Rutgers University

Capstone Design

Electrical and Computer Engineering

Computer Vision-Based 3D Reconstruction for Object Replication

Authors:
Ryan Cullinane
Cady Motyka
Elie Rosen

Supervisor:
Kristen Dana

February 23, 2013

1 Introduction

This is the Introduction to

2 Methods

3 Experimental Results

4 Discussion

5 Cost Analysis

Item	Cost	Purpose
Printbot LC	§549.00	3D Printer for developing reconstructed objects
Microsoft Kinect	§109.99	The most powerful sensor for its price, contains color filtering and depth filtering in one package
ikg 3mm ABS Spool	§46.00	Material that the 2D Printer uses for creating objects, spool is necessary for feeding into the printer
2x 1lb 3mm ABS	§36.00	Extra printing material, no extra spool required
Arduino Leonardo	§24.95	Microcontroller to plan orientation of object during reconstruction process
EasyDriver Stepper Motor Driver	§14.95	Allows the ability to control a stepper motion at lower voltages such as from an Arduino microcontroller
Stepper Motor with Cable	§14.95	Motor that can be controlled in ”steps” this allows a more precise method for orienting objects being reconstructed
Total:	§795.84	

6 Current Trends in Robotics and Computer Vision

One of the reasons that the Kinect has become so popular for computer vision projects is that it is a cheap, quick, and highly reliable for 3D measurements. Many researchers are beginning to look into the possibility of using this device to achieve everything from a 3D reconstruction of a scene to aiding in a SLAM algorithm. The fact that this device is so affordable, and so many new resources are available, makes the Kinect a viable device for conducting research in the field of robotics and computer vision.

The KinectFusion Project is slightly different than other projects that were using the Kinect; instead of using both the RGB cameras and the sensor, this project tracks the 3D sensor pose and preforms a reconstruction in real time using exclusively the depth data. This paper points out that depth cameras aren't exactly new, but the Kinect is a low-cost, real-time, depth camera that is much more accessible. The accuracy of the Kinect is called into questions, the point cloud that the depth data creates does usually contain noise and sometimes has holes where no readings were obtained. This project also considered the Kinect's low X/Y resolution and depth accuracy and fixes the quality of the images using depth super resolution. KinectFusiont also looks into using multiple Kinects to preform a 3D body scan; this raises more issues because the quality of the overlapping sections of the images is compromised.

Another KinectFusion Project is the Real-time 3D Reconstruction and Interaction, this project is impressive because the entire process is done using a moving depth camera. With this software, the user can hold a Kinect camera up to a scene, and a 3D construction would be made. Not only would the user be able to see the 3D Reconstruction, but they would be able to interact with it; for instance, if they were to throw a handful of spheres onto the scene, they would land on the top appropriate surfaces, and fall under appropriate objects following the rules of physics. To accomplish this, the depth camera is used to track the 3D pose and the sensor is used to reconstruct the scene.

A study shown in the Asia Simulation Conference in 2011 demonstrated that a calibrated Kinect can be combined with Structure from Motion to find the 3D data of a scene and reconstruct the surface by Multiview Stereo. This study proved that the Kinect was more accurate for this procedure than a SwissRanger SR-4000 3D-TOF camera and close to a medium resolution SLR Stere rig. The Kinect works by using a near-infrared laser pattern projector and an IR camera as a stereo pair to triangulate points in 3D space, then the RGB camera is used to reconstruct the correct texture to the 3D points. This RGB camera, which outputs medium quality images, can also be used for recognition. One issue this study found was that the resulting IR and Depth images were shifted. To figure out what the shift was, the Kinect recorded pictures of a circle from different distances. The shift was found to be around 4 pixels in the u direction and three pixels in the v direction. Even after the camera has been totally calibrated, there are a few remaining residual errors in the close range 3D measurements. An easy fix for this error was to we form a z-correction image of z values constructed as the pixel-wise mean of all residual images and then subtract that correction image from the z coordinates of the 3D image. [?] Though the SLR Stereo was the most accurate, the error e (or the Euclidean distance between the points returned by the sensors and points reconstructed in the process of calibration) of the SR-400 was much higher than the Kinect and the SLR. This study shows that the Kinect is possible cheaper and simpler alternative to previously used cameras and rigs in the computer vision field.

Another subject of research that is looking into using the Kinect is the simultaneous localization and mapping algorithm, used to create a 3D map of the world so that the robot can avoid collision with obstacles or walls. The SLAM problem could be solved using GPS if the robot is outside, but inside one needs to use wheel or visual odometry. Visual odometry determines the position and the orientation of the robot using the associated camera images, algorithms like Scale Invariant Feature Transformation (SIFT), used to find the interest points, and laser sensors, used to collect depth data. Since the Kinect has both the RGB camera and a laser sensor, this piece of technology is a good piece of hardware to use for robots computing the SLAM Algorithm. In the study conducted by the students in the Graduates School of Science and Technology, at Meiji University, they found that the Kinect worked well for this process for horizontal and straight movement, but they had errors when they tried to recreate an earlier experiment, this means that their algorithm successfully solves the initial problem, but accuracy fell over time. [?] They found that the issue was not with the Kinect, and that it could be solved using the Speed-Up Robust Feature algorithm (SURF) and Smirnov-Grubbs test to further improve the accuracy of their SLAM Algorithm. This study proved that the Kinect was a reasonable, inexpensive and non-special piece of equipment that is capable of preforming well in computer vision applications.

7

References

[1] Helmut Grabner Xiaofeng Ren-Kurt Konolige Andrea Fossati, Juergen Gall, editor. *Consumer Depth Cameras for Computer Vision Research Topics and Applications.* Springer-Verlag London, 2013.

[2] Satoshi Tanaka Soo-Hyun Park Jong-Hyun Kim, Kangsun Lee. *Advanced Methods, Techniques, and Applications in Modeling and Simulation.* Springer Japan, 2012.

[3] Jing Tong, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan *Scanning 3D Full Human Bodies Using Kinects.* IEEE Transactions on Visualization and Computer Graphics, 2012

[4] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbo *KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera.* ACM Symposium on User Interface Software and Technology, 2011

8 Appendix