Object Segmentation Using Deformable Shape Models

Jonathan Ludwig
MS Thesis Pre-Proposal
Department of Computer Science
Thesis Adviser: Dr. Roxanne Canosa

November 4, 2008

The purpose of this thesis is to develop and evaluate an improved deformable shape model and to construct a system to perform object segmentation on images using this model. In order to accomplish this we shall draw on previous research to evaluate models and algorithms which provide accurate results while maintaining invariance where important for robust object segmentation. This includes conducting experiments to evaluate different candidate local feature descriptors which are capable of representing regions of interest within an object. The orginal contribution of this thesis will be to develop an improved descriptor model which is capable of representing the shape of an object and is deformable to provide robustness against intra-class variation and invariance to in-plane rotation and scale as well as tolerance for some degree of rotation out-of-plane. The system must be able to construct a descriptor based on a set of training images and then be able to vary the parameters of the descriptor to match the appropriate location and region of the object, if it exists, in a separate set of testing images.

The task of effectively segmenting objects in images is an important area of interest in computer vision. The ability to distinguish the objects in an image from each other and the background is a critical step in several approaches for achieving larger goals in automated or computer-aided vision. For example, object detection, or the ability to detect if a given object is present in a scene and determine its location, typically relies on the ability to segment the pixels which belong to the object from the rest of the pixels of an image. The related task of object recognition, the ability to determine which objects or classes of objects are present in a scene, also relies on object segmentation. Object segmentation is task that humans accomplish very quickly and very well, however, it is difficult to do in an automated fashion. The problem is difficult for several reasons, key among these are the following: there may be a large amount of variation between different types of objects, as well as between objects of the same class; objects may have significantly different appearances under varying environmental factors, such as lighting; rotations and changes in scale and viewpoint also change the appearance of an object from scene to scene. It is also not clear what is the best representation to capture the definition of an object while maintaining appropriate flexibility.

Many different approaches have been explored to solve this problem. Researchers have used probabilistic models to attempt to distinguish foreground pixels from background pixels, predefined shapes or compound shapes which attempt to closely match the shape of an object, three-dimensional representations of an object, collections of sample patches of images of the object, and biologically inspired networks. Biederman's seminal work regarding recognition by components provides a psychological hypothesis of human recognition. Leibe and Schiele et. al. have published several works on constructing implicit shape models with image patches. Berg et. al. have worked toward providing robust shape matching using correspondences. Ferrari et. al. present object detection schemes using deformable shape models. Finally, many vision techniques rely on the detection of significant image regions and local descriptors, including work done by Lowe, Harris, and Gabor. This work will draw on research concerning deformable shapes, interest points, shape model descriptors, and statistical methods to achieve robust object segmentation and provide a foundation on which further work may build to accomplish larger goals such as object detection and recognition.