3D Room Mapping With a Hand-Held Time-of-Flight Camera

Tobias Grundström Anna Hjelmberg Mats Nilsson Fredrik Olsson

Linköping University

Introduction

The project goal was delivering a program which utilizes the Fotonic G-series depth camera to produce a 3D model of an indoor environment.

Problem Description

When capturing point clouds, every single scan will produce a point cloud which will have its own coordinate system, this system will differ from the global (reference) coordinate system, based on the location and direction of the camera. The problem of finding a transformation matrix which will transform all the point-clouds to the common global coordinate system is called *Point Cloud Registration*, which was the problem dealt with in this work.

Theory

Some brief explanations of methods and theory connected to this project:

Fast Point Feature Histogram: Algorithm that calculates a 3D feature space based on the relationships between neighboring points and their estimated surface normals.

Fast Global Registration: Algorithm that estimates the transformation that best aligns two point clouds, it uses FPFH features to find corresponding points.

Iterative Closest Point: Iterative method to minimize the difference between two point clouds, it uses the result from FGR as a starting point.

Flying pixels: Artifacts at depth edges which appears when a sensor captures reflecting signals from edges of a target object and the background.

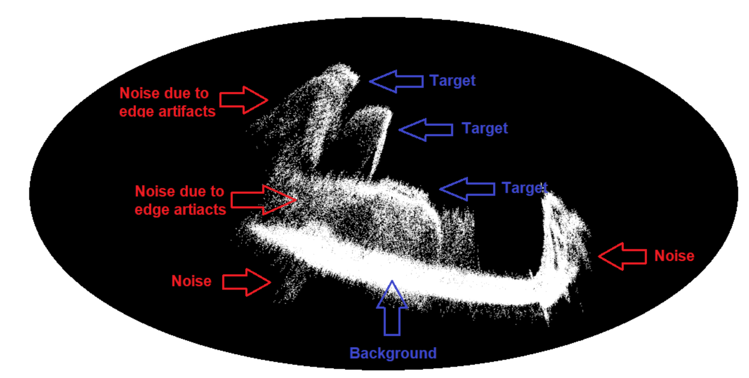
Multipath interference: Phenomenon which appears when a sensor detects waves via several paths.

Method

Data Collection

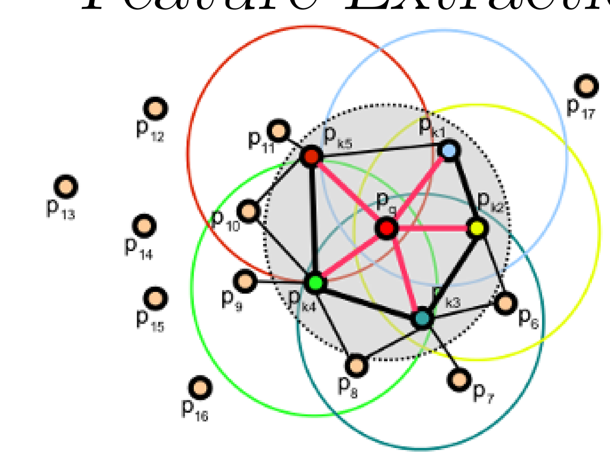
Images are captured using a Fotonic G-series depth camera (as seen in the figure). The data is then transformed into point clouds and downsampled using a voxel grid.

Outlier Removal



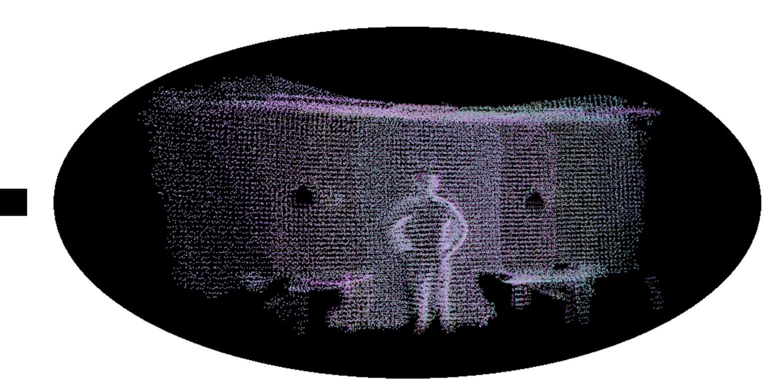
Outliers, such as flying pixels are removed using *Statistical Outlier Removal.*

Feature Extraction



The *Fast Point Feature Histograms* are calculated using point normals for each point.

Fast-Global-Registration

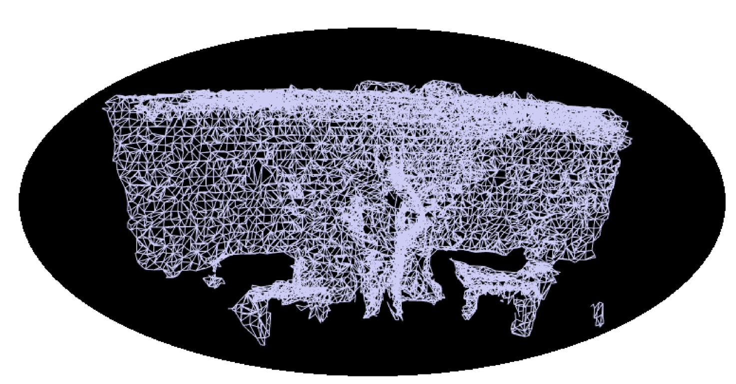


The registration is done using the *Fast Global Registration*, FGR, algorithm.

Iterative-Closest-Point

The transformation from FGR is then the initial value for the *Iterative Closest Point algorithm*, ICP, which tunes the registration.

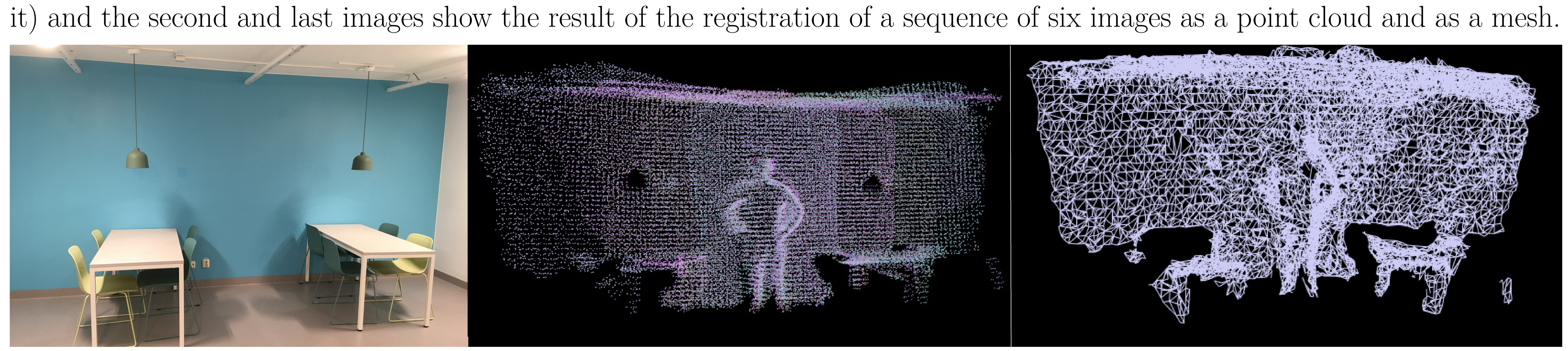
Mesh



When the registration is done a mesh is computed to visualize the result.

Results

Here follows the result from one of the scenes modeled. The first image shows the scene (with exception of the human standing in it) and the second and last images show the result of the registration of a sequence of six images as a point cloud and as a mesh.



Discussion

It is clear to see that there is a problem with multipath interference in our results, especially obvious at corners.

Another problem we came across was to find enough correspondences to get a satisfying registration when removing flying pixels. This is because many of the correspondences are found at these points even though they are fake.

We found that some environments were more difficult to model than others. The multipath interference error became more obvious and lead to a poorer result in environments where many edges and corners were present. On the other hand, there is important with objects and structures in the scene so that the FGR algorithm find enough correspondences between the image pairs. Some darker environments and some materials were more difficult to capture.

The overlap between the images was also important. The more overlap between the images in the sequence the better result of the registration.

Conclusion

This goal has been achieved but the result varies depending on what scene is captured and how it is captured. As said in the discussion section there are some factors that are important to get a good result, especially **the amount of overlap between each image pair** and **the physical environment of the scene**. Regarding the overlap between the images in the sequence, the conclusion is that the more overlap there is between each image pair, the better the registration. The physical environment is a somewhat more difficult factor. This includes corners, objects, materials, lighting and color variations between scenes.