Modeling Pigmented Materials for Realistic Image Synthesis

CHET S. HAASE and GARY W. MEYER
University of Oregon

This article discusses and applies the Kubelka-Munk theory of pigment mixing to computer graphics in order to facilitate improved image synthesis. The theories of additive and subtractive color mixing are discussed and are shown to be insufficient for pigmented materials. The Kubelka-Munk theory of pigment mixing is developed and the relevant equations are derived. Pigment mixing experiments are performed and the results are displayed on color television monitors. A paint program that uses Kubelka-Munk theory to mix real pigments is presented. Theories of color matching with pigments are extended to determine reflectances for use in realistic image synthesis.

Categories and Subject Descriptors: I.3.4 [Computer Graphics]: Graphics Utilities paint systems; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism color, shading, shadowing, and texture

General Terms: Algorithms, Experimentation, Human Factors

Additional Key Words and Phrases: Color matching, color science, color selection, illumination modeling, pigment mixing

1. INTRODUCTION

The illumination model in computer graphics is constantly being refined so that it provides an ever more accurate model of reality. However, no illumination model will work unless the objects in the scene have been properly described. This fact has been realized and the surface properties of objects have been painstakingly modeled, especially as they relate to reflection direction and strength. But a comprehensive approach to object modeling should also take into account the physical makeup of those surfaces. In particular, most manufactured objects in the real world have pigmented surfaces, such as painted items, plastic objects, and textiles. To accurately...
model the color of any pigmented object, the physical interactions within the
pigmented surface must be included.

A straightforward way of specifying color in a graphics scene is via the
common red, green, and blue (RGB) triplet. This approach is the easiest to
deal with in scene generation computations, since this specification is exactly
what is needed by the computer to generate colors for the screen output. But
this method is far removed from reality—the RGB method is only appropri-
ate for additive colorants, such as colored lights (e.g., the phosphors in the
monitor screen); nonemitting objects should be specified in terms of their
reflective and transmissive characteristics.

A second approach for specifying color is in cyan, magenta, and yellow
(CMY) space. This specification is in the domain of subtractive color synthe-
sis and accurately models the effect of light transmission through a colored
surface. This specification is especially important with respect to the color
reproduction industry. For instance, color printer manufacturers must deal
with translating images from the domain of computer monitors (in which all
colors are RGB) to the domain of hardcopy devices, in which the colors must
be specified in CMY coordinates (Stone, et al., 1988). This subtractive method
is also used in some screen-based painting applications, such as Small (1991)
and some commercially available programs, as a better attempt to model
realistic color mixing. However, as will be shown in this paper, this method is
insufficient for representing pigmented surfaces, since CMY color synthesis
works best for purely transmitting materials; pigmented surfaces have both
transmitting and reflecting characteristics.

The most accurate way of dealing with light in a synthetic imaging
application is on a wavelength-by-wavelength basis [8, 14, 19]. It is only in
this way that the more subtle illumination modeling and color calculation
problems can be handled. The spectral energy distributions of the light
sources in the environment must be given and the spectral reflectance,
transmittance, and absorptance of the surfaces with which these light sources
interact must be specified. Pigment mixing is an example of an optical
phenomenon which must be modeled in this way. The Kubelka–Munk theory
is used by the paint industry to predict the reflectance that will result when
two or more pigments are mixed. This theory has been discussed within a
computer graphics context [12] but has not been used to solve color synthesis
problems.

In this paper we introduce the Kubelka–Munk theory of pigment mixing
and apply that theory to computer graphics to facilitate correct-color calcula-
tion for realistic image generation. Section 2 develops some of the significant
differences in additive and subtractive color mixing and discusses the need
for different mixing theory for pigmented materials. Section 3 introduces and
develops the Kubelka–Munk theory of pigment mixing. Section 4 discusses
the pigment-mixing experiments carried out for this project. Section 5 intro-
duces the topic of color matching with respect to pigments; given a sample
color, it is possible to calculate a pigment mixture that will render an
approximation to this color. The final section provides a summary.
2. SUBTRACTIVE COLOR MIXING

The human visual system contains three types of daylight receptors, called cones. The spectral sensitivity functions of the three cone types are different, but they do overlap significantly. Because there are only three types of cones to sample the entire visible spectrum, it is possible for two different spectral energy distributions to produce the same three cone signals. When this happens the visual system is unable to distinguish between the color sensations evoked, and the spectral energy distributions are said to be metamers. This means that satisfactory color reproduction can be achieved without having to perfectly reconstruct the original spectral energy distribution. In fact, the number, shape, and position of the spectral sensitivity functions make it possible to reproduce most colors by manipulating only the short, medium, and long wavelength regions of the spectral energy distributions used to create the reproduction.

The most straightforward way of controlling the energy in these three different portions of the spectrum is to combine the light from three sources. This approach, which is called additive color synthesis, is used in the standard cathode ray tube of color televisions and monitors where three electron guns independently address the short, medium, and long (blue, green, and red) areas of the spectrum. For instance, in combining a light from the short wavelength region with one from the medium wavelength region, a spectral energy distribution is produced which elicits the visual sensation of cyan. The color television monitor gamut produced by additively mixing the light from the three red, green, and blue primaries is shown in Figure 1.

A more indirect way of controlling the energy in the three different spectral regions is to place a stack of three filters in front of a light source. This approach is called subtractive color synthesis, because information is removed from the spectral energy distribution of the light source. This ac-
accomplishes the same goal as additive color synthesis but does not allow completely independent control of the short, medium, and long wavelength regions of the spectrum. For instance, suppose two subtractive colorants are used—cyan and yellow—the cyan colorant transmits energy in the medium and short regions of the spectrum (similar to the cyan produced in the additive example above), and the yellow passes energy in the long and medium regions. By combining these colorants in subtractive color mixing (further explained below), we derive a colorant which controls energy in the medium region and we see a green color; essentially, the short and long spectral energy present in both components was subtracted from the final result. Again, the visual system is being stimulated by selectively manipulating particular regions of the visible spectrum.

The example of the colored filter given above is a very simplified view of subtractive color. Instead of talking about the long, medium, and short areas of the visible spectrum, we can quantify the characteristics of a given subtractive colorant on a wavelength-by-wavelength basis. In particular, at each wavelength of the visible spectrum, the colorant will allow some given percentage of light energy to pass through. This information, plotted graphically with wavelength versus the percentage of light passed through the material, gives us the transmittance curve for that filter. An example of such a curve is seen in Figure 2, where we see that more light energy toward the lower end of the spectrum is allowed to pass than light energy at the higher end. Thus, given a light source which emits equal energy at all wavelengths, the resulting color of the light passing through the example filter would be cyan.

These subtractive colorants are made from some clear material in which a colored substance is dissolved at a given concentration. The transmittance data discussed above assumes a given thickness of the colorant layer, \( h \), and a given concentration, \( c \), of the colorant. If we change either of these variables—the thickness or the colorant concentration—we would expect the nature of the transmittance of the filter to change as well. These transmittance changes are described by the Bouguer and Beer laws. Bouguer's law states that given a transmittance \( T \) at a particular wavelength and a thickness \( h \), the transmittance at a thickness of \( nh \) at that wavelength is \( T^n \). Similarly, Beer's law states that given a transmittance \( T \) at a particular colorant concentration \( c \) and a constant thickness, the transmittance at that wavelength is \( T^m \) where the concentration is \( mc \) [10]. The effect of varying the concentration of a hypothetical dye can be seen in Figure 2.

Changing the concentration or the thickness of a filter produces a locus of tristimulus values which is much different than the collection of tristimulus values created by varying the intensity of a light source that shines through a filter with constant transmittance characteristics. Given a filter that corresponds to unit concentration of the dye shown in Figure 2, Figure 3 depicts the locus of CIE \( XYZ \) tristimulus values produced by changing the intensity of the light passing through the filter. This locus is a straight line that projects to a single point on the unit plane. This means that a single pair of chromaticity coordinates can be used to describe the colors that are produced.
On the other hand, when the light source remains constant and the concentration of the filter varies, the locus of CIE XYZ tristimulus values shown in Figure 4 results. The projection of this space curve onto the unit plane produces a curved line segment. This indicates that the color produced by the filter is changing and no single pair of chromaticity coordinates can fully describe the situation.

Calculation of subtractive colorant mixtures is analogous to subtractive concentration calculations, as above. Whereas, in the case of varying concentrations, we multiply the transmittance at every wavelength by itself for every unit of concentration, subtractive mixtures require the multiplication of transmittances of the different colorants in the mixture. It is helpful to think of this process in terms of separate colored filters. Light passing through each filter is multiplied by the transmittance of that filter (taking into account whatever concentration of thickness that filter has), then the process is repeated for each additional filter (and color). In Figure 5 we see the effect of mixing the cyan dye from Figure 4 with two other dyes. The gamut produced by varying the individual concentrations of the colorants is substantially different from the additive color-television monitor gamut shown in Figure 1. The difference in shape between the additive and subtractive gamuts illustrates the linear versus nonlinear dichotomy that exists between these two approaches to color mixing.

Water color paints can be modeled using subtractive color mixing, because the paint is usually transparent enough to allow light to pass to the substrate.
Fig. 3. Locus of values in CIE XYZ space resulting from shining a light of varying intensity through a cyan filter with constant concentration. The intersection with the unit plane is a single point.

Fig. 4. Locus of values in CIE XYZ space resulting from shining a light of constant intensity through a cyan filter of varying concentration. The intersection with the unit plane is a curved line.

Fig. 5. Gamut produced in CIE XYZ space by mixing varying concentrations of cyan, yellow, and magenta dyes.

and reflect back through the paint appropriately colored with respect to the transmitting properties of the watercolor [10]. Pigmented materials such as artists' oil paints, on the other hand, do not work as well within the subtractive domain. The preceding discussion assumed that the materials involved were transparent and allowed light to pass through the substance,
Fig. 6. Light interactions within a pigmented material (after Evans, 1948).

selectively absorbing light energy at certain wavelengths; but pigmented solutions consist of opaque particles in a transparent medium, quite different from the completely transparent solutions of subtractive colorants.

To understand pigmented solutions, we must look at them on a particle level. Each pigment particle has both absorbing and reflecting properties. The absorption of a pigment particle is related to its transmittance, because absorption is equal to 1 minus the transmittance. But in addition to this absorption effect, there is the effect of selective reflectance, or scattering, of light energy. Scattering is the property that determines how much of the incident light is reflected from the surface of the pigment particle per wavelength. Each of these two factors (illustrated in Figure 6) can be described by spectral curves which describe the percentage of energy absorbed or scattered by the pigmented solution at each visible wavelength. Together these absorption and scattering properties produce spectral reflectance characteristics which create the color that we observe.

The problem in making calculations based on these properties is that each pigment particle in the solution has the properties of scattering and absorption. In the case of the transparent solutions discussed previously, the colorant was completely dissolved in the solution or material. But in the case of pigments, the pigment material is ground up finely and distributed throughout some filler, such as linseed oil in the case of artists' oil paints. Making calculations based on the independent properties of scattering and absorption for every pigment in the solution would require accounting for the size, density, and shape of every particle. One of the theories of scattering...
and absorption of materials, the Mie theory, uses this information to calculate various properties of individual pigment particles. But this level of detail is somewhat too complex for calculating information about pigment solutions [17]. Some other means to calculate the absorption and scattering for the overall material would be very beneficial. This, then, was the motivation for the work done by Kubelka and Munk.

3. KUBLERK-MUNK THEORY

Instead of looking at pigmented solutions on a particle level, Kubelka and Munk observed the effects of light energy in the entire solution. This approach incorporated the ideas behind the pigment particle scattering and absorption interactions, but allowed a much easier and more comprehensive calculation of entire pigmented systems. By using this approach, Kubelka and Munk came up with the equations that are still used in the pigment industry today.

Kubelka and Munk examined what happens as light traverses a thin layer of paint applied over a substrate. At any location in the paint, light from the surface is moving deeper into the material and light that has reflected from the substrate is travelling back toward the top of the film. A certain fraction, $K$, of the light travelling in each direction will be absorbed by the material, and another portion, $S$, will be scattered. Light from each direction that is scattered is assumed to contribute to the amount of light travelling in the opposite direction. A set of differential equations were written as a result of this analysis. A complete derivation of these equations is given in the Appendix.

For complete hiding (i.e., when the pigment layer is so thick that the substrate cannot be seen), the solution of the differential equations for the reflectance, $R_x$, is:

$$R_x = 1 + \frac{K}{S} - \sqrt{\left(\frac{K}{S}\right)^2 + 2\frac{K}{S}}.$$  \hspace{1cm} (1)

This is often rewritten as:

$$\frac{K}{S} = \frac{(1 - R_x)^2}{2R_x}.$$  \hspace{1cm} (2)

Using Equation 1, we can now derive the spectral reflectance of any pigmented material at complete hiding if we know its respective $K$ and $S$ values (and the spectral properties of the solution in which the pigments are immersed). These are the equations that were used for all of the work reported in this paper. Solutions to the differential equations can also be found for the case in which the spectral characteristics of the base material affect the spectral composition of the entire system [16].

We can find the properties of mixtures of pigmented solutions by using the fact that combinations of absorption and scattering are linear [9]:

$$K_M = \sum_{i=1}^{n} K_i c_i$$  \hspace{1cm} (3)

$$S_M = \sum_{i=1}^{n} S_i c_i$$  \hspace{1cm} (4)

and

$$\left(\frac{K}{S}\right)_M = \frac{\sum_{i=1}^{n} K_i c_i}{\sum_{i=1}^{n} S_i c_i}$$  \hspace{1cm} (5)

where

- $K_M$ absorption of pigment mixture
- $S_M$ scattering of pigment mixture
- $n$ number of pigments in mixture
- $c_i$ concentration of $i$th pigment in mixture by weight of dry pigment
- $K_i$ absorption of $i$th pigment
- $S_i$ scattering of $i$th pigment.

Each equation is calculated for each wavelength. The use of these equations in predicting mixtures and matches of colors is covered in later sections of this paper.

It is helpful at this point to look at the $K$ and $S$ values more closely in terms of their significance and their calculation. First of all, the units of $K$ and $S$ are not important for the purposes of these equations. As they are always used in conjunction with each other for the Kubelka-Munk equations, their significance is in their ratio. This simplifies the gathering of the absorption and scattering data significantly; instead of having to measure the actual units of $K$ and $S$ for any particular sample, we can use the equations above and spectral reflectance curves to derive unitless $K$ and $S$ values [15].

For example, to derive the $K$ and $S$ values for the titanium dioxide pigment used throughout this work, $S$ was set equal to 1 at all wavelengths and $K$ was determined by using a given spectral reflectance curve for $R$ [11] and Equation 2 above [15]. Once the $K$ and $S$ data for one pigment (titanium dioxide, in this case) has been calculated, the $K$ and $S$ for any given pigment may be derived, using the equations above and data for the spectral reflectances of various mixtures of the known pigment with the pigment in question. In the case of cadmium yellow medium, one of the primary pigments used in this project, two spectral curves were used [11]: one of cadmium yellow medium as a mass tone (i.e., a pure pigment) and one of cadmium yellow medium as a tint (i.e., mixed with titanium dioxide in a given ratio). By using Equations 2 and 5 above, and knowing the $K$ and $S$ values for titanium dioxide and the relative concentrations of the pigments in
the tinted sample, there were two equations in two unknowns, and the \( K \) and \( S \) values for cadmium yellow medium were solved for appropriately.

Again, the actual units of \( K \) and \( S \) are irrelevant as long as the relative proportions of the factors are correct. However, pigment data gathered from different sources must be used with caution. For instance, if the data for two different pigments were gathered from different sources (perhaps computed using two different sets of \( K \) and \( S \) data for titanium dioxide), the pigments should not be used together in the same equations. Unfortunately, the data required by the Kubelka–Munk equations must be extremely accurate, and gathering such data can be very difficult. The method of calculating \( K \) and \( S \) given above is very straightforward, but those \( K \) and \( S \) values will only be as good as the spectral curves on which they were based. Moreover, the data must be consistent across pigments. For example, if titanium dioxide is used as the tinting agent in the \( K \) and \( S \) calculations, any insufficiencies in the data for titanium dioxide will be reflected in unsatisfactory data for all other pigments that used these values in their own \( K \) and \( S \) calculations.

There are many areas in the process of data gathering which must be handled carefully. To begin with, the measurement of pigment data must be consistent with its proposed use. For instance, since this research uses oil paints for examples, we gathered data from oil paints. Spectral curves that describe the reflectances for those pigments in other solutions are insufficient for our purposes. Also, the mixing of the pigment compounds is important; the ratio of pigment to solution as well as the ratio of pigments in tints could affect the outcome. Application of the pigment compounds to the measurement background is also problematic; a layer of the pigmented solution which does not completely hide the background will be affected by that background color. Measuring the spectral curve for the mixture is another important area. The accuracy of the measuring device in addition to the suitability of the lighting and viewing conditions can strongly affect the outcome of the measurements.

Several simplifying assumptions were made in the theory and equations developed by Kubelka and Munk. First of all, the pigmented solution is treated as a uniform material, assuming complete dispersion of pigment in the solution and homogeneous density of pigment particles. In reality, various forms of deviations from this idea can occur, including flocculation (clumping) and floating of the pigment particles [21]. Second, diffuse lighting and viewing conditions are assumed, and there is no account for surface reflection. Finally, the pigment surface is assumed to be planar, with no light escaping from the edges of the surface. This case is seen in the above assumption that light energy is either undisturbed, absorbed, or scattered. These assumptions are discussed more fully in [4] and [21].

Further work has been done on the Kubelka–Munk equations and its shortcomings. An important example is the work of Saunderson [24], in which a formula was proposed which attempts to account for both external and internal surface reflection. This formula is used to adjust the measured reflectances from which \( K \) and \( S \) are determined. Given the Fresnel re-
The reflectance equation:

\[ k_1 = \left( \frac{n_2 - n_1}{n_2 + n_1} \right)^2 \]  \hspace{1cm} (6)

where

- \( n_1 \) is the refraction index of the external medium (i.e., the air).
- \( n_2 \) is the refraction index of the internal medium (i.e., the oil).

Saunderson's formula is as follows:

\[ R_t = k_1 + \frac{R_m - k_1}{1 - k_1 - k_2 + k_2 R_m} \]  \hspace{1cm} (7)

where

- \( R_t \) is the theoretical reflectance; spectral reflectance adjusted for use in determining \( K \) and \( S \).
- \( R_m \) is the measured reflectance; spectral reflectance measured by a spectrophotometer.
- \( k_1 \) is the front-surface reflectance of the film.
- \( k_2 \) is the internal reflectance of the film.

4. PIGMENT MIXING

Given a set of pigments with absorption and scattering data, we can calculate the absorption and scattering curves for any specified mixture of the pigments using Equations 3 through 5 in the previous section. The \( K \) and \( S \) curves for these pigments can then be used to derive the spectral curves for the mixtures using Equation 1. Finally, the colors of these mixtures can be displayed on the computer monitor by converting the spectral curve to CIE \( XYZ \) coordinates and then transforming into the \( RGB \) primaries of the monitor. This ability to calculate correct pigment mixtures and display those results accurately on the monitor means that we can now bring the world of realistic color selection and mixture into computer graphics. For instance, artists could use their experience with paints and palettes to select and mix colors on the computer just as they would in the studio; they are no longer limited to the \( RGB \) additive color selection methods that monitor technology has imposed.

The difference between organic and inorganic paint mixtures provides a good example of why it is critical to use Kubelka–Munk theory to predict pigment mixtures. This organic/inorganic distinction is an important one for artists; mixtures of white with inorganic paints tend to "gray" out much more than those with organic pigments. For example, mixtures of napthol red (an organic paint) with titanium dioxide and cadmium red (an inorganic paint) with titanium dioxide can be compared. Although the red colors look very similar in their masstone states, their respective mixtures with white are quite different. This difference can be seen in Figure 7 where the effect is...
demonstrated using real pigments. These differences are known to those who use paints; artists choose their palettes very carefully based on the types of colors they would like to produce from paint mixtures.

The existing algorithms that are used in computer graphics to simulate paint mixing are unable to capture this difference between organic and inorganic pigments. Commercially available electronic paint programs use an RGB instead of a spectral representation for the paints that are mixed. A weighted average based upon the convolution mask of the brush profile is taken between the RGB of the paint being applied and the RGB of the background. Oddy and Willis [22] have extended this approach to include a pigment medium, but it is still an RGB averaging technique. For efficiency, paint programs often use the non-gamma corrected RGB values, which means that the mixing cannot even be described as additive.

Given the RGB values for the pure cadmium red, napthol red, and titanium white pigments used in Figure 7, a pigment mixing simulation that applies this weighted average RGB approach leads to the result shown in Figure 8. Because there is no way to directly relate the pigment concentrations used to create Figure 8 to the proportions employed in the weighted average RGB approach, the color scales in this figure were computed by linearly interpolating between the RGB values for titanium white and the RGB values for each of the two pigments. The white pigment RGB to red pigment RGB proportions are identical for corresponding swatches in the two color scales, and the initial proportions were selected so that the begin-
Fig. 8. Result of using paint mixing approach employed by traditional painting programs to simulate the mixture of cadmium red and napthol red pigment with titanium white.

...ning colors in the real and the simulated cadmium red scales were close (an exact match was not possible). Note that because both pigments in their masstone states have almost identical RGB values, there is little difference between the cadmium red and the napthol red color scales in the simulation. Converting from an RGB to a CMY representation has no effect on this result because the transformation is linear. A wavelength-based approach that models the pigments as cyan, yellow, and magenta filters is also incapable of capturing the mixing differences between the two pigments.

In order to correctly model pigment mixing in computer graphics, a pigment-based approach, such as the Kubelka-Munk method, must be employed. This is the only way to include the absorption and scattering phenomena that take place in real paint films. An important step in this approach is to obtain data for the pigments which is suited to the purposes of the application. In our case, we were comparing our research results with artists' oil paints, so we needed to get data for the pigments in a medium of oil. Off-the-shelf oil paints can contain various ingredients in addition to the main pigment, so we used pure dry pigment and linseed oil.

To obtain K and S data for the cadmium red and napthol red pigments, it was first necessary to create masstone solutions for cadmium red, napthol red, and titanium white; this was done by mixing linseed oil with the pure pigments. Next, a tint with titanium white was created for each red pigment.
by combining measured amounts of dry pigment with oil. Spectral measurements were then taken of all masstones and tints (Figure 9). From the spectral curve for titanium white, we deduced $K$ and $S$ values for that pigment by setting $S$ equal to one at all wavelengths and solving for $K$. Using these absorption and scattering values and the spectral curves for the masstones and tints for each red pigment, the $K$ and $S$ curves were calculated as described in the previous section (Figure 10). Finally, given the data for all three pigments, the spectral curves for arbitrary concentrations were calculated using Equation 1.

The results of the Kubelka-Munk modeled canvas are seen in Figure 11. This example compares favorably with the actual paint mixtures in Figure 7. It is interesting to contrast this result with that shown for additive mixing in Figure 8. With both the real mixtures and the Kubelka-Munk modeled mixtures, there is a marked difference between the two stripes; cadmium red goes to white much faster than napthol red. But in the additive mixing example, both pigments proceed at about the same rate and the difference at any given concentration is not very significant. These differences are quantifiable, using a color difference formula. In Figure 12, we have plotted $\Delta E$ as measured in CIE $L^*a^*b^*$ coordinates for the Kubelka-Munk and additive pictures. We can see that the color differences between the pigment scales in the Kubelka-Munk experiment are much higher for the tints at smaller red pigment concentrations. But with the additive experiment, the color differ-

Fig. 10. K and S curves corresponding to napthol red pigment shown in Fig. 9

Fig. 11. Result of using Kubelka-Munk theory to simulate the mixture of cadmium red and napthol red pigment with titanium white.

ACM Transactions on Graphics, Vol. 11, No. 4, October 1992
ences remain relatively flat, showing that the mixing differences for the pigments were not picked up by the additive mixing approach.

Another experiment that was performed for paint mixing calculations involved modeling a color wheel which was painted using standard off-the-shelf paints and then modeled using data for the pigments in those paints. The color wheel (Figure 13) was thus constructed: the outer circle uses just the three primaries—alizarin red, prussian blue, and cadmium yellow medium—as marked on the canvas. These primaries were mixed in 50/50 proportions to derive secondaries, which lie halfway between each of the primaries. Each secondary was then mixed with each primary to derive the tertiaries, which complete the circle. The inner circle is composed of 50/50 mixtures of the primaries and secondaries with titanium dioxide, shown in the middle of the wheel.

The possible gamut in CIE $L^*a^*b^*$ space that can be produced using these four pigments was computed using Kubelka-Munk theory and is shown in Figure 14. For comparison, consider Figure 15 which shows the position in CIE $L^*a^*b^*$ space of the color television monitor gamut from Figure 1. Clearly the pigments are not capable of producing colors as saturated as those created by the monitor. It is also interesting to note that a very dark color is also not possible using the four pigments selected. A simulation of the canvas shown in Figure 13 was also done and is included as part of Figure 19 which is discussed in Section 6. It should be noted that both the gamut in Figure 14 and the canvas in Figure 19 were not computed using data that we measured ourselves but from masstone and tint curves published elsewhere [11]. As described in the previous section, such data must be used with care, because of differences that may exist in materials and measurement techniques.

There are many potential uses for this approach to color mixing. One of these applications is in computer painting programs; as the organic/inorganic mixing experiment pointed out, much more realistic results can be derived by using these pigment mixing techniques than by using current painting program mixing techniques. One such application that has been developed is an airbrush program that uses the Kubelka-Munk approach to color mixing to predict realistic mixing of colors on the painting canvas.

Instead of using arbitrary RGB colors for the painting tool, the user selects colors from among a palette of available pigments (Figure 16). This palette is limited only by the pigment data available to the program. When the user selects a color, the K and S data for that pigment is loaded into the airbrush. Painting then occurs in the usual paint program manner except that, where colors overlap on the canvas, their mixture is correctly computed and displayed. This is accomplished by having each pixel, as well as the brush, keep track of its own current K and S values. When any pixel is painted over, new K and S values are calculated using the data for the brush and the pixel and Equations 3 and 4. Then the reflectance of the pigment mixture is determined, the RGB tristimulus values are found, and the appropriate color is displayed in that pixel.

This painting program uses a K and S frame buffer technique which stores the current K and S values for each pixel at four specific wavelengths. These wavelengths were selected to compute tristimulus values from spectral energy distributions by using Gaussian quadrature [19]. Limiting the wavelength information is necessary for the purposes of real time interaction and

conservation of memory resources. Since the pigment reflectances are quite smooth, the errors caused by this low sampling rate are minimized. The program also manipulates the floating point $K$ and $S$ values in appropriate integer form to save memory resources. Future hardware developments may make it possible to avoid these optimization methods and still maintain good performance.

5. COLOR MATCHING
Most scene generation in computer graphics involves specification of object color by RGB triplets; this method is certainly an easy way to pick a color,
and it is also readily interpreted by the computer during scene computation. However, when color fidelity is critical and when rendering techniques, such as ray tracing and radiosity are being used, a wavelength based approach to specifying color is necessary. A database of spectral reflectances can be created [20] to facilitate the use of such curves in image synthesis. Colors available when such an approach is used are limited to the number of curves in the database. The number of possible colors can be increased, and the number of curves in the database can be decreased, by storing pigment data instead of individual reflectances. Pigment mixtures can be determined by

using the techniques described in the previous section. However, specifying color in terms of pigment concentrations is not very intuitive for most people; if a particular shade of red is imagined for some plastic object in the scene, the typical computer user would probably not be able to come up with appropriate pigment concentrations which would produce this particular red.

What is required is a method for interpreting the user's intuitive color designation and then producing a pigment specification. This problem is known in the pigment industry as "color matching." In standard color matching problems, a sample of color is presented and a match is computed which most closely approximates that color. For instance, if you want to repaint part of your house but do not have the original paint, the paint store could give you a mixture which approximates that color using potentially very different pigments than were used originally. What we need to do in computer graphics matching is analogous. For example, in creating a graphics image, the user may want to specify the color of a wall using RGB values (which could be textually or graphically entered, depending on what user interface tools were available for the scene description). Then the matching program would convert this color into concentrations of whatever pigments provided the best match.

From the preceding sections of this article it is clear that, given a set of pigments about which the proper data is known, we can predict the outcome of mixtures of those pigments. Therefore it is reasonable to assume that we can use these capabilities to calculate color matches with the pigments—that is, given certain data about an arbitrary color, it is possible to specify which pigments and concentrations are required to approximate the closest match possible to that color.

For this project, we wanted a program that would allow the user to specify a color in terms of its color space coordinates, either RGB or CIE XYZ. An important additional feature was to give the user a graphical interface for this specification, so that the user did not have to know the actual coordinates but could choose the color interactively and visually. The program would also allow the user to specify various parameters and constraints for the program to use in calculating the match, such as how close the match should be and which set of pigments to use. Given this color specification, the program would compute the best color match with the given set of pigments.

There are several possible approaches to this matching problem. They differ in the types of data required for the computations, the computational complexity involved, and the suitability of the results. The most straightforward approach to solving the problem is to iterate through the set of available pigments, calculate their mixtures in all possible combinations and concentrations, and choose the mixture that comes the closest to matching the sample. Although this brute-force method is certainly the slowest possible, it is also the most accurate; since all possible mixtures are checked, the result is the most correct. In addition, this method provides the greatest flexibility in terms of the data required for the calculations, because data about the sample can be provided in any suitable form. The calculations depend on the data only to the extent that it is needed to see how close a
mixture has come to matching the sample. The search time can be limited by various means, including: restricting the set of pigments being searched, halting calculations when a mixture comes within some specified tolerance, and having the program intelligently direct the search based on some criteria such as pigment suitability. Although these methods can shorten the search time significantly, a much more efficient method is required.

Another approach to color matching is spectrophotometric matching [2, 5, 17, 23, 27], in which the calculations attempt to find a match with a spectral curve that best fits the spectral curve of the given color. While the algorithm used for this approach is much more efficient than the brute-force approach described above, the data required for the calculations are not appropriate for our application. Since the color we would like to match may be specified only in terms of CIE XYZ or RGB, we do not necessarily have a characteristic spectral curve to match against. Clearly, another method is needed.

Tristimulus matching, a better approach, is the one taken in most of the color matching references used in this research, including [2, 3, 13, 17, 18]. Just as in spectrophotometric matching, this method intelligently narrows down the search using a steepest descent method of optimization to find the best match for a set of pigments. But this method also has the advantage of requiring that only the tristimulus coordinates of the sample color be matched. Tristimulus matching makes a linear approximation of the nonlinear functions involved in the Kubelka-Munk equations, so it may require several iterations to arrive at an answer (if a match is possible). Each iteration results in a new trial match which is then compared to the sample using some color difference formula. If the trial is not within a pre-specified tolerance, then a new match is calculated. If at any time the concentration values of any of the pigments become negative, then a match with these pigments is not possible.

While this method is adequate for many cases of color matching, it has some disadvantages. Most significantly, this approach as described in the literature is limited to a given set of three pigments plus white. To use the formulas as is, a general match from a larger set of pigments could only be derived by iterating through the set of pigments, coming up with the best match from all sets of three. But in a general color-matching case, this may be terribly inefficient; a large pigment database would require a substantial calculation time to find the best match. Furthermore, the user may not necessarily want to be limited to three pigments in the resulting color match. If a better match could be derived from a larger set of pigments, then it would be helpful to provide that match instead.

Another problem with the tristimulus method is that the amount of change required in the pigment concentrations for the next match is based on the absolute differences between the CIE XYZ coordinates of the current trial match and the sample match; a color difference formula is used to judge whether the trial is a good match, but CIE XYZ coordinates are used to derive each new trial. The color difference formula would be a better metric to use in deriving the trial matches as is done in Cogno et al. [7]. Finally, it is important to account for metamerism in any color match. Metamerism is the
degree to which a color changes in different lighting and viewing conditions. Usually, we do not want a color to appear very different across environments, so we would like to limit the degree of metamerism to some tolerance level.

A matching algorithm must therefore take into account an arbitrarily large set of pigments, make better use of the color difference formula, and have the ability to control the degree of metamerism. Its main goal, of course, is to find a match which is as close to the sample color as possible. In other words, the color difference should be equal to zero. Therefore, we need to minimize the color difference function.

The CIE $L^*a^*b^*$ color difference formula [28] is:

$$\Delta E = \left[ (\Delta a^*)^2 + (\Delta b^*)^2 + (\Delta L^*)^2 \right]^{1/2}$$

where

$$L^* = 116 \left( \frac{Y}{Y_0} \right)^{1/3} - 16$$

$$a^* = 500 \left[ \left( \frac{X}{X_0} \right)^{1/3} - \left( \frac{Y}{Y_0} \right)^{1/3} \right]$$

$$b^* = 200 \left[ \left( \frac{Y}{Y_0} \right)^{1/3} - \left( \frac{Z}{Z_0} \right)^{1/3} \right]$$

and

$X, Y, Z$ tristimulus values of the color in question

$X_0, Y_0, Z_0$ tristimulus values of the light source

$\Delta a^*, \Delta b^*, \Delta L^*$ differences in values between colorants.

First we note that minimizing this value can also be achieved by minimizing its square, $\Delta E^2$, thus we can eliminate the square root from Equation 8. To find the minimum for $\Delta E^2$, any of several numerical methods may be used. We used the method of steepest descent, which follows the negative of the gradient at any given point on the function to get as close to zero as possible. This is an iterative method, like the tristimulus method explained above, which takes guesses at minimum values and then refines these guesses by repeated application of the same equation. Both this method and the tristimulus method depend on some initial guess. Allen [1] describes an algorithm to derive an intelligent first guess based on spectrophotometric data. Lacking this type of data, we use an arbitrary guess for the first trial, as suggested by Billmeyer and Saltzman [6].

An important constraint which complicates the situation is that the concentrations of the pigments should remain positive. The tristimulus method, as described in the references, does not attempt to find the closest match. Instead, if a match within the specified tolerance cannot be found, its calculations return negative values for some pigment concentrations, indicating that a match is not possible. Our method, however, should be able to
obtain the closest match possible. This ability is especially important for what we are trying to accomplish—matching RGB colors with their closest subtractive approximations; many colors in RGB space may not have exact matches with any combination of pigments. Instead of using the positive concentration value as a constraint on the problem, though, we can rewrite the concentration value \( c_i \) as the square of some value \( q_i \). This will allow the equations to be solved for the optimal \( q_i \) values while the \( c_i \) values remain positive. Using this technique allows us to keep the values of the concentrations positive without adding a complicating constraint to the problem.

Now we can proceed with finding the gradient for \( \Delta E^2 \). This will be the direction vector toward the minimum. For the gradient we will need the partial derivatives of the formula with respect to the \( q_i \) of each of the pigments involved. We know that:

\[
\frac{\partial E^2}{\partial q_i} = \frac{\partial \Delta E^2}{\partial c_i} \times 2q_i. \tag{12}
\]

Next, by the Chain Rule:

\[
\frac{\partial \Delta E^2}{\partial c_i} = \frac{\partial \Delta E^2}{\partial X} \frac{\partial X}{\partial c_i} + \frac{\partial \Delta E^2}{\partial Y} \frac{\partial Y}{\partial c_i} + \frac{\partial \Delta E^2}{\partial Z} \frac{\partial Z}{\partial c_i}. \tag{13}
\]

These component partials are written as follows:

\[
\frac{\partial \Delta E^2}{\partial X} = \frac{1000 \Delta a^*}{3X^{2/3}X_D^{1/3}} \tag{14}
\]

\[
\frac{\partial \Delta E^2}{\partial Y} = \frac{-1000 \Delta a^* + 232 \Delta L^* + 400 \Delta b^*}{3Y^{2/3}X_D^{1/3}} \tag{15}
\]

\[
\frac{\partial \Delta E^2}{\partial Z} = \frac{-400 \Delta b^*}{3Z^{2/3}X_D^{1/3}} \tag{16}
\]

and

\[
\frac{\partial X}{\partial c_i} = k \sum_\lambda S_\lambda \tilde{x}_\lambda \frac{\partial R_\lambda}{\partial c_i}, \tag{17}
\]

\[
\frac{\partial Y}{\partial c_i} = k \sum_\lambda S_\lambda \tilde{y}_\lambda \frac{\partial R_\lambda}{\partial c_i}, \tag{18}
\]

\[
\frac{\partial Z}{\partial c_i} = k \sum_\lambda S_\lambda \tilde{z}_\lambda \frac{\partial R_\lambda}{\partial c_i}, \tag{19}
\]

where

\[
\frac{\partial R}{\partial c_i} = \frac{S_M K_i - K_M S_i}{S_M^2} \times \left( 1 - \sqrt{\frac{K_M}{S_M}} + \frac{1}{\sqrt{\frac{K_M}{S_M}} + 2\left(\frac{K_M}{S_M}\right)} \right). \tag{20}
\]
and

\[ k = \frac{100}{\sum \lambda S_\lambda \tilde{\lambda}} \]

\[ \tilde{\lambda}, \bar{\lambda}, \bar{\lambda} \]

observer matching functions

\[ S_\lambda \]

given light source

\[ R_\lambda \]
spectral curve of match

\[ K_M, S_M \]

\( K \) and \( S \) for the current match trial

\[ K_i, S_i \]

\( K \) and \( S \) for the \( i \)th pigment

\[ \lambda \]

wavelength; our method uses 16 wavelengths, spaced at 20 nm increments from 400 to 700 nm.

Given the current match trial, the gradient is just a vector containing the values for the left side of Equation 12. Some multiple of this gradient is added to each current \( q_i \), which are then used to derive new values for each \( c_i \). These new \( c_i \) values are used to calculate a new match trial. If the match is within a specified tolerance of the sample or if a match is not possible but the nearest minimum has been reached, then the program is finished. If not, the process is repeated.

The factor by which the gradient is multiplied has a severe effect on the calculation process, determining how fast an answer can be reached and how good that answer is. For our program, we based the multiplier on a scalar factor of the current color difference between the sample and the match; trials that were further away would take longer steps than those closer in. This method helps the program avoid overshooting the minimum and having to backtrack. The other important consideration for the multiplier is for those cases in which no perfect match is possible; in these cases, we would like to derive the closest minimum. The gradient multiplier helps this process. If the program ever overshoots the minimum, it is clear that the current step size is too large. So a scaling factor is input into the equation which decreases the gradient multiplier by half. Overshooting again causes another halving of the scalar multiplier. In this way, the closer the program gets to a minimum, whether it is a match or not, the smaller steps it will take. Execution then continues until either a match or a minimum has been found.

A further important enhancement can be made to this approach to account for metamerism. In order to keep the degree of metamerism as low as possible, the function to be minimized could be rewritten as some combination of equations of the form of Equation 8, where each equation was calculated for a different light source or viewing condition. Then minimizing this new formula would result in a match which was minimized across these different conditions. The effects of metamerism upon the result could be easily changed by weighting the different components of the equation appropriately. For example, let us suppose that we wished to achieve a match under CIE standard illuminant \( C \) but also wished to minimize metamerism for CIE standard illuminant \( A \). Let us further suppose that the metamerism component of the match was only \( \frac{1}{4} \) as important as the initial match. Then our minimizing equation would be \( \Delta E^2 = 4 \Delta E^2_C + \Delta E^2_A \).
The approach we took toward color matching provided two major advantages over the standard color matching techniques covered in the references. First, we minimize in terms of a color difference formula, so the results we get are the closest color match in terms of this formula. This is in contrast to the tristimulus approach, which minimizes over $X$, $Y$, and $Z$ and thus only achieves a minimum in tristimulus space, which does not necessarily correspond to a minimum color difference. Second, our formula admits an arbitrary number of pigments and attempts to find a suitable match from that set, regardless of how many pigments are to be used. This is helpful in many cases where either many pigments are required to make a match or the user does not know which limited set of pigments would be adequate. The tristimulus method formula, on the other hand, requires there to be a set number of pigments. In the usual form, this number is only three, though more could be added for different light sources [13, 17]; in this case, the user would have to know ahead of time which pigments were suitable or the machine would have to iterate through a larger set of pigments until an appropriate match were found.

There are some things about this approach, however, that would require further development for a full-featured color matching algorithm. First, this steepest descent approach is guaranteed to find a minimum, but it will not necessarily be the global minimum; i.e., if there are many pigments involved and the initial guess is far from the true global minimum, the equations might settle on a local minimum even though a better match might exist with some other combination. This problem can be easily addressed through pigment choice and through the initial guess, although a more automated way of achieving this is desirable. Also, the equation does not take into account the cost of the pigments, which is important in many color matching applications. Finally, there is currently no way to limit the number of pigments used in the resulting match, other than the total number of pigments involved; if the user inputs 100 pigments into the program, then the result could potentially be a mixture which uses a small concentration from each pigment.

We developed a program using the above techniques and specifications. The user inputs several pieces of information for the search to begin. To choose the color to be matched, the user can input color coordinates either textually using CIE $XYZ$ or $RGB$ coordinates or graphically (Figure 17). Next, the user may choose which pigments should be used in determining the match. The tolerance level, which is the color difference in CIE $L^*a^*b^*$ space below which the computer considers the match successful, can also be specified by the user. All of these selections are handled via a dialog box (Figure 17).

Using this information, the program uses the method described above, iterating through several trial matches until a minimum is found. Data for the closest match is displayed on the screen, including concentration values for each of the chosen pigments, the color difference, the resulting CIE $XYZ$ values, a color swatch of the match color, and windows displaying the spectral curve and color of each trial used in deriving the match (Figure 18).
6. CONCLUSION

The purpose of this work was to demonstrate the capabilities of realistic pigment modeling in the field of computer graphics. Techniques used in the pigment industry have been modified to work within the domain of computer graphics to facilitate more realistic computer color specification. The opening sections of this paper motivated and validated the approach taken here. The mixing techniques discussed in Section 4 provide the basis for better, more realistic painting programs and other tools which require mixing and displaying pigmented materials. Finally, the matching techniques discussed in Section 5 show the feasibility of combining intuitive color selection with realistic color computation in generating realistic computer images.

As a final example, consider the image in Figure 19. Using the matching and mixing programs described in this article we can derive spectral curves for arbitrary pigment mixtures. These curves can then be used to specify realistic colors for materials in computer generated scenes. Every color in the scene depicted in Figure 19 was described by a characteristic spectral curve. Most of the objects, except the $RGB$ monitor picture and the white canvas, use curves which represent real pigment mixtures calculated by our programs. For example, the color wheel on the canvas is based upon the pigments used to create the artwork shown in Figure 13. With the exception of the color television picture, notice that unrealistic highly saturated colors do not occur. The range of colors possible has been constrained in a realistic way by the selection of a limited number of pigments and by the accurate simulation of their mixing properties.
APPENDIX

The Kubelka-Munk equations for complete hiding given in Section 3 were used to perform the pigment mixing work described in this paper. They are derived here from the initial assumptions because this detail is not available in most of the widely available references.

The situation as it occurs in a given pigmented solution is shown in Figure 20. In this figure, the pigmented solution is a paint, applied in a uniform thickness of \( x \) over a substrate of reflectance \( R_s \). Let \( i \) be the incident light which is transmitted down through the paint, either directly from the source or by scattering of light off of pigment particles. Let \( j \) be the light which is returning to the surface, either by being reflected off of the substrate or scattered from pigment particles.

Suppose we are examining some very thin sub-layer of the paint, of thickness \( dx \). The amount of light that is lost from \( i \) through the layer is:

\[
(K + S)i \, dx
\]  (21)

and the amount lost from \( j \) is:

\[
(K + S)j \, dx
\]  (22)

where \( S \) is the scattering component and \( K \) is the absorption. Now we must add back to \( i \) and \( j \) that light which was scattered from light coming in the

other direction; i.e., the light that was scattered from the incident light is
added to the light returning to the surface and vice versa. Then we have:

\[
\begin{align*}
    di &= (K + S)i dx - Sj dx \quad (23) \\
    -dj &= (K + S)j dx - Si dx. \quad (24)
\end{align*}
\]

Note that since \(i\) and \(j\) represent light vectors traveling in opposite direc-
tions, the change of their signs through the film are opposite.

The solution to Equations 23 and 24 begins by setting \(a = (1 + K/S)\) and
reexpressing them as

\[
\begin{align*}
    \frac{di}{S dx} &= ai - j \quad (25) \\
    -\frac{dj}{S dx} &= aj - i. \quad (26)
\end{align*}
\]

Adding these two equations together and rearranging leads to

\[
\frac{idj - jdi}{i^2 S dx} = -2a \frac{j}{i} + \frac{j^2}{i^2} + 1. \quad (27)
\]

From the Quotient Rule, we get:

\[
\frac{d(j/i)}{S dx} = -2a \frac{j}{i^2} + \frac{j^2}{i^2} + 1. \quad (28)
\]

Setting \(r = j/i\) then gives us

\[
\frac{dr}{S dx} = r^2 - 2ar + 1 \quad (29)
\]
and therefore

\[ \int \frac{dr}{r^2 - 2ar + 1} = S \int dx. \quad (30) \]

We would like to know the value of the change in \( r \) as the thickness varies from 0 to some arbitrary thickness \( t \). At a thickness of 0, the value of \( r \) will simply be the reflectance of the substrate, \( R_{\mu} \). At a thickness of \( t \), the reflectance is equal to some \( R \). Using integration by partial fractions we have

\[
\int_{R_{\mu}}^{R} \frac{dr}{r^2 - 2ar + 1} = \frac{1}{2b} \int_{R_{\mu}}^{R} \frac{dr}{r - (a + b)} - \frac{1}{2b} \int_{R_{\mu}}^{R} \frac{dr}{r - (a - b)}
\]

\[
= \frac{1}{2b} \ln \left( \frac{R - a - b)(R_{\mu} - a + b)}{(R - a + b)(R_{\mu} - a - b)} \right)
\]

\[ (31) \]

where \( b = \sqrt{a^2 - 1} \). Therefore, from Equation 30

\[
\frac{1}{2b} \ln \left( \frac{R - a - b)(R_{\mu} - a + b)}{(R - a + b)(R_{\mu} - a - b)} \right) = St.
\]

\[ (32) \]

The particular thickness that we are interested in solving for is \( t = \infty \), as this thickness provides complete hiding of the substrate. We can solve for the reflectance at this thickness, \( R_{\infty} \), as \( t \to \infty \):

\[
\frac{\ln \left( \frac{R - a - b)(R_{\mu} - a + b)}{(R - a + b)(R_{\mu} - a - b)} \right)}{e^{2Stb}} = (R - a + b)(R_{\mu} - a - b).
\]

\[ (33) \]

For the following substitutions,

\[
R = R_{\infty}
\]
\[
R_{\mu} = 0
\]
\[
t = \infty
\]

we get:

\[
R_{\infty} = \frac{1}{a + \sqrt{a^2 - 1}}
\]

\[
= \frac{1}{1 + \frac{K}{S} + \sqrt{\left(1 + \frac{K}{S}\right)^2 - 1}}.
\]

\[ (35) \]

This form is one variation of the Kubelka-Munk formula for computing \( R_{\infty} \). The more common forms can be obtained from this by rearrangement and are given in the text as Equations 1 and 2.
ACKNOWLEDGMENTS

The authors would like to thank Robert Gamblin of Gamblin Oils for his help in understanding the world of artists' paints, Michael Kelly for writing the Airbrush pigment painting program shown in Figure 16, Christy Wiseman for creating the artwork displayed in Figure 16, Linda Peting for developing the color space visualization program that generated Figures 1, 4, 3, 5, and 14, Sohail Mohammad for modeling the scene in Slides 8 and 11, and David Collett for his mathematical assistance in obtaining the color matching method used in Section 5. The images shown in Slides 8, 11, and 19 were produced using the Radiance software package developed by Greg Ward at the Lawrence Berkeley Laboratories and modified by Roy Ramberg at the University of Oregon to handle light on a wavelength-by-wavelength basis.

REFERENCES


Received September 1991; revised March 1992; accepted June 1992.