A Bayesian Approach to Digital Matting

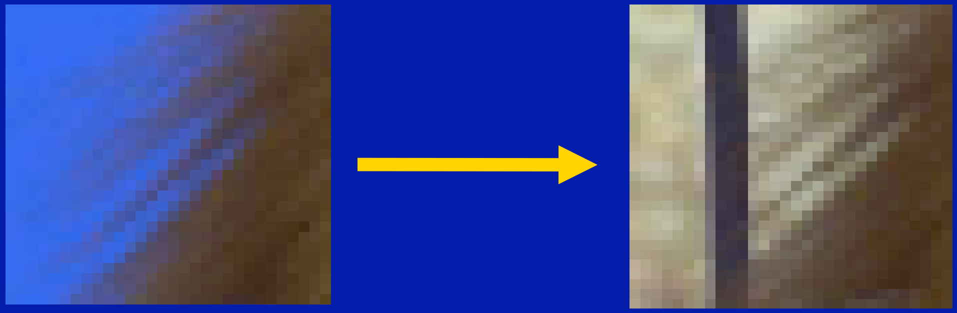
Yung-Yu Chuang Brian Curless
David H. Salesin   Richard Szeliski

[today's presentation by Jiayuan Meng]

Digital matting

- A foreground element is extracted from a background image by estimating a color and opacity for the foreground element at each pixel
- Alpha is the opacity value at each pixel

$$C = \alpha F + (1 - \alpha) B$$



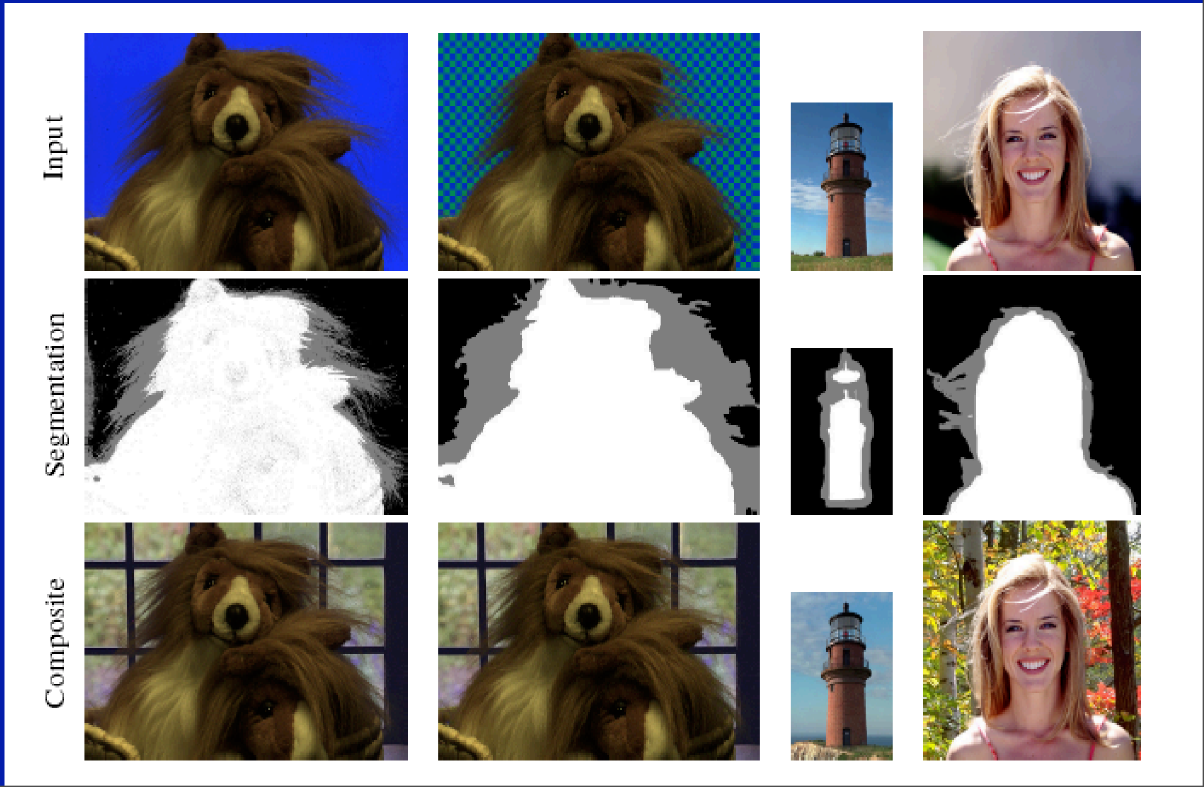
Baby approaches

- Blue screen
  - Use background of known color
  - Make certain assumptions about the colors in the foreground
  - Assumptions fine tuned

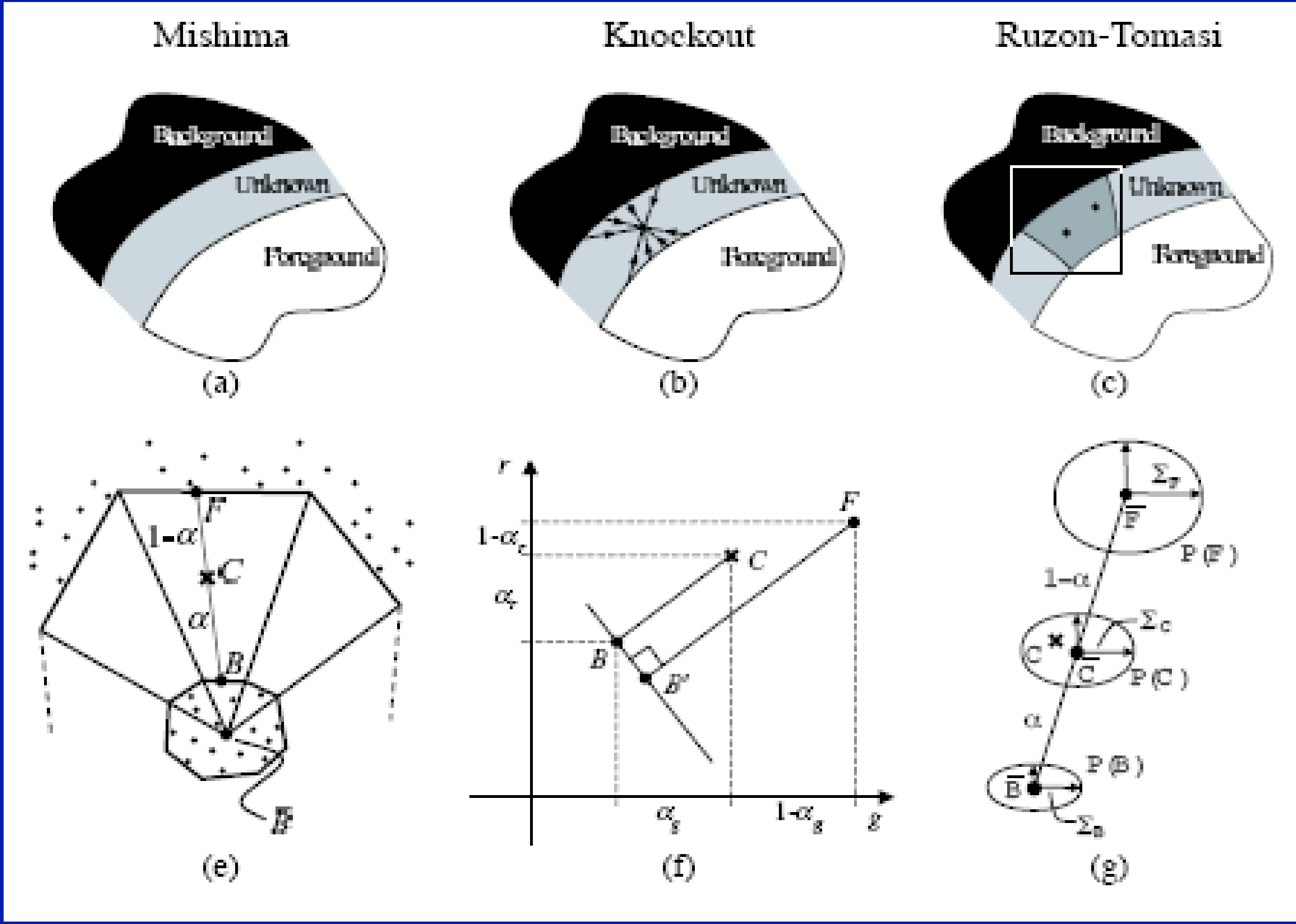
Master approaches …

- Mattes from natural background use statistics of known regions of foreground or background
  - - In order to estimate the foreground and background colors along the boundary
  - Once colors are known opacity value is uniquely determined

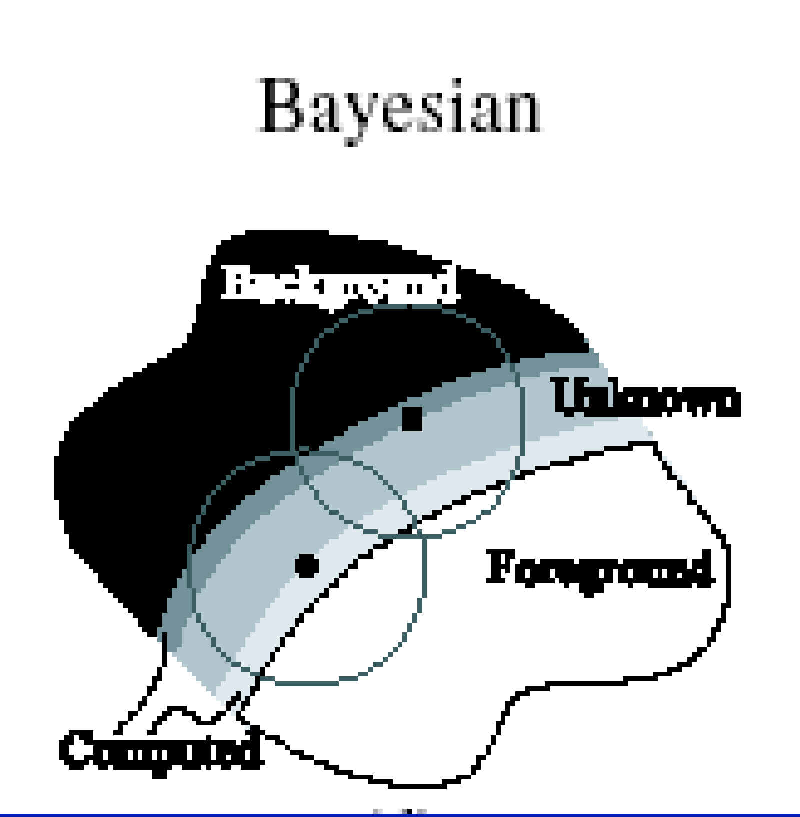
* Mishima
* Knockout
* Ruzon Tomasi



Matting process



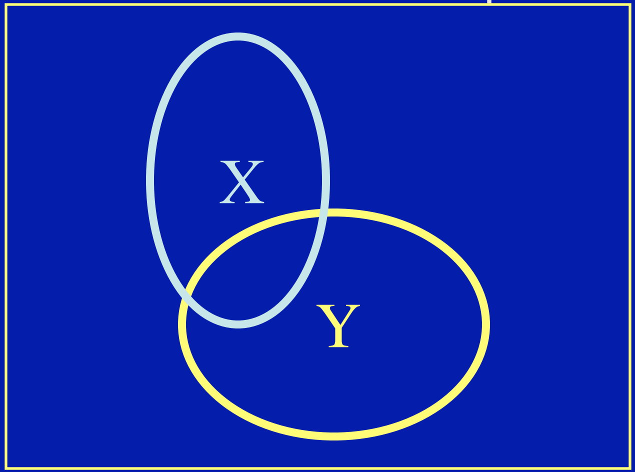
Build foreground and background probability distributions from a given neighborhood



$$P(F, B, \alpha \mid C)$$

Bayes rule

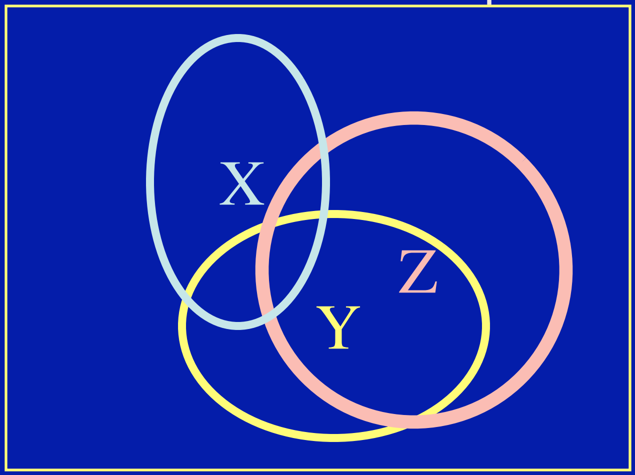
- In terms of our problem:



$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(Y \mid X)P(X)}{P(Y)}$$

Bayes rule

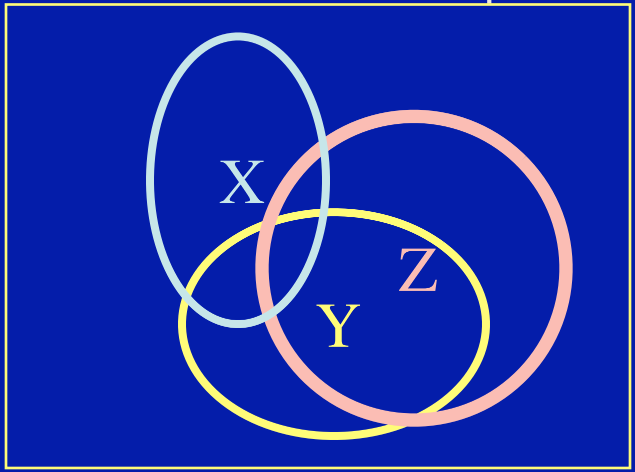
- In terms of our problem:



$$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(Y \mid X)P(X)}{P(Y)}$$

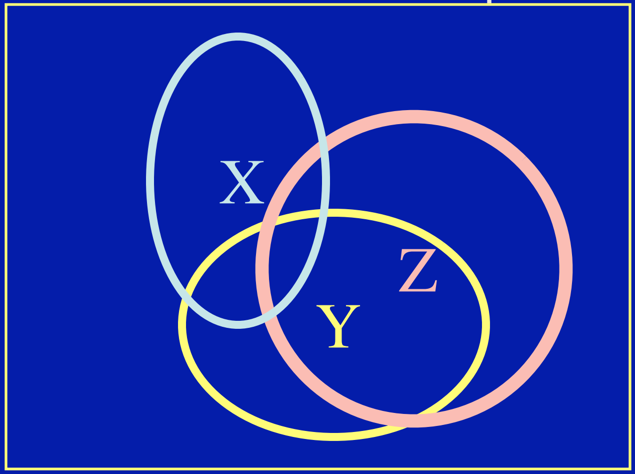
Bayes rule

- In terms of our problem:



Bayes rule

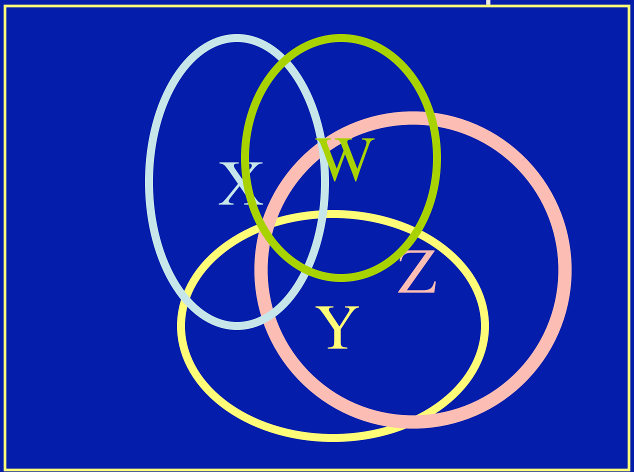
- In terms of our problem:



$$P(X,Z\mid Y)=\frac{P(Y\mid X,Z)P(X)P(Z)}{P(Y)}$$

Bayes rule

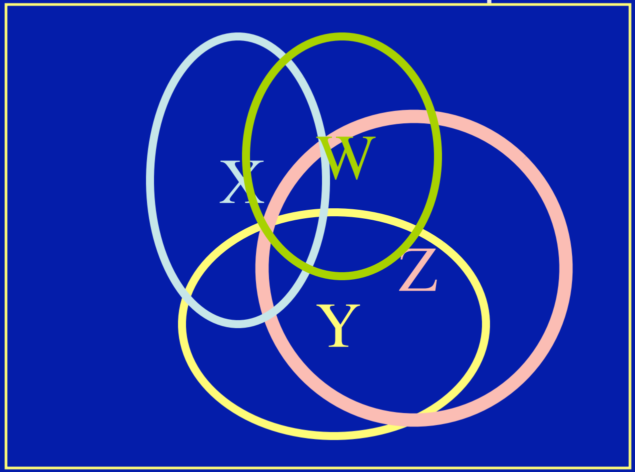
- In terms of our problem:



$$P(X,Z\mid Y)=\frac{P(Y\mid X,Z)P(X)P(Z)}{P(Y)}$$

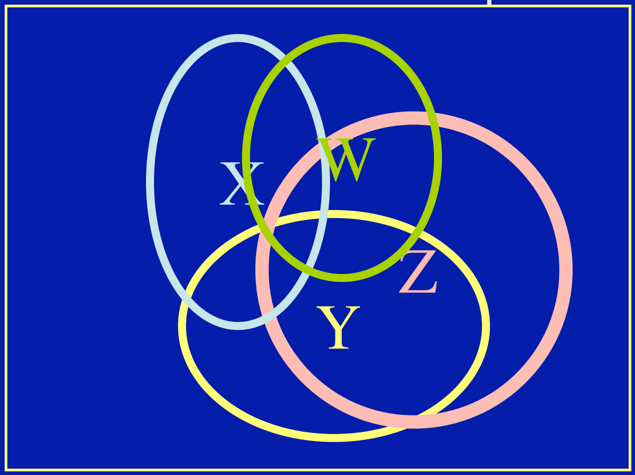
Bayes rule

- In terms of our problem:



Bayes rule

- In terms of our problem:



$$P(F,B,\alpha \mid C) = \frac{P(C \mid F,B,\alpha)P(F)P(B)P(\alpha)}{P(C)}$$

MAP estimation

- Try to find the most likely estimates for F, B, and α
  - Given the observation C
- Express as maximization over a probability distribution P
- then use Bayes's rule to express result as maximization over sum of log likelihoods: Equation (4)

$$\arg\max_{F,B,\alpha} P(F,B,\alpha \mid C) = \frac{\arg\max_{F,B,\alpha} P(C \mid F,B,\alpha)P(F)P(B)P(\alpha)}{P(C)}$$
$$= \arg\max_{F,B,\alpha} L(C \mid F,B,\alpha) + L(F) + L(B) + L(\alpha)$$

- Where L(·) is the log likelihood : L(·) = log P (·)
- α is assumed to be uniform distribution, so L (α)=const

The first term: L(C|F,B, α)

$$\arg\max_{F,B,\acute{a}} L(F,B,\acute{a} \mid C) + L(F) + L(B)$$

- Modeled by measuring the difference between the observed color and the color that would be predicted by the estimated F, B, and α:

  Gaussian/Normal Distribution:

$$L(C \mid F, B, \alpha) = -\left\| C - \alpha F - (1-\alpha)B \right\|^2 / \sigma_C^2$$

- This log-likelihood models error in the measurement of C and corresponds to a Gaussian probability distribution centered at

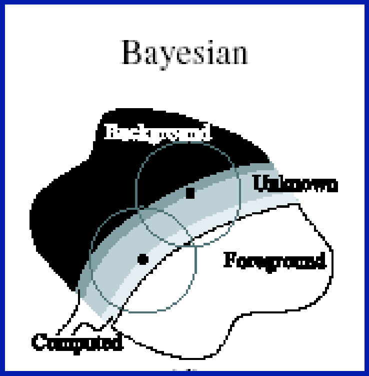
$$\overline{C} = \alpha F + (1-\alpha)B$$

- with standard deviation σ_C

Estimate the foreground term L(F)

$$\arg\max_{F,B,\alpha} L(F,B,\alpha \mid C) + L(F) + L(B)$$

- Use the spatial coherence of the image
- Build the color probability distribution using the known and previously estimated foreground colors within each pixel's neighborhood N.
- To more robustly model the foreground color distribution
  - We weight the contribution of each nearby pixel i in N according to 2 separate factors:
  1. weight the pixel's contributions by α^2_i which gives colors of more opaque pixels higher confidence
  2. use a spatial Gaussian falloff g_i, with $\sigma = 8$ to stress the contribution of nearby pixels over those that are further away
- The combined weight is then $\omega_i = \alpha^2_i g_i$



Estimate the foreground term L(F) $\arg\max_{F,B,á} L(F,B,á \mid C) + L(F) + L(B)$

- Given a set of foreground colors and their corresponding weights
  - First partition colors into several clusters using the method of Orchard and Bouman (1991)
- For each cluster, we calculate the weighted mean color and the weighted covariance matrix Σ_F

$$\overline{F} = \frac{1}{W} \sum_{i \in N} \omega_i F_i$$

$$\sum_F = \frac{1}{W} \sum_{i \in N} \omega_i (F_i - \overline{F})(F_i - \overline{F})^T$$

- Where $W = \sum_{i \in N} \omega_i$

Estimate the foreground term L(F)

- The log likelihoods for the foreground L(F) can then be modeled as being derived from an oriented elliptical Gaussian distribution, using the weighted covariance matrix as follows:

$$L(F) = -(F - \overline{F})^T \sum_F^{-1} (F - \overline{F}) \big/ 2$$

Estimating the log likelihood for L(B)

$\arg\max_{F,B,\acute{a}} L(F,B,\acute{a} \mid C) + L(F) + L(B)$

- The definition of the log likelihood for the background L(B) depends on which matting problem we are solving
- For natural image matting:
  - Use an analogous term to that of the foreground
  - Setting ω_i to $(1-\alpha)^2{}_i g_i$
  - and substituting B in place of F in every term of the previous equations

$$\overline{B} = \frac{1}{W} \sum_{i \in N} W_i B_i$$

$$\Sigma_B = \frac{1}{W} \sum_{i \in N} \omega_i (B_i - \overline{B})(B_i - \overline{B})^T$$

$$L(B) = -(B - \overline{B})^T \Sigma_B^{-1} (B - \overline{B}) / 2$$

Object Function

$$\mathrm{O}(F, B, \alpha) = -\left\| C - \alpha F - (1-\alpha) B \right\|^2 \Big/ \sigma_C^2 - (F - \overline{F})^T \sum\nolimits_F^{-1} (F - \overline{F})/2 - (B - \overline{B})^T \sum\nolimits_B^{-1} (B - \overline{B})/2$$

Break problem into 2 quadratic sub-problems $\arg\max_{F,B,\alpha} L(F,B,\alpha \mid C) + L(F) + L(B)$

- Because of the multiplications of α with F and B in the log likelihood L (C|F,B, α)
  - The function we are maximizing in (4) is not a quadratic equation in its unknowns
- To solve the equation efficiently
- In the first sub-problem
  - Assume that α is a constant
  - Under this assumption, taking the partial derivatives of (4) with respect to F and B and setting them equal to 0 gives:
  - equation (9)

$$\begin{bmatrix} \Sigma_F^{-1} + I\sigma^2/\sigma_C^2 & I\alpha(1-\alpha)/\sigma_C^2 \\ I\alpha(1-\alpha)/\sigma_C^2 & \Sigma_B^{-1} + I\alpha(1-\alpha)/\sigma_C^2 \end{bmatrix} \begin{bmatrix} F \\ B \end{bmatrix} = \begin{bmatrix} \Sigma_F^{-1}\overline{F} + C\alpha/\sigma_C^2 \\ \Sigma_B^{-1}\overline{B} + C(1-\alpha)/\sigma_C^2 \end{bmatrix}$$

- We can find the best parameters F and B by solving the 6X6 linear equation

The second sub-problem

- Assume that F and B are constant, yielding a quadratic equation in α.
- We arrive at the solution of this equation by projecting the observed color C onto the line segment FB in color space: equation (10)

$$\alpha = \frac{(C - B) \times (F - B)}{\left\| F - B \right\|^2}$$

- Where the numerator contains a dot product between two color difference vectors

The second sub-problem

- To optimize the overall equation (4) we alternate between
  - assuming that α is fixed to solve for F and B using (9)

$$\begin{bmatrix} \Sigma_F^{-1} + I\sigma^2/\sigma_C^2 & I\alpha(1-\alpha)/\sigma_C^2 \\ I\alpha(1-\alpha)/\sigma_C^2 & \Sigma_B^{-1} + I\alpha(1-\alpha)/\sigma_C^2 \end{bmatrix} \begin{bmatrix} F \\ B \end{bmatrix} = \begin{bmatrix} \Sigma_F^{-1}\overline{F} + C\alpha/\sigma_C^2 \\ \Sigma_B^{-1}\overline{B} + C(1-\alpha)/\sigma_C^2 \end{bmatrix}$$

  - And assuming that F and B are fixed to solve for α using (10)

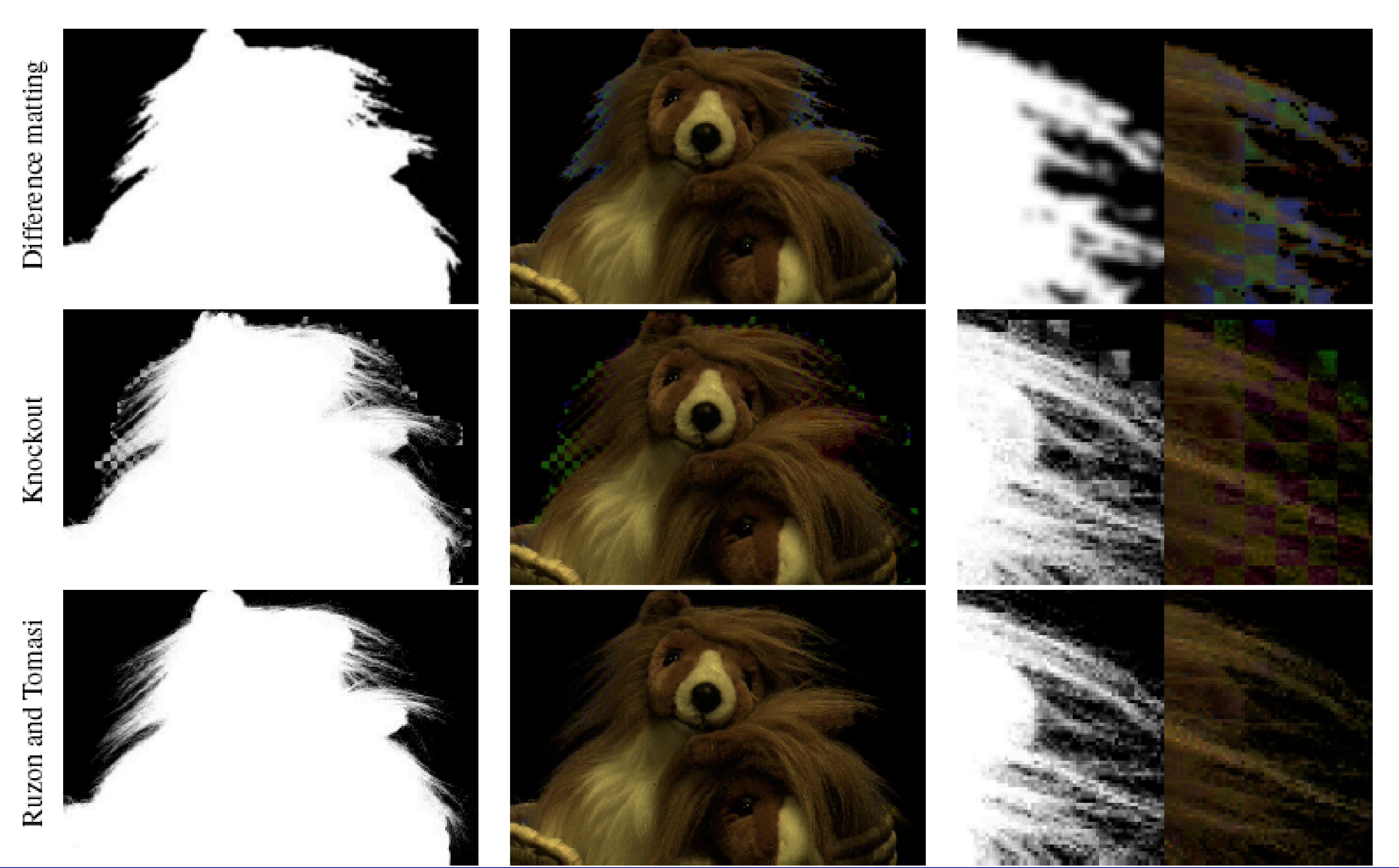
$$\alpha = \frac{(C-B)\times(F-B)}{\|F-B\|^2}$$

- To start the optimization, we initialize α with the mean α over the neighborhood of nearby pixels
- And then solve the constant -α equation (9)

When more than one foreground or background cluster

- We perform the above optimization procedure for each pair of foreground and background clusters and choose the pair with the maximum likelihood
- In this model, in contrast to a mixture of Gaussians model, assumes that the observed color corresponds to exactly pair of foreground and background distributions.
- In some cases, this model is likely to be the correct model
  - But we can conceive of cases where mixtures of Gaussians would be desirable
  - I.e. when 2 foreground clusters can be near one another spatially and thus can mix in color space
- Ideally, we would like to support a true Bayesian mixture model
- In practice, even with our simple exclusive decision model, we have obtained better results than the existing approaches

Comparison



Comparison of 2 natural images