CS231A Course Project Proposal
Fully automated image matting with Kinect

Karen Cheng
kycheng@gmail.com

Buu-Minh Ta
bmta@stanford.edu

Abstract

Future Distribution Permission

1. Proposal

1.1. Idea

We plan to use the depth map from the Kinect to perform image matting, ie to extract the foreground of an image, and replace the background with another background. Typically, matting algorithms rely only on color information. We believe the additional depth information provided by the Kinect could enhance the accuracy and efficiency of the algorithm by automating the trimap generation process.

1.2. Data collection

The data will be collected using a Kinect. Our data collection will consist of images that are easy to interpret and images that are more complex. An easy-to-interpret image would have a uniform background and a well-defined foreground, whereas a complex image would have a sophisticated background and a fuzzy foreground.

1.3. Algorithms

The process is broken down into three steps. First, we generate a trimap to provide an initial guess. Next, we apply the Bayesian matting algorithm proposed by Chuang et al. Finally, we replace the original background with a new background. To compute a trimap, we propose two methods that utilize the depth data provided by the Kinect. The simplest approach applies a threshold to the depth values in order to differentiate the foreground from the background. The second approach applies segmentation to the depth map. The segmentation methods we consider applying are Mean-shift, Ncuts and edge detection.

J. Finger et al. suggest using depth information to improve the classic Bayesian matting in their paper, but the algorithm is not fully detailed. We will explore this option to see how we can achieve better performance on the matting.

1.4. Papers and existing algorithms

- A Bayesian approach to digital matting, B. Curless, D. Salesin, R. Szeliski
- Video matting from depth maps, J. Finger, O. Wang
- A closed-form solution to natural image matting, A. Levin, D. Lischinski, Y. Weiss

1.5. Results evaluation

From our sample images, we will define a ground truth by manually extracting the foreground, and compare this to the results of our algorithm. We can measure the error as the number of false positives and false negatives with the ground truth. This test will be conducted on images of various levels of complexity, which will allow us to understand in which conditions the algorithm performs well.