A Stereo Confidence Metric Using Single View Imagery

Geoffrey Egnal, Max Mintz
GRASP Laboratory
University of Pennsylvania
{gegnal,mintz}@grasp.cis.upenn.edu

Richard P. Wildes
Centre for Vision Research
York University
wildes@cs.yorku.ca

Abstract

Although stereo vision research has progressed remarkably, stereo systems still need a fast, accurate way to estimate confidence in their output. In the current paper, we explore using stereo performance on two different images from a single view as a confidence measure for a binocular stereo system incorporating that single view. Although it seems counterintuitive to search for correspondence in two different images from the same view, such a search gives us precise quantitative performance data. Correspondences significantly far from the same location are erroneous because there is little to no motion between the two images. Using hand-generated ground truth, we quantitatively compare this new confidence metric with five commonly used confidence metrics. We explore the performance characteristics of each metric under a variety of conditions.

1 Introduction

1.1 Overview

We present a new method to diagnose where a stereo algorithm has performed well and where it has performed badly. All stereo systems estimate correspondences, but not all of these correspondences are correct. Many systems do not give an accurate estimate of how trustworthy their results are. Our new confidence method addresses many causes of stereo error, but it focuses on predicting stereo error caused by low-texture regions. This error source is particularly bothersome when operating in urban terrain or when the imagery contains large regions of sky.

The new confidence metric is based upon correspondence in images from a single view. Although it seems counterintuitive to search for correspondence in two different images from the same view, such a search provides valuable information. It gives us precise quantitative performance data; disparities significantly far from zero are erroneous because there is no motion between the two images. We use the term Single View Stereo (SVS) to refer to the disparity map produced by a stereo system applied to two images from one view, separated in time.

Specifically, we use single view stereo performance data as a confidence metric that predicts how a binocular stereo system which incorporates the single view would perform. The SVS output shows empirically where the stereo system has failed on the current scenery, and for this reason, SVS failure predicts failure pixelwise in the binocular case (see Figure 1). In practice, many stereo cameras are in motion, and the SVS disparity will not be zero at each pixel. We discuss how to modify the static SVS algorithm to accommodate images taken from a moving camera.

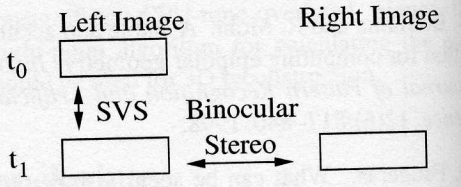

Figure 1: Single View Stereo. SVS searches for correspondence in two images from the same view. We use SVS failure as a confidence metric to predict binocular failure.

We compare the performance of SVS as a confidence metric with five commonly-used confidence metrics: (i) left/right consistency (LRC) predicts stereo error where the left-based disparity image values are not the inverse mapping of the right-based disparity image values. (ii) The matching score metric (MSM) directly bases confidence upon the magnitude of the similarity value at which the stereo matcher finds that left and right image elements match. (iii) The curvature metric (CUR) marks disparity points resulting from flat correlation surfaces with low confidence. (iv) The entropy-based confidence score (ENT) predicts stereo error at points where the left image has low image entropy. (v) The peak ratio (PKR) estimates error for those pixels with two similar match candidates. A brief comparison of the five metrics can be found in Table 1.

1.2 Previous Work

Unlike much research on stereo performance, we do not attempt to compare different stereo algorithms to see which is best [10, 5, 24, 29, 9]. Also, we do not attempt to find theoretical limits to stereo performance [4, 8, 19].

Instead, our current research deals with on-line confidence metrics which predict errors within seconds for a