University of Bucharest
Faculty of Mathematics and Computer Science
Department of Computer Science

Alexandra Diaconu
Bogdan Alexe
MSC in Computer Science

Computer Vision - Laboratory class 3

GrabCat - Foreground/background segmentation of cats in images

Objective

Image segmentation is a fundamental problem in computer vision where the goal is to group pixels with similar characteristics such as color, intensity, texture, semantic class. The result of image segmentation is a set of segments that collectively cover the entire image. Pixels grouped in the same segments are similar with respect to some characteristic. Adjacent regions are significantly different with respect to the same characteristic.

In this exercise class we will focus on the problem of image segmentation using low-level features like color. The goal is to produce results similar to the GrabCut algorithm..

GrabCat on raw pixels

Our initial task is to segment the images from the 'dataset' folder (images with cats) working with the raw pixels. Initially, a user draws a rectangle around the foreground region, such that the foreground region should be completely inside the rectangle (Figure 1). In a similar manner wrt GrabCut we follow an iterative procedure of building a color model for background and another one for foreground. Consequently, our GrabCat (Figure 2) algorithm consists of the following steps:

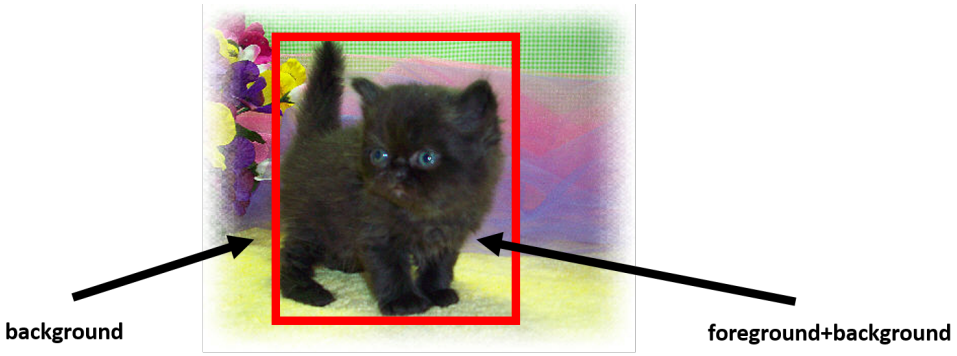

Figure 1: *Regions from the initial image after the user input*

Figure 2: *Original image (a), ground truth (b), our algorithm-GrabCat (c), GrabCut (d)*

- user inputs the rectangle. All pixels outside this rectangle are considered to model only the background whereas pixels inside the rectangle model background and foreground.
- do an initial assignment of all pixels, assigning label BG (background) to all pixels outside the rectangle and label FG (foreground) to all pixels inside the rectangle;
- cluster each set of pixels (BG and FG) using the K-means algorithm and obtain $2 \times K$ centroids, K centroids for the BG model and K centroids for the FG model.
- recompute the assignment of all pixels in the rectangle, my computing their Euclidian distance to the $2 * K$ centroids of the BG and FG models. Assign label BG if the minimum distance is found wrt a centroid from BG.
- update the cluster centroids using the updated set of pixels from BG and FG.
- repeat the last 3 steps for a number of iterations.

The folder 'ground-truth' contains binary masks for the input images (from 'dataset' folder). This enables computing precision and recall in order to evaluate the performance of GrabCat compared to the GrabCut implementations from OpenCV.

GrabCat on supepixels

The second task is to adapt the initial algorithm designed to work on raw pixels using the superpixels obtained via the SLIC algorithm. Superpixels contains groups of pixels similar

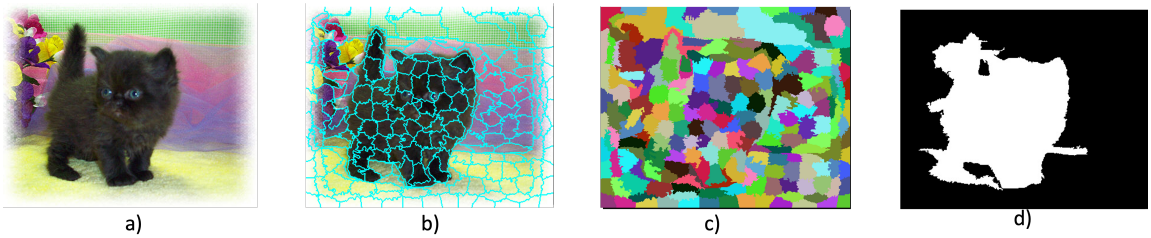

Figure 3: *Original image (a), superpixels (b), superpixels color labeling (c), GrabCat using superpixels (d)*

in color and texture and usually undersegment objects in an image (an object contains several superpixels). More importantly, superpixels adhere to object boundaries which makes them suitable for using in segmenting objects. The essential difference wrt initial implementation is that all pixels in a superpixel receive the same label (either BG or FG). We compute BG and FG color models by using for each superpixel its mean color value derived from the colors of all pixels contained in it (Figure 3).