}_{\sigma_j \in I}\{\|\sum_{j=1}^{n} \sigma_j S_j - S\|\}.$$

Let x_i, i = 1 to m and $x_1 = 1$, and y_j, j = 1 to n, be the such values that

$$\|\sum_{i=1}^{m} x_i E_i \sum_{j=1}^{n} y_j S_j - ES\| = \mathbf{Min}_{\epsilon_i,\sigma_j \in I}\{\|\sum_{i=1}^{m} \epsilon_i E_i \sum_{j=1}^{n} \sigma_j S_j - ES\|\}$$

with $\epsilon_1 = 1$. If all the product functions $E_i(\lambda)S_j(\lambda)$, i = 1 to m and j = 1 to n, are linearly independent, then there exist constants C and D so that

$$\|\sum_{i=1}^{m} x_i E_i - E\| \leq C(e_F + e_G)$$

and

$$\|\sum_{i=1}^{n} y_i S_i - S\| \leq D(e_F + e_G).$$

Proof:

Let u_i, i = 1 to m and $u_1 = 1$, and v_j, j = 1 to n, be the such values that

$$e_F = \|\sum_{i=1}^{m} u_i E_i - E\|$$

and

$$e_G = \|\sum_{j=1}^{n} v_j S_j - S\|.$$

Let

$$P(\lambda) = \sum_{i=1}^{m} u_i E_i(\lambda) - E(\lambda)$$

and

$$Q(\lambda) = \sum_{j=1}^{n} v_j S_j(\lambda) - S(\lambda).$$

Then $\|P\| = e_F$ and $\|Q\| = e_G$. Let

$$f(\lambda) = \sum_{i=1}^{m} x_i E_i(\lambda), \quad g(\lambda) = \sum_{j=1}^{n} y_j S_j(\lambda),$$

$$f'(\lambda) = \sum_{i=1}^{m} u_i E_i(\lambda) \text{ and } g'(\lambda) = \sum_{j=1}^{n} v_j S_j(\lambda).$$

Therefore,

$$\|fg - ES\| = \|fg - (f - P)(g' - Q)\|$$

$$= \|fg - fg' + fQ + g'P - PQ\|.$$

$$= \|fg - fg' - (-fQ - g'P + PQ)\|.$$

Since $\|fg - ES\| \leq c(e_F + e_G)$ for some constant c by **Corollary 5**, we have

$$\|fg - fg' - (-fQ - g'P + PQ)\| \leq c(e_F + e_G).$$

In addition,

$$\sum_{i=1}^{m} \epsilon_i E_i \qquad \text{and} \qquad \sum_{j=1}^{n} \sigma_j S_j$$

are bounded functions. Therefore, there exists d so that

$$\|\sum_{i=1}^{m} \epsilon_i E_i\| \leqslant d$$

and

$$\|\sum_{j=1}^{n} \sigma_j S_j\| \leqslant d.$$

As a result of **Corollary 2** we know that

$$\|fg - f'g'\| - \|(-f'Q - g'P + PQ)\| \leqslant \|fg - f'g' - (-f'Q - g'P + PQ)\|,$$

or

$$\|fg - f'g'\| \leqslant \|fg - f'g' - (-f'Q - g'P + PQ)\| + \|(-f'Q - g'P + PQ)\|.$$

Note that

$$\begin{aligned} &\|(-f'Q - g'P + PQ)\| \\ &\qquad \leqslant \|-f'Q\| + \|-g'P\| + \|PQ\| \\ &\qquad = \|f'Q\| + \|g'P\| + \|PQ\| \end{aligned}$$

Hence,

$$\begin{aligned} \|fg - f'g'\| &\leqslant c(e_F + e_G) + \|Qf'\| + \|Pg'\| + \|PQ\| \\ &\leqslant c(e_F + e_G) + \|Q\| \times \|f'\| + \|P\| \times \|g'\| + \|P\| \times \|Q\| \\ &\leqslant c(e_F + e_G) + d(e_F + e_G) + e_f e_G \\ &= (c + d)(e_F + e_G) + e_f e_G \end{aligned}$$

That is,

$$\|\sum_{i=1}^{m} \sum_{j=1}^{n} (x_i y_j - u_i v_j) E_i S_j\| \leqslant (c + d)(e_F + e_G) + e_f e_G$$

Let

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(x_iy_j - u_iv_j)E_iS_j = H(\lambda).$$

Then

$$\|H\| \leq (c + d)(e_F + e_G) + e_fe_G.$$

Define the vector function

$$V(\lambda) = [E_1(\lambda)S_1(\lambda), E_1(\lambda)S_2(\lambda), \ldots, E_n(\lambda)S_m(\lambda)].$$

The product function $E_i(\lambda)S_j(\lambda)$ will be in column $(i\text{-}1)n+j$, or $[(i\text{-}1)n+j]^{th}$ element of the vector function $V(\lambda)$. We can understand this by substituting $i = 1$ and $j = 1$. $(1 - 1)n + 1 = 1$ means E_1S_1 will be in column 1 of $V(\lambda)$. If $i = m$ and $j = n$, then $(i\text{-}1)n+j = (m\text{-}1)n+n = nm$. So E_mS_n will be in column nm of $V(\lambda)$. Choose nm λ values $\lambda_1, \ldots, \lambda_{nm}$ so that the *nm* vectors $V(\lambda_1), \ldots, V(\lambda_{nm})$ are linearly independent. Thus, the matrix $M = [V(\lambda_1),\ldots,V(\lambda_{nm})]^T$ is not singular, where $[V_1, \ldots, V_k]^T$ is the column vector of $[V_1, \ldots, V_k]$. This is possible because of the linear independency of the E_iS_j's. Substituting the nm values of λ_i's into the previous equations we have nm linearly independent equations, namely

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(x_iy_j - u_iv_j)E_i(\lambda_t)S_j(\lambda_t) = H(\lambda_t), \qquad t = 1 \text{ to } nm.$$

or

$$Mw = A,$$

where w is the column vector $[x_1y_1\text{-}u_1v_1, x_1y_2\text{-}u_1v_2, \ldots, x_my_n\text{-}u_mv_n]^T$ and A is the column vector $[H(\lambda_1), \ldots, H(\lambda_{nm})]^T$. Let $M^{-1} = (b_{i,j})$, where M^{-1} is the inverse matrix of M. Since those t_i's are predetermined,

$$b = \sum_{i=1}^{nm}\sum_{j=1}^{nm}|b_{i,j}|$$

is constant. After solving this system of equations for the coefficients of E_iS_j's, we have

$$w = M^{-1}A,$$

or

$$x_i y_j - u_i v_j = \sum_{l=1}^{nm} b_{(i-1)n+j,l} H(\lambda_l). \qquad i = 1 \text{ to } m \text{ and } j = 1 \text{ to } n.$$

So

$$|x_i y_j - u_i v_j| = |\sum_{l=1}^{nm} b_{(i-1)m+j,l} H(\lambda_l)|$$

$$\leqslant \sum_{l=1}^{nm} |b_{(i-1)m+j,l}| \times |H(\lambda_l)|$$

$$\leqslant \sum_{l=1}^{nm} |b_{(i-1)m+j,l}| \times \|H\|$$

$$\leqslant b[(c+d)(e_F+e_G)+e_F e_G].$$

Since $x_1 = u_1 = 1$,

$|y_j - v_j| \leqslant b[(c + d)(e_F + e_G) + e_F e_G]$, $\quad j = 1$ to n.

Hence,

$$\|g - S\| = \|g - (g' - Q)\|$$

$$\leqslant \|g - g'\| + \|Q\|$$

$$= \|\sum_{j=1}^{n} (y_j - v_j) S_j\| + e_G$$

$$\leqslant \sum_{j=1}^{n} |y_j - v_j| \|S_j\| + e_G$$

$$\leqslant (\sum_{j=1}^{n} \|S_j\|) b[(c + d)(e_F + e_G) + e_F e_G] + e_G.$$

From the above inequality it is obvious that there exists a constant D so that $\|g-S\| \leqslant D(e_F + e_G)$. And it is easy to show that there exists another constant C which allows the inequality $\|f-E\| \leqslant C(e_F + e_G)$ to hold.

6.4. Discussion

From **Theorem 5** in the previous section, we know that if the basis functions $E_i(\lambda)$'s for the spectral power distribution of the illumination and $S_j(\lambda)$'s for the surface spectral reflectances have the property that all $E_i(\lambda)S_j(\lambda)$ are linearly independent, a good approximation for the surface spectral reflectance and illumination can be achieved by using least squares fitting. This is useful when the spectral power distribution of the incoming light is known. In this case, any number of basis functions can be used to model illumination and surface spectral reflectances. If the basis functions have the above independency and the norm of the inverse matrix M^{-1} will not increase with the number of basis functions used, we can have errors in the modeling as small as we wish.

The idea underlying the extraction of illumination and surface spectral reflectance is to find the most plausible separation of illumination and surface reflectance from their product. In general, the functions $f(\lambda)$ and $g(\lambda)$ cannot be derived from their product $f(\lambda)g(\lambda)$, because $f(\lambda)g(\lambda) = [c(\lambda)f(\lambda)][g(\lambda)/c(\lambda)]$ for any non-zero function $c(\lambda)$. This separation can be done only if the illumination and the surface spectral reflectances are very different. The difference between the illumination and the surface spectral reflectances is represented in the linear independency of the functions $E_i(\lambda)S_j(\lambda)$. Since Judd's basis functions for daylight and Cohen's basis functions for surface spectral reflectance have the required property, the illumination and the surface spectral reflectance can be recovered from their product.

6.5. Implementation

The algorithm has been implemented using the first three of Cohen's vectors [Cohen 64] as basis functions for surface spectral reflectances and the first three of Judd's vectors [Judd 64] as basis functions for spectral power distribution of ambient light. If these particular basis functions are used, all $E_i(\lambda)S_j(\lambda)$ are independent. Therefore, the algorithm for solving the ϵ_i and σ_j's are programmed. Given the spectral power distribution of the incoming light, the ϵ_i and σ_j's, that is, the colours of the objects, can be recovered.

Let the visible spectum of wavelength be divided into v intervals $\lambda_0 < \lambda_1 < \ldots < \lambda_v$. In our program, we use samples at intervals of 10nm from 400nm to 700nm. Assume we know the spectral power distribution of the incoming light $I(\lambda)$. Then we will use the least squares method to minimize the equation

$$\sum_{l=0}^{v} \{[\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l)][\sum_{j=1}^{m} \sigma_j S_j(\lambda_l)] - I(\lambda_l)\}^2$$

The set of equations for the minimization are the following:

For k=1 to m,

$$\sum_{l=0}^{v} E_k(\lambda_l)[\sum_{j=1}^{m} \sigma_j S_j(\lambda_l)]\{(\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l))(\sum_{j=1}^{m} \sigma S_j(\lambda_l)) - I(\lambda_l)\} = 0.$$

For k=1 to n,

$$\sum_{l=0}^{v} S_k(\lambda_l)[\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l)]\{(\sum_{i=1}^{n} \epsilon_i E_i(\lambda_l))(\sum_{j=1}^{m} \sigma S_j(\lambda_l)) - I(\lambda_l)\} = 0.$$

A numerical algorithm is used to solve these kinds of systems of nonlinear equations. The mathematical library package *minpack* is used. After solving this set of nonlinear systems of equations we can obtain the solutions for the ϵ_i and σ_j's. Since the daylight spectral power distribution is very limited, we can have a very good guess

for the ϵ_i's, and hence for the σ_j's as the initial values for the input for the *minpack* programs. Given the solution, the approximate spectral power distribution of the illumination and the approximate surface spectral reflectance can be derived. However, the *minpack* program is very unstable. That is, for slightly different initial values, it will produce different solution. Thus, another program which produces much more stable results is used. The algorithm for this program uses the fact that the first m equations above are linear in ϵ_i and the last n equations are linear in σ_j. Hence, given an initial value for the ϵ_i's, we can use the last n equations to solve for σ_j's, then substitute the σ_j's into the first m equations we have another set of ϵ_i's. Iteratively we finally get the convergent solutions. The convergence can easily be proved since all the equations are continuous. Since the viability of our program has been proved for the 370 surface spectral reflectances recorded by Krinov and the daylights by Judd, the detail proof will here be omitted. For the theory concerning the convergence of our method, refer to [Miel 80].

6.6. Results

The results of the algorithm for Krinov's surface reflectances and Judd's daylights are promising. The errors produced by the least squares fittings range from 3.2 to 1.3 times the errors of the individual fitting of the finite-dimensional models. The average is only about twice the error of the individual fitting. This means that the error coefficient C in **Theorem 5** is about 2. With the input as in *fig.* 6-1 which is the spectram of 10000°K daylight illuminating heather, the program will find the illumination (*fig.* 6-3) and the surface reflectance (*fig.* 6-2). *fig.*6-2 plots the spectral reflectance result of the program using the Krinov's surface spectral reflectance (#53) of heather and Judd's 10000°K daylight. *fig.* 6-3 shows the resulting illumination as well as the actual 10000°K daylight.

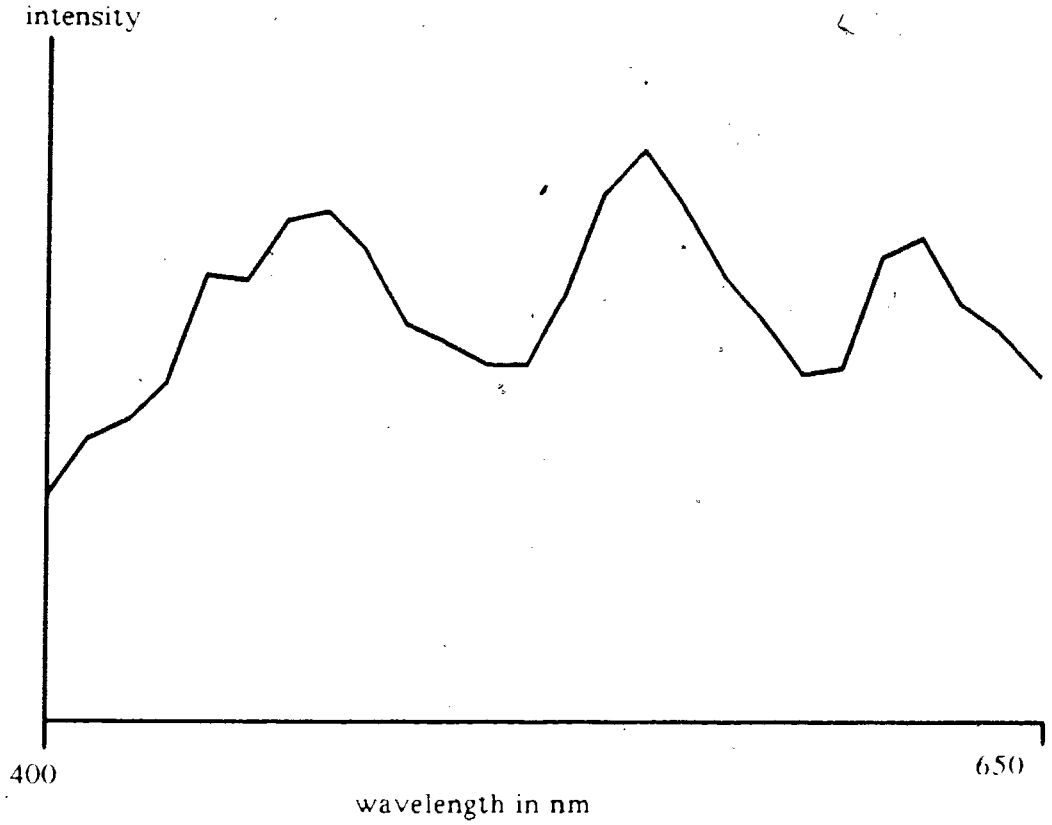


Figure 6-1: Spectrum of 10000°K daylight illuminating heather

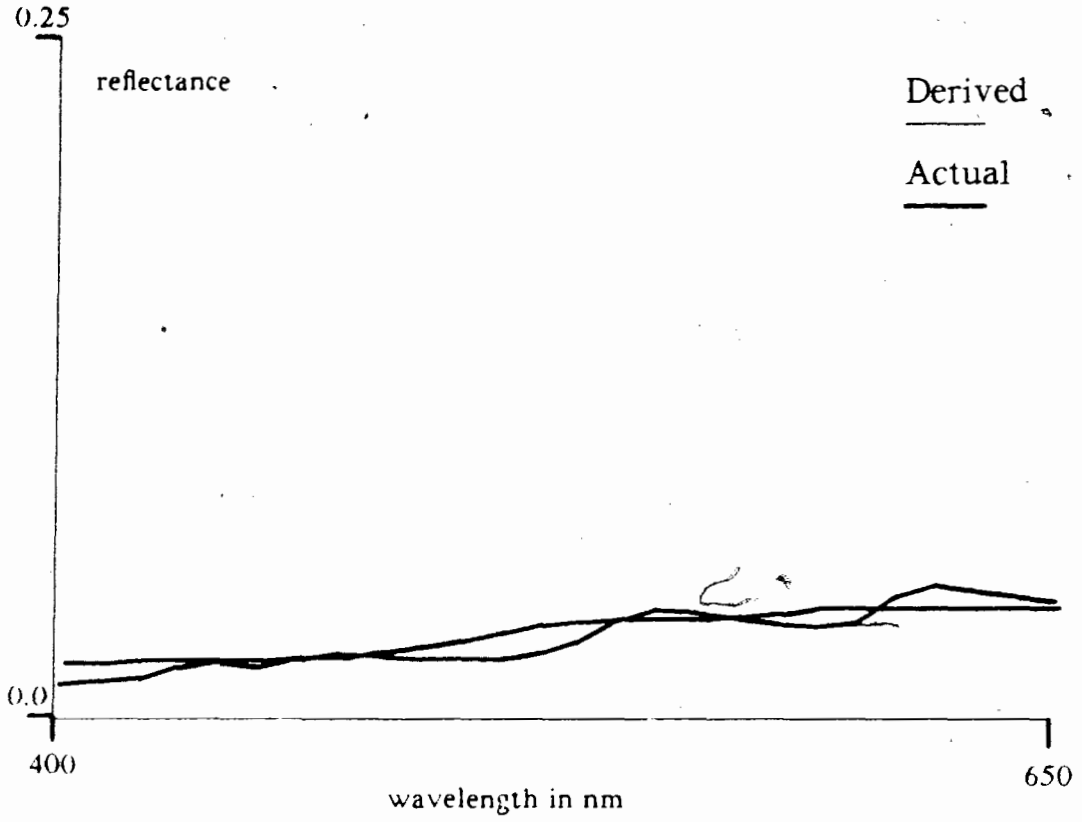


Figure 6-2: Reflectance curve of heather and the fitted curve from our program. The reflectance has been magnified four times.

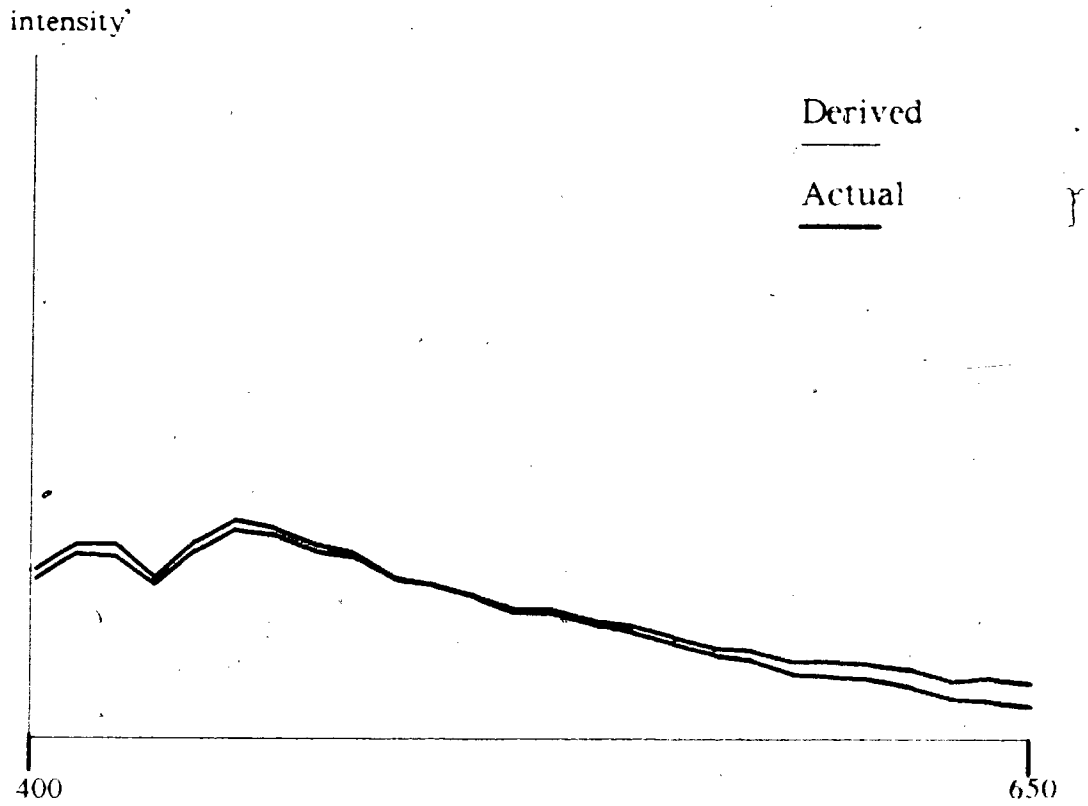


Figure 6-3: Result of the illumination from our program and the actual 10000°K daylight

Chapter 7

Physics and Anatomy of the Human Eye

In this chapter, we want to show the amount of chromatic aberration occurring inside the human eye. The eye's lens has a diameter of about 9 to 10 millimeters and a thickness of about 4 to 5 millimeters. In addition, the lens is about 3 millimeters from the surface of the cornea, and the sclera, choroid and retina together are approximately 1 millimeter thick. The diameter of the human eye is approximately 24 millimeter along the geometric axis [Fatt 78]. Therefore, the distance between the lens and the fovea (the central part of the receptors) is about 24 - 3 - 5/2 - 1 = 17.5 millimeters.

Experiments show that the chromatic aberration occurring inside the human eye is about 1.87 Diopter. If the shorter wavelength end of the visible spectrum is in focus, then the image distance of the other end of the visible spectrum will be S, where S is determined by the following equation:

$$\frac{1}{17.5mm} - \frac{1}{S} = 1.87 \text{ D}$$

or

$$\frac{1}{17.5mm} - \frac{1}{S} = \frac{1.87}{1000mm}$$

That is, $S \approx 18.09$mm approximately. If the radius of the iris opening of the eye is about 1mm (one fifth of the eye's lens), the area of the chromatic aberration effect will be a circle with radius

$$r = \frac{1\times(18.09-17.5)}{18.09}\text{mm} = 0.0326\text{mm}.$$

This area is large in comparison to the size of the rods and cones.

Since the distance between the centers of the cones is about 3μm or 35 seconds of arc [Cornsweet 70], there are about 11 cones along the radius of the circle showing the chromatic aberration effect when the pupil radius is 1mm. The data also implies that the distance between the lens and the fovea is about 17.67 mm. which is consistent with the distance calculated above. The distance may be as small as 20 seconds or 1.71 μm at the very center of the fovea [Cornsweet 70]. Thus, if the chromatic effect occurs at the very center of the fovea, as many as 19 cones along the radius of the circle can be affected. Since there are three classes of sensors inside the eye, there are at least 7 cones of the same class along the radius. With such a large sample size, it is possible to recover the approximate spectral power distribution of the incoming light. The above data provide an indication of the number of samples along the radius of the circle of chromatic aberration effect in the human visual system. In fact, the radius of the pupil is usually greater than 2mm. Hence, the above figures should be doubled.

As a result, we can conclude that the receptors in the human eye are sufficiently dense for measuring the spectral power distribution of the incoming light from chromatic aberration. Although one may expect to see coloured fringes at the edges of white objects, the reason that these fringes are not seen is that the human eye makes the necessary adaptations rather than that the aberration is insignificant [HOWARTH84].

Chapter 8

Conclusion

We can conclude that rich chromatic information can be extracted from chromatic aberration. One application of chromatic aberration is to help determine the colours of objects, *i.e.* to achieve colour constancy. We may also use the spectral information from chromatic aberration to solve further problems in computational vision, for instance, to find the depth of an object. The work presented in this paper may be extended by considering three-dimensional objects and gradual changes of illumination. Moreover, we may continue to explore the information provided by chromatic aberration in combination with other optic phenomena, such as diffraction, to achieve other tasks, such as measuring speed and motion on the basis of changes in colour.

All the calculations in this paper can be applied to "distortions" other than chromatic aberration. For example, in the absence of chromatic aberration, the diffraction effect at the edge of two regions can be used to find the difference of the spectral power distributions of the regions. This can be done because our calculations only rely on the spread function to be wavelength dependent. The best means for working on diffraction is to use a pinhole camera because this type of camera produces very little chromatic aberration. As the amount of chromatic aberration decreases, the amount of diffraction increases, while the spread function still remains wavelength dependent. Hence, ideally we can also get the difference of the spectral power distributions of two regions. Applying the finite-dimensional models we can achieve colour constancy.

References

[Bedford 57] Bedford, R.E. and Wyszecki, G.
Axial Chromatic Aberration of the Human Eye.
J. Opt. Soc. Am. 47:564-565, 1957.

[Borish 70] Borish, I.
Clinical Refraction.
Professional Press, Chicago, 1970.

[Brainard 85] Brainard, D. and Wandell, B.
An Analysis of the Retinex theory of Color Vision.
Technical Report, Stanford Applied Psychology Lab Tech Rep 1985-04
Stanford University, Calif., 1985.

[Cohen 64] Cohen, J.
Dependency of The Spectral Reflectance Curves of The Munsell Color Chips.
Psychon. Sci. 1:369-370, 1964.

[Cornsweet 70] Cornsweet, Tom N.
Visual Perception.
Academic Press, 1970.

[D'Zmura 65] D'Zmura, M. and Lennie, P.
Mechanisms of Color Constancy.
J. Opt. Soc. Am. 3(10):1662-1672, 1965.

[Emsley 52] Emsley, H. H.
Visual Optics.
Hatton Press, London, 1952.

[Fatt 78] Fatt, Irving.
Physiology of the Eye.
Butterworths, 1978.

[Fraleigh 87] Fraleigh, John B. and Beauregard, Raymond A.
Linear Algebra.
Addison-Wesley Publishing Company, 1987.

[Gershon 87] Gershon, R., Jepson, A.D. and Tsotsos, J.
From [R,G,B] to Surface Reflectance: Computing Color Constant Descriptors in Images.
In *Proceedings of IJCAI conference*. 1987.

[Gilmartin 85] Gilmartin, B. and Hogan, R.E.
The Magnitude of The Longitudinal Chromatic Aberration of The Human Eye Between 458 and 633 nm.
Vision Research 25:1747-1753, 1985.

[Howarth 86] Howarth, Peter Alan and Bradley, Arthur.
The Longitudinal Chromatic Aberration of The Human Eye, and Its Correction.
Vision Research 26:361-366, 1986.

[Judd 64] Judd, Deane B., MacAdam, David L. and Wyszecki, Gunter.
Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature.
J. Opt. Soc. Am. 54(8):1031-1040, August 1964.

[Juetz 85] Juetz, J.; Sincerbox, G.T.; Yung, B.H.
Spectral Range Sensing Using Dynamic Chromatic Aberration.
IBM Tech. Disclosure Bull 27(9):515, Feb. 1985.

[Klein 70] Klein, Miles V.
Optics.
John Wiley & Sons, Inc., 1970.

[Land 71] Land, E.H. and McCann, J.J.
Lightness and The Retinex Theory.
J. Opt. Soc. Am. 61:1-11, 1971.

[Land 74] Land, E.H.
The Retinex Theory of Color Vision.
Proc. R. Inst. 47:23-58, 1974.

[Land 83] Land, E.H.
Recent Advances In Retinex Theory and Some Implications for Cortical Computations: Color Vision and The Natural Image.
Proc. Natl. Acad. Sci. USA Physics 80:5163-5169, 1983.

[Lapicque 73] Lapicque, C.
La Formation des Images Retiniennes.
Revue d'Optique, 1973.

[Malacara 75] Malacara, D.
An Observation That Demonstrates The Chromatic Aberration of The Eye.
Boltin Del Instituto De Tonantzintla 1(5):301, 1975.

[Maloney 85] Maloney, Laurence T.
Computation Approaches To Color Constancy.
PhD thesis, Stanford University, 1985.
Applied Psychology Laboratory.

[Meyer-Arendt 72]
Meyer-Arendt, Jurgen R.
Introduction to Classical and Modern Optics.
Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1972.

[Miel 80] Miel, George.
An Updated Version of the Kantorovich Theorm for Newton's Method.
Technical Report, MRC Technical Summary Report #2125, University of Wisconsin, 1980.

[Pentland 87] Pentland, Alex Paul.
A New Sense for Depth of Field.
IEEE Trans. on Pattern Analysis and Machine Intelligence
PAMI-9(4):523-531, July 1987.

[Polack 22] Polack, A.
Bull. Soc. d'Ophtal.
France, 1922.

[Sivak 78] Sivak, J.G. and Bobier, C.W.
Chromatic Aberration of The Eye and Retinoscopy.
In *Proceedings of ICO-11 conference.* 1978.

[Wald 47] Wald, G. and Griffin, D. R.
The Change in Refractive Power of The Human Eye in Dim and Bright Light.
J. Opt. Soc. Am. 37:321-336, May, 1947.

[Wandell 87] Wandell, Brian A.
The Synthesis and Analysis of Color Images.
IEEE Transactions on Pattern Analysis and Machine Intelligence
PAMI-9(1):2-13, January 1987.