Blue Screen Matting

Alvy Ray Smith and James F. Blinn[1]
Microsoft Corporation

ABSTRACT

A classical problem of imaging—the *matting problem*—is separation of a non-rectangular foreground image from a (usually) rectangular background image—for example, in a film frame, extraction of an actor from a background scene to allow substitution of a different background. Of the several attacks on this difficult and persistent problem, we discuss here only the special case of separating a desired foreground image from a background of a constant, or almost constant, *backing* color. This backing color has often been blue, so the problem, and its solution, have been called *blue screen matting*. However, other backing colors, such as yellow or (increasingly) green, have also been used, so we often generalize to *constant color matting*. The mathematics of constant color matting is presented and proven to be unsolvable as generally practiced. This, of course, flies in the face of the fact that the technique is commonly used in film and video, so we demonstrate constraints on the general problem that lead to solutions, or at least significantly prune the search space of solutions. We shall also demonstrate that an algorithmic solution is possible by allowing the foreground object to be shot against *two* constant backing colors—in fact, against two completely arbitrary backings so long as they differ everywhere.

Key Words: Blue screen matte creation, alpha channel, compositing, chromakey, blue spill, flare, backing shadows, backing impurities, separating surfaces, triangulation matting.

CR Categories: I.3.3, I.4.6, J.5.

DEFINITIONS

A *matte* originally meant a separate strip of monochrome film that is transparent at places, on a corresponding strip of color film, that one wishes to preserve and opaque elsewhere. So when placed together with the strip of color film and projected, light is allowed to pass through and illuminate those parts desired but is blocked everywhere else. A *holdout matte* is the complement: It is opaque in the parts of interest and transparent elsewhere. In both cases, partially dense regions allow some light through. Hence some of the color film image that is being matted is partially illuminated.

The use of an *alpha channel* to form arbitrary compositions of images is well-known in computer graphics [9]. An alpha channel gives shape and transparency to a color image. It is the digital equivalent of a holdout matte—a grayscale channel that has full value pixels (for opaque) at corresponding pixels in the color image that are to be seen, and zero valued pixels (for transparent) at corresponding color pixels not to be seen. We shall use 1 and 0 to represent these two alpha values, respectively, although a typical 8-bit implementation of an alpha channel would use 255 and 0. Fractional alphas represent pixels in the color image with partial transparency.

We shall use “alpha channel” and “matte” interchangeably, it being understood that it is really the holdout matte that is the analog of the alpha channel.

The video industry often uses the terms “key” and “keying”—as in “chromakeying”—rather than the “matte” and “matting” of the film industry. We shall consistently use the film terminology, after first pointing out that “chromakey” has now taken on a more sophisticated meaning (e.g., [8]) than it originally had (e.g., [19]).

We shall assume that the color channels of an image are premultiplied by the corresponding alpha channel and shall refer to this as the *premultiplied alpha* case (see [9], [14], [15], [2], [3]). Derivations with non-premultiplied alpha are not so elegant.

THE PROBLEM

The mixing of several pictures—the *elements*—to form a single resulting picture—the *composite*—is a very general notion. Here we shall limit the discussion to a special type of composite frequently made in film and television, the *matte shot*. This consists of at least two elements, one or more *foreground objects* each shot against a special backing color—typically a bright blue or green—and a *background*. We shall limit ourselves to the case of one foreground element for ease of presentation.

The matting problem can be thought of as a perceptual process: the analysis of a complex visual scene into the objects that comprise it. A matte has been successfully *pulled*, if it in combination with the given scene correctly isolates what most humans would agree to be a separate object in reality from the other objects in the scene, that we can collectively refer to as the background. Note that this analysis problem is the reverse of classic 3D geometry-based computer graphics that synthesizes both the object and its matte simultaneously, and hence for which there is no matting problem.

There is also no matting problem of the type we are considering in the case of several *multi-film* matting techniques such as the sodium, infrared, and ultraviolet processes [6], [16]. These record the foreground element on one strip of film and its matte simultaneously on another strip of film.

[1] One Microsoft Way, Redmond, WA 98052-6399.
alvys@microsoft.com, blinn@microsoft.com.

The problem we address here is that of extracting a matte for a foreground object, given only a composite image containing it. We shall see that, in general, this is an underspecified problem, even in the case where the background consists of a single backing color. Note that a composite image contains no explicit information about what elements comprise it. We use the term "composite" to convey the idea that the given image is in fact a representation of several objects seen simultaneously. The problem, of course, is to determine, the objecthood of one or more of these objects. In the film (or video) world, the problem is to extract a matte in a *single-film* process—that is, one with no special knowledge about the object to be extracted, such as might be contained in a separate piece of film exposed simultaneously in a multi-film process.

Now a formal presentation of the problem: The *color* $C = [R\ G\ B\ \alpha]$ at each point of a desired composite will be some function of the color C_f of the foreground and color C_b of the new background at the corresponding points in the two elements. We have for convenience extended the usual color triple to a quadruple by appending the alpha value. As already mentioned, each of the first three *primary color coordinates* is assumed to have been premultiplied by the alpha coordinate. We shall sometimes refer to just these coordinates with the abbreviation $c = [R\ G\ B]$, for color C. For any subscript i, $C_i = [R_i\ G_i\ B_i\ \alpha_i]$ and $c_i = [R_i\ G_i\ B_i]$. Each of the four coordinates is assumed to lie on [0, 1]. We shall always assume that $\alpha_f = \alpha_b = 1$ for C_f and C_b—i.e., the given foreground and new background are opaque rectangular images.

The foreground element C_f can be thought of as a composite of a special background, all points of which have the (almost) constant backing color C_k, and a foreground C_o that is the foreground object in isolation from any background and which is transparent, or partially so, whenever the backing color would show through. We sometimes refer to C_o as the *uncomposited foreground color*. Thus $C_f = f(C_o, C_k)$ expresses the point-by-point foreground color as a given composite f of C_k and C_o. We shall always take $\alpha_k = 1$ for C_k.

We assume that f is the *over* function [9], $C_a + (1 - \alpha_a)\ C_b$, combining C_b with (premultiplied) C_a by amount α_a, $0 \le \alpha_a \le 1$. One of the features of the premultiplied alpha formulation is that the math applied to the three primary color coordinates is the same as that applied to the alpha coordinate. An alpha channel holds the factor α_a at every point in an image, so we will use channel and coordinate synonymously. This facilitates:

The Matting Problem

Given C_f and C_b at corresponding points, and C_k a known backing color, and assuming $C_f = C_o + (1 - \alpha_o)C_k$, determine C_o which then gives composite color $C = C_o + (1 - \alpha_o)C_b$ at the corresponding point, for all points that C_f and C_b share in common.

We shall call C_o—that is, the color, including alpha, of a foreground object—a *solution* to the Matting Problem. Once it is known at each point, we can compute C at each point to obtain the desired result, a composite over a new background presumably more interesting than a single constant color. We shall refer to the equation for C_f above as the Matting Equation. We sometimes refer to an uncomposited foreground object (those pixels with $\alpha_o > 0$) as an *image sprite*, or simply a *sprite*.

PREVIOUS WORK

Blue screen matting has been used in the film and video industries for many years [1], [6], [21] and has been protected under patents [17], [18], [19], [20] until recently. The most recent of these expired July, 1995. Newer patents containing refinements of the process still exist, however. Any commercial use of the blue screen process or extensions should be checked carefully against the extant patents—e.g., [22], [23], [24], [25], [5], [4].

An outstanding inventor in the field is Petro Vlahos, who defined the problem and invented solutions to it in film and then in video. His original film solution is called the *color-difference technique*. His video solution is realized in a piece of equipment, common to the modern video studio, called the Ultimatte®. It is essentially an electronic analog of the earlier color-difference film technique. He was honored in 1995 with an Academy Award for lifetime achievement, shared with his son Paul.

Vlahos makes one observation essential to his work. We shall call it the *Vlahos Assumption*: Blue screen matting is performed on foreground objects for which the blue is related to the green by $B_o \le a_2 G_o$. The usual range allowed by his technique is $.5 \le a_2 \le 1.5$ [20]. That this should work as often as it does is not obvious. We shall try to indicate why in this paper.

The Vlahos formula for α_o, abstracted from the claims of his earliest electronic patent [18] and converted to our notation, is

$$\alpha_o = 1 - a_1(B_f - a_2 G_f),$$

clamped at its extremes to 0 and 1, where the a_i are tuning adjustment constants (typically made available as user controls). We will call this the First Vlahos Form. The preferred embodiment described in the patent replaces B_f above with $\min(B_f, B_k)$, where B_k is the constant backing color (or the minimum such color if its intensity varies, as it often does in practice). In the second step of the Vlahos process, the foreground color is further modified before compositing with a new background by clamping its blue component to $\min(B_f, a_2 G_f)$.

A more general Vlahos electronic patent [20] introduces

$$\alpha_o = 1 - a_1(B_f - a_2(a_5 \max(r, g) + (1 - a_5)\min(r, g))),$$

where $r = a_3 R_f$, $g = a_4 G_f$, and the a_i are adjustment parameters. Clamping again ensures 0 and 1 as limiting values. We shall call this the Second Vlahos Form. Again the blue component of the foreground image is modified before further processing.

A form for α_o from a recent patent [4] (one of several new forms) should suffice to show the continued refinements introduced by Vlahos and his colleagues at Ultimatte Corp.:

$$\alpha_o = 1 - ((B_f - a_1) - a_2 \max(r, g) - \max(a_5(R_f - G_f), a_6(G_f - R_f))),$$

with clamping as before. They have continually extended the number of foreground objects that can be matted successfully.

We believe Vlahos et al. arrived at these forms by many years of experience and experiment and not by an abstract mathematical approach such as presented here. The forms we derive are related to their forms, as we shall show, but more amenable to analysis.

With these patents Vlahos defined and attacked several problems of matting: *blue spill* or *blue flare* (reflection of blue light from the blue screen on the foreground object), *backing shadows* on the blue screen (shadows of the foreground object on the backing, that one wishes to preserve as part of the foreground object), and *backing impurities* (departures of a supposedly pure blue backing screen from pure blue). We shall touch on these issues further in later sections.

Another contribution to matting [8] is based on the following thinking: Find a family of nested surfaces in colorspace that separate the foreground object colors from the backing colors. Each surface, corresponding to a value of α_o, is taken to be the set of colors that are the α_o blend of the foreground and backing colors.

See Fig. 4. The Primatte® device from Photron Ltd., based on this concept, uses a nested family of convex multi-faceted polyhedra (128 faces) as separating surfaces. We shall discuss problems with *separating surface models* in a later section.

THE INTRINSIC DIFFICULTY

We now show that single-film matting, as typically practiced in a film or video effects house, is intrinsically difficult. In fact, we show that there is an infinity of solutions. This implies that there is no algorithmic method for pulling a matte from a given foreground element. There must be a human—or perhaps someday a sufficiently intelligent piece of image processing software—in the loop who "knows" a correct matte when he (she or it) sees one, and he must be provided with a sufficiently rich set of controls that he can successfully "home in" on a good matte when in the neighborhood of one. The success of a matting machine, such as the Ultimatte or Primatte, reduces then to the cleverness of its designers in selecting and providing such a set of controls.

The argument goes as follows: We know that R_f is an interpolation from R_k to R_o with weight α_o, or $R_f = R_o + (1 - \alpha_o)R_k$, and that similar relations hold for G_f and B_f. This is $c_f = c_o + (1 - \alpha_o)c_k$ in our abbreviated notation. (We ignore the relation for α_f because it is trivial.) A complete solution requires R_o, G_o, B_o, and α_o. Thus we have three equations and four unknowns, an incompletely specified problem and hence an infinity of solutions, unsolvable without more information.

There are some special cases where a solution to the matting problem does exist and is simple.

SOLUTION 1: NO BLUE

If c_o is known to contain no blue, $c_o = [R_o \;\; G_o \;\; 0]$, and c_k contains only blue, $c_k = [0 \;\; 0 \;\; B_k]$, then

$$c_f = c_o + (1-\alpha_o)c_k = \begin{bmatrix} R_o & G_o & (1-\alpha_o)B_k \end{bmatrix}.$$

Thus, solving the $B_f = (1 - \alpha_o)\, B_k$ equation for α_o gives solution

$$C_o = \begin{bmatrix} R_f & G_f & 0 & 1-\dfrac{B_f}{B_k} \end{bmatrix}, \quad \text{if } B_k \neq 0.$$

This example is exceedingly ideal. The restriction to foreground objects with no blue is quite serious, excluding all grays but black, about two-thirds of all hues, and all pastels or tints of the remaining hues (because white contains blue). Basically, it is only valid for one plane of the 3D RGB colorspace, the RG plane.

The assumption of a perfectly flat and perfectly blue backing color is not realistic. Even very carefully prepared "blue screens" used in cinema special effects as backings have slight spatial brightness variations and also have some red and green impurities (backing impurities). A practical solution for brightness variations, in the case of repeatable shots, is this: Film a pass without the foreground object to produce a record of B_k at each point to be used for computing C_o after a second pass with the object.

We rather arbitrarily chose pure blue to be the backing color. This is an idealization of customary film and video practice (although one sees more and more green screens in video). We shall soon show how to generalize to arbitrary and non-constant backing colors and hence do away with the so-called backing impurities problem in certain circumstances.

SOLUTION 2: GRAY OR FLESH

The matting problem can be solved if c_o is known to be gray. We can loosen this claim to say it can be solved if either R_o or G_o equals B_o. In fact, we can make the following general statement:

There is a solution to the matting problem if R_o or $G_o = aB_o + b\alpha_o$, and if c_k is pure blue with $aB_k + b \neq 0$. To show this, we derive the solution C_o for the green case, since the solution for red can be derived similarly:

The conditions, rewritten in color primary coordinates, are:

$$c_f = \begin{bmatrix} R_o & aB_o + b\alpha_o & B_o + (1-\alpha_o)B_k \end{bmatrix}.$$

Eliminate B_o from the expressions for G_f and B_f to solve for α_o:

$$C_o = \begin{bmatrix} R_f & G_f & B_\Delta + \alpha_o B_k & \dfrac{G_f - aB_\Delta}{aB_k + b} \end{bmatrix}, \quad \text{if } aB_k + b \neq 0.$$

Here we have introduced a very useful definition $C_\Delta = C_f - C_k$.

The special case C_o gray clearly satisfies Solution 2, with $a = 1$ and $b = 0$ for both R_o and G_o. Thus it is not surprising that science fiction space movies effectively use the blue screen process (the color-difference technique) since many of the foreground objects are neutrally colored spacecraft. As we know from practice, the technique often works adequately well for desaturated (towards gray) foreground objects, typical of many real-world objects.

A particularly important foreground element in film and video is flesh which typically has color $[d \;\; .5d \;\; .5d]$. Flesh of all races tends to have the same ratio of primaries, so d is the darkening or lightening factor. This is a non-gray example satisfying Solution 2, so it is not surprising that the blue screen process works for flesh.

Notice that the condition $G_o = aB_o + b\alpha_o$, with $^2/_3 \le a \le 2$ and $b = 0$, resembles the Vlahos Assumption, $B_o \le a_2 G_o$. In the special case $b = 0$, our derived expression for α_o can be seen to be of the same form as the First Vlahos Form:

$$\alpha_o = 1 - \frac{1}{B_k}\left(B_f - \frac{1}{a} G_f \right).$$

Thus our B_k is Vlahos' $1/a_1$ and our a is his $1/a_2$. Careful reading shows that $B_k = 1/a_1$ is indeed consistent with [18]. By using these values, it can be seen that Vlahos' replacement of B_f by $\min(B_f, a_2 G_f)$ is just his way of calculating what we call B_o.

The next solution does not bear resemblance to any technique used in the real world. We believe it to be entirely original.

SOLUTION 3: TRIANGULATION

Suppose c_o is known against two different shades of the backing color. Then a complete solution exists as stated formally below. It does not require any special information about c_o. Fig. 1(a-d) demonstrates this *triangulation solution*:

Let B_{k_1} and B_{k_2} be two shades of the backing color—i.e., $B_{k_1} = cB_k$ and $B_{k_2} = dB_k$ for $0 \le d < c \le 1$. Assume c_o is known against these two shades. Then there is a solution C_o to the matting problem. N.B., c_{k_2} could be black—i.e., $d = 0$.

The assumption that c_o is known against two shades of B_k is equivalent to the following:

$$c_{f_1} = \begin{bmatrix} R_o & G_o & B_o + (1-\alpha_o)B_{k_1} \end{bmatrix}$$
$$c_{f_2} = \begin{bmatrix} R_o & G_o & B_o + (1-\alpha_o)B_{k_2} \end{bmatrix}.$$

The expressions for B_{f_1} and B_{f_2} can be combined and B_o eliminated to show $\alpha_o = 1 - \dfrac{B_{f_1} - B_{f_2}}{B_{k_1} - B_{k_2}}$, where the denominator is not 0 since the two backing shades are different. Then

$$R_o = R_{f_1} = R_{f_2} \qquad G_o = G_{f_1} = G_{f_2} \qquad B_o = \frac{B_{f_2} B_{k_1} - B_{f_1} B_{k_2}}{B_{k_1} - B_{k_2}}$$

completes the solution.

No commonly used matting technique asks that the foreground object be shot against two different backgrounds. For computer controlled shots, it is a possibility but not usually done. If passes of a computer controlled camera are added to solve the problem of nonuniform backing mentioned earlier, then the triangulation solution requires four passes.

Consider the backing shadows problem for cases where the triangulation solution applies. The shadow of a foreground object is part of that object to the extent that its density is independent of the backing color. For a light-emitting backing screen, it would be tricky to perform darkening without changing the shadows of the foreground objects. We will give a better solution shortly.

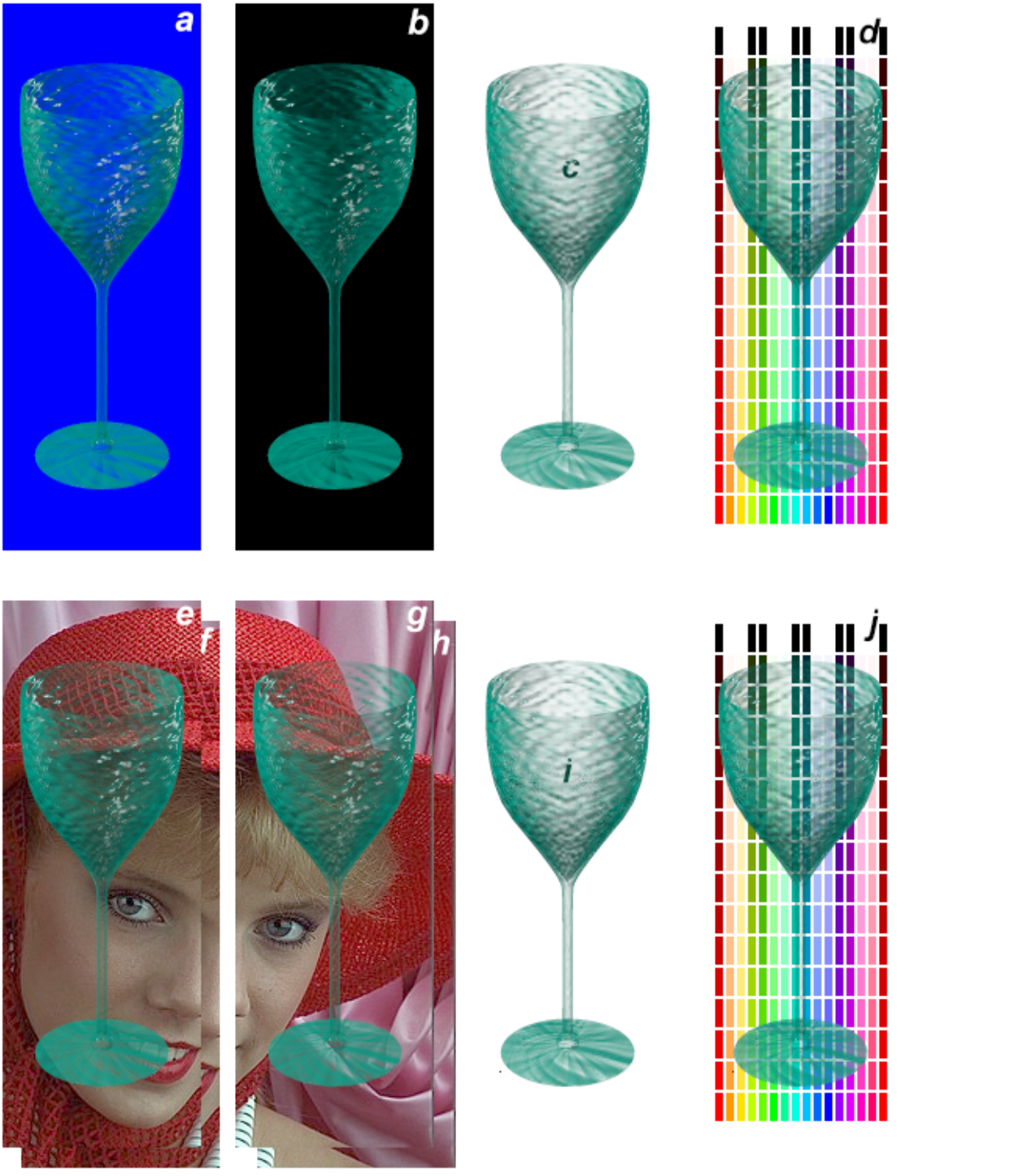

Figure 1. Ideal triangulation matting. (a) Object against known constant blue. (b) Against constant black. (c) Pulled. (d) Composited against new background. (e) Object against a known backing (f). (g) Against a different known backing (h). (i) Pulled. (j) New composite. Note the black pixel near base of (i) where pixels in the two backings are identical and the technique fails.

GENERALIZATIONS

The preceding solutions are all special cases of the generalization obtained by putting the Matting Equation into a matrix form:

$$C_o\begin{bmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ -R_k & -G_k & -B_k & t_4 \end{bmatrix} = \begin{bmatrix} R_\Delta & G_\Delta & B_\Delta & T \end{bmatrix},$$

where a fourth column has been added in two places to convert an underspecified problem into a completely specified problem. Let

$$\bar{t} = \begin{bmatrix} t_1 & t_2 & t_3 & t_4 \end{bmatrix}.$$

The matrix equation has a solution C_o if the determinant of the 4x4 matrix is non-0, or

$$t_1R_k + t_2G_k + t_3B_k + t_4 = \bar{t}\cdot C_k \neq 0.$$

Standard linear algebra gives, since $\alpha_\Delta = 0$ always,

$$\alpha_o = \frac{T - (t_1R_\Delta + t_2G_\Delta + t_3B_\Delta)}{\bar{t}\cdot C_k} = \frac{T - \bar{t}\cdot C_\Delta}{\bar{t}\cdot C_k} = 1 - \frac{\bar{t}\cdot C_f - T}{\bar{t}\cdot C_k}.$$

Then $c_o = c_\Delta + \alpha_o c_k$ by the Matting Equation.

Thus Solutions 1 and 2 are obtained by the following two choices, respectively, for $\bar{t}$ and T, where the condition on $\bar{t}\cdot C_k$ is given in parentheses:

$\bar{t} = [0 \quad 0 \quad 1 \quad 0]$; $\quad T = 0$; $\quad (B_k \neq 0)$

$\bar{t} = [0 \quad -1 \quad a \quad b]$; $\quad T = 0$; $\quad (-G_k + aB_k + b \neq 0)$.

The latter condition reduces to that derived for Solution 2 by the choice of pure blue backing color—i.e., $G_k = 0$. We state the general result as a theorem of which these solutions are corollaries:

Theorem 1. There is a solution C_o to the Matting Problem if there is a linear condition $\bar{t}\cdot C_o = 0$ on the color of the uncomposited foreground object, with $\bar{t}\cdot C_k \neq 0$.

Proof. $T = 0$ in the matrix equation above gives $\alpha_o = 1 - \frac{\bar{t}\cdot C_f}{\bar{t}\cdot C_k}$. ■

The Second Vlahos Form can be shown to be of this form with a_1 proportional to $1/(\bar{t}\cdot C_k)$. Geometrically, Theorem 1 means that all solutions C_o lie on a plane and that C_k does not lie on that plane.

Solution 3 above can also be seen to be a special case of the general matrix formulation with these choices and condition, where by extended definition $C_{\Delta_i} = C_{f_i} - C_{k_i}$, i = 1 or 2:

$$\bar{t} = \begin{bmatrix} 0 & 0 & 1 & -B_{k_2} \end{bmatrix}; \quad T = B_{\Delta_2}; \quad (B_{k_1} - B_{k_2} \neq 0),$$

with $C_k = \begin{bmatrix} 0 & 0 & B_{k_1} & 1 \end{bmatrix}$ and right side $\begin{bmatrix} R_{f_1} & G_{f_1} & B_{\Delta_1} & B_{\Delta_2} \end{bmatrix}$.

The condition is true by assumption. This solution too is a corollary of a more general one, C_{k_1} not restricted to a shade of blue:

Theorem 2. There is a solution C_o to the Matting Problem if the uncomposited foreground object is known against two distinct backing colors C_{k_1} and C_{k_2}, where C_{k_1} is arbitrary, C_{k_2} is a shade of pure blue, and $B_{k_1} - B_{k_2} \neq 0$.

Proof. This is just the matrix equation above with $\bar{t}$ and T as for Solution 3, but with C_k generalized to $\begin{bmatrix} R_{k_1} & G_{k_1} & B_{k_1} & 1 \end{bmatrix}$ and the right side of the matrix equation being $\begin{bmatrix} R_{\Delta_1} & G_{\Delta_1} & B_{\Delta_1} & B_{\Delta_2} \end{bmatrix}$.

Thus, as for Solution 3, $\alpha_o = \frac{B_{\Delta_2} - B_{\Delta_1}}{B_{k_1} - B_{k_2}} = 1 - \frac{B_{f_1} - B_{f_2}}{B_{k_1} - B_{k_2}}$. ■

The following generalization of Theorem 2 utilizes all of the C_{k_2} backing color information. Let the sum of the color coordinates of any color C_a be $\Sigma_a = R_a + G_a + B_a$.

Theorem 3. There is a solution C_o to the Matting Problem if the uncomposited foreground object is known against two distinct backing colors C_{k_1} and C_{k_2}, where both are arbitrary and $\Sigma_{k_1} - \Sigma_{k_2} = (R_{k_1} - R_{k_2}) + (G_{k_1} - G_{k_2}) + (B_{k_1} - B_{k_2}) \neq 0$.

Proof. Change $\bar{t}$ and T in the proof of Theorem 2 to

$$\bar{t} = \begin{bmatrix} 1 & 1 & 1 & -\Sigma_{k_2} \end{bmatrix}; \quad T = \Sigma_{\Delta_2}.$$

This gives $\bar{t}\cdot C_o = \Sigma_o - \alpha_o\Sigma_{k_2} = \Sigma_{f_2} - \Sigma_{k_2}$, which is exactly what you get by adding together the three primary color equations in the Matting Equation, $C_o - \alpha_o C_{k_2} = C_{\Delta_2}$. The solution is

$$\alpha_o = \frac{\Sigma_{\Delta_1} - \Sigma_{\Delta_2}}{\Sigma_{k_1} - \Sigma_{k_2}} = 1 - \frac{\Sigma_{f_1} - \Sigma_{f_2}}{\Sigma_{k_1} - \Sigma_{k_2}}$$

$$= 1 - \frac{(R_{f_1} - R_{f_2}) + (G_{f_1} - G_{f_2}) + (B_{f_1} - B_{f_2})}{(R_{k_1} - R_{k_2}) + (G_{k_1} - G_{k_2}) + (B_{k_1} - B_{k_2})},$$

$c_o = c_{\Delta_1} + \alpha_o c_{k_1} = c_{f_1} - (1 - \alpha_o)c_{k_1}$, or $c_o = c_{f_2} - (1 - \alpha_o)c_{k_2}$. ■

The conditions of Theorem 3 are quite broad—only the *sums* of the primary color coordinates of the two backing colors have to differ. In fact, a constant backing color is not even required. We have successfully used the technique to pull a matte on an object against a backing of randomly colored pixels and then against that same random backing but darkened by 50 percent. Fig. 1(e-j) shows another application of the technique, but Fig. 2 shows more realistic cases. See also Fig. 5.

The triangulation problem, with complete information from the two shots against different backing colors, can be expressed by this non-square matrix equation for an overdetermined system:

$$C_o\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ -R_{k_1} & -G_{k_1} & -B_{k_1} & -R_{k_2} & -G_{k_2} & -B_{k_2} \end{bmatrix} =$$

$$\begin{bmatrix} R_{\Delta_1} & G_{\Delta_1} & B_{\Delta_1} & R_{\Delta_2} & G_{\Delta_2} & B_{\Delta_2} \end{bmatrix}.$$

The Theorem 3 form is obtained by adding the last three columns of the matrix and the last three elements of the vector.

The standard least squares way [7] to solve this is to multiply both sides of the equation by the transpose of the matrix yielding:

$$C_o\begin{bmatrix} 2 & 0 & 0 & -(R_{k_1} + R_{k_2}) \\ 0 & 2 & 0 & -(G_{k_1} + G_{k_2}) \\ 0 & 0 & 2 & -(B_{k_1} + B_{k_2}) \\ -(R_{k_1} + R_{k_2}) & -(G_{k_1} + G_{k_2}) & -(B_{k_1} + B_{k_2}) & \Lambda \end{bmatrix} =$$

$$\begin{bmatrix} R_{\Delta_1} + R_{\Delta_2} & G_{\Delta_1} + G_{\Delta_2} & B_{\Delta_1} + B_{\Delta_2} & \Gamma \end{bmatrix}$$

where $\Lambda = R_{k_1}^2 + G_{k_1}^2 + B_{k_1}^2 + R_{k_2}^2 + G_{k_2}^2 + B_{k_2}^2$ and

$\Gamma = -(R_{k_1}R_{\Delta_1} + G_{k_1}G_{\Delta_1} + B_{k_1}B_{\Delta_1} + R_{k_1}R_{\Delta_1} + G_{k_1}G_{\Delta_1} + B_{k_1}B_{\Delta_1})$.

Inverting the symmetric matrix and multiplying both sides by the inverse gives a least squares solution C_o if the determinant of the matrix, $4((R_{k_1} - R_{k_2})^2 + (G_{k_1} - G_{k_2})^2 + (B_{k_1} - B_{k_2})^2)$, is non-0. Thus we obtain our most powerful result:

Theorem 4. There is a solution C_o to the Matting Problem if the uncomposited foreground object is known against two arbitrary backing colors C_{k_1} and C_{k_2} with nonzero distance between them—

$(R_{k_1} - R_{k_2})^2 + (G_{k_1} - G_{k_2})^2 + (B_{k_1} - B_{k_2})^2 \neq 0$ (i.e., distinct).

The desired alpha α_o can be shown to be one minus

$$\frac{(R_{f_1} - R_{f_2})(R_{k_1} - R_{k_2}) + (G_{f_1} - G_{f_2})(G_{k_1} - G_{k_2}) + (B_{f_1} - B_{f_2})(B_{k_1} - B_{k_2})}{(R_{k_1} - R_{k_2})^2 + (G_{k_1} - G_{k_2})^2 + (B_{k_1} - B_{k_2})^2}.$$

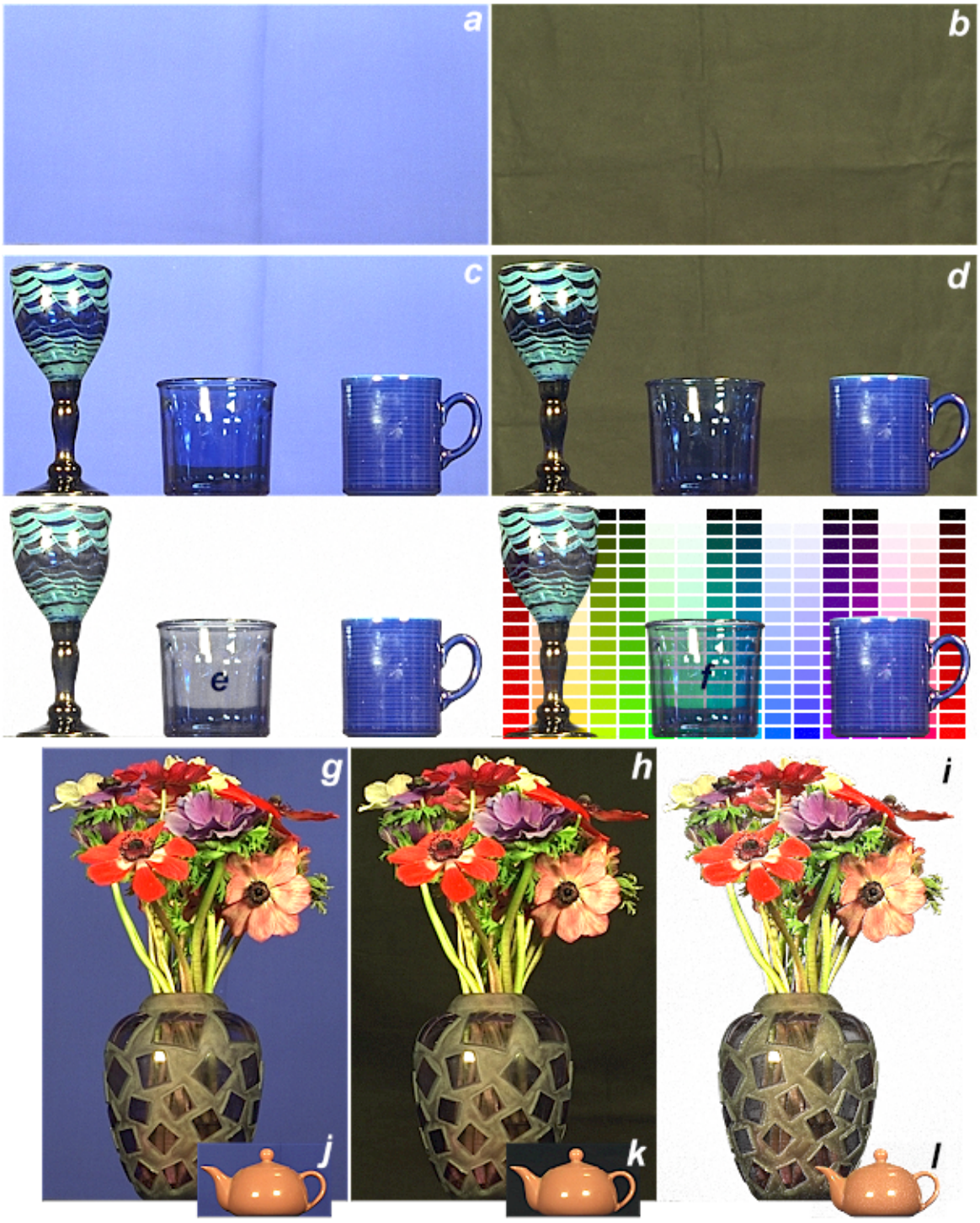

Figure 2. Practical triangulation matting. (a-b) Two different backings. (c-d) Objects against the backings. (e) Pulled. (f) New composite. (g-i) and (j-l) Same triangulation process applied to two other objects (backing shots not shown). (l) Object composited over another. The table and other extraneous equipment have been "garbage matted" from the shots. See Fig. 5.

The Theorem 3 and 4 expressions for α_o are symmetric with respect to the two backings, reflected in our two expressions for c_o (in the proof of Theorem 3).

Theorems 2 and 3 are really just special cases of Theorem 4. For Theorem 2, the two colors are required to have different blue coordinates. For Theorem 3, they are two arbitrary colors that do not lie on the same plane of constant Σ. In practice we have found that the simpler conditions of Theorem 3 often hold and permit use of computations cheaper than those of Theorem 4.

Theorem 4 allows the use of *very* general backings. In fact, two shots of an object moving across a fixed but varied background can satisfy Theorem 4, as indicated by the lower Fig. 1 example. If the foreground object can be registered frame to frame as it moves from, say, left to right, then the background at two different positions can serve as the two backings.

Notice that the Theorem 3 and 4 techniques lead to a backing shadows solution whereas simple darkening might not work. The additional requirement is that the illumination levels and light-emitting directions be the same for the two backing colors so that the shadows are the same densities and directions.

The overdetermined linear system above summarizes all information about two shots against two different backing colors. A third shot against a third backing color could be included as well, replacing the 4×6 matrix with a 4×9 matrix and the 1×6 right-hand vector with a 1×9 vector. Then the same least squares solution technique would be applied to find a solution for this even more overdetermined problem. Similarly, a fourth, fifth, etc. shot against even more backing colors could be used. An overdetermined system can be subject to numerical instabilities in its solution. We have not experienced any, but should they arise the technique of singular value decomposition [11] might be used.

IMPLEMENTATION NOTES

The Fig. 1(a-d) example fits the criteria of Theorem 2 (actually the Solution 3 special case) perfectly because the given blue and black screen shots were manufactured by compositing the object over perfect blue and black backings. As predicted by the theorem, we were able to extract the original object in its original form, with only small least significant bit errors. Similarly Fig. 1(e-j) illustrates Theorem 3 or 4.

Fig. 2 is a set of real camera shots of real objects in a real studio. Our camera was locked down for the two shots required by Theorem 3 and 4 plus two more required for backing color calibrations as mentioned before. Furthermore, constant exposure was used for the four shots, and a remote-controlled shutter guarded against slight camera movements. The results are good enough to demonstrate the effectiveness of the algorithm but are nevertheless flawed from misregistration introduced during the digitization process—pin registration was not used—and from the foreground objects having different brightnesses relative one another, also believed to be a scanning artifact.

Notice from the Theorem 3 and 4 expressions for α_o that the technique is quite sensitive to brightness and misregistration errors. If the foreground colors differ where they should be equal, then α_o is lowered from its correct value of 1, permitting some object transparency. In general, the technique tends to err towards increased transparency.

Another manifestation of the same error is what we term the "fine line" problem. Consider a thin dark line with bright surroundings in an object shot against one backing, or the complement, a thin bright line in a dark surround. Such a line in slight misregister with itself against the other backing can differ dramatically in brightness at pixels along the line, as seen by our algorithm. The error trend toward transparency will cause the appearance of a fine transparent line in the pulled object.

The conclusion is clear: To effectively use triangulation, pin-registered filming and digitization should be used to ensure positional constancy between the four shots, and very careful monitoring of lighting and exposures during filming must be undertaken to ensure that constant brightnesses of foreground objects are recorded by the film (or other recording medium).

Since triangulation works only for non-moving objects (excluding rigid motions, such as simple translation), it should be possible to reduce brightness variations between steps of the process due to noise by averaging several repeated shots at each step.

A LOWER BOUND

The trouble with the problems solved so far is that the premises are too ideal. It might seem that the problems which have Solutions 1 and 2, and Theorem 1 generalizations, are unrealistically restrictive of foreground object colors. It is surprising that so much real-world work approaches the conditions of these solutions. Situations arising from Solution 3, and Theorems 2-4 generalizations, require a doubling of shots, which is a lot to ask even if the shots are exactly repeatable. Now we return to the general single-background case and derive bounds on α_o that limit the search space for possible solutions.

Any C_o offered as solution must satisfy the physical limits on color. It must be that $0 \le R_o \le \alpha_o$ (since R_o is premultiplied by α_o) and similarly for G_o and B_o. The Matting Equation gives $R_f = R_o + (1 - \alpha_o)R_k$. The inequalities for R_o applied to this expression give

$$(1-\alpha_o)R_k \le R_f \le (1-\alpha_o)R_k + \alpha_o ,$$

with the left side being the expression for $R_o = 0$ and the right for $R_o = \alpha_o$. Similar inequalities apply to G_f and B_f. Fig. 3 shows all regions of valid combinations of α_o, R_f, G_f, and B_f using equality in the relationship(s) above as boundaries. The color c_k for this figure is taken to be the slightly impure blue [.1 .2 .98].

The dashed vertical lines in Fig. 3 represent a given c_f—in this figure, [.8 .5 .6]. The dotted horizontal lines represent the minimum α_o for each of R_f, G_f, and B_f which gives a valid R_o, G_o, and B_o, respectively. Let these three α_o's be called α_R, α_G, and α_B. Since only one α_o is generated per color, the following relationship must be true:

$$\alpha_o \ge \max(\alpha_R, \alpha_G, \alpha_B).$$

We shall call the α_o which satisfies this relationship at equality $\alpha_{\min}$, and any $\alpha_o \ge \alpha_{\min}$ will be called a valid one. Notice that although the range of possible α_o's is cut down by this derivation, there are still an infinity of valid ones to choose from, hence an infinity of solutions.

If $R_f > R_k$, as in the Fig. 3 example, then α_R corresponds to $R_o = \alpha_o$, the right side of the inequalities above for R_f and α_o. If $R_f < R_k$ then α_R corresponds to $R_o = 0$, the left side. Thus

$$\alpha_R = \begin{cases} 1-\dfrac{R_f}{R_k}, & \text{if } R_f < R_k \\ \dfrac{R_\Delta}{1-R_k}, & \text{if } R_f > R_k \\ 0, & \text{if } R_f = R_k \end{cases} .$$

In the example of Fig. 3, $\alpha_{\min} \approx .78$. For the special case of pure blue backing, $\alpha_{\min} = \max(R_f, G_f, 1 - B_f)$. So long as a valid α_o

exists, a foreground object color can be derived from the given c_f by $c_o = c_\Delta + \alpha_o c_k$ as before.

AN UPPER BOUND

Tom Porter pointed out (in an unpublished technical memo [10]) that an upper bound could also be established for α_o, by taking lessons from Vlahos.

The Vlahos Assumption, when valid, has $B_o \le a_2 G_o$. The rearrangement of the Matting Equation above for the green channel is

$$G_o = G_f - (1-\alpha_o)G_k \ .$$

Another rearrangement, this time for the blue channel, gives us

$$\alpha_o = 1 + \frac{B_o - B_f}{B_k} \le 1 + \frac{a_2 G_o - B_f}{B_k} \ .$$

Combining these two, by substituting the equation for G_o into the inequality for α_o and solving, gives

$$\alpha_o \le 1 - \frac{B_f - a_2 G_f}{B_k - a_2 G_k} \ ,$$

clamped to [0, 1] if necessary. Recall that $.5 \le a_2 \le 1.5$ typically. Let α_o at equality be $\alpha_{\max}$. Then, in our Fig. 3 example, $a_2 = 1$ yields $\alpha_{\max} \approx .87$, which constrains the possible solutions a bit more: $.78 \le \alpha_o \le .87$.

BLUE SPILL

Vlahos tackled the very important blue spill (blue flare) problem of backing light reflecting off the foreground object in [19]. He solved it for an important class of objects, bright whites and flesh tones, by making what we call the Second Vlahos Assumption: Foreground objects have $\max(B_o - G_o\ , 0) \le \max(G_o - R_o, 0)$. If this is not true, the color is assumed to be either the backing color or flare from it. Object transparency is taken, as before, to be proportional to $B_o - G_o$, and this distinguishes the two cases.

Our statement of the Matting Problem needs to be altered to include the blue spill problem. Our current model says that the foreground color C_f is a linear combination of the uncomposited foreground object color C_o and the backing color C_k, $C_f = C_o + (1 - \alpha_o)C_k$. The Extended Matting Problem would include a term C_s for the backing spill contribution. For example, it might be modeled as a separate foreground object, with its own alpha α_s, in linear combination with the desired foreground object color C_o: $C_f = C_s + (1 - \alpha_s)(C_o + (1 - \alpha_o)C_k)$. Now the problem becomes the more difficult one of determining both C_s and C_o from the given information C_f and C_k.

A simplification is to assume that the spill color is the same as the backing color, $C_s = \alpha_s C_k$. Thus $C_f = (1 - \alpha_s)C_o + (1 - \alpha_o + \alpha_o\alpha_s)C_k$. For brevity, let $C_{\Delta_s} = C_\Delta / (1-\alpha_s)$. Then this spill model can be put into a matrix equation of the same form as before (but notice the $\alpha_s = 1$ singularity):

$$C_o \begin{bmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ -R_k & -G_k & -B_k & t_4 \end{bmatrix} = \begin{bmatrix} R_{\Delta s} & G_{\Delta s} & B_{\Delta_s} & T \end{bmatrix}.$$

Hence, since $\alpha_{\Delta_s} = 0$ always, the solutions are of the same form as before: $\alpha_o = \dfrac{T - \bar{t} \cdot C_{\Delta s}}{\bar{t} \cdot C_k}$ and $c_o = c_{\Delta_s} + \alpha_o c_k$. This does not solve the problem since α_s is still unknown. We shall not pursue the spill problem further here but recommend it for future research.

SEPARATING SURFACE PROBLEMS

Fig. 4 illustrates the separating surface approach to the general matting problem. A single plane of colorspace is shown for clarity. A family of three separating surfaces for different values of α_o have been established between the body of backing colors C_k and the body of object colors C_o. A given foreground color C_f is shown at the point of intersection with the $\alpha_o = .5$ locus along the straight line through object colors *A* and *B*.

The Vlahos (or Ultimatte) matting solutions can be cast into the separating surface model. In the First Vlahos Form (as well as in our Solutions 1 and 2 and Theorem 1), each dotted line of Fig.

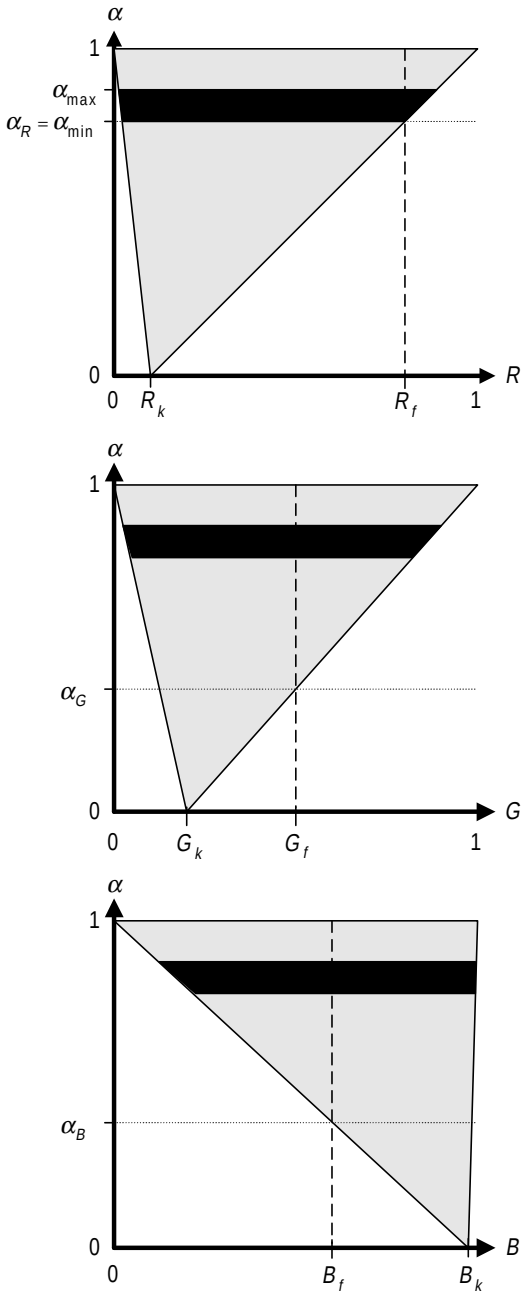

Figure 3. Shaded areas show solution space. Black areas are constrained by upper and lower alpha limits to valid alphas for the given foreground color. Valid alphas for C_o lie along intersection of C_f (dashed lines) with black areas.

4 would simply be a straight line (a plane in RGB). In the Second Vlahos Form it would be a line with two straight segments (two polygons sharing an edge in RGB). The third form simply adds a third segment (polygon) to this shape. The Primatte solution extends this trend to many (up to 128) segments (faces of a convex polyhedron).

Fig. 4 illustrates a general problem with the separating surface model. All mixtures of *A* with the backing color will be correctly pulled if they indeed exist in the foreground object. However, all mixtures of *B* with the backing will not be correctly pulled because they have been disguised as mixtures of *A*.

Another problem is that it is not always possible to have foreground object colors disjoint from backing colors. Another is the assumption that a locus of constant α_o is a surface rather than a volume, connected rather than highly disconnected, and planar or convex.

SUMMARY

The expiration of the fundamental Vlahos patents has inspired us to throw open the very interesting class of constant color matting problems to the computer graphics community. Thus one of our purposes has been to review the problems of the field—the general one of pulling a matte from a constant color shot plus related subproblems such as blue spill, backing impurities, and backing shadows.

The mathematical approach introduced here we believe to be more understandable than the ad hoc approach of the Vlahos patents, the standard reference on blue screen matting. Furthermore, we believe that the treatment here throws light on why the process should work so well as it does in real-world applications (gray, near-gray, and flesh tones), surprising in light of the proof herein that the general problem has an infinity of solutions. Consistent with the lack of a general algorithmic solution is the fact that human interaction is nearly always required in pulling a matte in film or video.

Our principal idea is that an image from which a matte is to be pulled can be represented by a model of two images, an uncomposited foreground object image (a sprite) and a backing color image, linearly combined using the alpha channel of the foreground object. Our main results are deduced from this model. In each case, the expression for the desired alpha channel α_o is a function of the two images in the model, C_f, the given image—a composite by our model—and C_k, the given backing image. This may be compared to the Vlahos expressions for alpha which are functions of the given image C_f only.

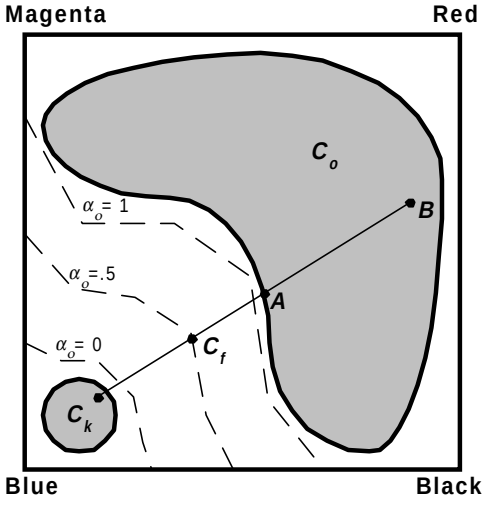

Figure 4. A slice through a (non-convex) polyhedral family of surfaces of constant alpha separating backing colors from foreground object colors. Given color C_f will be interpreted as *A* with alpha of .5 whereas the object might actually be *B* with an alpha of .25.

We have introduced an algorithmic solution, the triangulation solution, by adding a new step to the blue screen process as usually practiced: Another shot of the foreground object against a second backing color. This *multi-background* technique cannot be used for live actors or other moving foreground objects because of the requirement for repeatability. Whenever it is applicable, however, it is powerful, the only restriction on the two backings being that they be different pixel by pixel. Hence the backing colors do not even have to be constant or pure—the backing impurities problem does not exist. However, to solve the backing shadows problem, illumination level and direction must be the same for both backings, particularly important if they are generated by light emission rather than reflection.

We have bounded the solution space for the general non-algorithmic problem, a new extension to the Vlahos oeuvre. Hopefully, this will inspire further researches into this difficult problem. See the Vlahos patents (including [4] and [5]) for further inspiration.

We have touched on the blue spill (blue flare) problem and suggest that additional research be aimed at this important problem. We have sketched a possible model for this research, generalizing the idea of the given image being a composite of others. In particular, we propose that the idea of modeling blue spill by an additional blue spill image, with its own alpha, might lead to further insight.

Finally, we have briefly reviewed the modeling of the matting problem with separating surface families (cf. [8]), shown how to cast the Vlahos work in this light, and discussed some problems with the general notion. We urge that this class of solutions be further explored and their fundamental problems be elucidated beyond the initial treatment given here.

ACKNOWLEDGMENTS

To Tom Porter for his alpha upper limit. To Rob Cook for an early critical reading of [12] and [13] on which much of this paper is based. To Rick Szeliski for use of his automatic image registration software. To Jack Orr and Arlo Smith for studio photography and lighting.

REFERENCES

[1] BEYER, W. Traveling Matte Photography and the Blue Screen System. *American Cinematographer*, May 1964, p. 266. The second of a four-part series.

[2] BLINN, J. F. Jim Blinn's Corner: Compositing Part 1: Theory. *IEEE Computer Graphics & Applications*, September 1994, pp. 83-87.

[3] BLINN, J. F. Jim Blinn's Corner: Compositing Part 2: Practice. *IEEE Computer Graphics & Applications*, November 1994, pp. 78-82.

[4] DADOURIAN, A. Method and Apparatus for Compositing Video Images. U. S. Patent 5,343,252, August 30, 1994.

[5] FELLINGER, D. F. Method and Apparatus for Applying Correction to a Signal Used to Modulate a Background Video Signal to be Combined with a Foreground Video Signal. U. S. Patent 5,202,762, April 13, 1993.

[6] FIELDING, R. *The Technique of Special Effects Cinematography*. Focal/Hastings House, London, 3rd edition, 1972, pp. 220-243.

[7] LANCZOS, C. *Applied Analysis*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1964, pp. 156-161.

[8] MISHIMA, Y. A Software Chromakeyer Using Polyhedric Slice. Proceedings of NICOGRAPH 92 (1992), pp. 44-52 (Japanese). See http://206.155.32.1/us/primatte/whitepaper.

[9] PORTER, T. and DUFF, T. Compositing Digital Images. Proceedings of SIGGRAPH 84 (Minneapolis, Minnesota, July 23-27, 1984). In *Computer Graphics* 18, 3 (July 1984), pp. 253-259.

[10] PORTER, T. Matte Box Design. Lucasfilm Technical Memo 63, November 1986. Not built.

[11] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., and FLANNERY, B. P. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1988, p. 59.

[12] SMITH, A. R. Analysis of the Color-Difference Technique. Technical Memo 30, Lucasfilm Ltd., March 1982.

[13] SMITH, A. R. Math of Mattings. Technical Memo 32, Lucasfilm Ltd., April 1982.

[14] SMITH, A. R. Image Compositing Fundamentals. Technical Memo 4, Microsoft Corporation, June 1995.

[15] SMITH, A. R. Alpha and the History of Digital Compositing. Technical Memo 7, Microsoft Corporation, August 1995.

[16] VLAHOS, P. Composite Photography Utilizing Sodium Vapor Illumination. U. S. Patent 3,095,304, May 15, 1958. Expired.

[17] VLAHOS, P. Composite Color Photography. U. S. Patent 3,158,477, November 24, 1964. Expired.

[18] VLAHOS, P. Electronic Composite Photography. U. S. Patent 3,595,987, July 27, 1971. Expired.

[19] VLAHOS, P. Electronic Composite Photography with Color Control. U. S. Patent 4,007,487, February 8, 1977. Expired.

[20] VLAHOS, P. Comprehensive Electronic Compositing System. U. S. Patent 4,100,569, July 11, 1978. Expired.

[21] VLAHOS, P. and TAYLOR, B. Traveling Matte Composite Photography. *American Cinematographer Manual*. American Society of Cinematographers, Hollywood, 7th edition, 1993, pp. 430-445.

[22] VLAHOS, P. Comprehensive Electronic Compositing System. U. S. Patent 4,344,085, August 10, 1982.

[23] VLAHOS, P. Encoded Signal Color Image Compositing. U. S. Patent 4,409,611, October 11, 1983.

[24] VLAHOS, P., VLAHOS, P. and FELLINGER, D. F. Automatic Encoded Signal Color Image Compositing. U. S. Patent 4,589,013, May 13, 1986.

[25] VLAHOS, P. Comprehensive Electronic Compositing System. U. S. Patent 4,625,231, November 25, 1986.

Figure 5. A composite of nine image sprites pulled from studio photographs using the triangulation technique shown in Fig. 2.