Introduction to Computer Vision Course

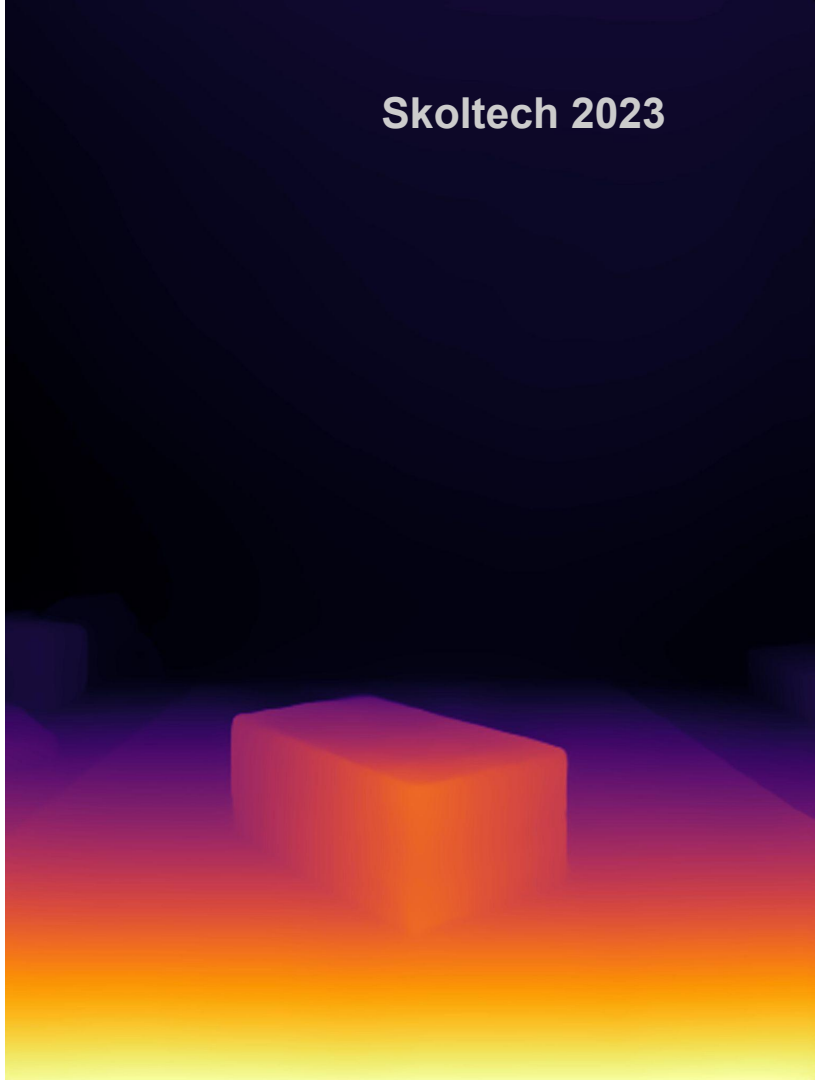
Skoltech 2023

OBJECT SIZE RECOGNITION

TEAM

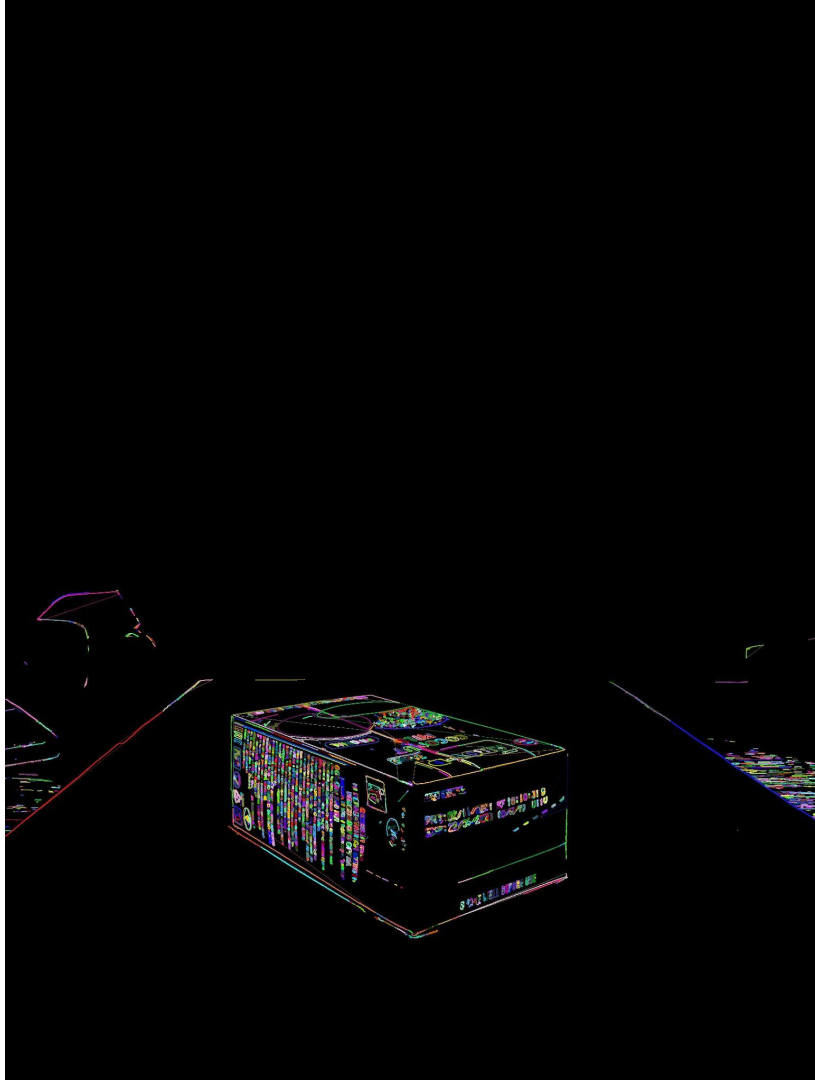
Alexander Tutunin
MSc-1 Energy Systems

Anuar Daurenov
MSc-1 Energy Systems



Contents

- Problem statement
- Our pipeline
- Our methods
- Our data
- Our results



Problem statement

Manual sorting of packages by size

Impact:

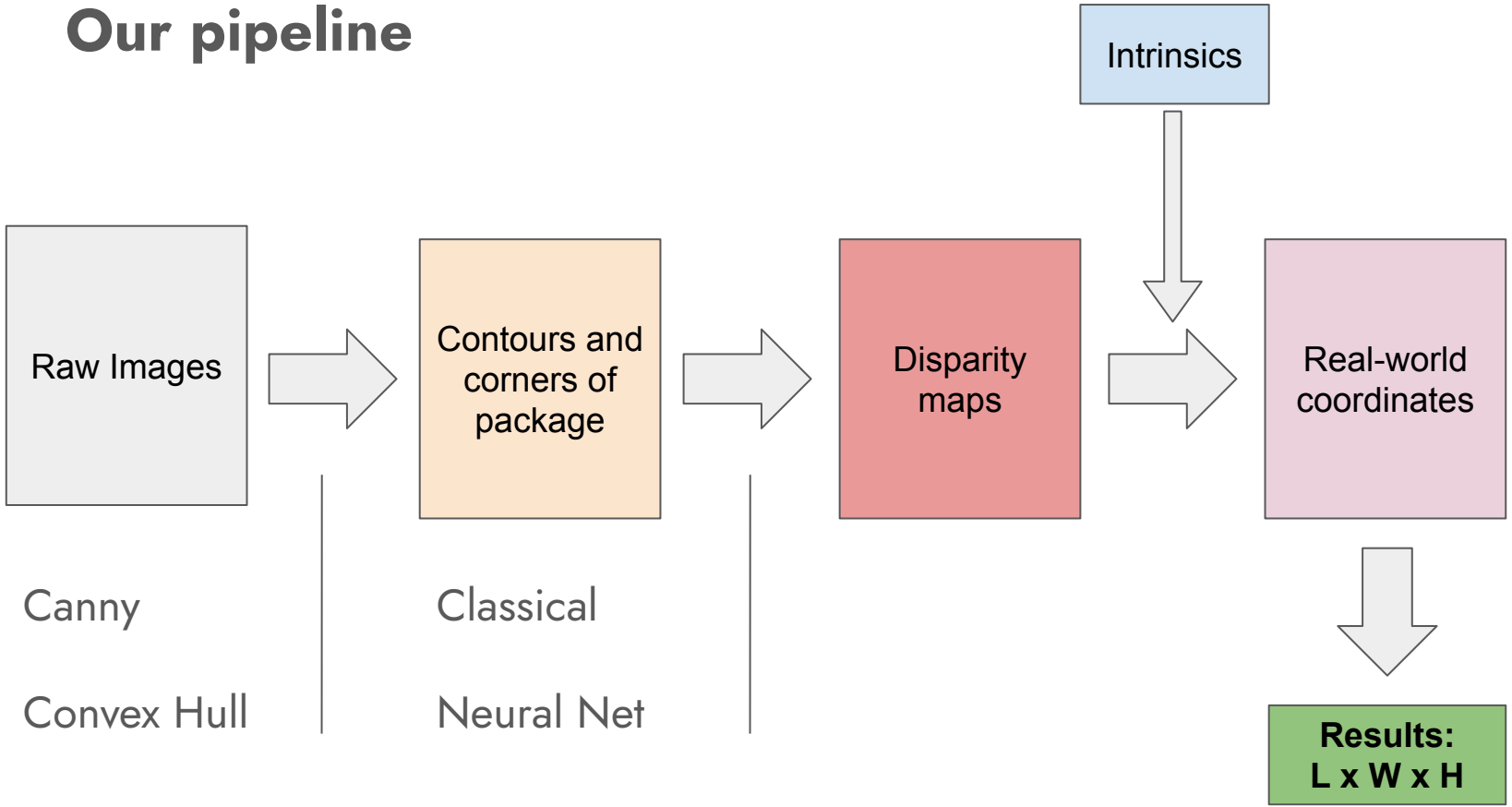
- Money loss (salary)
- Time loss (low-efficient work)

Who suffers from?

- Logistics companies
- Post Offices

Our solution: Automated system for automated package size recognition

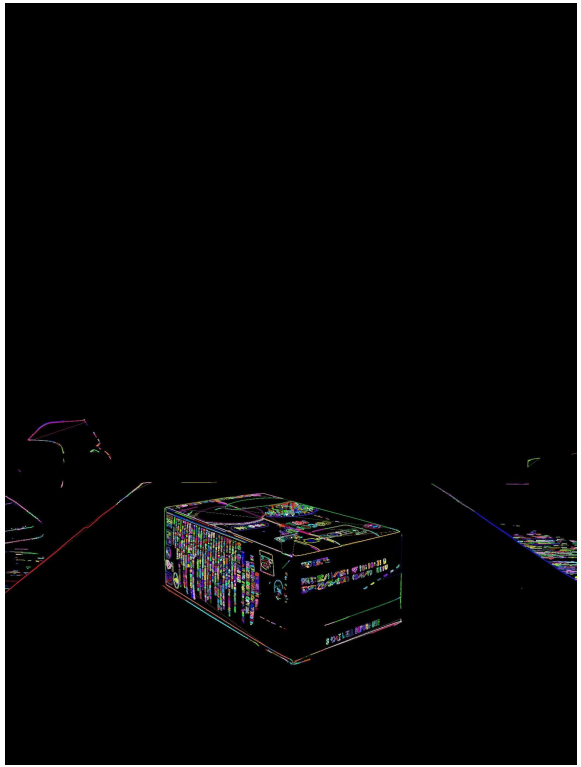
Our pipeline



Our methods

Corners detection:

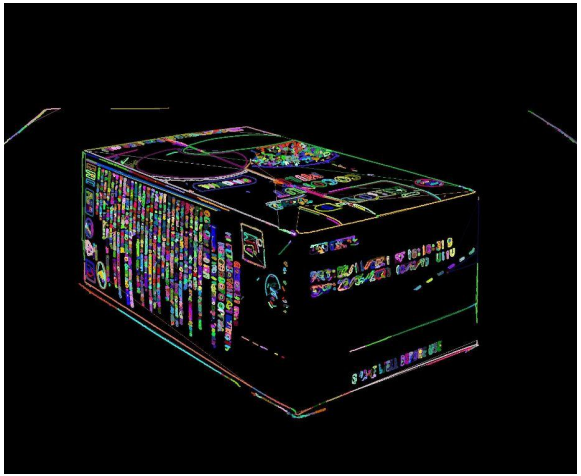
- **Canny** was used to detect all edges and contours

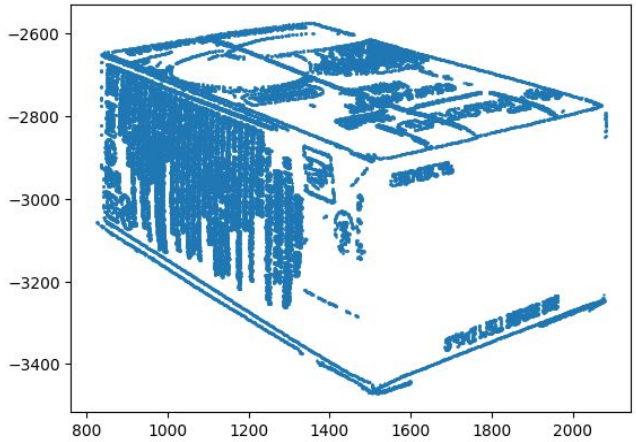


Our methods

Corners detection:

- **Canny** was used to detect all edges and contours
- **DBScan** was used for clusterization



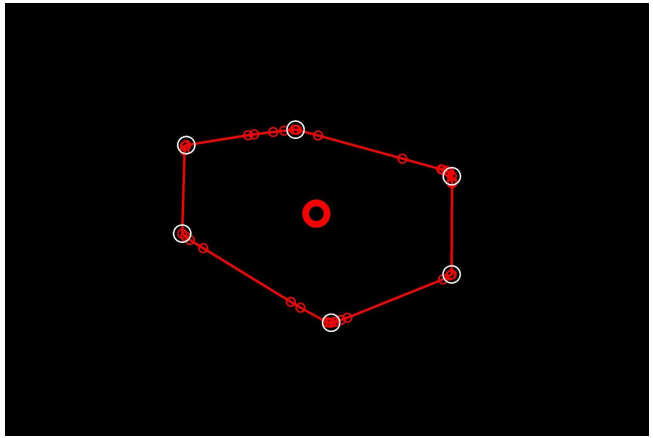


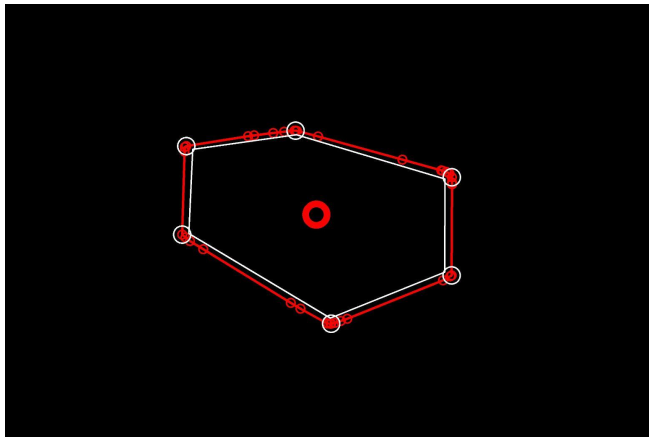
Our methods

Corners detection:

- **Canny** was used to detect all edges and contours
- **DBScan** was used for clusterization
- **cv2.convexHull** was used to obtain convex hulls

After this: get corners





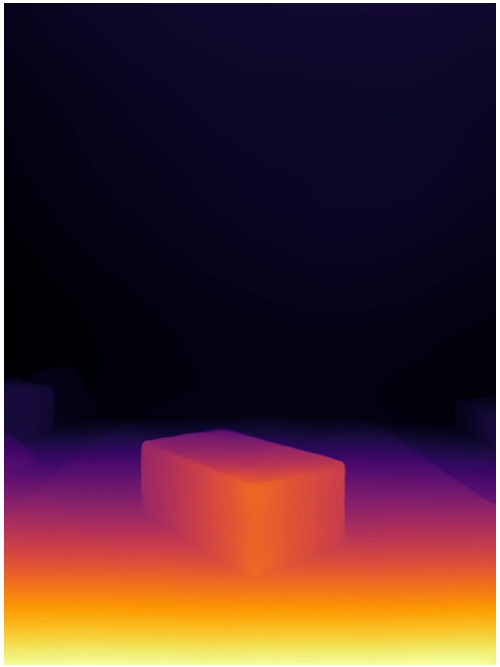
Our methods

Depth Estimation:

- cv2.StereoBM as classical
- MiDaS as NN

They give **disparity**

StereoBM
2 images are required



MiDaS
Only 1 image is required

Our methods

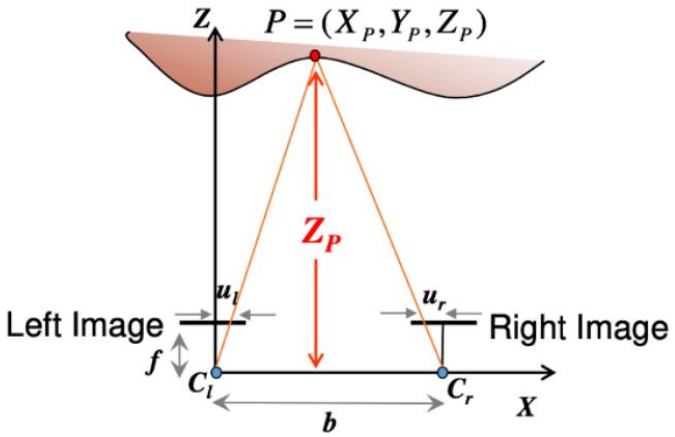
Depth Estimation:

- **cv2.StereoBM as classical**
- **MiDaS as NN**

They give disparity

To convert from disparity to depth:

- Calculate *f* - obtain camera matrix
- Measure *b*:
  our case *b = 16* mm



Baseline = distance between the optical centers of the two cameras

$$\frac{f}{Z_P} = \frac{u_l}{X_P}$$

$$\frac{f}{Z_P} = \frac{-u_r}{b - X_P}$$

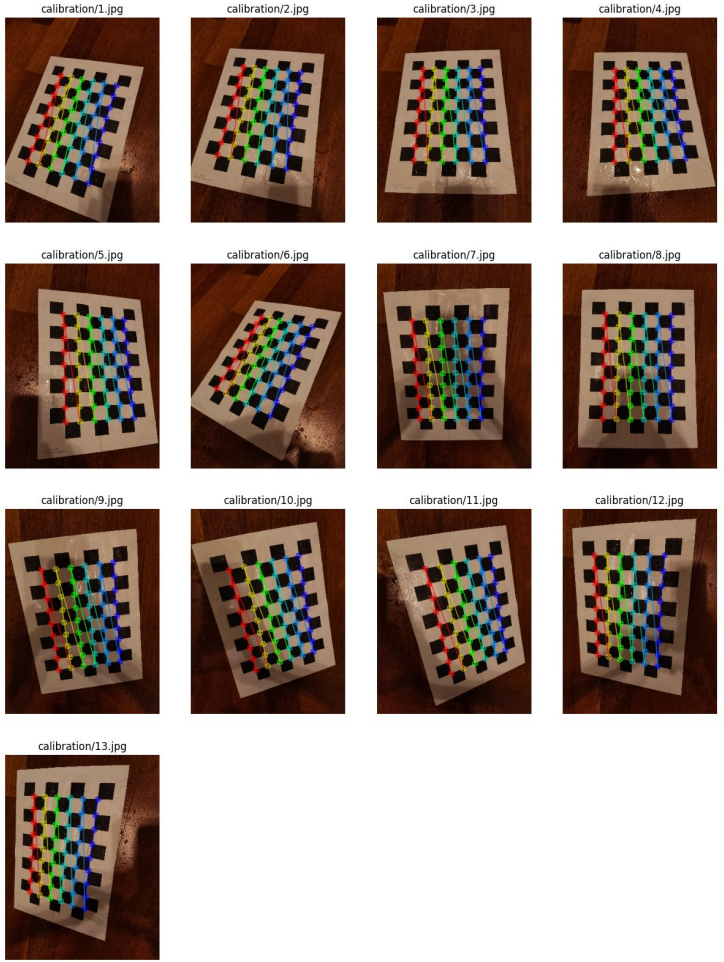
$$Z_P = \frac{bf}{u_l - u_r}$$

Disparity
difference in image location of the projection of a 3D point on two image planes

Our methods

Get Intrinsics:

- cv2.calibrateCamera to obtain camera intrinsics
- Chessboard procedure was used

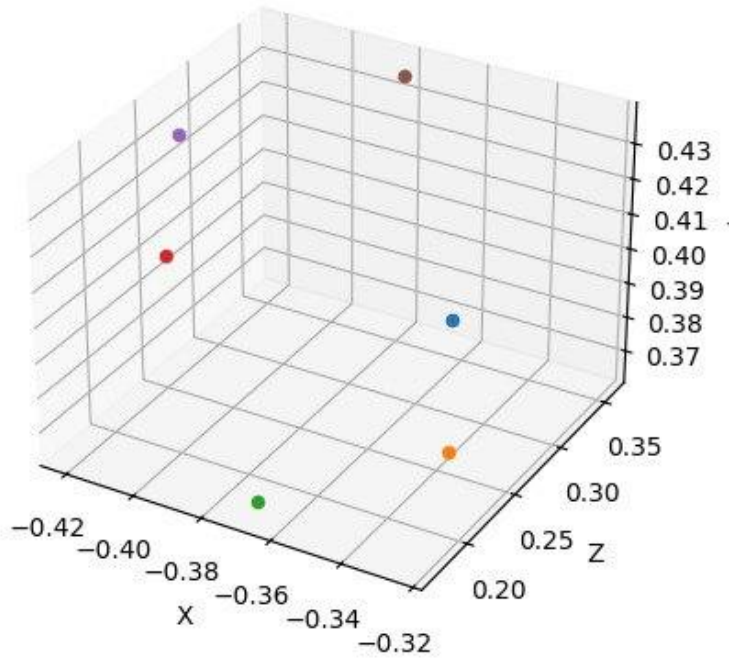


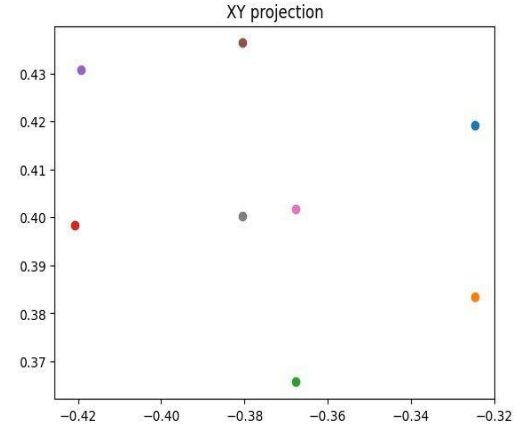
Our data

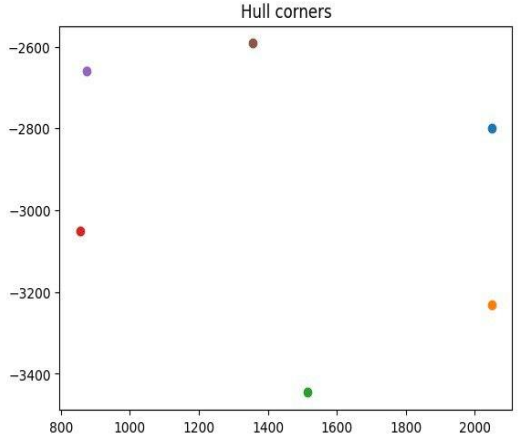
StereoBM as triangulation requires at least 2 images

Our dataset is small and contains 2 images that was taken as stereo camera sees

Our results


0.43
0.42
0.41
0.40
0.39
0.38
0.37
0.35
0.30
0.25
Z
0.20
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32
X


XY projection
0.43
0.42
0.41
0.40
0.39
0.38
0.37
−0.42
−0.40
−0.38
−0.36
−0.34
−0.32


Hull corners
−2600
−2800
−3000
−3200
−3400
800
1000
1200
1400
1600
1800
2000

Our results

```
coords_rw:
[[-0.3246397  -0.41929248  0.2071523 ]
 [-0.3247203  -0.3833839   0.20475729]
 [-0.3676742  -0.36567783  0.16828957]
 [-0.4207823  -0.39827687  0.2537193 ]
 [-0.4193317  -0.43071043  0.26418027]
 [-0.38040724 -0.43633667  0.35872453]]
distances:
[0.035988454, 0.059062995, 0.105742164, 0.034109715, 0.102398165]
```

	Ours	Truth
Width	0.0360	0.040
Length	0.0590	0.060
Height	0.1057	0.105