Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Computer Vision and Geometry Lab

Computer Vision Lecture

Assignment 4 Report

Autumn Term 2022

Author:
Pascal Lieberherr

Contents

Chapter 1

Report

1.1 RANSAC

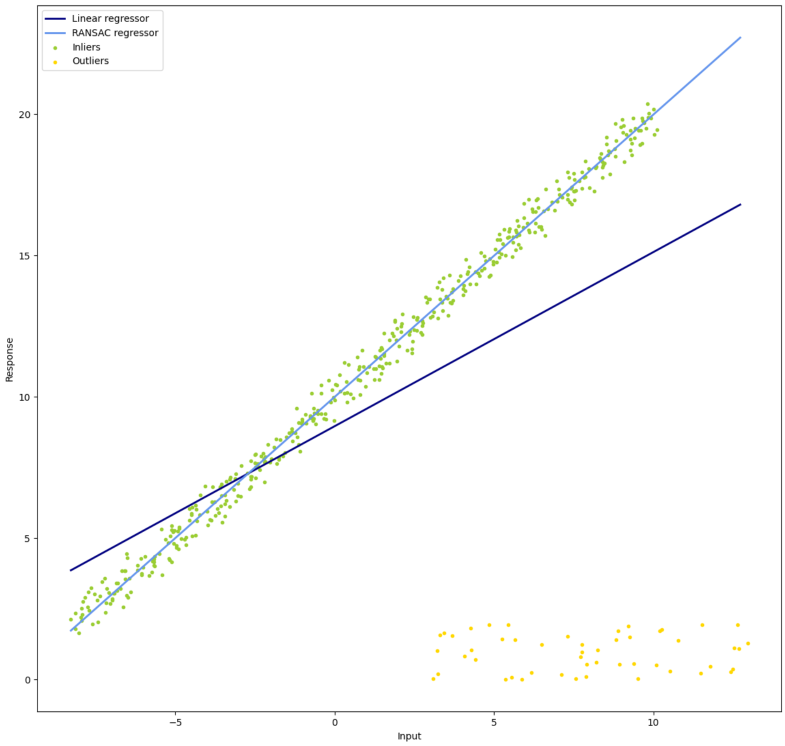

Figure 1.1: Results visualized.

k ground truth: 1
b ground truth: 10
k least squares: 0.6159656578755459
b least squares: 8.96172714144364

k ransac: 0.9987449792570882
b ransac: 9.997009758791009

1.2 Multi-View Stereo

Note: I only trained for 3 epochs instead of 4 due to time constraints. I resumed the training twice. Please find more tensor board screenshots in the folder images. I only inserted the one for the first epoch here in the report.

Equation of corresponding pixel

For the equation see fig. 1.2. The detailed derivation can be found in the appendix.

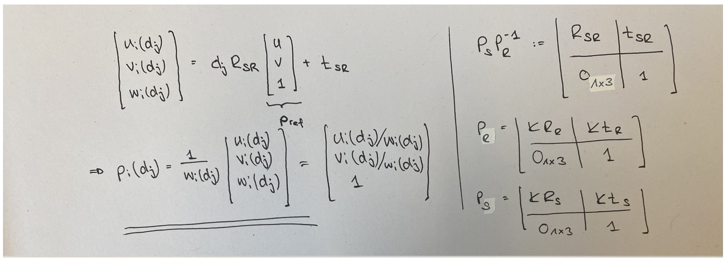

Figure 1.2: The projection equation

Tensorboard

Figure 1.3: The tensorboard for test and train. For a higher resolution find the image in the folder images.

Geometric Consistency

The geometric consistency filtering detects outliers and returns a mask that filters out the outliers. The geometric consistency gets the depth estimates for the source and the reference view plus the extrinsics and intrinsics for both the ref and src view. Based on this information it first projects a pixel in the ref view to the 3D point with the current depth estimate. This 3D point is then projected onto the image plance of the src view. From there it's again projecte to 3D but with the dept estimate of the src view. Now we again project this 3D point to the ref view. What we get out of this is the dist between the original pixel in the reference view and the reprojected pixel. In addition we also have a new depth information for the original pixel in the ref view named depth reprojected. Provided both, the distance between original and reprojected pixel and the difference of the depth ref and reprojected depth are small enough, the depth estimate is considered to be an inlied. The function returns a mask to filter out the outliers. The fig fig. 1.4 illustrates all projections.

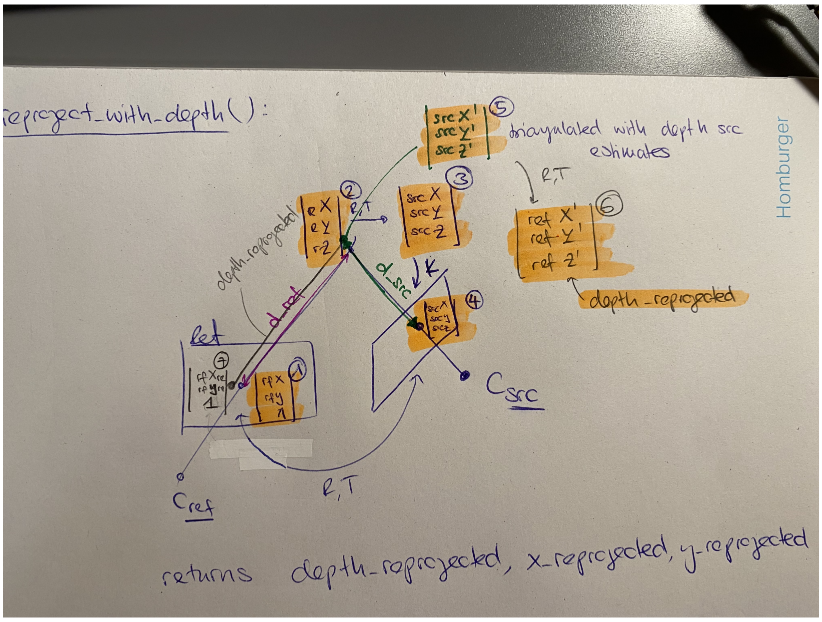

Figure 1.4: Reprojection with depth

Inverse Depth Sampling

Inverse depth sampling is more suitable for large-scale scenes. Because even when the depth samples are uniformly distributed (in the uniform sampling case) along the depth dimension they are not uniformly distributed along the epipolar line in the source view. Thus, when we sample from the inverse range the points far away are less often sampled than the close points. This is supposed to lead to a more uniform distribution along the epipolar line in the source view. The drawing below illustrates the uneven distribution along the epiploar line for uniform sampling.

Robust to Occlusions

No, simply taking the average is not a robust method in presence of occlusions. It could be that in one view sth. is occluded. This would negatively affect the matching similarity.

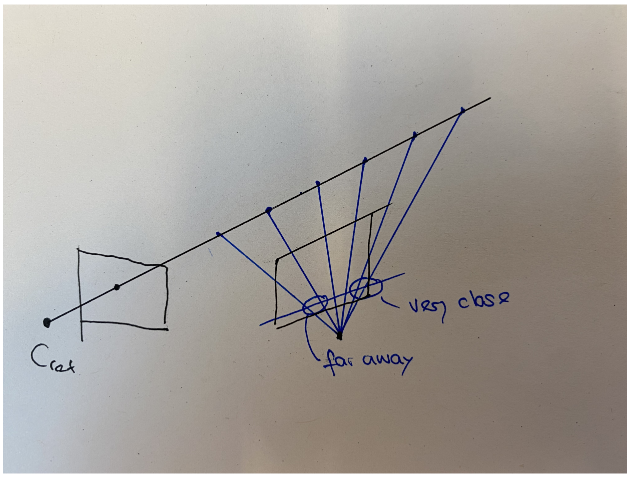

Figure 1.5: Uniform sampling effect.

MVS Results: 3D Point Clouds

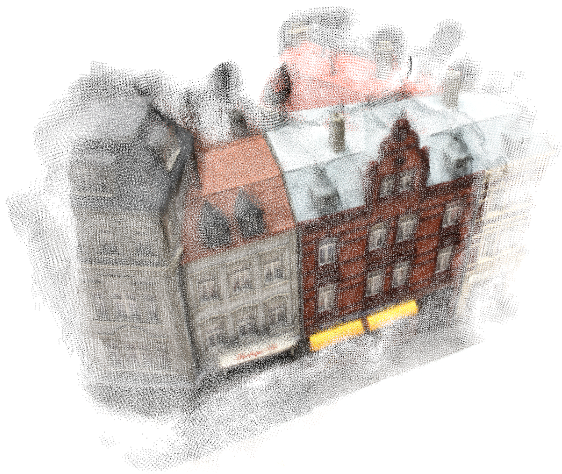

Figure 1.6: mvs point cloud visualization 001.

Figure 1.7: mvs point cloud visualization 009.

1.3 Appendix

1.3.1 Derivation of projection from ref view to src view

1.) transform each pixel p_i in ref view to src view

1.) Project image coordin. in ref view to 3D

$$\begin{bmatrix}\lambda u_s\\ \lambda v_s\\ \lambda\cdot 1\\ 1\end{bmatrix} = P_R \begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix} \quad \text{with} \quad \lambda = d_i$$

$$\Rightarrow P_R^{-1}\begin{bmatrix}u_s\cdot\lambda\\ v_s\cdot\lambda\\ 1\cdot\lambda\\ 1\end{bmatrix} = \begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

$$P_R^{-1}\begin{bmatrix}du_s\\ dv_s\\ d\\ 1\end{bmatrix} = \begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

2.) project 3D point to source view

$$\begin{bmatrix}u_R\\ v_R\\ w_R\\ 1\end{bmatrix} = P_S\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}$$

insert result from 2.)

$$\Rightarrow \begin{bmatrix}u_R\\ v_R\\ w_R\\ 1\end{bmatrix} = \underbrace{P_S P_R^{-1}}_{P_{SR}}\begin{bmatrix}du_s\\ dv_s\\ d\\ 1\end{bmatrix}$$

$$\hookrightarrow \left[\begin{array}{c|c} R_{SR} & t_{SR}\\ \hline 0 & 1\end{array}\right]$$

$$\Rightarrow \begin{bmatrix}u_R\\ v_R\\ w_R\end{bmatrix} = R_{SR}\begin{bmatrix}du_s\\ dv_s\\ d\end{bmatrix} + t_{SR}$$

$$= \begin{bmatrix}u_R\\ v_R\\ w_R\end{bmatrix} = d\cdot R_{SR}\begin{bmatrix}u_s\\ v_s\\ 1\end{bmatrix} + t_{SR}$$

$\Rightarrow$ normalize to $w_R = 1$

$$\Rightarrow \begin{bmatrix}u_R/w_R\\ v_R/w_R\\ 1\end{bmatrix}$$

normalize to $[-1, 1]$

$$\begin{bmatrix}\bar{u}_R\\ \bar{v}_R\\ 1\end{bmatrix} = \begin{bmatrix}w/2 & 0 & -1\\ 0 & h/2 & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}u_R/w_R\\ v_R/w_R\\ 1\end{bmatrix}$$

Figure 1.8: Derivation of the projection of a pixel from reference view to source view.