Limiting the Search Range of Correlation Stereo Using Silhouettes

Geoffrey Egnal, Max Mintz and Kostas Daniilidis
GRASP Laboratory
University of Pennsylvania
Email: {gegnal,mintz,kostas}@grasp.cis.upenn.edu

Abstract

We present a new approach to combine two approaches to three-dimensional reconstruction: silhouette-based and correspondence-based approaches. The two approaches have complementary costs and benefits. Silhouette-based approaches deliver volumetric descriptions which often have very few outliers, but they cannot reconstruct concave surfaces. Correspondence-based approaches give surface descriptions with sub-pixel accuracy, but their search range either allows outliers or falls short of the correct match. We show that a combination of the two can deliver fine-grained accuracy with few outliers. Our specific implementation uses the silhouette reconstruction as prior data to center and bound a stereo search process. We explore the different performance characteristics of the combination and its two component methods qualitatively and quantitatively using real imagery.

1 Introduction

Silhouette and correspondence reconstruction approaches have both been successfully used for visual reconstruction. Despite their success, there is still a need for more accurate reconstruction. Aside from inherent limitations to the accuracy of the methods, some situations present different scenery and allow different camera configurations than is typically found for the two approaches. For instance, a system might only have 4-5 cameras, but which can be placed far apart. Moreover, some images in a system may have high texture while others have low texture. In some images, it may be easier than others to segment the foreground against the background. For these reasons, we present a new method that combines the two approaches. We show that the combination gives a more accurate reconstruction and operates under a larger range of camera configurations than either approach.

The first approach, based on correspondence, searches for matching points in two or more images. Using the resulting disparity information between the images, the method triangulates to build a 3D surface model of the scene. Correspondence approaches operate under two basic scenarios: uncontrolled environments and controlled environments. In uncontrolled environments, a single moving platform usually holds all cameras. Since the cameras are closely spaced relative to the scenery, correspondence approaches work well. In controlled environments, systems use correspondence to achieve highly accurate surface reconstructions. Some of these environments only permit a few camera locations; others are highly textured. In such cases, correspondence seems an ideal choice. Even though correspondence can produce high resolution reconstructions, often to sub-pixel accuracy, the approach does have problems. Among them, if the search range is too small, the correspondence search will not find the correct match, but if the range is too large, the matching process may find a similar match far away from the actual match.

The second approach uses silhouettes to build a volumetric model of objects of interest. Silhouettes from multiple cameras can each be projected out to the real world, forming a conic which bounds the possible volume of the object of interest. The intersection of these conics is a 'visual hull' around the actual 3D shape of the object of interest. The silhouette approach generally operates in controlled environments, where multiple cameras can be placed at large separations. When the objects of interest are easily segmented against the background and when there are many permissible camera positions, silhouette approaches are ideal. A benefit of silhouette approaches is that they exhibit few outliers because segmentation algorithms usually produce contiguous silhouettes. However, the silhouette approach does not provide insight into concave surfaces because concave surfaces are invisible to silhouettes. This fault ultimately limits the accuracy of the silhouette approach, especially when the system has few cameras.

We present a method combining the correspondence-based and silhouette-based approaches. The method uses the silhouette model from at least one camera as a starting point for the correspondence process for each pixel. The prior model limits the correspondence search and prevents outliers, but also keeps the search range large enough so that