

NASA-TM-86844 19860001310

The Synthesis and Analysis of Color Images

Brian A. Wandell

September 1985

LIBRARY COPY
OCT 16 1985
LANGLEY RESEARCH CENTER
LIBRARY, NASA
HAMPTON, VIRGINIA

NASA
National Aeronautics and
Space Administration



The Synthesis and Analysis of Color Images

Brian A. Wandell, Stanford University, Stanford, California

September 1985



National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

N86-10777#

The Synthesis and Analysis of Color Images

Introduction

The two primary physical factors that influence the color appearance of objects in a still image are the spectral power distribution of the ambient light and the surface spectral reflectance of the objects in the image. The information one wishes to display in a synthetic image, and the information available to analyze in during image acquisition, is called the *color signal*. The color signal is the product of the light and surface functions. In this paper I will report on a common representational framework for the synthesis and analysis of color images based upon the synthesis and analysis of the color signal.

In the *synthesis* of color images we begin with specifications of the surface spectral reflectance function and ambient light spectral power distribution. From these we calculate the expected color signal. We then seek the intensities of the color display guns such that the display will have the same effect on the human visual system as the actual color signal would have. In the *analysis* of color images the sensor responses in the acquired image are determined by the color signal. We then seek a method of recovering the underlying surface spectral reflectance and ambient light spectral power distribution functions that gave rise to the color signal.

The methods for synthesis of color images are fairly straightforward. The application of the synthesis methods is most significant for the problem of accurate rendering of color images. The methods for analysis of color images are useful both for the correct rendering of color images, and for the more general vision problem of identifying surface properties of objects from image data. Color image analysis can be applied to color rendering in the important case in which the spectral sensitivities of the sensors used in a camera or film are not within a linear transformation of the spectral sensitivities of the photopigments in the human eye. In this rather common situation, it is impossible to use the information in the camera to infer what the response of the human eye would have been had it been at the same spot as the camera. If one can use the camera information in order to analyze the color image into the surface and light functions, however, the estimates may be used to construct a version of the scene that will appear correct to the human viewer. The potentially greatest significance of color image analysis for computer vision generally is that it offers an estimate of a surface property -- the surface spectral reflectance function -- from image data.

In the next section I will develop the basic notation and matrix operations used in the synthesis and analysis of color images. I will then describe the procedure for color synthesis, showing how we may relate the surface spectral reflectance functions, ambient light spectral power distribution, human photoreceptor responses, and display intensities on a monitor. In the third section I will treat the analysis of color images.

I will describe a version of Maloney and Wandell's (1985; Wandell and Maloney, 1984; Maloney, 1984) method for recovering the spectral power distribution of the ambient light and surface spectral reflectance functions of the objects from the sensor responses to the color signal.

The basic theoretical material will be presented using general notation. The basic results apply to a wide class of representations of the ambient light and surface reflectance functions. In the final section I will turn to the problem of selecting specific representations of these functions that are efficient with respect to a wide class of naturally occurring objects and lights.

List of Symbols

Symbol	Symbol Definition
λ_n	Sample wavelength value
$C^x(\lambda_n)$	Color signal at location x
$E^x(\lambda_n)$	Ambient light spectral power distribution at x
$E_i(\lambda_n)$	Basis function of ambient light representation
ϵ_i	Scalar weight of i^{th} ambient light basis function
ϵ	Vector of ambient light weights
Ω_E	Ambient lighting matrix with respect to wavelength basis
Λ_E	Ambient lighting matrix with respect to arbitrary basis functions
$S^x(\lambda_n)$	Surface reflectance at x
$S_j(\lambda_n)$	Basis function of surface reflectance
σ_j^x	Scalar weight of j^{th} surface basis function at x
σ^x	Vector of surface weights
$R_k(\lambda_n)$	Spectral sensitivity of k^{th} sensor
ρ_k^x	Scalar response of k^{th} sensor at x
ρ^x	Vector of sensor responses at x
π	Best perpendicular to the sensor response vectors
$D(E)$	Dimensionality of the ambient light representation
$D(S)$	Dimensionality of the surface representation

Preliminary considerations

The asymmetry between light and surface

Suppose that we acquire information at sample wavelengths λ_n for $n = 1, N$. The color signal is obtained by multiplying the spectral power distribution of the ambient light at x , $E^x(\lambda_n)$, times the spectral reflectance function of the surface spectral reflectance at x , $S^x(\lambda_n)$. We denote the color signal at point x as $C^x(\lambda_n) = E^x(\lambda_n)S^x(\lambda_n)$.

In developing the mathematical formulation for the synthesis of color images, we can retain symmetry between the role played by the surface and light functions without penalty. This symmetry is evident in the definition of the color signal, $C^x(\lambda_n) = E^x(\lambda_n)S^x(\lambda_n)$, and can be maintained throughout (cf. Sallstrom, 1973; Buchsbaum, 1980). In most images, however, there is a physical asymmetry between the surface spectral reflectance and the ambient light spectral power distribution (cf. Horn, 1974; Oppenheimer and Schafer, 1975). The asymmetry arises since the *rate of spatial variation* of the ambient light spectral power distribution, $E^x(\lambda_n)$, is generally lower than the rate of spatial variation in the surface spectral spectral reflectance function, $S^x(\lambda_n)$. The slow spatial variation of the spectral power distribution makes it possible to represent the ambient light at lower spatial resolution than the surface functions, permitting some additional computational savings.

The most significant consequence of the difference in the rate of spatial variation, however, occurs for the *analysis* of color images. The uniqueness results (Maloney, 1984) for the recovery of the surface and light spectral power distributions from the color signal indicate that we cannot uniquely recover the light and surface functions if both vary freely over space. Essentially unique recovery is possible only if one of these functions varies more slowly over space than the other. We adopt the assumption that spatial variation of the ambient light is slower than the spatial variation of the surface reflectance function (cf. Horn, 1974; Oppenheimer and Schafer, 1975).

To evaluate the plausibility of this idea, the reader might notice that in natural scenes in which the ambient light varies more rapidly spatially than the underlying surface -- for example when viewing a slide projected on a screen -- the human visual system interprets the spatial variation as if it were due to a surface variation, and not from its true cause, the ambient light variation. The analysis we will use here also assumes that there is an asymmetry between light and surface functions: In any local region of the image we assume there is only spatial variation in the surface reflectance function, and no spatial variation in the ambient light spectral power distribution. We will indicate this assumption in our notation by suppressing the superscript x when we refer to the parameters of the spectral power distribution of the ambient light. Thus,

we will write $E(\lambda_n)$ instead of $E^x(\lambda_n)$, so that in a small region of the image $C^x(\lambda_n) = E(\lambda_n)S^x(\lambda_n)$.

I emphasize that this assumption does not imply that the ambient light remains constant over the entire image, but only that the spectral power distribution remains constant over a local region of the image in which there is significant spatial variation in the surface spectral reflectance function.

Synthesis of color images

Introduction

In generating color images we wish to determine the appropriate display gun intensities so that the display will have the same *visual effect* as the actual color signal, $C^x(\lambda_n)$. In their important work on color standards, the members of the CIE have spent more than 50 years describing various methods and representations for achieving this result. The various methods and representations (using XYZ values, YIQ values, etc.) are logically equivalent, but have sometimes been adopted because of odd technical limitations in the computation. For purposes of the analysis here, I prefer to use a calculation advocated by Cornsweet (1970) based on the photoreceptor absorptions of the human eye. This is the method most closely based on the physical factors governing color perception. Although the method is logically equivalent to the several methods suggested by the CIE, and has an excellent physical interpretation, no standard photoreceptor representation has been adopted by the CIE. The estimates by Smith and Pokorny (1975) are widely used at present.

Basic equations: wavelength representation

To determine the display gun intensities that will appear the same as the color signal to the human eye, we first compute the effect that the color signal, $C^x(\lambda_n)$, would have on the three types of photoreceptors in the human eye. Suppose that the sensitivity of the receptors in the eye are described by the three sensitivity functions $R_k(\lambda_n)$ for $k = 1, 3$. If we have measurements of the surface reflectance function and spectral power distribution function at N wavelength values, for λ_n , $n = 1, N$, then the strength of the k^{th} receptor response at location x is

$$\rho_k^x = \sum_{n=1}^N C^x(\lambda_n) R_k(\lambda_n) \quad (1)$$

If there are N sample points along the wavelength dimension, the calculation requires $2N$ multiplications per pixel. To compute the three receptor responses requires $6N$ multiplications.

The effect of a setting the intensity values of the three color guns on the receptors can be calculated as follows. Suppose that the red, green and blue display guns have spectral emission functions $G_j(\lambda_n)$ for $j = 1, 3$. Denote the intensity of the j^{th} gun as I_j . When the display gun intensities at pixel x are set to the vector of value $I^x = (I_1^x, I_2^x, I_3^x)$ the response of the k^{th} class of photoreceptors will be

$$\rho_k^x = \sum_{j=1}^3 I_j^x \sum_{n=1}^N G_j(\lambda_n) R_k(\lambda_n) \quad (2)$$

Equation 2 can be written as a matrix equation of the form

$$\rho^x = \Gamma I^x \quad (3)$$

where the k, j^{th} entry of the matrix Γ is

$$\sum_{n=1}^N G_j(\lambda_n) R_k(\lambda_n) \quad (4)$$

The gun intensities that are visually equivalent to the color signal may be determined by first solving equation 1 for the values of ρ^x and then solving equation 3 for I^x . Two issues arise immediately in computing the display gun intensities.

First, the solution may have negative values. This will occur when the required color image is outside of the gamut of colors that can be displayed by the particular device. I will discuss a method of treating this problem later in the paper when the theory has been more fully developed.

Second, there is the issue of computational efficiency. The entries of the matrix Γ are fixed by the display gun spectral emission functions and the human photoreceptor sensitivity curves. This matrix need only be calculated once and therefore does not represent a significant computational burden. The calculation of the receptor responses, ρ^x to a color signal $C^x(\lambda_n)$ must be made for every point in the image. This is the most expensive part of the computation in the synthesis of color images. The cost of this computation is governed by the number of sample points along the wavelength dimension for the functions, $E^x(\lambda_n)$ and $S^x(\lambda_n)$. Since the amount of computation is proportional to the sampling rate, N , one method of limiting the computation is to reduce the sampling rate, N .

A second option is to transform the representation of the surface spectral reflectance and ambient light spectral power distribution functions from a wavelength representation to a more efficient representation. Great savings in efficiency can be obtained because the wavelength functions describing naturally occurring lights and surfaces are highly correlated at nearby wavelengths. To improve computational efficiency we must seek representations of the spectral power distributions of the lights and spectral reflectance functions of the surfaces that are de-correlated. We will consider specific representations later in this paper. The general effect of applying such transformations is analyzed in the following section.

Basic equations: transformed representation

In order to efficiently encode ambient light and surface vectors we will seek alternative representations, related to the wavelength representation by a linear transformation. These take the form

$$E(\lambda_n) = \sum_{i=1}^{D(E)} \epsilon_i E_i(\lambda_n) \quad (5)$$

and

$$S^x(\lambda_n) = \sum_{j=1}^{D(S)} \sigma_j^x S_j(\lambda_n) \quad (6)$$

where $D(E)$ is the chosen dimensionality of the light representation and $D(S)$ is the dimensionality of the surface representation. Generally in order to obtain efficient representations we will seek sets of functions $E_i(\lambda_n)$ or $S_j(\lambda_n)$ that are independent in the sense that

$$\sum_{n=1}^N E_i(\lambda_n) E_j(\lambda_n) = 0 \quad (7)$$

when $i \neq j$. We refer to these sets of functions $E_i(\lambda_n)$ and $S_j(\lambda_n)$ as *basis functions*. The usual wavelength representation is equivalent to choosing a vector of

weights with respect to the set of functions $E_i(\lambda_n)$ such that

$$E_i(\lambda_n) = 1 \text{ if } i = n \text{ and } 0 \text{ otherwise} \quad (8)$$

and $S_j(\lambda_n)$ such that

$$S_j(\lambda_n) = 1 \text{ if } j = n \text{ and } 0 \text{ otherwise} \quad (9)$$

We can express the relationship between the vector of wavelength sample measurements that characterize the ambient light, $E(\lambda_n)$, and the column vector of weights that characterize the transformed representation of the ambient light, $\epsilon = (\epsilon_1, \epsilon_2, \dots)$ as

$$E(\lambda_n) = E\epsilon \quad (10)$$

The columns of the matrix E are equal to the basis vectors $E_i(\lambda_n)$. Similarly we can express the relationship between the vector of sample measurements of the surface reflectance that characterizes the surface with respect to wavelength, and the vector of weights that characterize the surface with respect to the transformed representation $\sigma^x = (\sigma_1^x, \sigma_2^x, \dots)$ as

$$S^x(\lambda_n) = S\sigma^x \quad (11)$$

The columns of the matrix S are the basis vectors $S_j(\lambda_n)$. To compute the matrix required to transform from one representation to another one may use the linear equations 10 and 11 to compute the vector in either wavelength or the transformed domain. When the vectors in the two domains are linearly independent and of the same dimension, a transformation is always possible without loss of information. To improve computational efficiency we would like to choose the dimensionality in the transform domain to be smaller than the dimensionality in the original representation ($D(E) < M$ and $D(S) < N$), but with little loss of precision. When we choose a smaller dimensionality in the transformed domain than in the wavelength domain, the equations may have no exact solution. In that case we will always choose *the least squares solution* to equations 10 and 11 as our representation.

Relationship among physical quantities

In order to compute the response of a receptor ρ_k^x from the surface and light functions we use the equation:

$$\rho_k^x = \sum_{n=1}^N E(\lambda_n) S^x(\lambda_n) R_k(\lambda_n) \quad (12)$$

1

Using matrix notation we can express the relationship between the sensors, the ambient light, and the surface reflectance function in the form

$$\rho^x = \Omega_E S^x \quad (13)$$

When the surface is represented with respect to the wavelength basis terms, $S_i(\lambda_n) = 1$ if $i = n$ and 0 otherwise, then kj^{th} entry of the matrix Ω_E is

$$E(\lambda_j) R_k(\lambda_j) \quad (14)$$

The relationship between the surface, the light, and the sensor responses may be characterized as a simple matrix product. The surface vector is mapped by a linear transformation into the vector of sensor responses. The linear transformation is determined by two factors: the ambient light, $E(\lambda_n)$, and the receptor sensitivities,

¹ In general, the effective surface reflectance at a point in the image may depend upon the geometry of the image. For example, specular surface reflections depend strongly upon the angles between the lighting source, the surface, and the viewer. Although we do not specifically indicate such dependencies, we assume that the surface reflectance function $S^x(\lambda_n)$ will generally depend upon the viewing geometry. Since our analysis here is for a single fixed image, we simply assume that the surface spectral reflectance has been properly specified for the particular viewing geometry. An analysis of the effects of specularly at the physical level may be found in Torrance and Sparrow (1967; Torrance et al, 1966). An analysis of the effects of incorporating more sophisticated specular effects in computer generated images may be found in Cook and Torrance (1981). I will consider the issue of specularly again in the section on the selection of basis functions.

$R_k(\lambda)$. Across viewing conditions for a particular camera, or for the human eye, the receptor sensitivities are unchanging and known. Thus the only unknown factor that governs the linear transformation, Ω_E is the ambient light spectral power distribution. For that reason we make E explicit when we write down the matrix representing the linear transformation. We will call matrices such as Ω_E *lighting matrices*. We reserve the symbol Ω for the special case in which the lights and surfaces are represented with respect to the sampled values in the wavelength domain.

The relationship between the surface, lights and sensors takes on a parallel form when we transform the surface representation to a new set of basis vectors. If we use

$$S^x(\lambda_n) = \sum_{j=1}^N \sigma_j^x S_j(\lambda_n) \quad (15)$$

then matrix equation becomes

$$\rho^x = \Omega_E S \sigma^x = \Lambda_E \sigma^x \quad (16)$$

where the columns of the matrix S are simply the N vectors $S_j(\lambda_n)$. The matrix S is fixed and we have incorporated the change of basis of the surface representation into the lighting matrix to define a new lighting matrix, $\Lambda_E = \Omega_E S$. The k_j^{th} entry of Λ_E is now

$$\sum_{n=1}^N E(\lambda_n) S_j(\lambda_n) R_k(\lambda_n) \quad (17)$$

This formula generalizes the formula in equation 14 for which the surface representation is which $S_j(\lambda_n) = 1$ if $j = n$ and 0 otherwise. We will use the expression 17 to define the entries of the lighting matrix in general.

The matrix Λ_E can be further decomposed into a set of fixed primitive matrices that depend only upon the choice of basis functions. First consider the decomposition of the lighting matrix with respect to the wavelength basis. We may

express the lighting matrix as the sum of elementary matrices

$$\Omega_E = \sum_{n=1}^N E(\lambda_n) \Omega_n \quad (18)$$

The k_j^{th} entry of matrix Ω_n is

$$S_j(\lambda_n) R_k(\lambda_n) \quad (19)$$

If we express the spectral power distribution of the ambient light with respect to a different set of basis vectors, $E_i(\lambda_n)$ then decomposition of Λ_E becomes

$$\Lambda_E = \sum_{i=1}^{D(E)} \epsilon_i \Lambda_i \quad (20)$$

where the k_j^{th} entry of Λ_n is

$$\sum_{n=1}^N E_i(\lambda_n) S_j(\lambda_n) R_k(\lambda_n) \quad (21)$$

The expression in 21 is with respect to an arbitrary set of basis vectors. This generalizes the expression in equation m in which the basis vectors for the wavelength representation are $E_i(\lambda_n) = 1$ if $i = n$ and 0 otherwise.

Summary of results

The set of matrix operations that capture the relationship between these various quantities may be summarized as follows. Using the rule that the effect of the display on the receptor responses should equal the effect of the surfaces and light on the

receptors we may write

$$\Gamma I^x = \Lambda_E \sigma^x \quad (22)$$

Inverting Γ and using the expansion of the lighting matrix from equation 20 we have

$$I^x = \Gamma^{-1} \Lambda_E \sigma^x \quad (23)$$

If we wish to make the role of the ambient light parameters explicit we may also write this equation in the form

$$I^x = \left[\sum_{n=1}^{D(E)} \epsilon_n \Gamma^{-1} \Lambda_n \right] \sigma^x \quad (24)$$

Equation 23 expresses the complete relationship between the surface spectral reflectances, light spectral power distribution, and display gun intensities.

We may assess the computational burden as follows. In the typical color display we can control the intensities of three color guns, thus the vector I^x is three dimensional. The matrix Γ^{-1} is 3 by 3 and its values are fixed by the display phosphors and human photopigment sensitivities. The selection of the basis functions to represent the surfaces and lights completely fixes the matrices Λ_n (equation 21). These matrices are 3 by $D(S)$, so that the product of $\Gamma^{-1} \Lambda_n$ is 3 by $D(S)$, and may be pre-computed. They do not represent a significant computational burden.

Thus we see that the amount of computation depends primarily on the dimensionality of the surface representation $D(S)$. This number plays the same role as the number of sample points on the wavelength dimension, N , when the equations are written with respect to the wavelength representation. Significant computational savings can be achieved when the representation with respect to an alternative set of basis functions permits us to use a value of $D(S)$ that is significantly smaller than the value of N .

Display gamut limitations. When the display monitor cannot generate the desired signal, one of the required gun intensities, I^x , will be negative. In general such an outcome cannot be entirely avoided since the span of colors obtainable by present technology is certainly less than the span of colors obtainable choosing arbitrary

surface and light functions. In general, the range of display intensities that may be required is determined by the range spanned by the columns of the matrix

$$\Gamma^{-1} \Lambda_E \quad (25)$$

Once the surface and lighting basis functions are chosen, the only freedom left in determining the span of the columns of this matrix are the weights of the ambient light spectral power distribution, ϵ_i .

By suitable choice of the light and surface basis functions, however, the number of instances in which these gamut errors occur can be greatly limited (cf. Buchsbaum and Gottschalk, 1984). In particular, if these functions are restricted to be bandlimited, the corresponding color signal will be bandlimited and generally fall within a restricted color gamut.

In representing objects outside of the color gamut of the display one may choose two methods. First, one may simply adjust the individual point outside of the gamut, for example by setting the negative value to zero. This method distorts that color representation of that point alone. This solution may be adequate, depending upon the significance of the color of that particular portion of the image.

An alternative method of solution is to adjust all of the display values together, attempting to move all of the colors to within the display gamut. This causes some error at each point in the display, but preserves the relationship among the different objects. A reasonable method of making this type of adjustment is to adjust the value of the lighting vector, ϵ_i , in order to avoid negative display intensities.

The advantage of this type of correction derives from the psychophysical observation that human observers are reasonably good at discounting the appearance of the ambient light when judging the surface appearance of objects (cf. Helson, 1938). This method of color correction in the synthesis of images will be evaluated more fully elsewhere.

The display of color camera data. When the synthesis of the color image on the screen depends upon data acquired from a color camera, rather than data generated by an internal model, the relationship between the camera responses and the desired display values may be rather complex. The difficulty arises when the camera sensors are not within a linear transformation of the human receptor sensitivities. In this case (cf. Horn, 1984) no simple calculation relating the camera responses to the desired display values is possible.

The methods described in the next section -- on image analysis -- may be used to decide upon the correct color display values. The idea is the following: Even if the camera sensors are not within a linear transformation of the human photoreceptor responses, their response can still be used to estimate the surface spectral reflectances and light spectral power distributions in the image. These estimated values can then be used to determine the display output. I describe the procedure for using camera data to estimate the surface and the light parameters in the next section. I will return to the problem of correcting camera output at the end of the paper.

The analysis of color images

In this section I will treat the analysis of image data. Beginning with the sensor responses acquired in a camera, ρ^x , we ask what we can learn about the surface reflectance functions and the ambient light spectral power distribution. The development in this section is based on work done by Maloney and me (Maloney, 1984; Wandell and Maloney, 1984; Maloney and Wandell 1985).

The key equation that relates the response of the sensors in a camera system to the ambient light spectral power distribution and surface reflectance functions is equation 16 which we now write as

$$\rho^x = \Lambda_E \sigma^x \quad (26)$$

The known values in the acquired image are on the left hand side of the equation, and the unknown values are on the right hand side of the equation. An extensive mathematical analysis of the recovery procedure is contained in Maloney's (1984) dissertation. The main results are summarized in Maloney and Wandell (1985). I describe the computational method used in the Stanford color analysis package here.

Conditions for recovery: spatial variation

We begin by considering what can be recovered from the color signal when the surface and light functions are both free to vary over space. Recall that the response of the sensors, ρ^x , is completely determined by the color signal, which is given by $C^x(\lambda_n) = E^x(\lambda_n) S^x(\lambda_n)$. To evaluate the uniqueness of recovery note that we can introduce an arbitrary function of both space and wavelength, $G^x(\lambda_n)$ into the equation for the color signal so that.

$$C^x(\lambda_n) = E^x(\lambda_n) S^x(\lambda_n) \quad (27)$$

and

$$C^x(\lambda_n) = (E^x(\lambda_n) G^x(\lambda_n)) \left(\frac{S^x(\lambda_n)}{G^x(\lambda_n)} \right) \quad (28)$$

The uniqueness of the recovery can be characterized, then, by describing the limitations on the function $G^x(\lambda_n)$. Clearly, without some limitation on the structure of the surface and light functions unique recovery is impossible. In order to obtain unique recovery we introduce restrictions on the variation of the light and surface functions. The consequences of the restrictions I describe here are thoroughly analyzed by Maloney (1984).

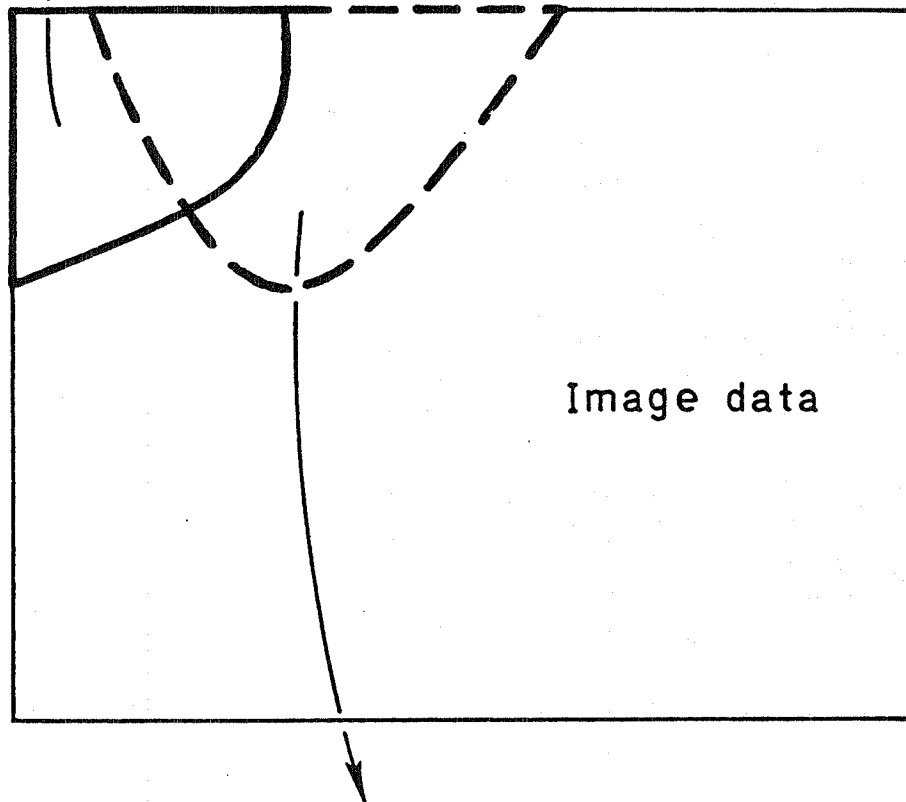
A condition on spatial variation. Suppose the scalar intensity of both the ambient light and surface reflectances are free to vary at each point. If we scale the ambient light at x by a factor $g(x)$, and we scale the surface reflectivity by a factor $\frac{1}{g(x)}$, then the color signal remains unchanged. This emphasizes that we cannot discriminate among the alternative decompositions of the signal into $E^x(\lambda_n)$ and $S^x(\lambda_n)$ versus $E^x(\lambda_n)g(x)$ and $\frac{S^x(\lambda_n)}{g(x)}$.

To permit a unique recovery of the light and surface functions, therefore, we adopt an assumption about the spatial properties of the light and surface functions. We assume that the rate of variation over space of the surface reflectance is greater than the rate of variation over space of the ambient light.

The basic premise of the analysis is illustrated in figure 1. We analyze the image sequentially by dividing it into a set of overlapping spatial areas. Within each of these local areas we analyze the image under the assumption that the ambient light is spatially uniform. We choose the image areas with two considerations in mind. First, we choose the areas so that within each one there is significant variation in the color signal. This permits us to perform that local estimates of surfaces and lights. Second, we choose the areas with a high degree of overlap so that we can detect slow variations of the spectral power distribution of the ambient light. The complete estimate of the ambient light function is determined by combining the separate estimates of the (piecewise constant) ambient light into a smoother, interpolated function. The precise amount of variation necessary in the spectral reflectance function of the surface reflectance functions within each of the areas is specified in the next section in which I describe the recovery method.

In each local region, then, the spatial variation in the color signal is assumed to be due entirely to the surface reflectance, and not at all to the ambient light. With this assumption, the estimate of the spectral power distribution of the ambient light is unique up to a single scalar, $\alpha E(\lambda_n)$ and $\frac{1}{\alpha} S^x(\lambda_n)$, with weak conditions. The value of α is fixed for all points within each region of the analysis. Further, since the regions may be chosen to overlap, the requirement of consistency at points of overlap leaves us with only a single unknown scalar for the entire image. By convention we choose α to normalize the total power in the ambient light to unity. We must remember during our subsequent analyses that the uniqueness is obtained by

First estimate of ambient light
from this region



Next estimate from this overlapping region

Figure 1. The analysis of the image data proceeds by selecting regions of the image with enough complexity in order to form independent estimates of the surface and the light. Within each region indicated in the figure there is enough data to permit an estimate of the surfaces and light functions within that region. As we move from region to region the light and surface functions estimates are forced to be consistent at the points of overlap.

assumption, and not imposed by the data. We should draw no conclusions that depend only on the absolute intensity, or local spatial uniformity, of the ambient light.

A condition on the number of color sensors. Let us now consider the second important condition that is necessary for recovery. This is a condition on the number of color sensors necessary for recovery.

Suppose that the dimensionality of the surface representation, $D(S)$, is equal to the number of sensor classes, that is the dimensionality of ρ^x . In that case, again, no unique solution of the equations across space is possible. To see this, let the reader select any vector whatsoever to represent the weights of the vector of the ambient light, E . This choice specifies the lighting matrix, Λ_E . With this lighting vector chosen, and the data ρ^x known, we may invert the lighting matrix and solve for the surface vectors, σ^x . This solution provides a completely acceptable account of the image data. For each arbitrary choice of the lighting parameters we can generate a set of surface vectors that are consistent with the observed sensor responses. It follows that when the number of dimensions in the surface representation equals (or exceeds) the number of sensor classes, we do not have enough data to obtain a unique solution.

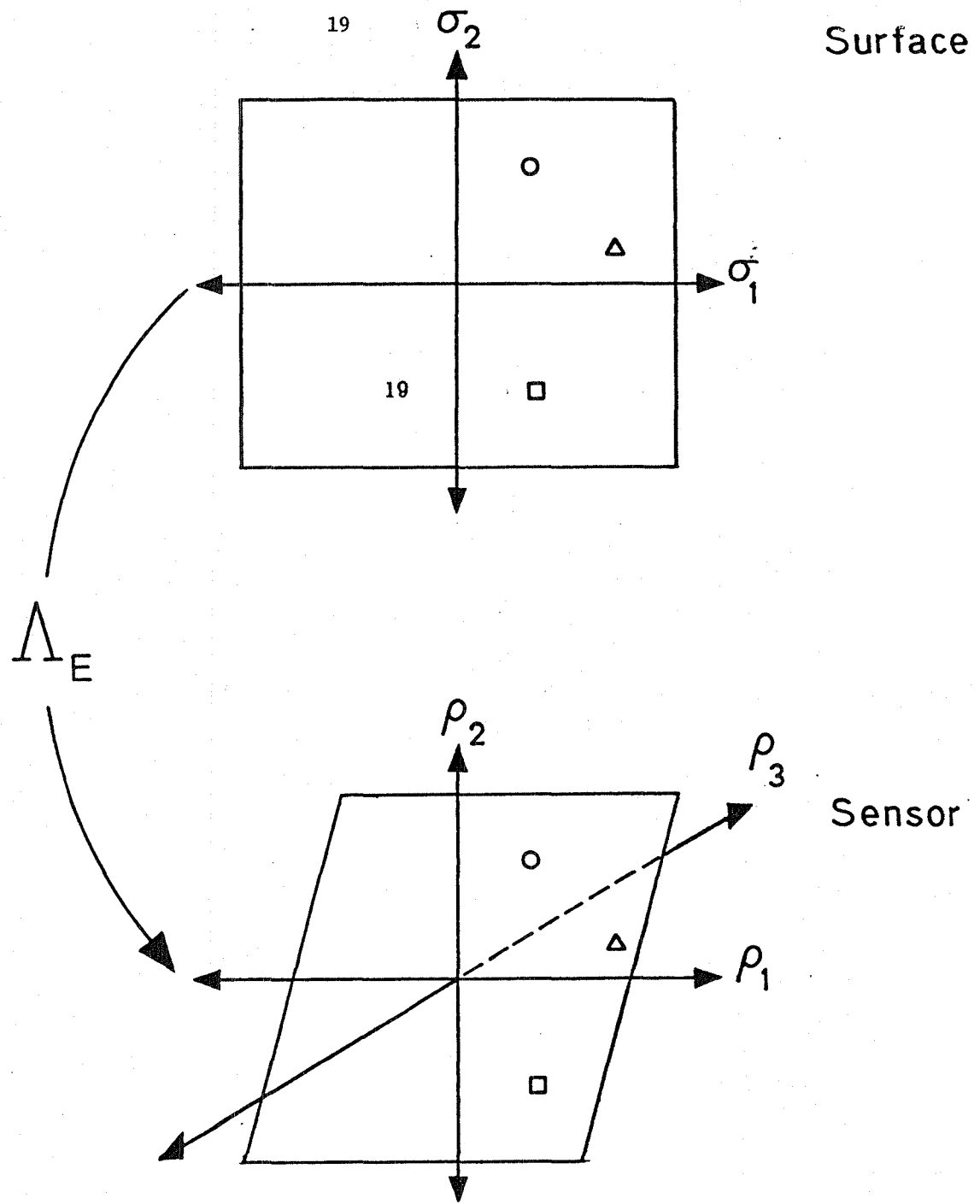
This is the first important difference between the synthesis of images and the analysis of images. When we *synthesize images* we can use a surface representation with dimensionality equal to or greater than the number of independent display guns. In order to analyze an acquired color image we must use a surface representation whose dimensionality is strictly less than the number of sensor classes.

The recovery method

From equation 16

$$\rho^x = \Omega_E S \sigma^x = \Lambda_E \sigma^x \quad (29)$$

we see that when the dimensionality of the surface vector σ^x is less than the dimensionality of the receptor vector ρ^x the observed data will be restricted to a subspace of the receptor responses. For example, when the set of surface vectors is a two-dimensional space, the image of the surface vectors under a fixed light will span a plane within the three-dimensional space of sensor catches. This is illustrated in figure 2. The plane spanned in receptor space is determined by the columns of the matrix Λ_E . The columns of Λ_E are determined entirely by the lighting parameters, ϵ . It follows, therefore, that the identity of the subspace is determined by the values of the light parameters, and the position within the sub-space is determined by the surface parameters, σ^x .



SUB-SPACE ANALYSIS

Figure 2. The surface vectors are represented in a two-dimensional space, shown at the top of the figure. The effect of the light is to act as a linear transformation that maps these two-dimensional vectors into a sub-space of the three-dimensional space of receptor responses.

We can use this fact to achieve recovery in a three step process. First we identify the sub-space which contains the set of receptor responses. Second, we use the identity of the sub-space to estimate the lighting vector, ϵ . Finally, we use the lighting parameters to determine the pseudo-inverse of the lighting matrix, $\Lambda \bar{E}^{-1}$. This is the final step in the problem, since with this matrix we can determine the estimated surface parameters, σ^x , from the sensor responses, ρ^x .

I will now describe the implementation of each of the surface and light recovery steps used in the Stanford color analysis package. This will permit me to describe the points at which failures can occur.

Identifying the sub-space. The first step in the recovery procedure is to attempt to identify the sub-space that contains the set of sensor vectors, ρ^x . For image analysis, we will generally wish to use a surface and light representation whose dimensionality is as large as possible in order to minimize the error in the analysis. Thus, we will assume that $D(E)$ and $D(S) + 1$ equal the number of sensors since any other assumption will produce an inferior fit. With this assumption, we expect the data to fall in a sub-space of the sensor response that can be uniquely characterized by the unit normal vector, perpendicular to all of the vectors in the sub-space. Call the unit normal vector, perpendicular to the sub-space, π . The dot product of this vector with each of the sensor vectors, ρ^x will be zero.

In fact, of course, the data are not likely to fall precisely within a sub-space, but rather to span the sensor response space. Our strategy, however, is to find the vector that is "most perpendicular" to the data from the sensor vectors, under the assumption that perturbations outside of the sub-space are due to noise. To find the most perpendicular vector we form the matrix, Δ , whose rows consists of the sensor values in a region of the image. We then solve the homogeneous linear equation

$$\Delta\pi = 0 \tag{30}$$

where 0 is the zero vector.

In any particular image the best perpendicular vector, π , may not be very perpendicular to the data set. That is, the sensor values may not be well confined to a proper sub-space. This represents the first potential failure of the recovery method.

The principal difficulty in this case is that the dimensionality of the surfaces and lights that generate the image are actually equal to or greater than the number of sensors in the camera. If this is the case nothing can be done and accurate recovery is impossible using this particular sensing apparatus. The recovery procedure then has the virtue that it warns the user that accurate recovery cannot take place since the data are not well-confined to a sub-space. I discuss the selection of good basis functions in

the next section.

Notice that in order for the subspace to be properly defined, there must be enough independent surfaces to define the sub-space. In general the required number is $D(E) - 1$ (Maloney, 1984). In practice we require the spatial partitions of the image to be large enough so that each over the overlapping areas in figure 0 has enough surface variation to permit a solution.

Finally, notice that if the relative spectral power distribution is constant, but the absolute level of the ambient light varies, the data will continue to fall within a sub-space defined by the relative spectral power distribution. Variation in the power of the ambient light may occur due to the geometry of point source illuminations, shadowing, or other factors. This type of spatial variation does not pose a problem for the recovery procedure, but it emphasizes the fact that absolute levels of the surface spectral reflectance and ambient light distributions, in particular for widely separated spatial locations, are not uniquely recovered.

As an alternative to solving the homogeneous linear equation 30 in order to find the perpendicular vector, one may perform a singular value decomposition of the data matrix, R . The direction corresponding to the smallest singular value would then be selected as the best normal direction to the data.

Recovering the light vector. We now wish to recover the light vector, ϵ , from the vector perpendicular to the sensor data, π . We can accomplish this by expanding equation 30 in the following way.

First, notice that the requirement that π be perpendicular to each observed sensor vector may be put as the set of equations

$$\pi^t \rho^x = \pi^t (\Lambda_E \sigma^x) = 0 \quad (31)$$

where π^t is the transpose of π , and equation 31 holds for every surface vector, σ^x , within the model. We may take advantage of the symmetric role played by the light and the surface at each point in the image to re-write the matrix product $\Lambda_E \sigma^x$ as

$$\Lambda_E \sigma^x = \Lambda_S \epsilon \quad (32)$$

where the ik^{th} entry of Λ_S is

$$\sum_{n=0}^N E_i(\lambda_n) S^x(\lambda_n) R_k(\lambda_n) \quad (33)$$

It follows that for any surface within the linear model

$$\pi^t \Lambda_S \epsilon = 0 \quad (34)$$

For each independent surface in the model, we have a linear equation in ϵ . We may form a matrix, Π , whose i^{th} row is $\pi^t \Lambda_{S_i}$ and solve for epsilon using

$$\Pi \epsilon = 0 \quad (35)$$

In this way we can determine the unit vector ϵ whose dimensionality is $D(E) \leq D(S) + 1$.

Solving for surfaces once the light is known. Once the lighting vector is known, the remainder of the problem is straightforward. The dependency of the lighting matrix, Λ_E , on the lighting vector is made explicit in equation 17. This matrix may be explicitly computed and its pseudo-inverse computed using conventional techniques. The surface vectors can then be computed using

$$\Lambda_E^{-1} \rho^x = \sigma^x \quad (36)$$

Two difficulties may arise at this stage in the analysis. First, the estimated Λ_E may not be one to one. In this case there will be no unique solution for the surfaces unless the dimensionality of the surface representation is reduced.

Second, the recovered surface weights, σ^x , while consistent with the model may in fact give rise to estimates of the surface reflectance functions that are not physically realizable, that is that contain negative values of surface reflectivity. When this happens broadly throughout an image, it is an indication that the linear model for the lights and surfaces is not providing a very good fit to the actual lights and surfaces. This too requires re-evaluating the choice of basis functions used in the estimation procedure.

Noise analysis

Two categories of error can enter into the calculation. The first source of error is measurement noise introduced at the level of the sensor responses. The second source of error is failures of the linear models of lights and surfaces to properly characterize the surfaces and lights. The problem of sensor noise is relatively straightforward. If we model sensor noise by a vector η^x added to true sensor vector, ρ_0^x

$$\rho^x = \rho_0^x + \eta^x \quad (37)$$

then the error analysis will depend upon the distribution of the noise vector, η^x . Generally, for this source of noise the typical assumptions of independence of the noise vectors across the image are adequate. The analysis presented in the previous sections is the maximum likelihood calculation under the assumption that the entries of the noise vector are Gaussian and independent.

Model errors. If the dimensionality of the linear model representations of lights and surfaces, $D(E)$ or $D(S)$, are allowed to become arbitrarily large, then model error can be made arbitrarily small. As the dimensionality of the linear representations are restricted to provide increasing efficiency, then additional error will be introduced by the model approximations to the actual surface and light functions.

Suppose that the dimensionality of the surface and light functions required to produce an essentially error-free representation of the data are M and N . To reduce the size of the computation we will seek linear models in which $D(E) < M$ and $D(S) < N$. The sensor data will be correctly modeled by an equation of the form

$$\rho^x = \sum_{n=1}^M \epsilon_n \Lambda_n \sigma^x \quad (38)$$

in which σ^x is an N dimensional vector. The analysis, however, will be based upon

the model equation

$$\rho_0^x = \sum_{n=1}^{D(E)} \epsilon_n \Lambda_n \sigma^x . \quad (39)$$

in which ρ^x is a $D(S)$ dimensional vector. The difference between the measured sensor data (equation 38) and the best model estimates (equation 39) primarily depends on the ambient light and the surfaces in the image. The error terms that express the difference between these two equations may be described by an additional vector with contributions from various sources, as illustrated in figure 3. The distribution of this error vector, then, depends upon the surfaces in the image as well as the ambient light falling on the surfaces. Since we cannot specify the sampling distribution of the actual surfaces without some more knowledge of the specific task of the visual sensor (a robotic inspection system will sample a different class of surfaces from a color correction circuit in a home video recorder) the distribution of the error signal cannot be specified without further information concerning the class of materials and ambient lights that the system will encounter.

The efficient representation of surfaces and lights

Introduction

The choice of the linear models for the ambient light and surface spectral reflectance terms is crucial in building an efficient computation in the synthesis of images, and it is crucial in permitting one to use a camera with only a few color sensors in the analysis of image data. In this section I will review the selection of basis functions that permit a more efficient representation and computation of the lighting and surface parameters.

Fourier Basis

The principal means available to us for reducing the computational burden is to reduce the dimensionality of the surface spectral reflectance, N . Were the surface spectral reflectance functions and the ambient light spectral power distributions essentially bandlimited functions, then from the sampling theorem we know that specifying these functions at, say, four sampling points along the wavelength dimension would uniquely determine the identity of the function.

Stiles and Wyszecki (1962; Stiles, Wyszecki and Ohta, 1977) performed an early and important analysis of surfaces reflectance functions in which they suggested that these functions are lowpass functions of wavelength in the visible range. Maloney (1984) has put forth the intriguing hypothesis -- based on the physical chemistry of surface reflectances -- that these functions are generally forced to be lowpass in the visible range.

There is wide informal agreement that natural surfaces and lights are generally smoothly varying. There are certain exceptions, such as fluorescent light and the surface reflectance functions of the rare earth metals, but the fact that the overwhelming majority of functions are smoothly varying suggests that it may be possible to treat them as bandlimited functions, and represent them as the weighted sum of several low frequency terms across the visual spectrum.

Specular reflectance. When using a low frequency representation the functions $S_j(\lambda_n)$ are illustrated in panel (b) of figure 4. The function S_1 is the the DC component of the spectral reflectance function, and the functions S_2 and S_3 would represent the sinusoidal and cosinusoidal components at one cycle across the visible region. Notice that when this representation is chosen the weight of the first dimension, S_1 , is a mixture of the specular reflectance at that point in the image, added together with the DC component of the matte portion of the surface reflectance function. Using such a representation, then, it is relatively straightforward to systematically vary the specularity. In practice this can be achieved by simply manipulating the value of the first dimension of the surface reflectance function vector.

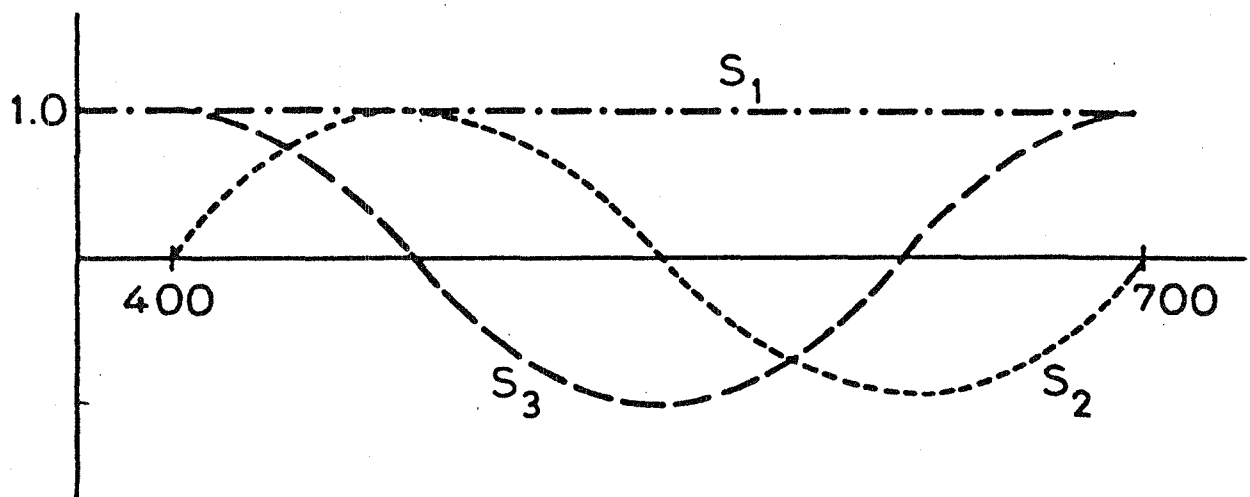
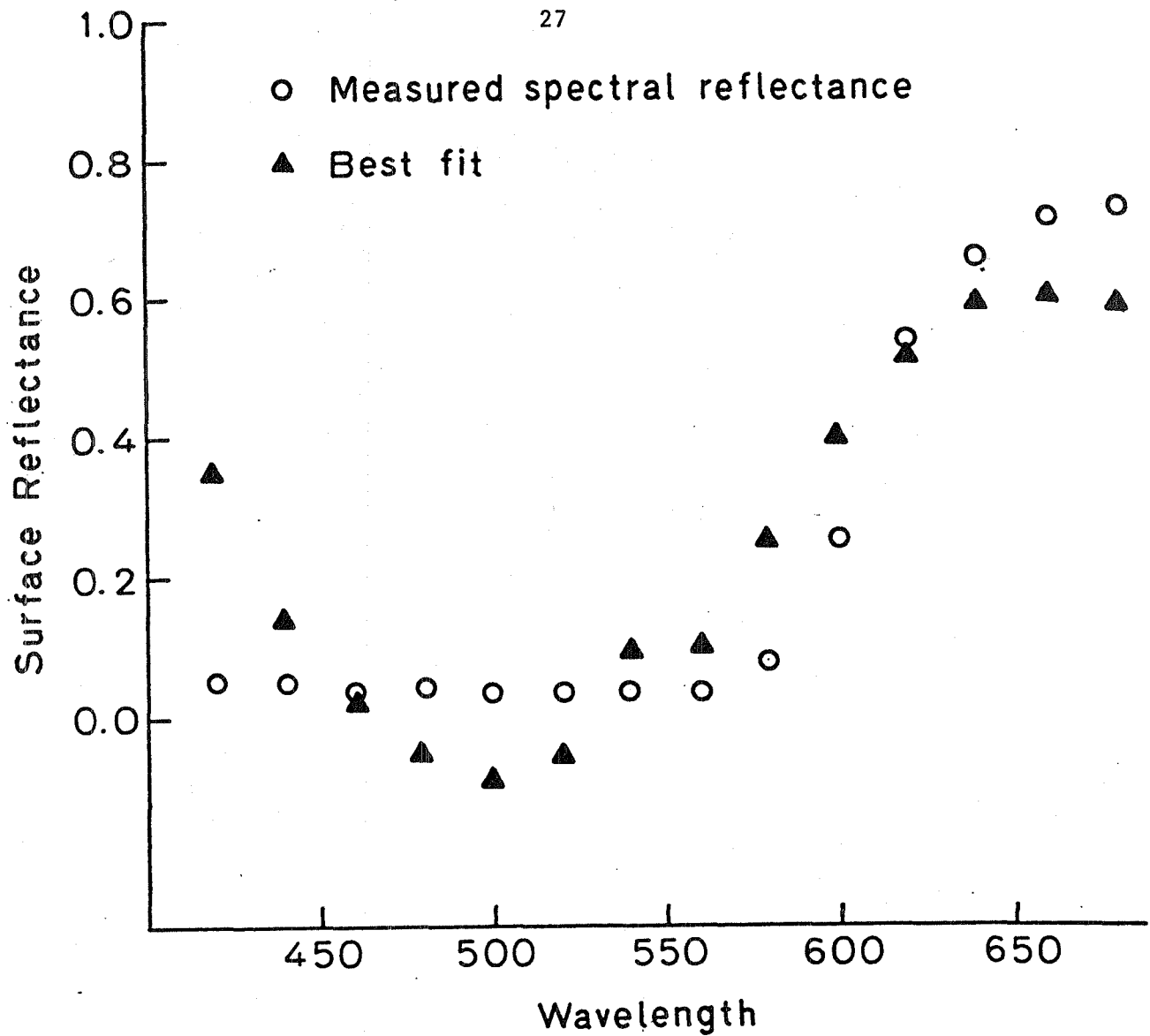


Figure 4. (a) The surface reflectance function (smooth line) and best-fitting estimate (dashed line) of the Munsell chip with the largest residual error (worst fit) of all 462 chips. (b) The three Fourier basis functions used to fit the surface reflectance functions of the Munsell chip samples.

In figure 5 I have analyzed how well the Munsell chips, a widely used set of color surfaces, can be described as bandlimited functions. The figure is derived as follows. First, I used the spectral reflectance functions of the Munsell chips measured by Nickerson (1943)² to calculate the least-squares fit of Fourier basis terms to the surface spectral reflectance functions of each of the Munsell chips, measured from 400 nm to 700 nm. I then normalized the root mean squared error (rmse) of the best fit by the vector length of the surface reflectance function. This is the square root of the proportion of variance accounted for. On the vertical axis I have plotted the normalized rmse of the median of the 462 Munsell chips available in our data as a function of the number of Fourier basis terms. In figure 4 panel (a) I have plotted the worst of the 462 fits to the sample data, along with the three basis functions.

Principal Components Analysis

The Fourier representation is efficient for representing spectral reflectance functions and ambient light because of the high degree of correlation across the wavelength spectrum. It has the further advantage of identifying specular reflectance with the DC component of the surface reflectance function.

An alternative method for selecting the representation, which is particularly appropriate when the set of surfaces and lights are known in advance, is to perform a principal component analysis of the distribution of surface reflectance functions and spectral power distributions. The principal component analysis -- sometimes called the Karhunen-Loeve transformation -- explicitly seeks that set of basis functions that simultaneously minimizes the correlation among dimensions and identifies those dimensions that are most descriptive of the data set.

A principal components analysis was applied to the spectral power distribution of ambient daylights by Judd, MacAdam and Wyszecki (1964; Dixon 1978) and to the surface reflectances of the Munsell chips by Cohen (1964).

The principal components in Judd et al. (1964) were computed at an early stage in the technical development of these methods. The reported functions are not precisely orthogonal. Unfortunately, the original data have been misplaced so that we cannot re-analyze the original data. Maloney (personal communication) is preparing a new analysis of a large set of such data.

Relationship to physical processes. It is interesting to note that the linear model for the variations in the ambient daylight spectral power distribution is not a reasonable model for the physical processes that give rise to the observed spectral power distributions (compare the argument in Witkin and Pentland, 1984). For example, blackbody radiators form a one parameter non-linear family of spectral

² I thank the Munsell corporation for making these measurements available.

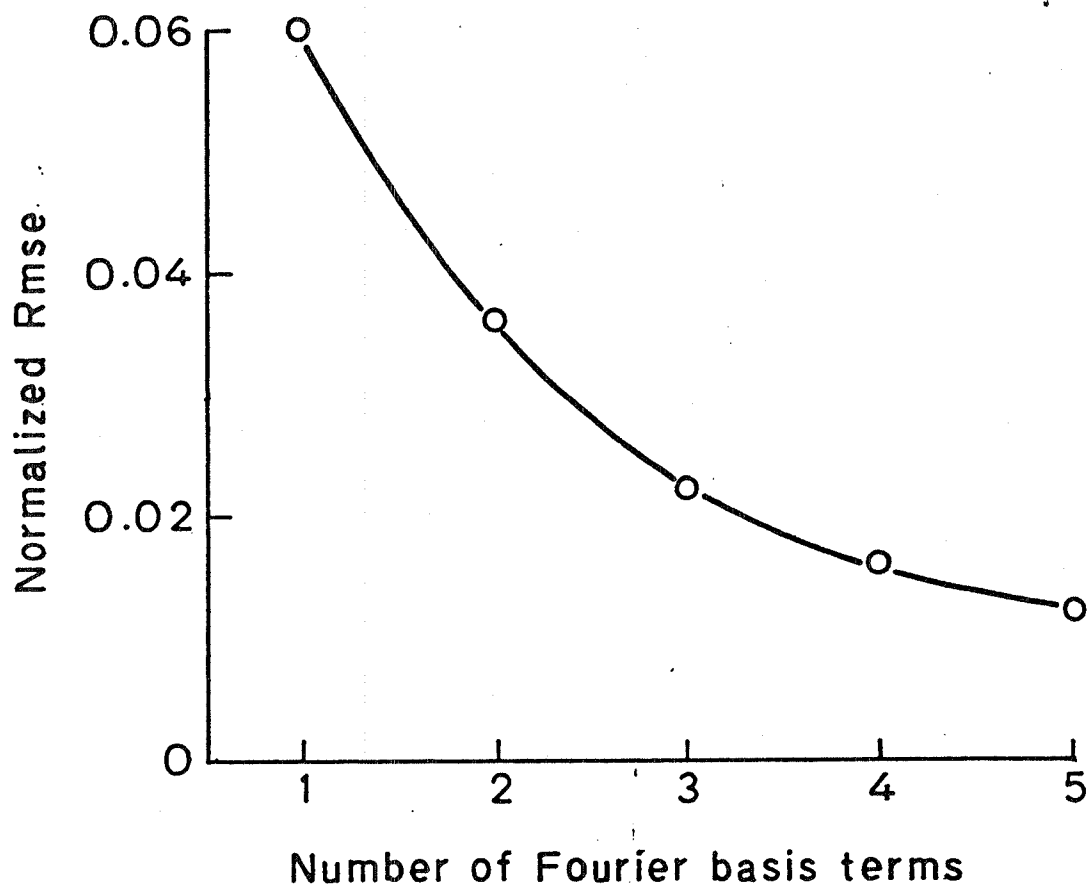


Figure 5. Each of 462 Munsell chip surface reflectance functions was fit by the sum of sinusoidal and cosinusoidal functions. This figure plots the normalized RMSE of the median of 462 Munsell chip fits as a function of the number of Fourier basis terms used in the fit. The percent of variance accounted for may be computed as $\sqrt{1 - y}$ where y is the normalized RMSE.

power distribution curves. Yet Maloney has shown (personal communication) that the set of principal components of the natural daylights also provides a rather good approximation to the spectral power distribution seen in blackbody radiators. The physical effects that cause the fluctuation in the spectral power distribution of daylight include such factors as Rayleigh scattering and the absorption and scattering by atmospheric moisture. Were one to provide a serious physical model of these effects, the linear model would never be selected. Normally such processes are modeled by a multiplicative filtering action and non-linear molecular actions. The linear model used here is not closely tied to the physics of the imaging environment, but rather is simply a useful mathematical method. Its usefulness depends on our ability to accurately describe the wavelength functions of the desired quantities. The fact that the underlying physical processes are not well modeled by the linear sum of different physical processes is not relevant to the performance of the algorithm.

The Color camera revisited. In figure 6 I indicate how the analysis and synthesis calculations developed in this paper can be put together in order to obtain a visually satisfactory rendering of an image sensed by a color camera. As Horn (1984) has pointed out, if the camera sensor spectral sensitivities are within a linear transformation of the spectral sensitivities of the human eye, the problem may be easily solved. The more difficult -- and common -- case is when the camera sensors are not related to the visual pigment sensitivities by a linear transformation.

In that case we may proceed as indicated in figure .6 We use the camera sensor responses and the procedure in the analysis section to form an estimate of the ambient light spectral power distribution, E . We then calculate estimates of the surface reflectances,

$$\sigma^x = \Lambda_E^{-1} \rho^x \quad (6)$$

Using these estimates of the surface and ambient light distributions we set the display gun intensities using the synthesis calculations in equation 23, which I repeat here.

$$I^x = \left[\sum_{n=1}^{D(E)} \epsilon_n \Gamma^{-1} \Lambda_n \right] \sigma^x \quad (7)$$

To the extent that the camera data permit an accurate recovery of the surface and light functions, the colors in the synthesized image will appear as if the observer had been at the position of the camera.

COLOR RENDERING PROCEDURE

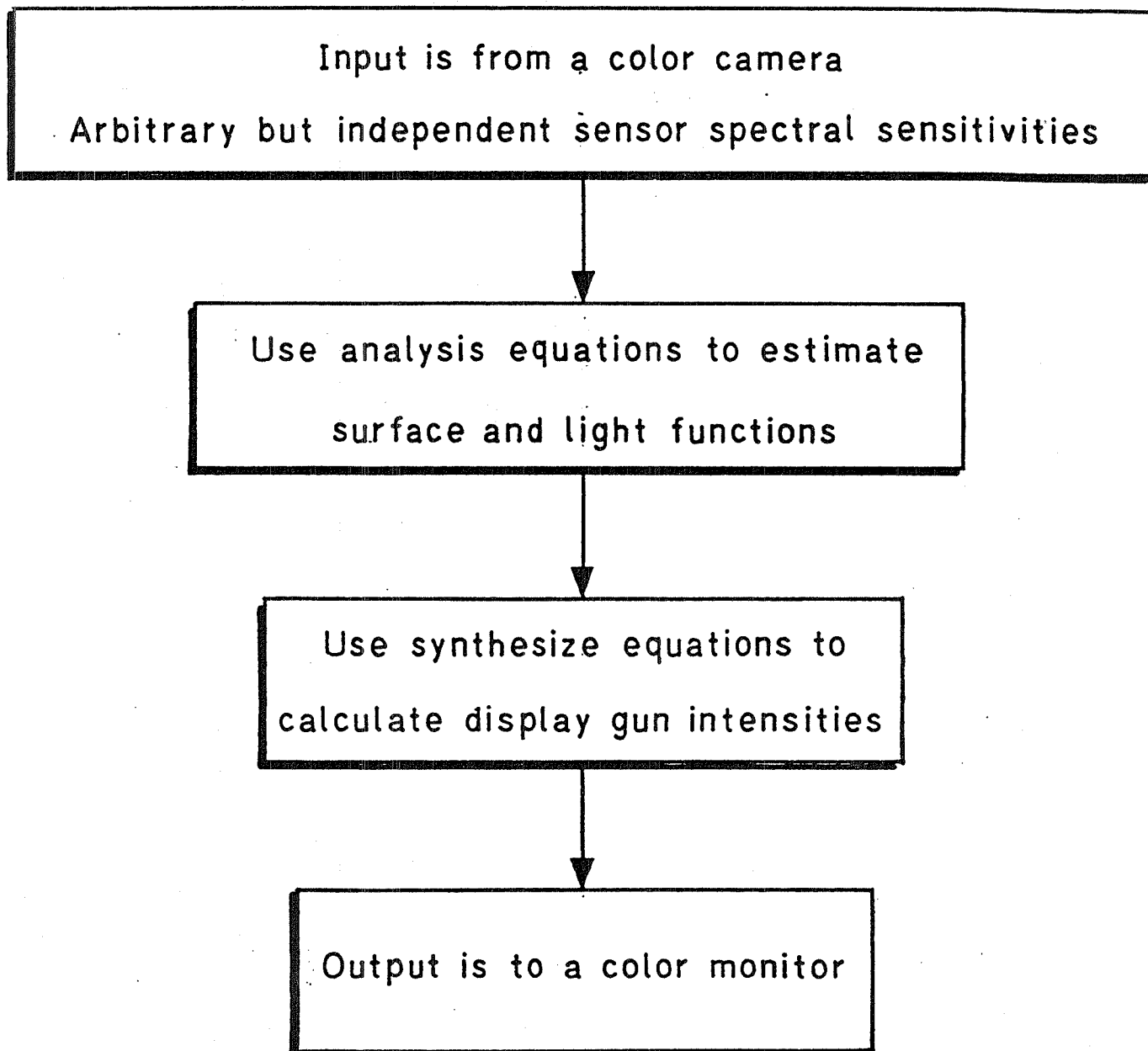


Figure 6. Flow diagram of how the analyses and synthesis procedures may be used together to convert the outputs of a camera with arbitrary sensor spectral to display an image so that it will appear to an observer as if he were at the camera position.

Conclusions

I have described algebraic methods for the synthesis and analysis of color images. The methods are based on the physical factors that give rise to the color information. The reliance on the common physical processes permits us to treat several different problems in the theory of color images from a common framework. In particular, the equations for selecting display gun intensities, and the equations for analyzing color images, are both based on a common formula, equation 29.

Maloney and Wandell's (1985) method for analyzing the color signal into its surface and light components, has applications to several different problems in color science. The analysis of the image data into surface and light components permits us to correct the acquired sensor signal for the ambient light. Having this information permits superior color balancing of the acquired image than was previously available. The analysis further permits us to use color data acquired by cameras whose sensors are not simply related to the human photopigments in order to display images with a reasonable color appearance. Finally, the method provides a quantitative method for estimating surface properties of objects from image data. The ability to use estimates of surface properties -- rather than gray-level intensity images which confound surface and light information -- should become an important aid in the development of visual inspection systems that are robust with respect to the ambient lighting.

References

- Buchsbaum, G. & Gottschalk, A. Chromaticity coordinates of frequency-limited functions. *J. Opt. Soc. Am.*, 1984, 1, 885-887.
- Cohen, J. Dependency of the spectral reflectance curves of the Munsell color chips. *Psychon. Sci*, 1964, 1, 369-370.
- Cook, R.L. & Torrance, K.E. A reflectance model for computer graphics. *Computer Graphics*, August 1981, 15, 307-316.
- Cornsweet, T. N. *Visual Perception*. New York: Academic Press, 1970.
- Helson, H. Fundamental problems in color vision. I. The principle governing changes in hue, saturation and lightness of non-selective samples in chromatic illumination. *J. Exp. Psychol.*, 1938, 23, 439-476.
- Horn, B. K. P. Determining lightness from an image. *Computer Graphics and Image Processing*, 1974, 3, 277-299.
- Horn, B. K. P. Exact reproduction of colored images. *Computer Vision, Graphics and Image Processing*, 1984, 26, 135-167.
- Judd, D. B., MacAdam, D. L., & Wyszecki, G. W. Spectral distribution of typical daylight as a function of correlated color temperature. *J. Opt. Soc. Am.*, 1964, 54, 1031.
- Oppenheim, A. V. & Schaffer, R. W. . In (Ed.), *Digital Signal Processing* : Prentice-Hall, 1975.
- Smith, V. & Pokorny, J. Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Res.*, 1975, 15, 161-171.
- Stiles, W.S., Wyszecki, G., & Ohta, N. Counting metameric object-color stimuli using frequency-limited spectral reflectance functions. *J. Opt. Soc. Am.*, 1977, 67, no. 6, 779-784.
- Stiles, W.S. & Wyszecki, G.W. Counting metameric object colors. *J Opt. Soc. Amer.*, 1962, 52, 313-327.
- Torrance, K. E. & Sparrow, E. M. Theory for off-specular reflection from roughened surfaces. *J. Opt Soc. Am.*, 1967, 57, 1105-1114.

Torrance, K. E., Sparrow, E. M., & Birkebak, R. C. Polarization, direction distribution, and off-specular peak phenomena in light reflected from roughened surfaces. *J. Opt Soc. Am.*, 1966, 56, 916-925.

Witkin, A. & Pentland, A. Perception and Generic Modeling. *Paper presented at J. Opt. Soc. Am.*, 1984, P6.

1. Report No. NASA TM-86844		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle THE SYNTHESIS AND ANALYSIS OF COLOR IMAGES				5. Report Date September 1985	
				6. Performing Organization Code	
7. Author(s) Brian A. Wandell				8. Performing Organization Report No. 85419	
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035				10. Work Unit No.	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546				13. Type of Report and Period Covered Technical Memorandum	
				14. Sponsoring Agency Code 506-57-21	
15. Supplementary Notes Point of Contact: Andrew B. Watson, Ames Research Center, M/S 239-3, Moffett Field, CA 94035 (415)694-5419 or FTS 464-5419					
16. Abstract I describe a method for performing the synthesis and analysis of digital color images. The method is based on two principles. First, image data are represented with respect to the separate physical factors, surface reflectance and the spectral power distribution of the ambient light, that give rise to the perceived color of an object. Second, the encoding is made efficient by using a basis expansion for the surface spectral reflectance and spectral power distribution of the ambient light that takes advantage of the high degree of correlation across the visible wavelengths normally found in such functions. Within this framework, the same basic methods can be used to <u>synthesize image data</u> for color display monitors and printed materials, and to <u>analyze image data</u> into estimates of the spectral power distribution and surface spectral reflectances. The method can be applied to a variety of tasks. Examples of applications include the color balancing of color images, and the identification of material surface spectral reflectance when the lighting cannot be completely controlled.					
17. Key Words (Suggested by Author(s)) Visual perception Human factors Color vision			18. Distribution Statement Unclassified - unlimited Subject category - 53		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 36	
				22. Price* A03	

End of Document