SAPIENZA
UNIVERSITÀ DI ROMA

Position tracking in 3D scenes using high order primitives

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea Magistrale in MARR

Candidate

Edoardo Ghini
ID number 1616301

Thesis Advisors

Prof. Nome Cognome
Dr. Nome Cognome

Co-Advisor

Dr. Nome Cognome

Academic Year 2018/2019

Thesis defended on 31 October 2019
in front of a Board of Examiners composed by:

Prof. Nome Cognome (chairman)
Prof. Nome Cognome
Dr. Nome Cognome

Position tracking in 3D scenes using high order primitives
Master thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Version: October 10, 2019

Author's email: edoardo.ghini@gmail.com

Dedicated to nessuno

Abstract

fill with abstract

Acknowledgments

grazie

Contents

Chapter 1

Introduction

1.1 General idea

1.2 Problem statement

1.3 Previous work

1.4 State of the art

Chapter 2

Development

2.1 Dataset Processing

2.1.1 Sensor used

The data used are coming from an open sourced group of datasets, namely the Kitty Vision Benchmark Suite [ref kitti]. Among the others the datasets of interest for this project are built with a 360[degree symb] Velodyne Laserscanner, called lidar going forth. This kind of sensor is based on an array of high precision laser sensors which rotate at considerable speed around the vertical axis, thus covering a field of view of 360 degree. Each laser sensor computes the distance covered by its laser beam and registers the depth of the surface hit by the beam. Consequently the ranges can be aggregated and after a full rotation of the array of scanners a complete range based image of the environment could be reconstructed. Since there is a delay between each measurement in which the sensor rotates and, if its base is moving, is affected by an external rototranlation is needed a procedure to recover the distortion of the cloud of ranges. This is possible only with a reasonable extimation of the external isometrical movement that has affected the sensor or better of the velocity or the higher order derivatives of this movement, an example of this practice is explained in [ref to loam].

2.1.2 Dataset Composition

After the extraction of the ranges between the sensor and the surfaces of the enviroment, each range is converted in a three-dimensional coordinate. Thogether this coordinates form a cloud of points which will be used as initial data for the feature extraction procedure.

2.1.3 Spherical Conversion

Since the cloud is coming from an high precision sensor it can contain a considerable amount of points. Hence a spatially ordered reorganisation of the cloud could improve the performances. The redistribution considered is, as in [ref spherical depth image paper], spherically shaped this allows the points of the cloud to be described in spherical coordinates.

For this reason the whole cloud is projected in spherical coordinates. During this projection, it could be possible that, since the cloud points in certain areas are copious and they are also coming from distortion correction, some points lay in similar azimuthal and elevational coordinates, if this is the case then only the nearest point to the sensor is mantained and all the others will be ignored going forward.
This heuristic and the choice of the desired resolution of the sphere in terms of number of rows and columns will determine the result of the unprojection of the cloud.
In particular the result will be a matrix of points in which each point has been inserted in a determinated row and column accordingly to its elevation and azimuth. Consequently this structure will be exploited since it will be useful to find with a linear complexity all the points near a given one and in this particular case is enough to consider the points found in the matrix with coordinates that are adjacent to the coordinates of the desired point.

[mat formule cartesian to spherical]

2.1.4 Flat-surface removal

In order to reduce the overall complexity of the following computation blocks the cloud is stripped away from the redundant information: each point which does not carry relevant information about the relevantaracteristics of the environment should be discarded. This is the case for the portions of the environment which are scarce of edges and peculiar shapes, namely the floor and the floor and the ceiling if present or in general all the surfaces that lack of verticality. Therefore similarly as how is done in [ref to flat-surf paper] an algorithm of selection is used.

2.2 Features extraction

2.2.1 Normal Computation

Integral image

2.2.2 Clustering

2.2.3 Eigenvalue-based characteristics

Every potential feature is characterized, among the other things, by the eigenvalues of its [omega] matrix. There are structured informations that can be gained by relatively simple manipulation of this eigenvalues.
As it has been done in [ref to semantic3d weinmann], every feature when is created will also compute eight different traits :
(mat 8 traits linearity, planarity etc.) Two of these distinctive traits have been considered a consistent indicator of the nature of the feature that they are describing. Thus each feature that does not respect two upper bounds, one about the scattering and the other regarding the change of curvature, will be filtered out.

2.3 Motion tracking

2.3.1 Motion model

2.3.2 Correspondance finder

Solver

2.3.3 Aligner

Solver

2.3.4 Tracker

Chapter 3

Results

3.1 Flat surface removal experiment

3.2 Normal computation experiment

3.3 Clustering experiment

3.4 Full traking experiment