CS6640 Computational Photography

15. Matting and compositing

Final projects

- **Flexible group size**
- **This weekend: group yourselves and send me:**

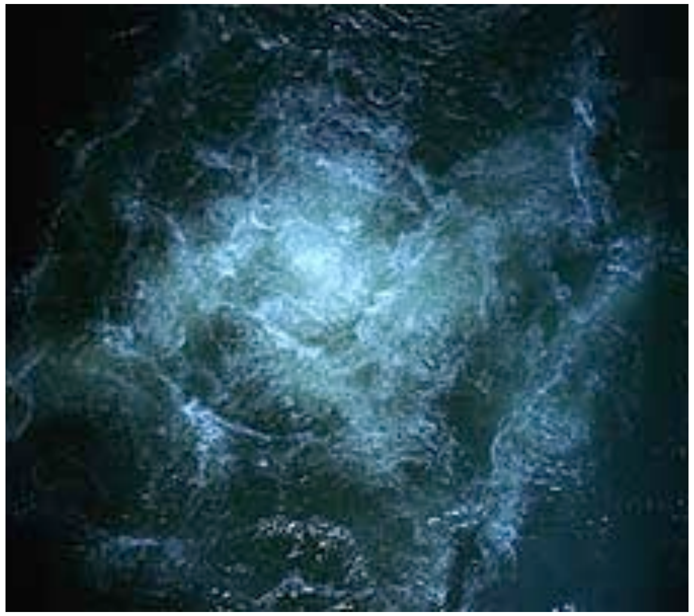
  a one-paragraph description of your idea if you are fixed on one

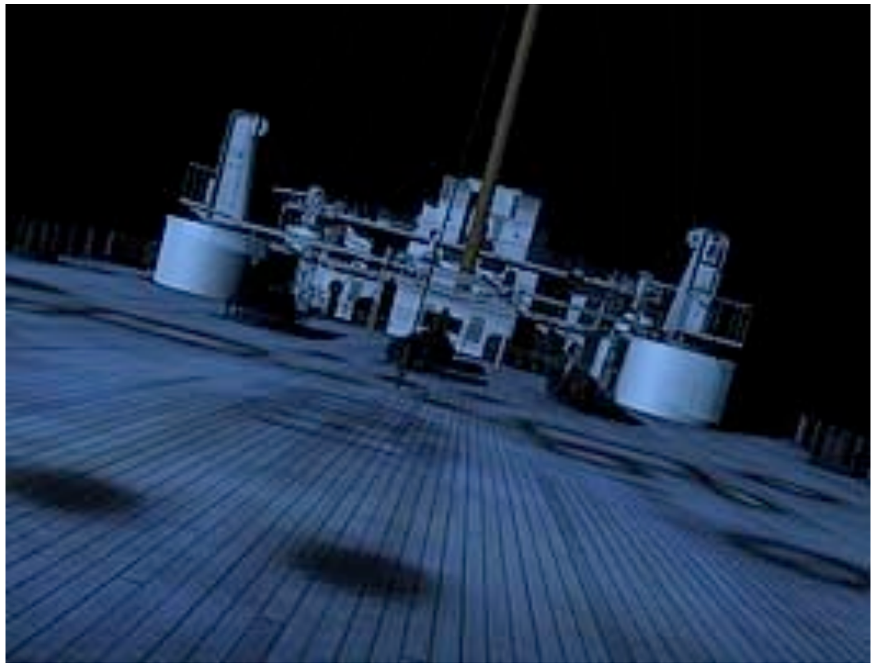
  one-sentence descriptions of 3 ideas if you are looking for one
- **Next week: project proposal**

  one-page description

  plan for mid-project milestone
- **Before thanksgiving: milestone report**
- **December 5 (day of scheduled final exam): final presentations**

Compositing





[*Titanic* ; DigitalDomain; vfxhq.com]

Foreground and background

- How we compute new image varies with position

[Chuang et al. / Corel]

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

Binary image mask

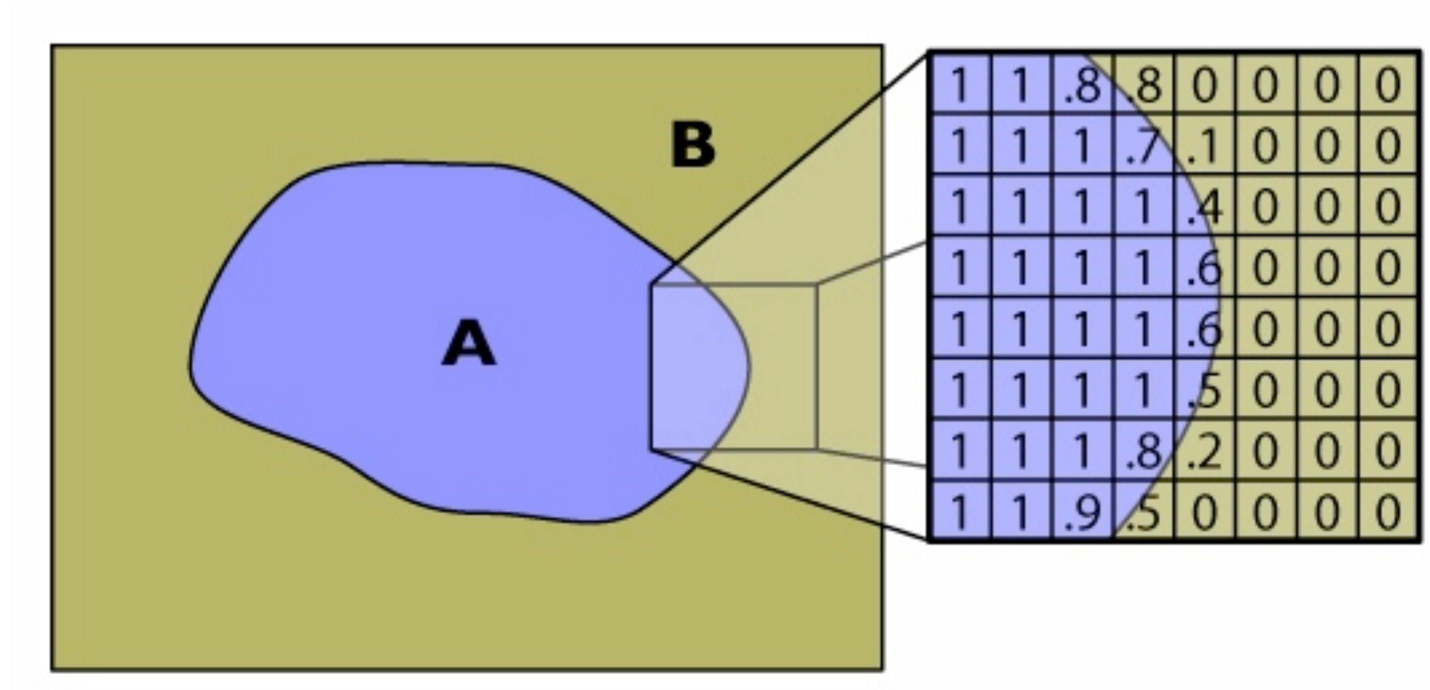
- First idea: store one bit per pixel
  - answers question “is this pixel part of the foreground?”

[Chuang et al. / Corel]

  - causes jaggies similar to point-sampled rasterization
  - same problem, same solution: intermediate values

Partial pixel coverage

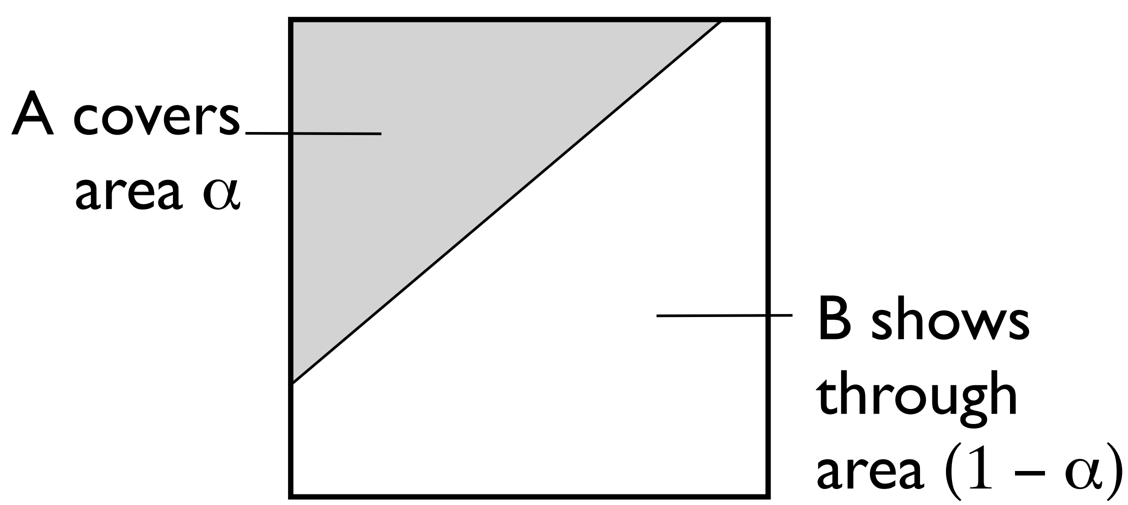
- The problem: pixels near boundary are not strictly foreground or background



  - 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 α



$$C = A \textbf{ over } B$$

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

  – 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

[Chuang et al. / Corel]

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 *α*) and *B*.**

The Hobbit promotional image

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**

  $$\alpha = 1 - \mathrm{clamp}(a_1(C_b - a_2 C_g)) \leftarrow \text{for a blue background (bluescreen)}$$

  and other more complicated schemes

- **More principled approach: Bayesian matting**

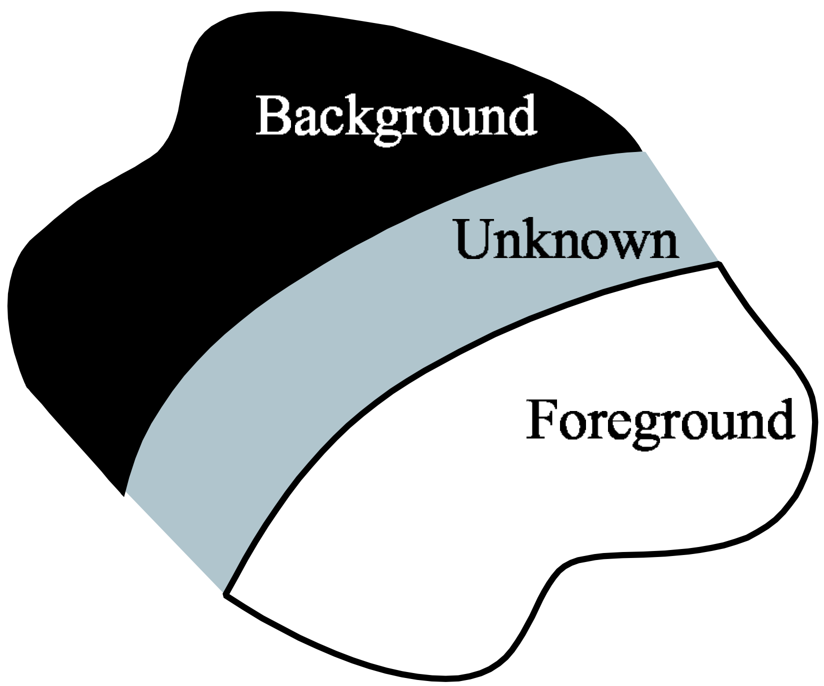
  based on statistical models for colors of F and B

  compute per-pixel statistical estimate of each pixel's *F* and *α*

Formula from [Smith & Blinn 1996]

Trimap

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



[Chuang et al. 2001]

Estimating the matte

refresher

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: $p(a|b)p(b) = p(a)p(b|a)p(a)$

- **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, α assumed independent

  multivariate gaussians for F, B

  α assumed uniform

A Bayesian Approach to Digital Matting

Yung-Yu Chuang[1] Brian Curless[1] David H. Salesin[1,2] Richard Szeliski[2]

[1]Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195
[2]Microsoft Research, Redmond, WA 98052
E-mail: {cyy, curless, salesin}@cs.washington.edu szeliski@microsoft.com
http://grail.cs.washington.edu/projects/digital-matting/

Abstract

This paper proposes a new Bayesian framework for solving the matting problem, i.e. extracting a foreground element from a background image by estimating an opacity for each pixel of the foreground element. Our approach models both the foreground and background color distributions with spatially-varying sets of Gaussians, and assumes a fractional blending of the foreground and background colors to produce the final output. It then uses a maximum-likelihood criterion to estimate the optimal opacity, foreground and background simultaneously. In addition to providing a principled approach to the matting problem, our algorithm effectively 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 its *alpha*, and the opacity image, taken as a whole, is referred to as the *alpha matte* or *key*. Fractional opacities (between 0 and 1) are important for transparency and motion blurring of the foreground element, as well as for partial coverage of a background pixel around the foreground object's boundary.

Matting is used in order to *composite* the foreground element into a new scene. Matting and compositing were originally developed for film and video production [4], where they have proven invaluable. Nevertheless, "pulling a matte" is still somewhat of a black art, especially for certain notoriously difficult cases such as thin whisps of fur or hair. The problem is difficult because it is inherently underconstrained: for a foreground element over a single background image there are in general an infinite number of interpretations for the foreground's color versus opacity.

In practice, it is still possible to pull a satisfactory matte in many cases. One common approach is to use a background image of known color (typically blue or green) and make certain assumptions about the colors in the foreground (such as the relative proportions of red, green, and blue at each pixel); these assumptions can then be tuned by a human operator. Other approaches attempt to pull mattes from natural (arbitrary) backgrounds, using statistics of known regions of foreground or background in order to estimate the foreground and background colors along the boundary. Once these colors are known, the opacity value is uniquely determined.

In this paper, we survey the most successful previous approaches to digital matting—all of them fairly *ad hoc*—and demonstrate cases in which each of them fails. We then introduce a new, more principled approach to matting, based on a Bayesian framework. While no algorithm can give perfect results in all cases (given that the problem is inherently underconstrained), our Bayesian approach appears to give improved results in each of these cases.

2. Background

As already mentioned, matting and compositing were originally developed for film and video production. In 1984, Porter and Duff [8] introduced the digital analog of the matte—the *alpha channel*—and showed how synthetic images with alpha could be useful in creating complex digital images. The most common compositing operation is the *over* operation, which is summarized by the *compositing equation*:

$$C = \alpha F + (1 - \alpha)B\,, \tag{1}$$

where C, F, and B are the pixel's composite, foreground, and background colors, respectively, and α is the pixel's opacity component used to linearly blend between foreground and background.

The matting process starts from a photograph or set of photographs (essentially composite images) and attempts to extract the foreground and alpha images. Matting techniques differ primarily in the number of images and in what *a priori* assumptions they make about the foreground, background, and alpha.

Blue screen matting was among the first techniques used for live action matting. The principle is to photograph the subject against a constant-colored background, and extract foreground and alpha treating each frame in isolation. This single image approach is underconstrained since, at each pixel, we have three observations and four unknowns. Vlahos pioneered the notion of adding simple constraints to make the problem tractable; this work is nicely summarized by Smith

[Chuang et al. 2001]

Bayesian matting math

refresher

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: $p(a|b)p(b) = p(b|a)p(a)$

$$p(F, B, \alpha \,|\, C) = p(C \,|\, F, B, \alpha)p(F, B, \alpha)$$

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

prob. of *α* | multivariate normal dist. | covariance matrix

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

multivariate isotropic normal dist. | variance (of image noise)

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

what to maximize | bilinear in *α* and (*F*,*B*) | (+ ~~log~~ *k*)

uses a procedure of alternating linear system solves for *α* and for (*F*,*B*)

Defining priors for *F* and *B*

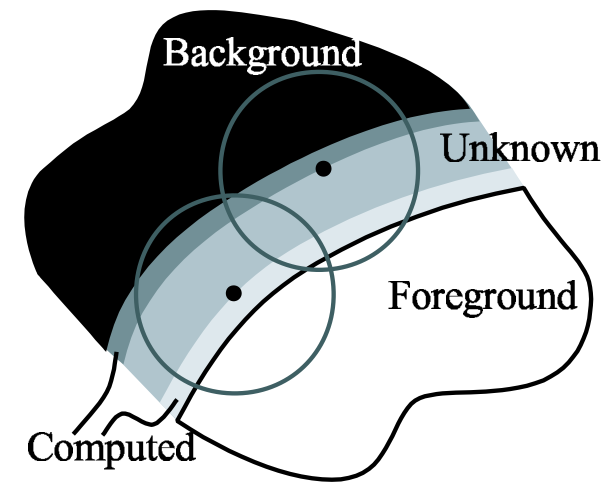
- **Use the weighted covariance of a region of the image around the pixel being solved**

$$(\Sigma_F)_{ij} = \sum_k w_k (F_{k,i} - \bar{F}_i)(F_{k,j} - \bar{F}_j) \bigg/ \sum_k w_k$$

color channels *i* and *j*

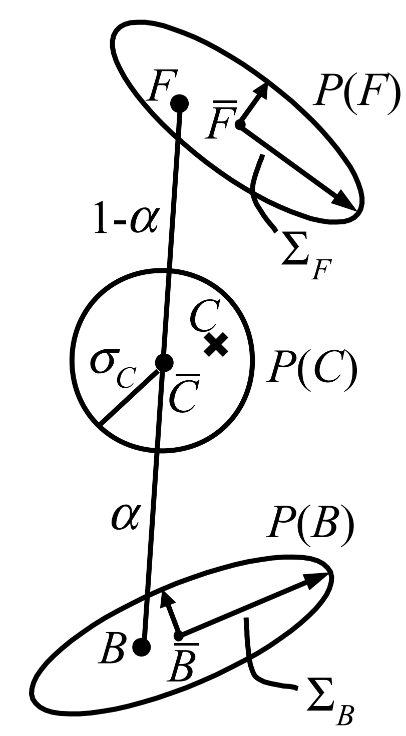
nearby pixels *k*

depends on distance and known *α*



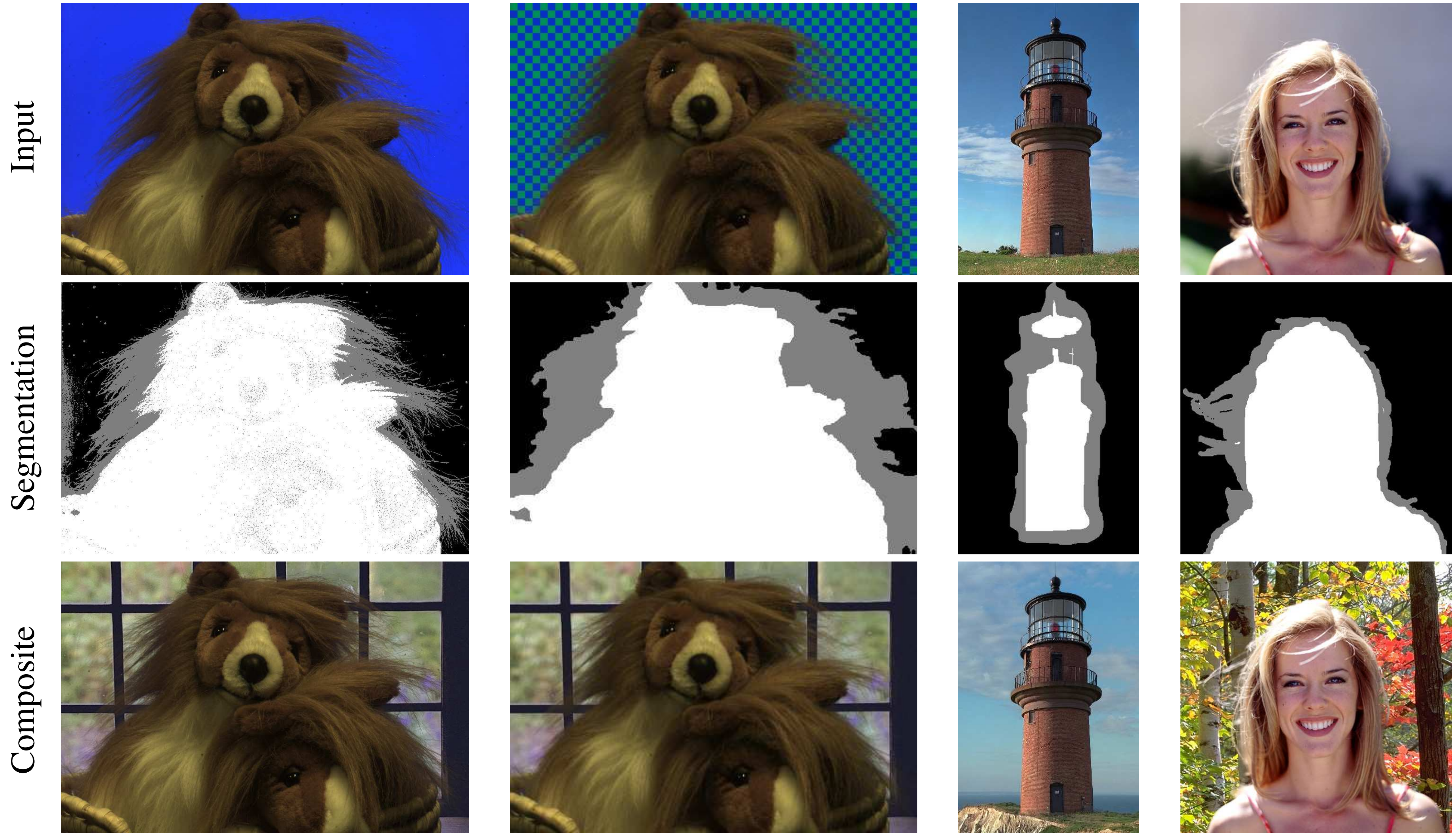
(d)

- **Solve the problem by marching inward from the edges of the "unknown" area**



[Chuang et al. 2001]

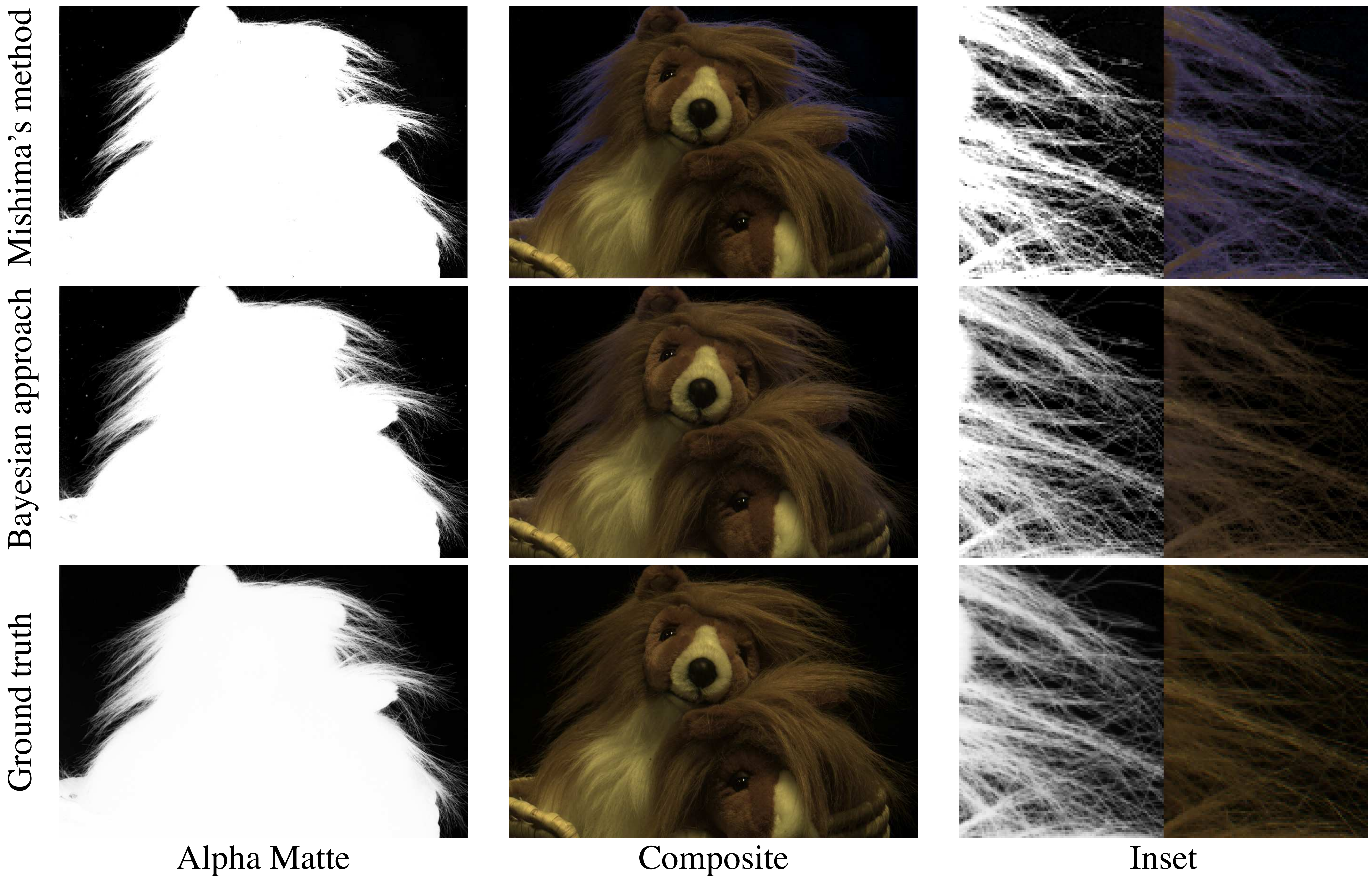
Bayesian matting results



[Chuang et al. 2001]

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 specifi ed 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



[Chuang et al. 2001]

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

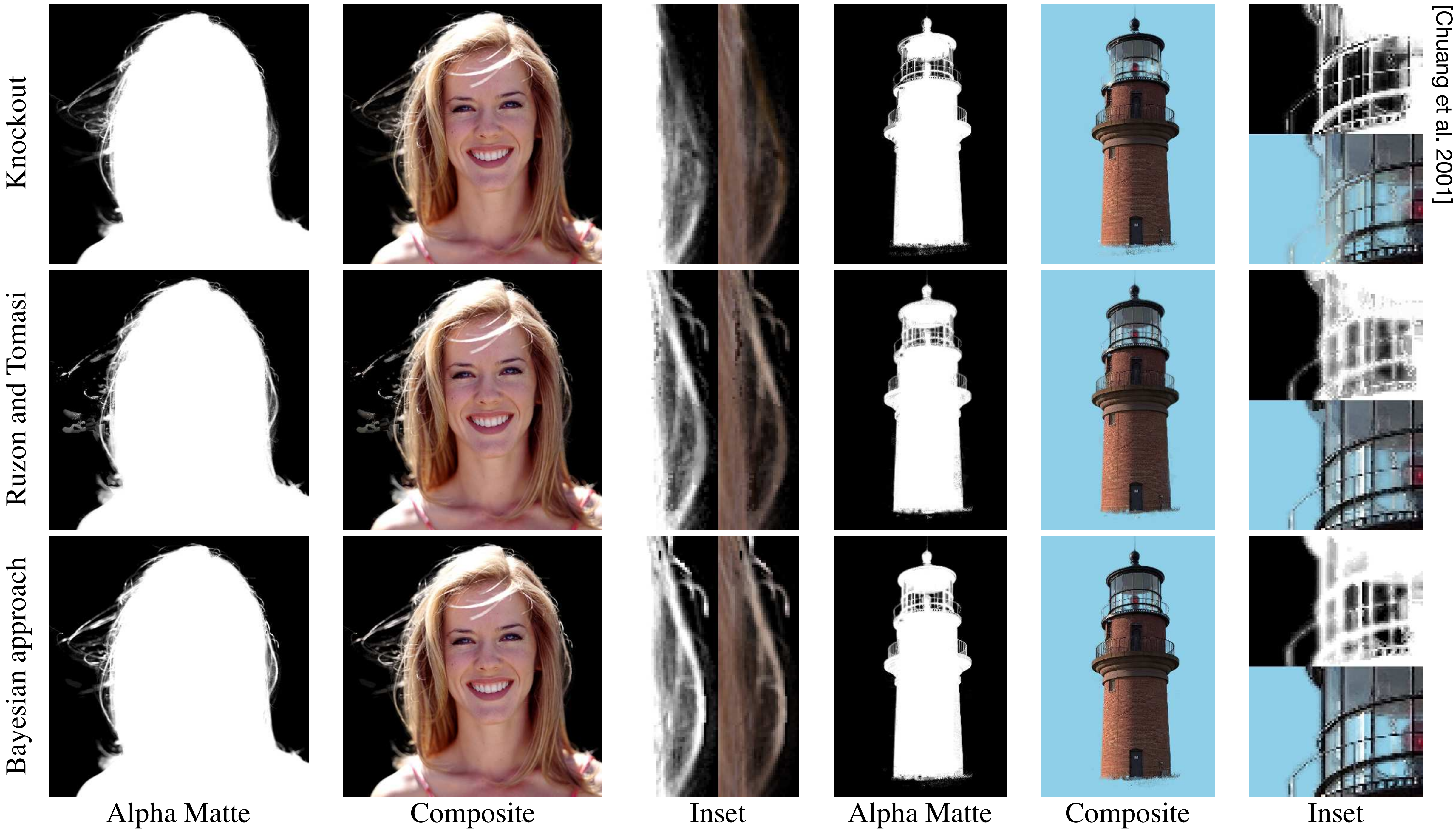


Figure 5 Natural image matting. These two sets of photographs correspond to the rightmost two columns of Figure 2, and the insets show both a close-up of the alpha matte and the composite image. For the woman's hair, Knockout loses strands in the inset, whereas Ruzon-Tomasi exhibits broken strands on 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.

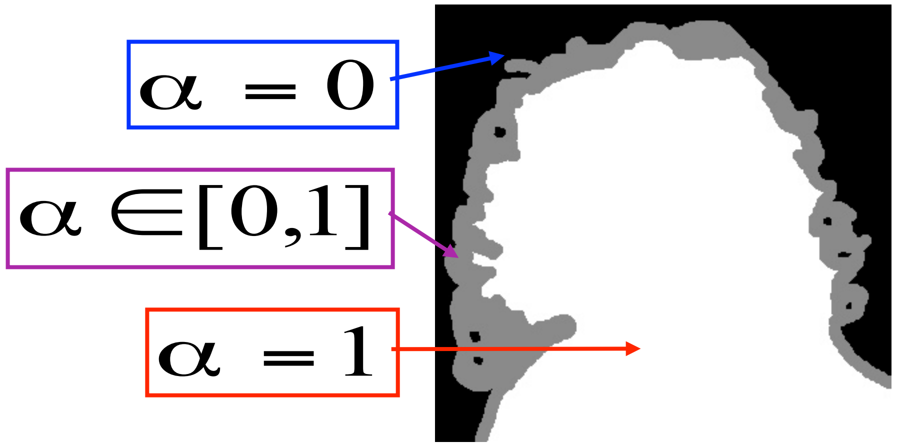
[Chuang et al. 2001 website]

[Chuang et al. 2001 website]

Closed form matting

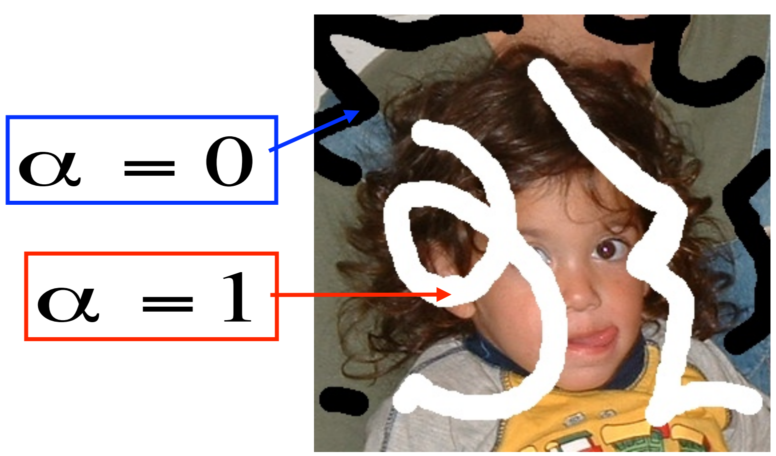
(blackboard)

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

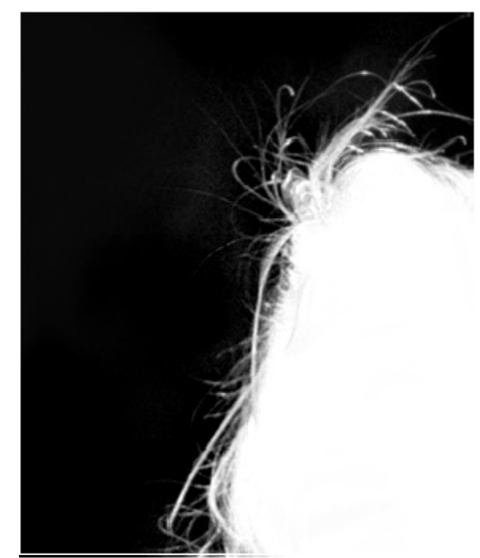
Problems with trimap based approaches

- Iterate between solving for F,B and solving for α
- Accurate trimap required

Input Scribbles

Bayesian matting from scribbles

Good matting from scribbles



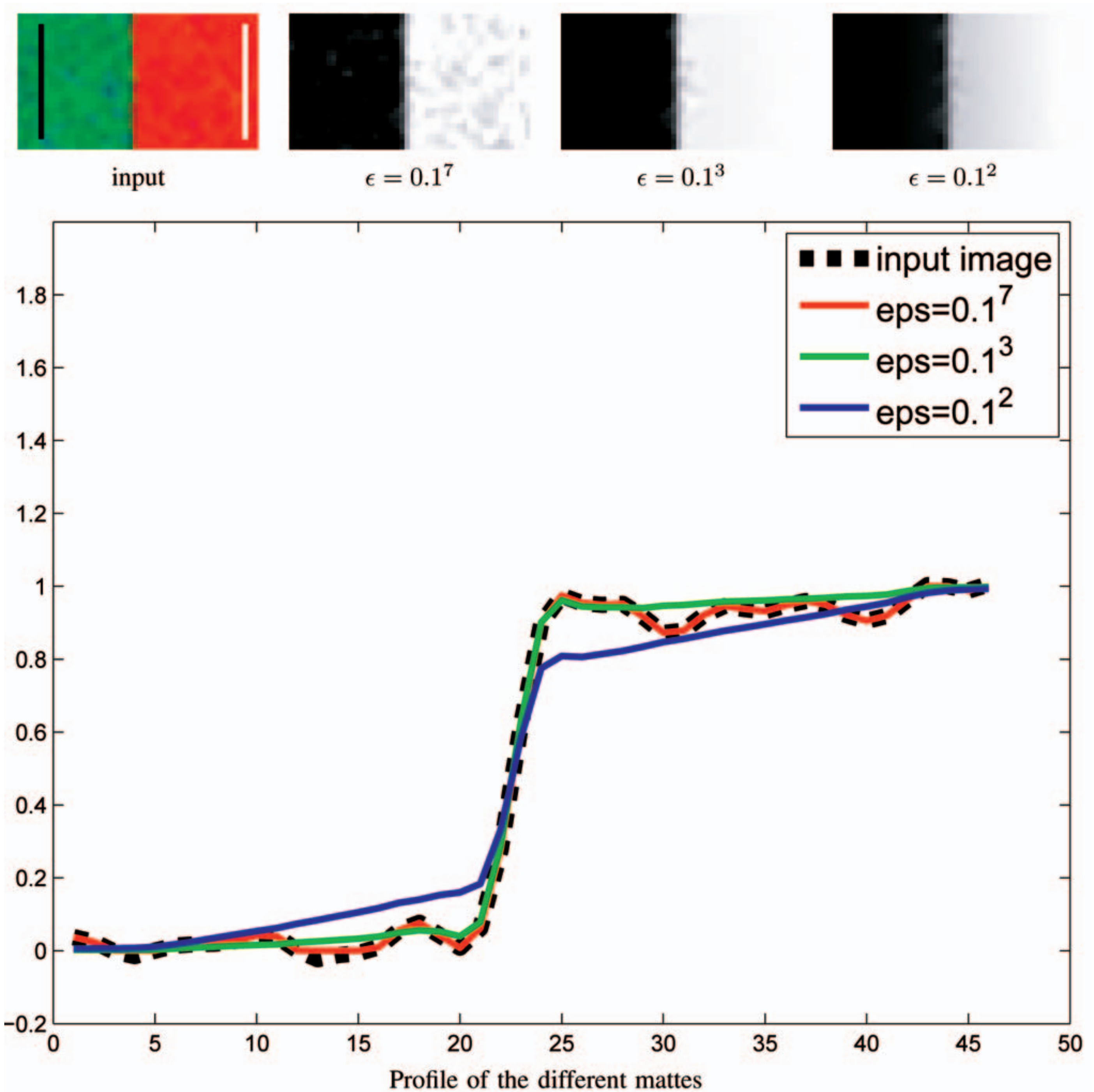
(Replotted from Wang&Cohen)

slide by Anat Levin, Weizmann Institute of Science

Closed-form matting results

[Levin et al. 2008]

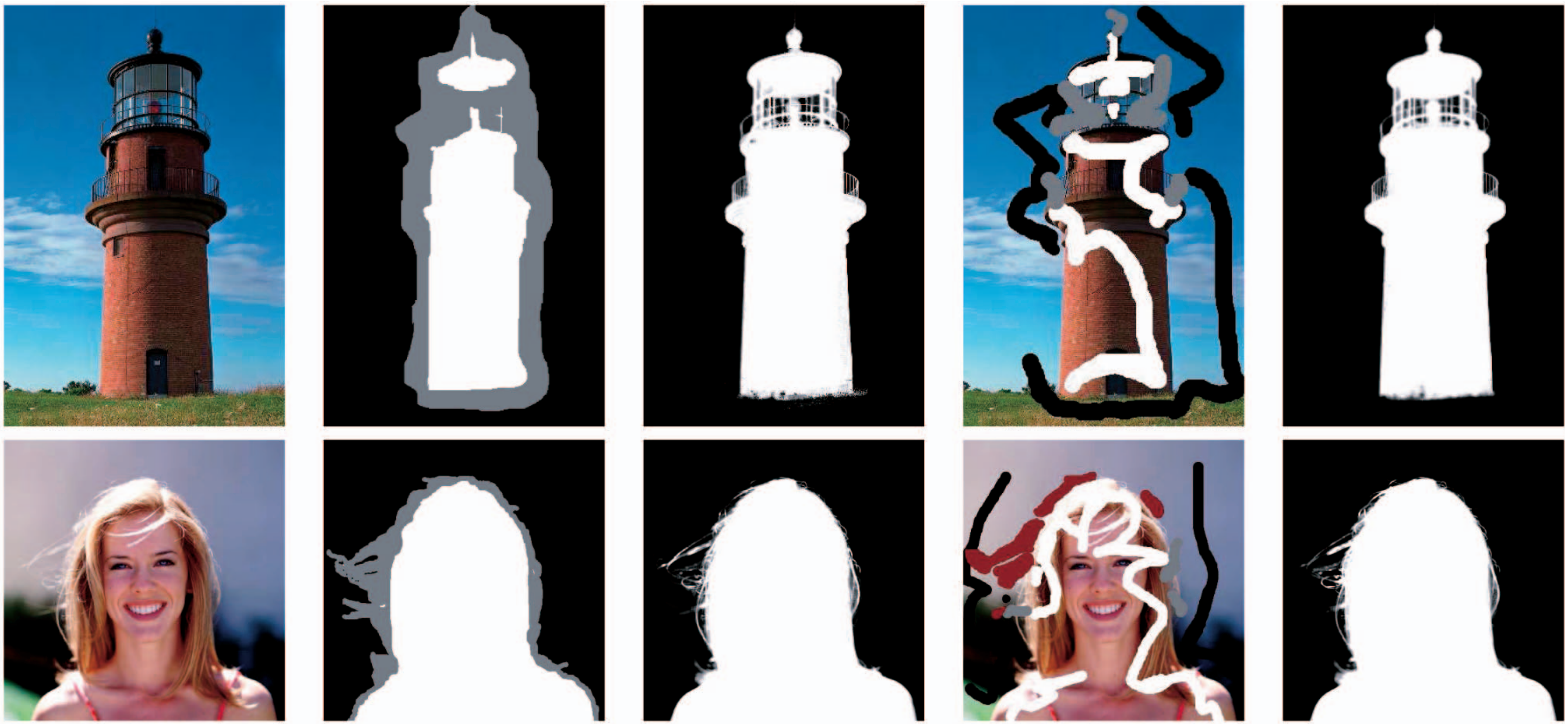
Effect of ε



[Levin et al. 2008]

Fig. 6. Computing a matte using different ϵ values.

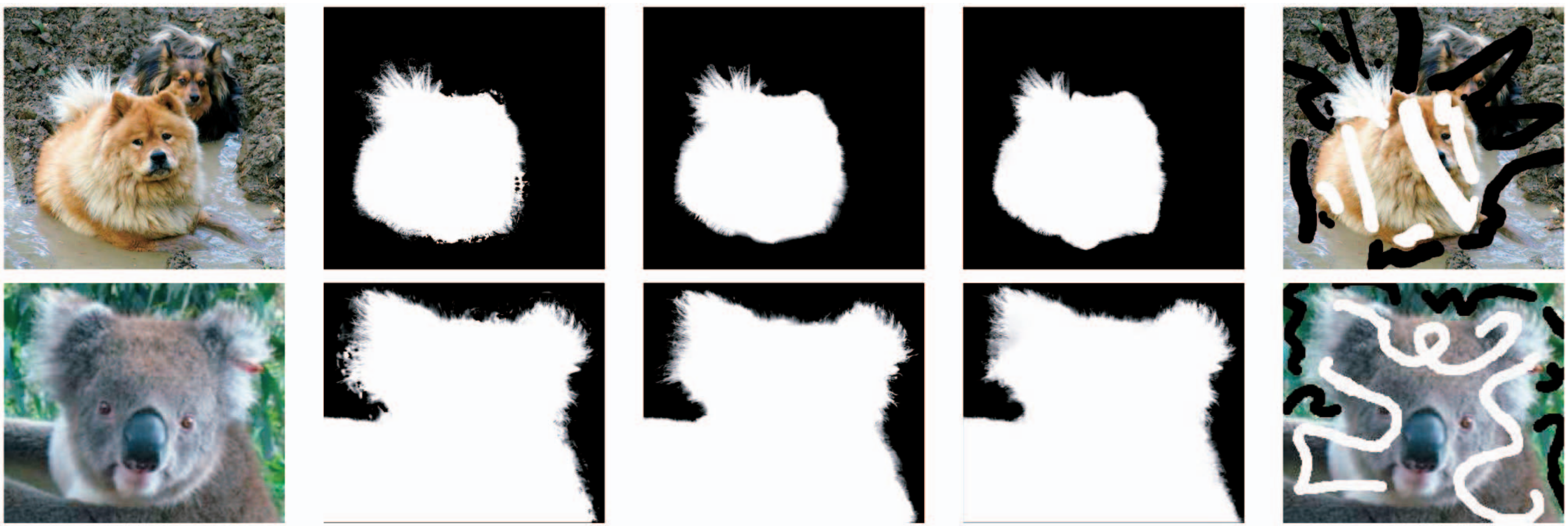
Closed-form matting results



input Bayesian Closed-form

[Levin et al. 2008]

Closed-form matting results



input Bayesian Poisson Closed-form

[Levin et al. 2008]

Bibliography

Y.Y. Chuang, B. Curless, D.H. Salesin, & R. Szeliski, **A bayesian approach to digital matting**, *CVPR 2001*.

A. Levin, D. Lischinski, & Y. Weiss, **A Closed-Form Solution to Natural Image Matting**, *PAMI* 30:2 (2008).

C. Rother, V. Kolmogorov, and A. Blake, **Grabcut: Interactive foreground extraction using iterated graph cuts**, *SIGGRAPH 2004*.

M.A. Ruzon & C. Tomasi, **Alpha estimation in natural images**, *CVPR 2000*.

A.R. Smith & J.F. Blinn, **Blue screen matting**, *SIGGRAPH 1996*.

J. Sun, J. Jia, C.K. Tang, and H.Y. Shum, **Poisson matting**, *SIGGRAPH 2004*.

J. Wang & M.F. Cohen, **Optimized color sampling for robust matting**, *CVPR 2007*.