Instance Segmentation based Semantic Matting for Compositing Applications

Guanging Hu Centre for Intelligent Machines, McGill University Montreal, Canada bonniehu@cim.mcgill.ca

James J. Clark Centre for Intelligent Machines, McGill University Montreal, Canada clark@cim.mcgill.ca





Instance 1

Composite 1

Composite 1 and 2

Figure 1: Our approach first generates coarse instance masks that are used to create trimaps. Then, the trimaps and the original image are used by the image matting network to produce an alpha matte for each instance. Finally, the alpha mattes are used for image compositing.

Abstract—Image compositing is a key step in film making and image editing that aims to segment a foreground object and combine it with a new background. Automatic image compositing can be done easily in a studio using chromakeying when the background is pure blue or green. However, image compositing in natural scenes with complex backgrounds remains a tedious task, requiring experienced artists to handsegment. In order to achieve automatic compositing in natural scenes, we propose a fully automated method that integrates instance segmentation and image matting processes to generate high-quality semantic mattes that can be used for image editing task.

Our approach can be seen both as a refinement of existing instance segmentation algorithms and as a fully automated semantic image matting method. It extends automatic image compositing techniques such as chroma-keying to scenes with complex natural backgrounds without the need for any kind of user interaction. The output of our approach can be considered as both refined instance segmentations and alpha mattes with semantic meanings. We provide experimental results which show improved performance results as compared to existing approaches.

Keywords-image matting; instance segmentation;

I. INTRODUCTION

Image compositing is a key technique in movie production and image editing. It combines visual elements from different sources into one image. The process of extracting visual elements from a source typically involves precise extraction of foreground objects from the background using either manual rotoscopy, where an artist traces the object to be extracted, or by an automatic chroma-keying when the object is in front of a uniformly colored background (e.g. green-screen). In this work, we present a method for automatic foreground object extraction that can work even with complex backgrounds. Our method combines instance segmentation and image matting processes, which allows for multiple foreground objects to be classified and segmented and extracted from the background. The semantic labels provided by the instance segmentation process provides a way to automatically extract objects of different types. For example, we can extract all people in the scene from the background, leaving other types of foreground objects behind. Or, we can obtain multiple mattes, one for each instance, and these can be used as desired in later compositing operations.

Object segmentation is considered as one of the most important and complex tasks in Computer Vision. There have been many recent advances in segmentation algorithms based on learning-based methods, in large part due to a number of publicly available datasets that provide human-