

# Interactive Foreground Extraction using Iterated Graph Cuts

Sasank Chilamkurthy  
Tharun Kumar Reddy  
Rajeev Puppala

April 30, 2015

## 1 Introduction

In this paper, we describe a method to extract foreground out of a image with help of a very simple interaction from user. This is based on the papers [1] and [2].

## 2 Image Segmentation by Graph Cut

In this section, general segmentation framework we have used is detailed. Denote the given input image as an array  $\bar{z} = (z_1, z_2, \dots, z_N)$ . The segmentation of the image is expressed as an array of "opacity" values  $\bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)$ . Generally for soft segmentation  $\alpha_n \in [0, 1]$  and for hard segmentation  $\alpha_n \in \{0, 1\}$ . In this paper, we deal with hard segmentation with 0 representing background and 1 representing foreground.

Let the model parameters for color/intensity model of the background and foreground be denoted by  $\theta_\alpha$  for  $\alpha$  equal to 0 and 1 respectively. Under this model, let likelihood function be given by  $f(z|\theta_\alpha)$ . Let us denote all the parameters of the model with  $\bar{\theta}$ . So, our objective is to find  $\bar{\alpha}$  given model parameters  $\bar{\theta}$ .

This can be achieved by maximizing the following with respect to  $\bar{\alpha}$  :

$$E(\bar{\alpha}|\bar{\theta}, \bar{z}) = U(\bar{\alpha}|\bar{\theta}, \bar{z}) + V(\bar{\alpha}, \bar{z})$$

Here,

$$U(\bar{\alpha}|\bar{\theta}, \bar{z}) = \sum_{n=1}^N -\log f(z_n|\theta_{\alpha_n})$$

is called data term or unary term. Observe that this is negative log likelihood of the image data  $\bar{z}$ . The term

$$V(\bar{\alpha}, \bar{z}) = \gamma \sum_{(m,n) \in C} \text{dis}(m, n) \mathbf{1}_{\alpha_m \neq \alpha_n} \exp -\beta(z_m - z_n)^2$$

is called smoothness term with  $\beta$  and  $\gamma$  suitably chosen constants.  $C$  is set of all neighbouring pixels and  $\text{dis}(m, n)$  is Euclidean distance between pixel  $m$  and  $n$ . This term can be seen as contribution of the prior on  $\bar{\alpha}$  that segmentation tends to be smooth. Thus our problem is

$$\underset{\bar{\alpha} \in \{0,1\}^N}{\text{maximize}} \quad E(\bar{\alpha}|\bar{\theta}, \bar{z}) \quad (1)$$

Note that this problem is a discrete optimization problem. However, efficient exact solution exists as this problem can be posed in the form of Graph Cuts. A standard minimum cut algorithm [2, 3] can be employed to solve this.

Now that we have seen general framework of segmentation, we'll next discuss the models of the background and foreground we have used.

### 3 Background/Foreground Intensity Models

In this section, we deal with models and estimation of model parameters  $\bar{\theta}$ . This is where user interaction comes into play. Our user draws a rectangle roughly around the desired foreground. We use this to estimate model parameters.

#### 3.1 Histograms

If the image  $\bar{z}$  is gray image, then it is conceivable to use histograms as intensity models. Indeed, if we wish to use this model, we convert the input colour image to a gray image and build histograms, one for background and another for foreground. However, we will lose discriminatory power color image provides.

#### 3.2 GMMs

Unlike above, we cannot hope to build a histogram for three dimensional color data. So instead, we fit a Gaussian mixture model of  $k$  components to each of the foreground and background.

#### 3.3 General Iterative Algorithm

In the first iteration, model parameters  $\bar{\theta}$  are obtained from user supplied rectangle foreground and background. These are used to segment the image and this segmentation is in turn used to improve the parameter estimates. This process is repeated until convergence or some maximum iterations are hit. This is beneficial as our segmentation estimate keeps on improving. We detail this iterative algorithm:

1. **User Interaction** : Set  $T_1 = \{z_i | i \in \text{user drawn rectangle}\}$  and  $T_0 = T_1^c$ .
2. **Parameter Estimation** : Estimate parameters  $\theta_0$  and  $\theta_1$  from  $T_0$  and  $T_1$  respectively.

3. **Segmentation** : Solve

$$\bar{\alpha} = \underset{\bar{\alpha} \in \{0,1\}^N}{\text{maximize}} \quad E(\bar{\alpha} | \bar{\theta}, \bar{z})$$

and update  $T_k = \{z_i | \alpha_i = k\}$ ,  $k = 0, 1$

4. **Check for Convergence** : If convergence is not reached or maximum iterations is not hit, go to step 2. Otherwise return  $\alpha$ .

## 4 Experiments

We present results from experiments from the above two models.

### 4.1 Histogram Model

## References

- [1] Rother, Carsten, Vladimir Kolmogorov, and Andrew Blake. "Grabcut: Interactive foreground extraction using iterated graph cuts." *ACM Transactions on Graphics (TOG)* 23.3 (2004): 309-314.
- [2] Boykov, Yuri Y., and M-P. Jolly. "Interactive graph cuts for optimal boundary & region segmentation of objects in ND images." *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. Vol. 1. IEEE, 2001.
- [3] Kolmogorov, Vladimir, and Ramin Zabini. "What energy functions can be minimized via graph cuts?." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.2 (2004): 147-159.