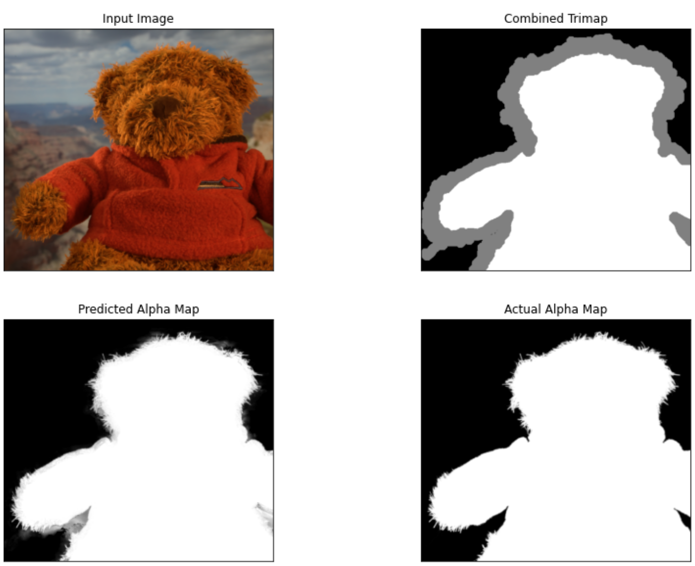
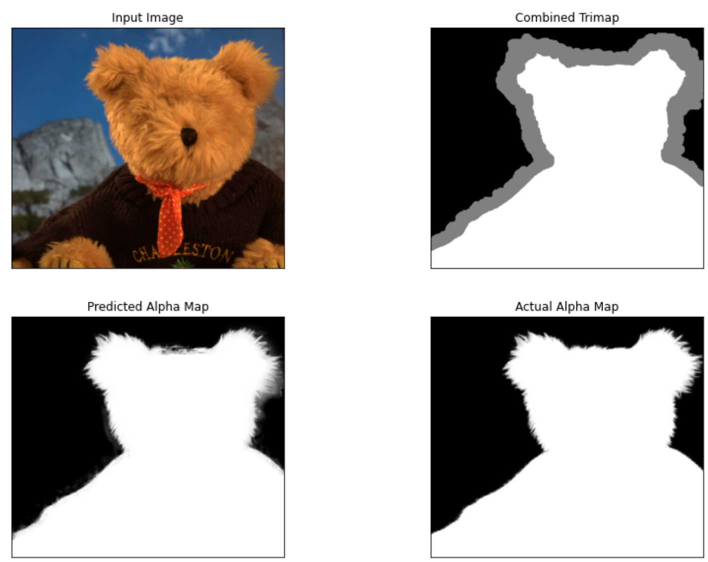
Bayesian Matting

Some example output

Error - ~91

Error - ~142





Error - ~29

Error - ~ 47

(Error - Sum of absolute differences between ground truth and predicted alpha scaled to 0-1)

Maths behind algorithm

We need to find the alpha, foreground, and background given the image to find the output. So in the probability sense, we need to find values of alpha, F and B that maximize its probability given image

Maximize $P\left(\alpha, F, B|C\right)$

Now applying Bayes theorem and law of total probability, we get

$$P\left(\alpha, F, B|C\right) = \frac{P\left(C|\alpha, F, B\right).P\left(\alpha, F, B\right)}{P\left(C\right)}$$

Since we are already given the image, we have a fixed value of P(C), which cannot be changed. So, the numerator is what we would like to maximize. It is quite intuitive that alpha, F and B would be independent of each other. Thus, we can separate each of them. The paper assumes that the first numerator term would be a gaussian pdf. And thus, taking logarithm would be better as it would remove the exponential, and it is monotonically the same sign as exponential. Thus the aim is to maximize the following

$$L\left(\alpha, F, B|C\right) = L\left(C|\alpha, F, B\right) + L\left(\alpha\right) + L\left(F\right) + L\left(B\right)$$

From the Gaussian assumption (difference between the actual image and matted image), we have the first term as

$$C_{pred} = \alpha F + \left(1 - \alpha\right) B$$

Above is the matting equation whose difference is modeled as gaussian. Thus, the closer the value to zero, the more is the value of the log and farther it is, the value is reduced exponentially.

$$L\left(C|\alpha, F, B\right) = -\frac{1}{\sigma_c^2}||C - \alpha F - \left(1 - \alpha B\right)||^2$$

For the remaining L(F) and L(B), since they are weighted values, the paper suggests using elliptical gaussian pdf whose log-likelihood is given by

$$L\left(F\right) = -\frac{1}{2}\left(F - \overline{F}\right)^T \cdot \Sigma_F^{-1} \cdot \left(F - \overline{F}\right) \quad \text{and}$$

$$L\left(B\right) = -\frac{1}{2}\left(B - \overline{B}\right)^T \cdot \Sigma_B^{-1} \cdot \left(B - \overline{B}\right)$$

The paper has kept L(alpha) as constant and left it to vary for future works. Hence, that can be ignored. Here the mean F and covariance matrix of F is for a particular cluster and the same for B.

(the final equation after taking sum gets non-quadratic and thus requires numerical analysis to solve - described on next page)

Clustering is done to partition colors using the Orchard and Bouman method. A point to note is that the numerical method is run for each combination of clusters of F and B.

The direct solution to the equation is unknown because the equations are non-trivial (non-quadratic derivatives). We would have to use numerical methods to solve them. Following are the steps that are used:

1. Initialized alpha with mean of all the known alpha pixels in the neighborhood.
2. I kept this value constant and used the following matrix to solve for F and B

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

   .

3. Now, F and B calculated from the above 6x6 (each element is 3x3) matrix are kept constant, and new alpha is calculated.

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

4. The log-likelihood is calculated using this newly calculated alpha, F and B, and the iteration repeats (2-4) to maximize log-likelihood. (Formulas written at the start) for a fixed number of iterations.

Once the alpha is known for a particular image, the matting equation can predict the background of any image with reasonable accuracy.