Range Image Segmentation Based on Differential Geometry and Refined by Relaxation Labeling

Gerhard Roth ‡ † Martin D. Levine ‡

† On leave from the Division of Electrical Engineering
of the National Research Council of Canada
‡ Computer Vision and Robotics Laboratory
McGill Research Center for Intelligent Machines
McGill University, Montréal, Québec

Abstract

Concepts from the field of differential geometry are used to categorize each pixel in a range image as being a member of one of eight classes depending on the sign of the Gaussian and Mean curvature of the pixel's local neighbourhood. Since the local curvature estimates used to provide a class label for each pixel are very noise sensitive, these labels are made consistent with each other by a relaxation labeling process that assumes the range image is a discrete sampling of a piecewise smooth surface. The local neighbourhoods used during the curvature estimation and relaxation process are shifted away from discontinuities in order to produce more accurate results. Connected pixels with the same label can be grouped together to produce a segmentation of the range image into regions of homogeneous curvature. The result is a segmentation of the range image into a small number of regions which have the same curvature class. This segmentation can then be used as input to a matching processes in an object recognition system or to a viewpoint integration process in an object reconstruction system.

1. Introduction

One of the key problems in computer vision is that of segmentation. To segment an image is to extract from it a symbolic description that is useful for later processing. Segmentation has long been attempted on intensity images [13]. The difficulty with these approaches is that the contents of such images vary considerably with the ambient lighting and the surface reflectivity of the objects in the scene. These dependencies are removed if the direct depth from the sensor to the objects in the scene is available since this information is a function only of the scene geometry. Such depth data is called range data and can be obtained by active sensors, such as lasers and structured light [8], or by passive sensors such as stereo vision [7, 10]. The passive methods have not lived up to expectations, and as a result interest has increased in active range finding methods. These produce a dense map of the range of the objects in the scene to the sensor, which is appropriately called a range image.

We segment range images obtained from a laser rangefinder [12] by classifying each pixel into one of eight classes depending on the sign of the local Gaussian and Mean curvature [4, 9]. These curvatures are differential geometric quantities that completely characterize the surface locally in the sense that knowing them at each point enables the surface to be reconstructed. They are also invariant to viewpoint and have been used by a number of researchers as a criterion for grouping pixels together into regions [1, 2, 5]. However, there are two difficulties with this approach. The first problem is that since the Gaussian and Mean curvatures are based on second order derivatives computed with discrete data, they are very noise sensitive. If the assumption is made that a range image is a discrete sampling of an underlying piecewise smooth surface, then the curvature should not change dramatically over a small area. This simple principle enables relaxation compatibilities to be defined between pixels of different curvature classes in a neighbourhood [14, 3]. These compatibilities are then used to make initial curvature estimates more consistent with their neighbours by a relaxation labeling process. The second problem is that the curvatures are not correct if calculated across discontinuities; that is, the surface is not smooth but piecewise smooth. In order to avoid such errors, the windows from which these initial local curvatures are calculated must be shifted away from discontinuities with the degree of the shift depending on the strength of the discontinuity. This same shifting process also occurs for the window that defines the neighbours of a pixel during the relaxation process so that regions separated by a discontinuity do not affect each other. The result is a simple and efficient method of producing a coarse segmentation of the range image into regions of homogeneous curvature which can then be used as input to higher level processes, such as matching in an object recognition system [16], or viewpoint integration in an object reconstruction system.

2. Computing H and K For Range Images

Since a range image is taken from a single viewpoint it is not a general surface in $\mathbf{R}^3$. Instead it is a graph, sometimes called a Monge patch. This means that the surface S defined by this graph can be written as follows:

$$S = \left\{ (x,y,z) : \; x, \; y, \; z = f(x,y) \; where \; (x,y) \in D \subseteq \mathbf{R}^2 \right\} \tag{1}$$

From this definition of the surface the following formulae for the Gaussian (K) and Mean (H) curvature can be computed