Robotics and Autonomous Systems 74 (2015) 309–317

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

# Non-parametric calibration for depth sensors

Maurilio Di Cicco *, Luca Iocchi, Giorgio Grisetti

*Department of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, I-00185, Rome, Italy*

## HIGHLIGHTS

- Automatic calibration method for removing distortions of RGBD sensors.
- Non-parametric distortion model adequate for any depth sensor.
- The procedure is easy to apply and can be done by non-expert users.
- It does not require the use of any other software or of any calibration device.
- Freely available at http://easy_depth_calibration.dis.uniroma1.it.

ARTICLE INFO

*Article history:*
Available online 3 September 2015

*Keywords:*
Calibration
Mobile robots
Depth camera

ABSTRACT

RGBD sensors are commonly used in robotics applications for many purposes, including 3D reconstruction of the environment and mapping. In these tasks, uncalibrated sensors can generate poor quality results. In this paper we propose a quick and easy to use approach to estimate the undistortion function of RGBD sensors. Our approach does not rely on the knowledge of the sensor model, on the use of a specific calibration pattern or on external SLAM systems to track the device position. We compute an extensive representation of the undistortion function as well as its statistics and use machine learning methods for approximation of the undistortion function. We validated our approach on datasets acquired from different kinds of RGBD sensors and using a precise 3D ground truth. We also provide a procedure for evaluating the quality of the calibration using a mobile robot and a 2D laser range finder. The results clearly show the advantages in using sensor data calibrated with the method described in this paper.



## 1. Introduction

Depth sensors, like Microsoft Kinect or ASUS Xtion, are major technological achievements and gave a new impulse to a wide range of 3D applications in robotics such as SLAM [1–3], super-resolution mapping [4], object recognition and many others. Unfortunately, these devices suffer from a large systematic distortion that is hard to model. Neglecting this distortion has substantial effects on the algorithms that use depth data. As shown in Fig. 1, uncalibrated sensors produce data with a systematic noise that cannot be easily compensated by standard SLAM algorithms.

Kinect-like RGBD sensors operate on a stereo principle, where one of the cameras is replaced by a light source and the other camera senses the light pattern reflected by the scene. The depth is recovered by a proprietary algorithm developed by PrimeSense. A third camera is then used to augment the depth layer with RGB information. In principle, calibrating this sensor would require to estimate the intrinsic and extrinsic parameters of the three cameras in the system. However, the lack of knowledge about the pattern matching algorithm used to determine the stereo correspondences makes the application of these parameters not straightforward. Even assuming that the matching is always correct, one would have to first determine the disparity by applying the nominal parameters to the measured depths then use the estimated parameters to recover an improved depth estimate. Still it remains unclear how to estimate the intrinsics of the IR projector used as light emitter. Furthermore, we would lose the hardware acceleration provided by the PrimeSense algorithms running on the device.

In this paper, we propose an approach to compute an *undistortion function* that is able to convert a distorted depth image onto another one where the systematic distortion is removed. The undistortion function maps every possible depth measurement to a multiplying factor that is the ratio between the true depth and the measured one, as proposed by Teichman et al. [5].

In our approach we experimented two different kinds of function approximators, computed by the corresponding machine

* Corresponding author.
*E-mail addresses:* dicicco@dis.uniroma1.it (M. Di Cicco), iocchi@dis.uniroma1.it (L. Iocchi), grisetti@dis.uniroma1.it (G. Grisetti).

http://dx.doi.org/10.1016/j.robot.2015.08.004
0921-8890/

**Fig. 1.** 3D reconstruction of a home environment from depth data acquired with an ASUS Xtion mounted on a Turtlebot mobile robot. Left: the result using uncalibrated data. Right: the output of the same algorithm when using data calibrated with our procedure. The systematic error in the calibration resulted in noisy open loop estimates that result visible even with correct loop closings.

learning regression methods: a K-Nearest Neighbor (KNN) and an Artificial Neural Network (ANN). The training data for these models are extracted automatically from a set of RGBD images looking at a scene with a main planar surface at different distances.

The calibration procedure presented in this paper is fully automatic, after a set of depth images looking at a planar surface has been taken. The acquisition of these data typically requires less than 1 min and the calibration approach is usually implemented off-line. The application of our calibration procedure does not require specific expertise and it can be performed by a non-expert user with just a few recommendations about how to operate the device. Moreover, if the sensor is mounted on a mobile robot, the acquisition procedure can be performed in a fully automatized way by implementing a behavior that slowly moves the robot towards a wall. Being so quick and easy, our approach can be the preliminary step of a mission. Finally, since our approach does not specialize on a specific sensor model, it can be used for any 3D sensor.

Compared to our previous work [6], this paper:

- presents a deeper analysis of imaging depth sensors that allows to better characterize different distortion phenomena;
- extends the previous approach with a routine to adjust for systematic depth dependent noise, thus resulting in increased sensor accuracy;
- proposes a wider set of quantitative experiments.

We tested our approach with different devices, such as Kinect and ASUS Xtion in different versions, and the results of the experimental evaluation show that our procedure is effective and accurate.

The software implementing the described method is available through the website http://easy_depth_calibration.dis.uniroma1.it, which contains also easy-to-follow instructions for its use, datasets and further results, as well as the description of how to reproduce the results reported in this paper.

## 2. Related work

Calibration of RGBD sensors has been investigated in previous work as described in this section.

Yamazoe et al. [7] proposed a traditional approach of parametric calibration of the Kinect by using a checker board. Similarly, Fuchs et al. [8] presented a general analytical model to calibrate ToF cameras. In contrast to these approaches, we do not use any external calibration device and we do not rely on a specific sensor model. This is particularly useful since, as Smisek et al. stated [9], different Kinect devices shows different radial distortion patterns that may not be well approximated by the same model. The same behavior has been observed also in Xtion devices. Since we do not require a sensor model, our method can be easily applied to a broader range of sensors where distortion is not regular across the image.

While the above mentioned methods are aiming at calibrating the sensor parameters to lessen its systematic distortion, Nguyen et al. [10] focus on characterizing the distortion as a function of both distance and angle of a Kinect on the observed surface and derive a noise model to filter depth maps. Our approach instead does not rely on a parametric model to deal with the sensor's distortion.

A notable approach for computing an undistortion function of a depth sensor has been proposed by Teichman et al. [5] and it is known as *Calibrating, localizing, and mapping, simultaneously* (CLAMS). This approach relies on a dense representation of the undistortion function. In contrast to our approach, to compute this function the CLAMS system relies on a working SLAM module. By exploiting the fact that the distortion grows with the distance [11] and that it is neglectable for distances below 2 m, CLAMS repeatedly executes the SLAM algorithm several times. At each run, only the short ranges are used to track the position of the camera and to determine the ground truth of the distances. Thus the overall procedure is iterative and, at each round, SLAM is executed taking into account the most recent calibration. After the SLAM rounds are terminated, the multipliers are estimated for the full range of the sensor. To achieve acceptable results, the approach requires around 3 min of recorded data and one round of calibration, given the SLAM trajectory takes around 10 min. The shortcoming of this approach is that it requires a working SLAM module that operates on data acquired on a scene suitable for the calibration and rich in features for SLAM. Executing a full calibration with CLAMS is reported to be an overnight procedure. In contrast to this approach, our method simplifies the extraction of the undistorted depths, by exploiting the fact that pixel in the middle of the image are usually affected by a small systematic error and by fitting a plane in this region. Our method does not require any external software module and it is much faster as it takes a few milliseconds per image. As already mentioned, the input of our method is just a set of depth images capturing a large wall at different distances.

Basso et al. [12] proposed a method that is closely related to ours, where they estimate a per-pixel undistortion map *and* recover the calibration parameters of both RGB and Depth camera. The distortion of each pixel is assumed to be locally continuous, and a piecewise linear approximation of this function is computed from a few samples obtained by capturing a planar surface at a set of known depths. Based on this initial undistortion map, the authors calibrate the intrinsics and extrinsic parameters of the depth–RGB pair by means of a checkerboard. Similarly, Canessa et al. [13] propose a method for calibrating a depth and RGB pair by using a checkerboard with the RGB and IR camera of the sensor. To this end they disable the IR projector and