15. Matting and compositing
Final projects

• Flexible group size

• This weekend: group yourselves and send me:
  a one-paragraph description of your idea if you are fixed on one
  one-sentence descriptions of 3 ideas if you are looking for one

• Next week: project proposal
  one-page description
  plan for mid-project milestone

• Before thanksgiving: milestone report

• December 5 (day of scheduled final exam): final presentations
Compositing
Foreground and background

• How we compute new image varies with position

• Therefore, need to store some kind of tag to say what parts of the image are of interest
Binary image mask

• First idea: store one bit per pixel
  – answers question “is this pixel part of the foreground?”

  – causes jaggies similar to point-sampled rasterization
  – same problem, same solution: intermediate values
Partial pixel coverage

• The problem: pixels near boundary are not strictly foreground or background

– how to represent this simply?
– interpolate boundary pixels between the fg. and bg. colors
Alpha compositing

- Formalized in 1984 by Porter & Duff
- Store fraction of pixel covered, called $\alpha$

\[ C = A \text{ over } B \]

\[ r_C = \alpha_A r_A + (1 - \alpha_A) r_B \]
\[ g_C = \alpha_A g_A + (1 - \alpha_A) g_B \]
\[ b_C = \alpha_A b_A + (1 - \alpha_A) b_B \]

- this exactly like a spatially varying crossfade

- Convenient implementation
  - 8 more bits makes 32
  - 2 multiplies + 1 add per pixel for compositing
Alpha compositing—example
Creating alpha mattes

• Compositing is ubiquitous in film production
  merge separately shot live action
  merge visual effects with live action
  merge visual effects from different studios/renderers

• Also useful in photography, graphic design
  composite photos [wired cover]
  photos as non-rectangular design elements [newsweek cover]

• The alpha channel can be called a “matte”
  (dates from matte paintings, painted on glass to allow backgrounds to show through when photographed)

• Getting a matte for a photographic source is tricky
  and getting it right is crucial to good results
  leads to hours and hours of manual pixel-tweaking
Matting

- Someone has computed $C = F$ over $B$ and lost $F$ and $B$, and we are supposed to recover $F$ (including $a$) and $B$.

When you can arrange it, it’s much easier if $B$ is some very unlikely color…
Strategy

• Simple approaches used for analog and early digital chroma-key devices

\[ \alpha = 1 - \text{clamp}(a_1(C_b - a_2C_g)) \quad \text{for a blue background (bluescreen)} \]

and other more complicated schemes

• More principled approach: Bayesian matting

based on statistical models for colors of F and B

compute per-pixel statistical estimate of each pixel’s F and \( \alpha \)
Trimap

- Someone has to specify which part is supposed to be extracted
- Trimap: label pixels as definitely F, definitely B, or not sure
Estimating the matte

• Applying the pattern of MAP estimation:

\[
p(F, B, \alpha | C) = p(C | F, B, \alpha)p(F, B, \alpha)
\]

what we want to maximize (likelihood)

what we have a model for (probability)

need some assumptions here (prior)

• Bayesian matting:

gaussian noise model for probability of \(C\)

\(F, B, \alpha\) assumed independent multivariate gaussians for \(F, B, \alpha\) assumed uniform

joint distribution: \(p(a, b)\)
marginal distribution (projection): \(p(a) = \int_b p(a, b)\)
conditional distribution (slice): \(p(a|b) = p(a, b)/p(b)\)
Bayes: \(p(a|b)p(b) = p(a)p(b|a)p(a)\)

A Bayesian Approach to Digital Matting

Yang Yi Chuang1 Brian Carlin2 David H. Salesin3,4 Richard Szeliski2
1Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195
2Microsoft Research, Redmond, WA 98052
Email: {yangyi, carlin}@cs.washington.edu, szeliski@microsoft.com
http://graphics.stanford.edu/projects/digital-matting

Abstract

This paper proposes a new Bayesian framework for solving the matting problem, i.e., a splitting of an input image into foreground and background images by estimating the support for each pixel of the foreground element. Our approach models both the foreground and background as distributions with underlying sets of parameters, and assumes a functional blending of the foreground and background to produce the final output. If there are no mismatched pixels or noise in the input, then the solution is unique, and both foreground and background images can be recovered exactly. In practice, there are always mismatched pixels or noise, so we formulate the matting problem as an optimization problem, and present an algorithmic framework that handles objects with intricate boundaries, such as hair strands and fur, and provides an improvement over existing techniques for these difficult cases.

1. Introduction

In digital matting, a foreground element is extracted from a background image by estimating a color and opacity for the foreground element at each pixel. The opacity value at each pixel is typically called alpha, and the intensity image, taken as a whole, is called the alpha channel or key. Discovered by Neider and Porter (1984) and refined by a number of investigators, the alpha channel is not a new idea, but it has become a standard reference in the computer graphics community. In its most general form, the alpha channel is a parameter that is used to control the degree of transparency of a foreground element.

Matting is used in order to incorporate the foreground element into a new scene. Matting and compositing were originally developed for film and television production. In 1958, Porter and Dufn (1) introduced the digital matting of the alpha channel—denoted here by synthetic images to which alpha could be used in creating complete digital images. The most important early system was the 3D computer graphics system, which is implemented by the compositor operation.

\[
C = C_1 - (1 - \alpha)C_2
\]

where \(C_1\) and \(C_2\) are the pixel's composite foreground and background colors, respectively, and \(\alpha\) is the pixel's spatially-varying component used to locally blend between foreground and background.

The matting process starts from a photograph or set of photographs (referred to as the alpha channel) and determines how synthetic images to which alpha could be used in creating complete digital images. The most important early system was the 3D computer graphics system, which is implemented by the compositor operation.

The basic idea is to use a background-based technique for the alpha channel, and a foreground-based technique for the foreground element. The principle is to use the alpha channel to define the shape and color of the foreground, and the foreground to impose the constraints on the alpha channel.

As a refresher for the key terms used for the rest of the meeting, the principles are to use the alpha channel to define the shape and color of the foreground, and the foreground to impose the constraints on the alpha channel. The key terms are:

- **Foreground**: the object that needs to be extracted.
- **Background**: the scene that the foreground is set against.
- **Alpha channel**: a grayscale image that indicates the transparency of each pixel, with white representing full transparency and black representing full opacity.

[Chuang et al. 2001]
Bayesian matting math

\[ p(F, B, \alpha | C) = p(C | F, B, \alpha)p(F, B, \alpha) \]

\[ p(F, B, \alpha) = kN(F - \bar{F}, \Sigma_F)N(B - \bar{B}, \Sigma_B) \]

\[ p(C | F, B, \alpha) = N(C - [\alpha F + (1 - \alpha)B], \sigma_C) \]

\[ 2 \log p(F, B, \alpha | C) = \frac{[C - B - \alpha(F - B)]^2}{\sigma_C} + (F - \bar{F})^T \Sigma_F (F - \bar{F}) + (B - \bar{B})^T \Sigma_B (B - \bar{B}) \]

uses a procedure of alternating linear system solves for \( \alpha \) and for \((F,B)\)
Defining priors for $F$ and $B$

- Use the weighted covariance of a region of the image around the pixel being solved

$$\begin{align*}
(\Sigma_F)_{i,j} &= \sum_k w_k (F_{k,i} - \bar{F}_i)(F_{k,j} - \bar{F}_j) \bigg/ \sum_k w_k \\
\text{color channels } i \text{ and } j \quad &\quad \text{depends on nearby pixels } k \quad \text{distance and known } \alpha
\end{align*}$$

- Solve the problem by marching inward from the edges of the “unknown” area
Bayesian matting results

Figure 2 Summary of input images and results. Input images (top row): a blue-screen matting example of a toy lion, a synthetic “natural image” of the same lion (for which the exact solution is known), and two real natural images, (a lighthouse and a woman). Input segmentation (middle row): conservative foreground (white), conservative background (black), and “unknown” (grey). The leftmost segmentation was computed automatically (see text), while the rightmost three were specified by hand. Compositing results (bottom row): the results of compositing the foreground images and mattes extracted through our Bayesian matting algorithm over new background scenes. (Lighthouse image and the background images in composite courtesy Philip Greenspun, http://philip.greenspun.com. Woman image was obtained from Corel Knockout’s tutorial, Copyright © 2001 Corel. All rights reserved.)
Bayesian matting results

Figure 3  Blue-screen matting of lion (taken from leftmost column of Figure 2). Mishima’s results in the top row suffer from “blue spill.” The middle and bottom rows show the Bayesian result and ground truth, respectively.
Bayesian matting results

**Figure 5** Natural image matting. These two sets of photographs correspond to the rightmost two columns of Figure 2, and the insets show both a close-up of the alpha matte and the composite image. For the woman’s hair, Knockout loses strands in the inset, whereas Ruzon-Tomasi exhibits broken strands on the left and a diagonal color discontinuity on the right, which is enlarged in the inset. Both Knockout and Ruzon-Tomasi suffer from background spill as seen in the lighthouse inset, with Knockout practically losing the railing.
Closed form matting

(blackboard)
Previous approaches

The trimap interface:

• Bayesian Matting (Chuang et al, CVPR01)
• Poisson Matting (Sun et al SIGGRAPH 04)
• Random Walk (Grady et al 05)

Scribbles interface:

• Wang&Cohen ICCV05
Problems with trimap based approaches

- Iterate between solving for F, B and solving for $\alpha$
- Accurate trimap required

(Replotted from Wang & Cohen)
Closed-form matting results

[Image 231x167 to 791x933]
Effect of $\epsilon$

Fig. 6. Computing a matte using different $\epsilon$ values.
Closed-form matting results

Fig. 15 presents compositing examples using our algorithm for some images from the previous experiments. We show compositing both over a constant background and over natural images.

Fig. 16 shows an example (from [19]) where Wang and Cohen's method fails to extract a good matte from sparse scribbles due to color ambiguity between the foreground and the background. The same method, however, is able to produce an acceptable matte when supplied with a trimap. Our method produces a cleaner but also imperfect matte from the same set of scribbles, but adding a small number of additional scribbles results in a better matte. (To produce this result, we applied clamping of alpha values as described in Section 6.)

Fig. 17 shows another example (a close-up of the Koala image from [17]), where there is an ambiguity between foreground and background colors. In this case, the matte produced by our method is clearly better than the one produced by the Wang-Cohen method. To better understand why this is the case, we show an RGB histogram of representative pixels from the F and B scribbles. Some pixels in the background fit the foreground color model much better than the background one (one such pixel is marked red in Fig. 17b and indicated by an arrow in Fig. 17d). As a result, such pixels are classified as foreground with a high degree of certainty in the first stage. Once this error has been made, it only reinforces further erroneous decisions in the vicinity of that pixel, resulting in a white clump in the alpha matte.

Since our method does not make use of global color models for F and B, it can handle ambiguous situations such as that in Fig. 17. However, there are also cases where our method fails to produce an accurate matte for the very same reason. Fig. 18 shows an actress in front of a background with two colors. Even though the black B scribbles cover both colors, the...
Closed-form matting results

Fig. 15 presents compositing examples using our algorithm for some images from the previous experiments. We show compositing both over a constant background and over natural images.

Fig. 16 shows an example (from [19]) where Wang and Cohen's method fails to extract a good matte from sparse scribbles due to color ambiguity between the foreground and the background. The same method, however, is able to produce an acceptable matte when supplied with a trimap. Our method produces a cleaner but also imperfect matte from the same set of scribbles, but adding a small number of additional scribbles results in a better matte. (To produce this result, we applied clamping of alpha values as described in Section 6.)

Fig. 17 shows another example (a close-up of the Koala image from [17]), where there is an ambiguity between foreground and background colors. In this case, the matte produced by our method is clearly better than the one produced by the Wang-Cohen method. To better understand why this is the case, we show an RGB histogram of representative pixels from the foreground and background scribbles. Some pixels in the background fit the foreground color model much better than the background one (one such pixel is marked red in Fig. 17b and indicated by an arrow in Fig. 17d). As a result, such pixels are classified as foreground with a high degree of certainty in the first stage. Once this error has been made, it only reinforces further erroneous decisions in the vicinity of that pixel, resulting in a white clump in the alpha matte. Since our method does not make use of global color models for foreground and background, it can handle ambiguous situations such as that in Fig. 17. However, there are also cases where our method fails to produce an accurate matte for the very same reason. Fig. 18 shows an actress in front of a background with two colors. Even though the black background scribbles cover both colors, the LEVIN ET AL.: A CLOSED-FORM SOLUTION TO NATURAL IMAGE MATTING


