



SCHOOL OF COMPUTER AND
COMMUNICATION SCIENCES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Computer Vision Laboratory
Unseen Spacecraft Pose Estimation

Baseline solution by implementing a machine learning
framework with target models included

Bachelor's Thesis in Computer Science

Author: JÉRÉMY CHAVEROT
Supervisor: Dr. MATHIEU SALZMANN
Advisors: Dr. ANDREW PRICE, PhD. CHEN ZHAO
Semester: Fall 2023

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the advisors.

Lausanne, Switzerland, 05.01.24

JÉRÉMY CHAVEROT

Acknowledgments

Before delving into the topic, I'd like to express some acknowledgments. First and foremost I must thank my advisors Andrew Price and Chen Zhao for accepting my request to take part in a semester project under their guidance. I am grateful to have been able to practice my skills with them, and can only hope that the feeling is mutual. Moreover I would also like to thoroughly thank my friends and family for supporting me in my academic journey, despite a rather unstable start in my studies.

Abstract

Contents

Acknowledgments	ii
Abstract	iii
1. Introduction	1
1.1. Problem Formulation	1
1.1.1. The settings	1
1.1.2. The goal	1
1.2. The work environment: Scitas Izar	1
2. Gen6D: Formal Description	2
2.1. Overview of the Network	2
2.2. Detection	2
2.3. Viewpoint Selection	2
2.4. Pose Refinement	2
3. Implementation of the model	3
3.1. Data Loader	3
3.2. From Quaternions to Rotation Matrices	3
4. Experimental Results and Analysis	6
4.1. Reference and Query Images	6
4.2. Evaluation Metrics	6
4.3. Vizualisation and Quantitative Evaluation	6
5. Ways of improvements	13
5.1. Specialized spacecraft training set	13
5.2. Improved object detection algorithms	13
5.3. Robustness to occlusion	13
6. Conclusion	14
Abbreviations	15
A. Python Scripts	16
B. Scitas Izar Setup Tutorial	21
Bibliography	22

1. Introduction

Test ref to Listing A.1. Test ref to Listing A.3 Test ref to Gen6D [1]

1.1. Problem Formulation

1.1.1. The settings

1.1.2. The goal

1.2. The work environment: Scitas Izar

Test to refer to the video from 3B1B: [2].

```
1 #!/bin/bash
2 #SBATCH --chdir /scratch/izar/jchavero
3 #SBATCH --partition=gpu
4 #SBATCH --qos=gpu_free
5 #SBATCH --gres=gpu:2
6 #SBATCH --nodes=1
7 #SBATCH --ntasks-per-node=1
8 #SBATCH --cpus-per-task=1
9 #SBATCH --mem 16G
10
11 echo STARTING AT 'date'
12
13 echo "Loading modules"
14 module load gcc openmpi py-torch py-torchvision cuda
15
16 echo "Launching the virtual environment"
17 source ~/opt/izar1/venv-gcc/bin/activate
18
19 echo "Navigating to the directory and executing the task"
20 cd ~/Gen6D
21 python eval.py --cfg configs/gen6d_pretrain.yaml --object_name spacecraft/hubble
22
23 echo FINISHED AT 'date'
```

Listing 1.1: Bash script `execute.sh` to run a machine learning model on Scitas Izar EPFL. While the overall structure remains consistent, this script is specific to Gen6D's architecture, further discussed later.

Then to run the script we use the following command:

```
1 $ sbatch execute.sh
```

Listing 1.2: Linux command to run the bash script.

2. Gen6D: Formal Description

2.1. Overview of the Network

2.2. Detection

2.3. Viewpoint Selection

2.4. Pose Refinement

3. Implementation of the model

3.1. Data Loader

abstract base classes (ABC) each and every abstract method

3.2. From Quaternions to Rotation Matrices

In Gen6D, the model represents the ground truth and estimated poses using the format $\mathbf{P} = (\mathbf{R}, \mathbf{t})$. Here, \mathbf{R} denotes the rotation matrix and \mathbf{t} is the translation vector. The thing is, in the SPACECRAFT dataset, the poses have the format $\mathbf{P} = (\mathbf{q}, \mathbf{t})$, where \mathbf{q} is a *quaternion*. We use quaternions for three-dimensional rotation calculations because they offer several benefits over rotation matrices. Notably, quaternions are more compact, requiring only four elements to be stored compared to nine for a matrix. Additionally, they are more efficient when composing rotations thanks to their algebraic properties.

Despite the benefits of quaternions mentioned earlier, other datasets frequently represent poses using a combination of rotation matrices and translation vectors. Therefore, to align with Gen6D's pose format, this section will focus on converting quaternions into rotation matrices [4].

Quaternions were first introduced by the Irish mathematician W. R. Hamilton in 1843 as an extension of the complex numbers [3]. In the first place, we provide the definition of a *quaternion*: it is the sum of a scalar q_0 and a vector $\mathbf{q} = (q_1, q_2, q_3)$, that is,

$$\underline{\mathbf{q}} \stackrel{\text{def.}}{=} q_0 + \mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}.$$

In the above, \mathbf{i} , \mathbf{j} and \mathbf{k} denote the three unit vectors of the canonical basis for the set of all ordered triples of real numbers \mathbb{R}^3 . The set of quaternions is denoted by the 4-space \mathbb{H} .

The quaternion addition is component-wise. Regarding the product of two quaternions, it is essential to first outline the foundational rule established by Hamilton:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} = -1.$$

We derive the following multiplication table:

\times	1	\mathbf{i}	\mathbf{j}	\mathbf{k}
1	1	\mathbf{i}	\mathbf{j}	\mathbf{k}
\mathbf{i}	\mathbf{i}	-1	\mathbf{k}	$-\mathbf{j}$
\mathbf{j}	\mathbf{j}	$-\mathbf{k}$	-1	\mathbf{i}
\mathbf{k}	\mathbf{k}	\mathbf{j}	$-\mathbf{i}$	-1

Let $(\underline{\mathbf{p}}, \underline{\mathbf{q}}) \in \mathbb{H}^2$, we are now able to present the multiplication of $\underline{\mathbf{p}}$ and $\underline{\mathbf{q}}$:

$$\underline{\mathbf{p}}\underline{\mathbf{q}} = \underbrace{p_0q_0 - \mathbf{p} \cdot \mathbf{q}}_{\text{scalar part}} + \underbrace{p_0\mathbf{q} + q_0\mathbf{p} + \mathbf{p} \times \mathbf{q}}_{\text{vector part}}.$$

In the preceding, we recall that the binary operation $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} : (\mathbf{p}, \mathbf{q}) \mapsto \mathbf{p} \cdot \mathbf{q} \stackrel{\text{def.}}{=} \mathbf{p}^\top \mathbf{q}$ is the *dot product*, and the other

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 : (\mathbf{p}, \mathbf{q}) \mapsto \mathbf{p} \times \mathbf{q} \stackrel{\text{def.}}{=} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \end{vmatrix} = [\mathbf{p}]_{\times} \mathbf{q}$$

is the *cross product*, and the operator $[\mathbf{p}]_{\times}$ yields the transformation matrix that when multiplied from the right with a vector \mathbf{q} gives $\mathbf{p} \times \mathbf{q}$.

Several additional definitions are essential before we can begin the conversion of quaternions to rotation matrices.

Let $\underline{\mathbf{q}} = q_0 + \mathbf{q}$ be a quaternion. The *complex conjugate* of $\underline{\mathbf{q}}$, denoted by $\underline{\mathbf{q}}^*$, is given by the map

$$\mathbb{H} \rightarrow \mathbb{H} : \underline{\mathbf{q}} \mapsto \underline{\mathbf{q}}^* \stackrel{\text{def.}}{=} q_0 - \mathbf{q}.$$

The *norm* of a quaternion $\underline{\mathbf{q}}$, denoted $|\underline{\mathbf{q}}|$, is the distance obtained from the map

$$\mathbb{H} \rightarrow \mathbb{R}^+ : \underline{\mathbf{q}} \mapsto |\underline{\mathbf{q}}| \stackrel{\text{def.}}{=} \sqrt{\underline{\mathbf{q}}^* \underline{\mathbf{q}}}.$$

Note that a quaternion whose norm is 1 is referred to as a *unit quaternion*. The *reciprocal* of a quaternion is defined as the map

$$\mathbb{H}^* \rightarrow \mathbb{H}^* : \underline{\mathbf{q}} \mapsto \underline{\mathbf{q}}^{-1} \stackrel{\text{def.}}{=} \frac{\underline{\mathbf{q}}^*}{|\underline{\mathbf{q}}|^2}.$$

We observe that if $\underline{\mathbf{q}}$ is a unit quaternion, we simply have $\underline{\mathbf{q}}^{-1} = \underline{\mathbf{q}}^*$. Furthermore, the subsequent proposition is accepted: if $\underline{\mathbf{q}}$ is a unit quaternion, there exists a unique $\theta \in [0, 2\pi]$ such that

$$\underline{\mathbf{q}} = q_0 + \mathbf{q} = \cos \frac{\theta}{2} + \mathbf{u} \sin \frac{\theta}{2},$$

where the unit vector \mathbf{u} is defined as $\mathbf{u} \stackrel{\text{def.}}{=} \frac{\mathbf{q}}{|\mathbf{q}|}$.

Quaternion Rotation Operator Let $\underline{\mathbf{q}} \in \mathbb{H}$ be a unit quaternion, and let $\mathbf{v} \in \mathbb{R}^3$ be a vector. The action of the $L_{\underline{\mathbf{q}}}$ function

$$\mathbb{R}^3 \rightarrow \mathbb{R}^3 : \mathbf{v} \mapsto L_{\underline{\mathbf{q}}}(\mathbf{v}) \stackrel{\text{def.}}{=} \underline{\mathbf{q}}\mathbf{v}\underline{\mathbf{q}}^*$$

on \mathbf{v} is equivalent to a rotation of the vector through an angle θ about the axis of rotation \mathbf{u} .

At last, we can proceed with our conversion task: our aim is to find a 3×3 rotation matrix \mathbf{R} , such that

$$\begin{cases} L_{\mathbf{R}}(\mathbf{v}) \stackrel{\text{def.}}{=} \mathbf{R}\mathbf{v} \\ L_{\mathbf{R}}(\mathbf{v}) = L_{\underline{\mathbf{q}}}(\mathbf{v}), \end{cases}$$

which means we wish to obtain an expression for \mathbf{R} by manipulating $L_{\underline{\mathbf{q}}}$, utilizing principles of linear algebra and vector calculus. We consider the vector \mathbf{v} as a quaternion with a zero scalar component:

$$\begin{aligned} L_{\underline{\mathbf{q}}}(\mathbf{v}) &= \underline{\mathbf{q}}\mathbf{v}\underline{\mathbf{q}}^* \\ &= (q_0 + \mathbf{q})(0 + \mathbf{v})(q_0 - \mathbf{q}) \\ &= (\underbrace{-\mathbf{q} \cdot \mathbf{v}}_{\text{scalar part}} + \underbrace{q_0\mathbf{v} + \mathbf{q} \times \mathbf{v}}_{\text{vector part}})(q_0 - \mathbf{q}) \\ &= q_0(-\mathbf{q} \cdot \mathbf{v}) - (q_0\mathbf{v} + \mathbf{q} \times \mathbf{v}) \cdot (-\mathbf{q}) \\ &\quad + (-\mathbf{q} \cdot \mathbf{v})(-\mathbf{q}) + q_0(q_0\mathbf{v} + \mathbf{q} \times \mathbf{v}) \\ &\quad + (q_0\mathbf{v} + \mathbf{q} \times \mathbf{v}) \times (-\mathbf{q}) \\ &= \underbrace{-q_0(\mathbf{q} \cdot \mathbf{v}) + q_0(\mathbf{q} \cdot \mathbf{v})}_{\text{scalar part}} \\ &\quad + \underbrace{\mathbf{q}(\mathbf{q} \cdot \mathbf{v}) + q_0^2\mathbf{v} + q_0(\mathbf{q} \times \mathbf{v}) + \mathbf{q} \times (q_0\mathbf{v} + \mathbf{q} \times \mathbf{v})}_{\text{vector part}} \\ &= \mathbf{q}(\mathbf{q}^T \mathbf{v}) + q_0^2\mathbf{v} + q_0(\mathbf{q} \times \mathbf{v}) + \mathbf{q} \times (q_0\mathbf{v} + \mathbf{q} \times \mathbf{v}) \\ &= (\mathbf{q} \otimes \mathbf{q} + q_0^2 \mathbf{I}_{3 \times 3} + 2q_0 [\mathbf{q}]_{\times} + [\mathbf{q}]_{\times}^2) \mathbf{v} \end{aligned}$$

In the above, \otimes stands for the *outer product* and $\mathbf{I}_{3 \times 3}$ is the *identity matrix*. Since $L_{\mathbf{R}}(\mathbf{v}) = L_{\underline{\mathbf{q}}}(\mathbf{v})$, we can identify \mathbf{R} as $(\mathbf{q} \otimes \mathbf{q} + q_0^2 \mathbf{I}_{3 \times 3} + 2q_0 [\mathbf{q}]_{\times} + [\mathbf{q}]_{\times}^2)$. We develop and simplify \mathbf{R} to find its final expression:

$$\begin{aligned} \mathbf{R} &= \mathbf{q} \otimes \mathbf{q} + q_0^2 \mathbf{I}_{3 \times 3} + 2q_0 [\mathbf{q}]_{\times} + [\mathbf{q}]_{\times}^2 \\ &= \begin{bmatrix} q_1^2 & q_1 q_2 & q_1 q_3 \\ q_2 q_1 & q_2^2 & q_2 q_3 \\ q_3 q_1 & q_3 q_2 & q_3^2 \end{bmatrix} + q_0^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + 2q_0 \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \\ &= 2 \begin{bmatrix} (q_0^2 + q_1^2) - \frac{1}{2} & q_1 q_2 - q_0 q_3 & q_1 q_3 + q_0 q_2 \\ q_1 q_2 + q_0 q_3 & (q_0^2 + q_2^2) - \frac{1}{2} & q_2 q_3 - q_0 q_1 \\ q_1 q_3 - q_0 q_2 & q_2 q_3 + q_0 q_1 & (q_0^2 + q_3^2) - \frac{1}{2} \end{bmatrix} \end{aligned}$$

This completes our derivation of the rotation matrix \mathbf{R} . The Python code can be found in Listing A.3.

4. Experimental Results and Analysis

4.1. Reference and Query Images

4.2. Evaluation Metrics

To appreciate the quality of the estimations, the most widely used pose error functions are the Average Distance of Model Points (ADD) and the Average Closest Point Distance (ADD-S) metrics, both introduced by Hinterstoisser et al. [5]. For an object model \mathcal{M} , we compute the average distance to the corresponding model point. Therefore the error of an estimated pose $\hat{\mathbf{P}} = (\hat{\mathbf{R}}, \hat{\mathbf{T}})$ w.r.t. the ground truth pose $\bar{\mathbf{P}} = (\bar{\mathbf{R}}, \bar{\mathbf{T}})$ is calculated as follows:

$$^1e_{\text{ADD}}(\hat{\mathbf{P}}, \bar{\mathbf{P}}, \mathcal{M}) \stackrel{\text{def.}}{=} \text{avg}_{\mathbf{x} \in \mathcal{M}} \left\| \bar{\mathbf{P}}\mathbf{x}^* - \hat{\mathbf{P}}\mathbf{x}^* \right\|_2 \quad (4.1)$$

$$= \text{avg}_{\mathbf{x} \in \mathcal{M}} \left\| (\bar{\mathbf{R}}\mathbf{x} + \bar{\mathbf{T}}) - (\hat{\mathbf{R}}\mathbf{x} + \hat{\mathbf{T}}) \right\|_2 \quad (4.2)$$

When the model \mathcal{M} has symmetries that leads to indistinguishable views, the error is computed as the average distance to the closest model point:

$$e_{\text{ADD-S}}(\hat{\mathbf{P}}, \bar{\mathbf{P}}, \mathcal{M}) \stackrel{\text{def.}}{=} \text{avg}_{\mathbf{x}_1 \in \mathcal{M}} \min_{\mathbf{x}_2 \in \mathcal{M}} \left\| \bar{\mathbf{P}}\mathbf{x}_1^* - \hat{\mathbf{P}}\mathbf{x}_2^* \right\|_2 \quad (4.3)$$

$$= \text{avg}_{\mathbf{x}_1 \in \mathcal{M}} \min_{\mathbf{x}_2 \in \mathcal{M}} \left\| (\bar{\mathbf{R}}\mathbf{x}_1 + \bar{\mathbf{T}}) - (\hat{\mathbf{R}}\mathbf{x}_2 + \hat{\mathbf{T}}) \right\|_2 \quad (4.4)$$

It's important to point out that $e_{\text{ADD-S}}$ is more lenient compared to e_{ADD} , and should only be applied in cases where there is a definite presence of symmetry in the object and the estimated pose is already notably precise. Otherwise, using $e_{\text{ADD-S}}$ becomes irrelevant since the estimation would be unfairly advantaged. In the illustrations below, we consistently provide both metrics, however it is up to the reader to assess the relevance of $e_{\text{ADD-S}}$ based on the observed spacecraft.

4.3. Vizualisation and Quantitative Evaluation

¹In this context, the vector \mathbf{x}^* represents a vector that has been extended by appending a 1, specifically for the purpose of matrix multiplication.

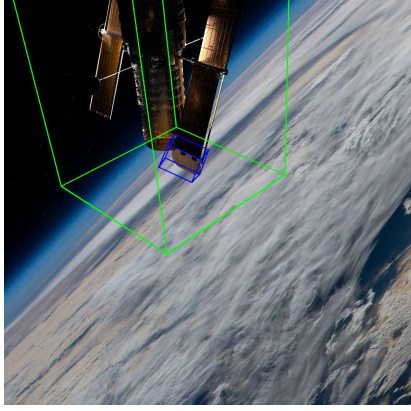


Figure 4.1: Hubble Space Telescope with earth rendered background, 1024x1024 first query image

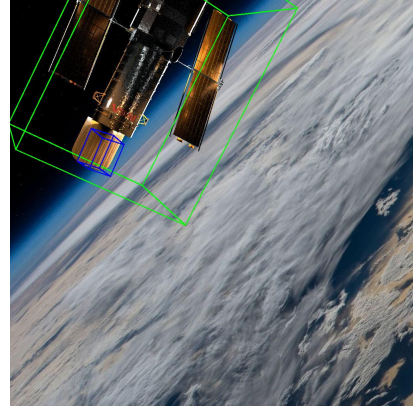


Figure 4.2: Hubble Space Telescope with earth rendered background, 1024x1024 second query image

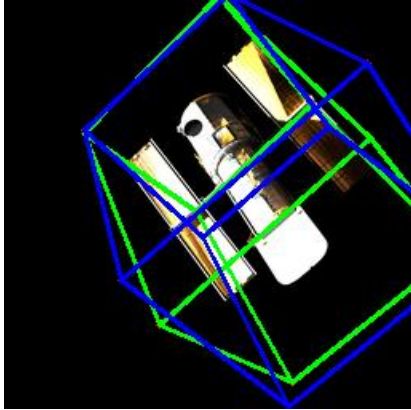


Figure 4.3: Hubble Space Telescope, no background, 256x256 query image, $e_{\text{ADD}} = 2.925$, $e_{\text{ADD-S}} = 1.183$

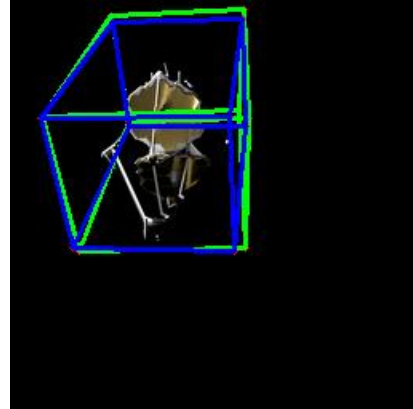


Figure 4.4: James Webb Space Telescope, no background, 256x256 query image, $e_{\text{ADD}} = 1.415$, $e_{\text{ADD-S}} = 0.808$

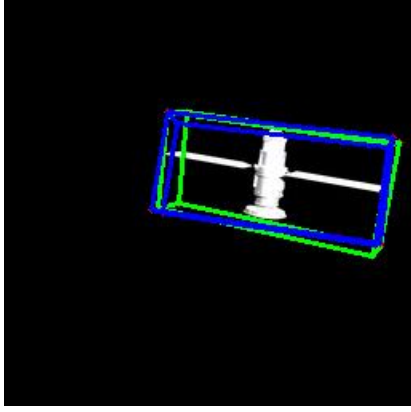


Figure 4.5: Cosmos Link, no background, 256x256 query image, $e_{\text{ADD}} = 1.718$, $e_{\text{ADD-S}} = 0.383$

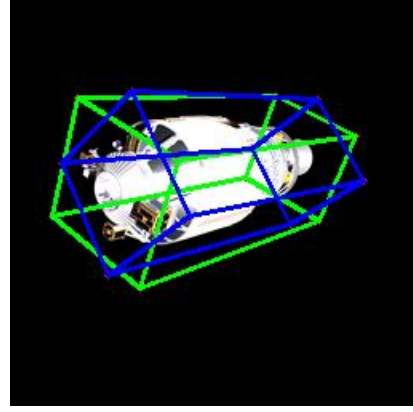


Figure 4.6: Rocket Body, no background, 256x256 query image, $e_{\text{ADD}} = 1.713$, $e_{\text{ADD-S}} = 0.252$

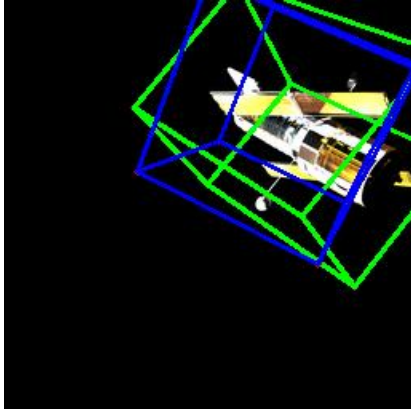


Figure 4.7: Hubble Space Telescope, no background, 256x256 query image, $e_{\text{ADD}} = 6.514$, $e_{\text{ADD-S}} = 1.571$

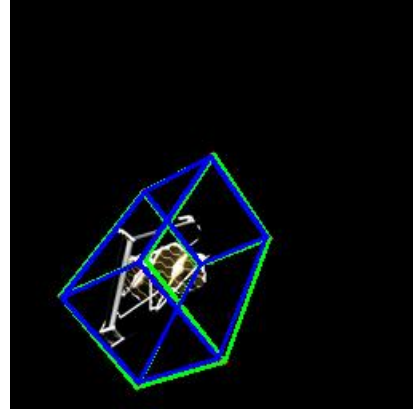


Figure 4.8: James Webb Space Telescope, no background, 256x256 query image, $e_{\text{ADD}} = 2.224$, $e_{\text{ADD-S}} = 1.261$

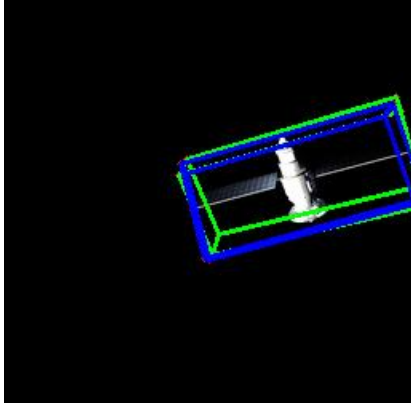


Figure 4.9: Cosmos Link, no background, 256x256 query image, $e_{\text{ADD}} = 1.925$, $e_{\text{ADD-S}} = 0.377$

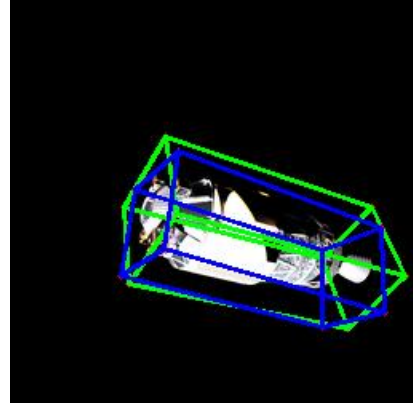


Figure 4.10: Rocket Body, no background, 256x256 query image, $e_{\text{ADD}} = 1.982$, $e_{\text{ADD-S}} = 0.501$

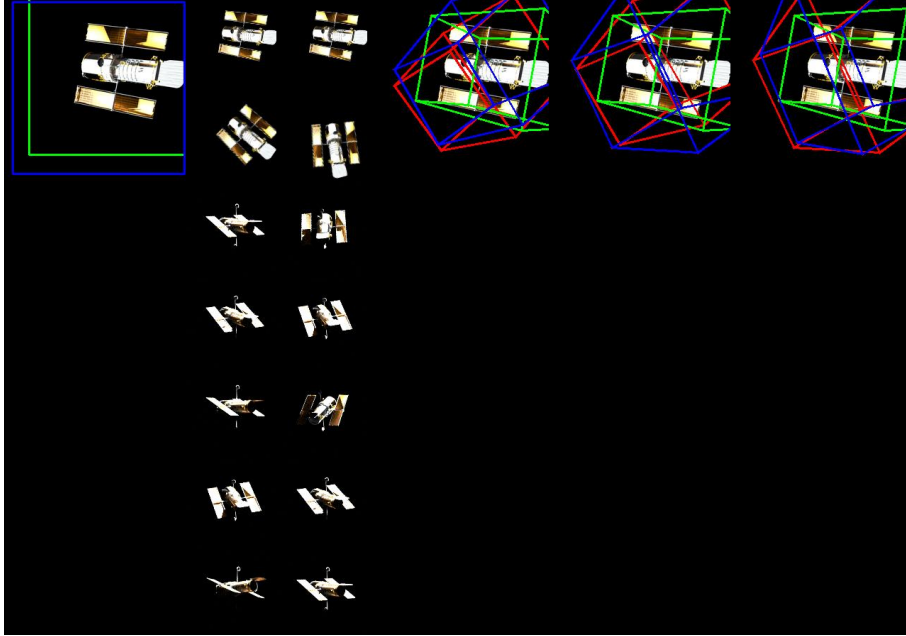


Figure 4.11: Hubble Space Telescope, no background, intermediary result, $e_{\text{ADD}} = 9.577$, $e_{\text{ADD-S}} = 5.196$

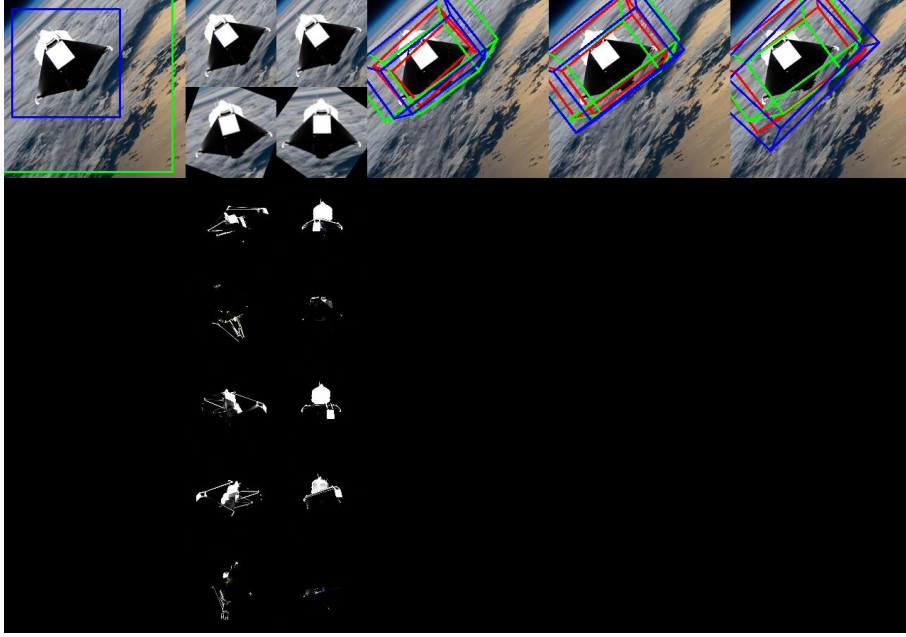


Figure 4.12: James Webb Space Telescope, with earth rendered background, intermediary result, $e_{\text{ADD}} = 10.934$, $e_{\text{ADD-S}} = 4.317$

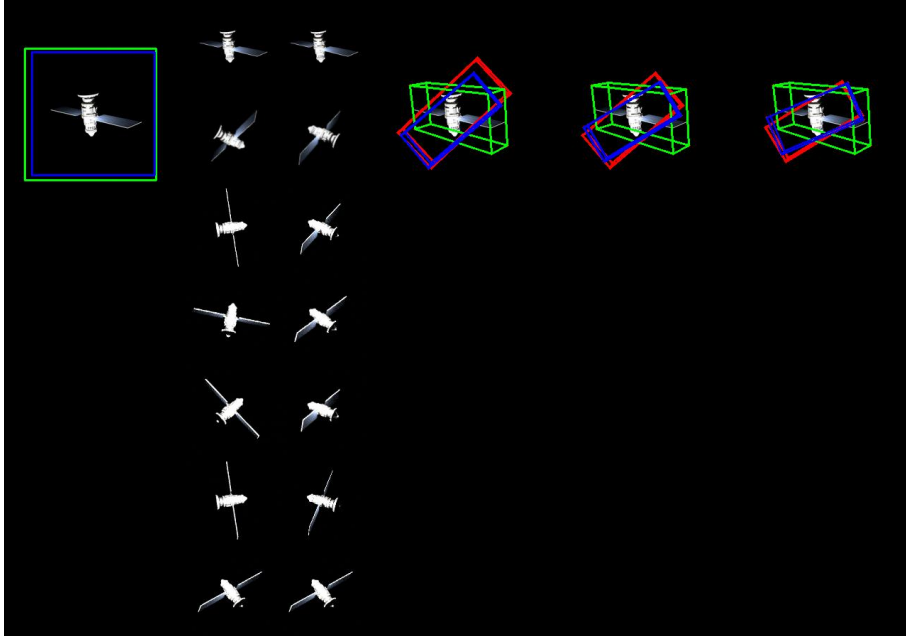


Figure 4.13: Cosmos Link, no background, intermediary result, $e_{\text{ADD}} = 11.094$, $e_{\text{ADD-S}} = 6.127$

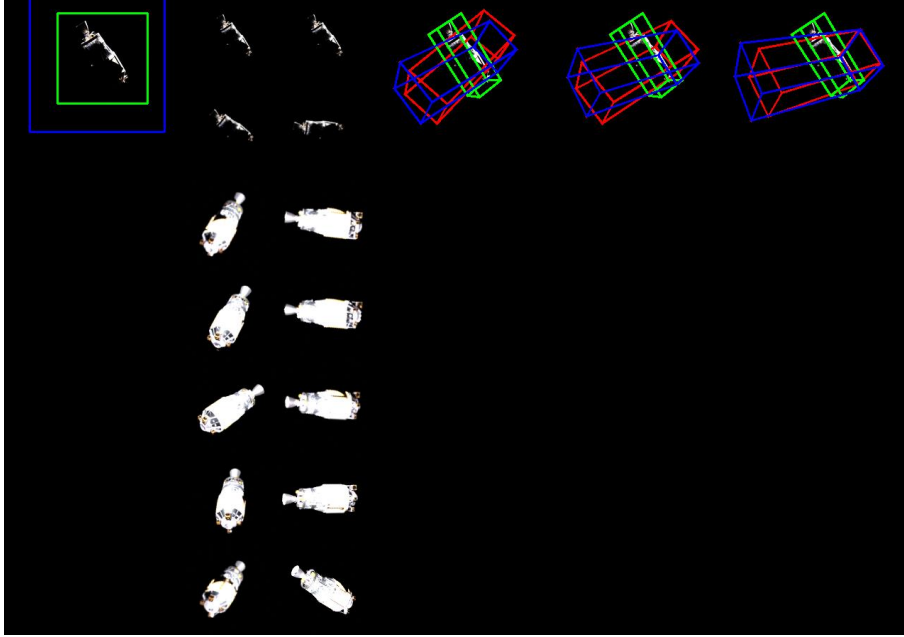


Figure 4.14: Rocket Body, no background, intermediary result, $e_{\text{ADD}} = 29.335$, $e_{\text{ADD-S}} = 17.743$

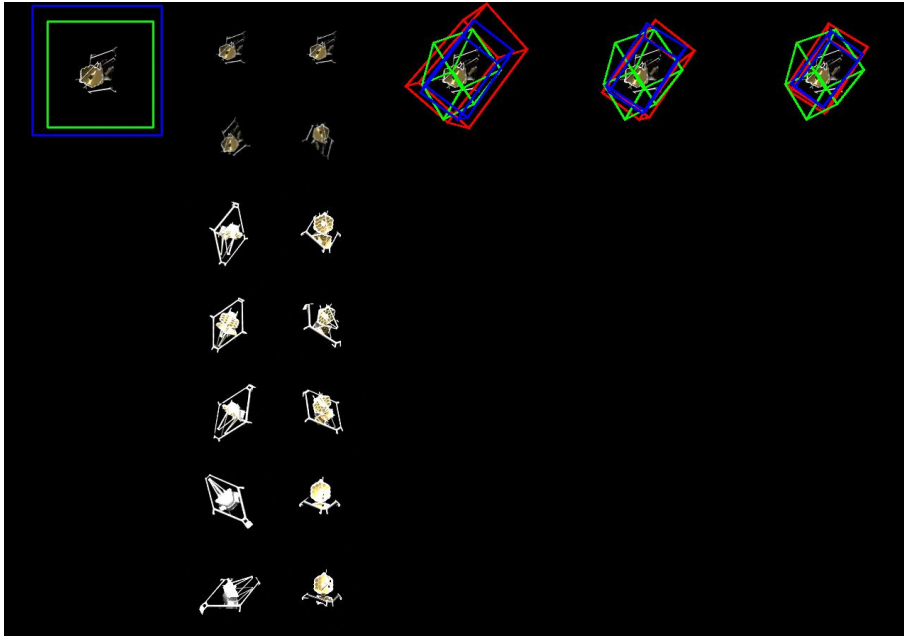


Figure 4.15: James Webb Space Telescope, with no background, intermediary result, $e_{\text{ADD}} = 21.983$, $e_{\text{ADD-S}} = 12.358$

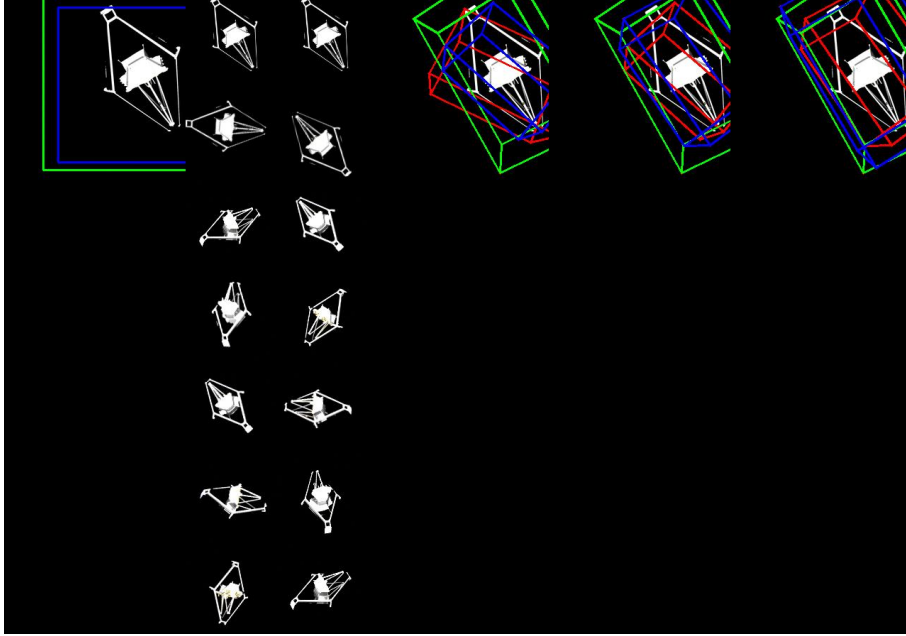


Figure 4.16: James Webb Space Telescope, with no background, intermediary result,
 $e_{\text{ADD}} = 1.060$, $e_{\text{ADD-S}} = 0.556$

5. Ways of improvements

5.1. Specialized spacecraft training set

5.2. Improved object detection algorithms

Rely more on the 3D model (for now only the size) and the segmented images, would optimize for symmetric and irregular shaped spacecrafts

5.3. Robustness to occlusion

6. Conclusion

Limitations Acknowledgments My personal contribution

Abbreviations

ADD Average Distance of Model Points

ADD-S Average Closest Point Distance

A. Python Scripts

```
1 """
2 Author:      Jeremy Chaverot
3 Date:        November 29, 2023
4 Description: Create the files val.txt, train.txt and test.txt according to a test
               percentage
5 """
6
7 import os
8 import sys
9 import random
10
11
12 if __name__ == "__main__":
13
14     # Check if the correct number of arguments is provided
15     if len(sys.argv) != 3:
16         print("Usage: python format.py <object_name> <test_percentage>")
17         sys.exit(1)
18
19     object = sys.argv[1]
20     test_percentage = float(sys.argv[2])
21
22     if (test_percentage < 0 or 1 < test_percentage):
23         print("Wrong value for the variable <test_percentage>. Should be between 0 and 1
24             included.")
25         sys.exit(1)
26
27     # Get a list of all files in the folder
28     all_files = os.listdir(f'data/SpaceCraft/{object}/images')
29
30     # Filter the list to include only image files and exclude MacOS temporary files
31     image_files = [file for file in all_files if file.lower().endswith(('.jpg')) and not
32         file.startswith('.') ]
33
34     # Get the number of images in the folder
35     num_images = len(image_files)
36
37     # Iterate through each image and apply the transformation
38     with open(f'data/SpaceCraft/{object}/train.txt', 'w') as train, open(f'data/SpaceCraft
39         /{object}/test.txt', 'w') as test:
40         for image_file in image_files:
41             rand = random.random()
42             image_path = 'SpaceCraft/hubble/images/' + image_file
43             if (rand < test_percentage):
44                 test.write(image_path + '\n')
45             else:
46                 train.write(image_path + '\n')
47
48     print(f"Done splitting {num_images} images in train.txt and test.txt")
```

Listing A.1: Python script format.py to randomly generate the training set and the test set based on a specified probability. Should be run from Gen6D's root folder.

```
1 """
2 Author:      Jeremy Chaverot
3 Date:        November 20, 2023
4 Description: Transform every images of a folder into jpg format.
5 """
6
7 import os
```

```
8 import sys
9 from PIL import Image
10
11
12 def transform_image(image_path):
13     img = Image.open(image_path)
14     new_image_path = image_path.split('.')[0] + '.jpg'
15     img.save(new_image_path)
16
17
18 if __name__ == "__main__":
19
20     # Check if the correct number of arguments is provided
21     if len(sys.argv) != 2:
22         print("Usage: python to_jpg.py </path/to/your/images>")
23         sys.exit(1)
24
25     folder_path = sys.argv[1]
26
27     # Get a list of all files in the folder
28     all_files = os.listdir(folder_path)
29
30     # Filter the list to include only image files and exclude MacOS temporary files
31     image_files = [file for file in all_files if file.lower().endswith(('png', '.jpg', '.jpeg', '.gif', '.bmp')) and not file.startswith('.')]
32
33     # Get the number of images in the folder
34     num_images = len(image_files)
35
36     # Iterate through each image and apply the transformation
37     for image_file in image_files:
38         image_path = os.path.join(folder_path, image_file)
39         transform_image(image_path)
40         os.remove(image_path)
41
42     print(f"Number of images transformed into .jpg: {num_images}")
```

Listing A.2: Python script to_jpg.py to transform every images of a specified folder into jpg format.

```
1 """
2 Author:      Jeremy Chaverot
3 Date:        November 20, 2023
4 Description: Transform a txt file with quaternions and the translation vector into multiple
5              npy files containing the rotation matrix concatenated with the translation vector.
6 """
7 import numpy as np
8 import sys
9 import os
10
11
12 def quaternion_to_matrix(Q, translation):
13     """
14         Covert a quaternion and translation into a full three-dimensional augmented
15         rotation matrix.
16
17         Input
18         :param Q: A 4 element array representing the quaternion (q0, q1, q2, q3).
19         :param translation: A 3 element array representing the translation (x, y, z).
20
21         Output
22         :return: A 3x4 element matrix representing the 3D rotation matrix concatenated with
23                 the translation vector.
24     """
25
26     # Extract the values from arguments
27     q0 = Q[0]
28     q1 = Q[1]
29     q2 = Q[2]
```

```

28     q3 = Q[3]
29
30     x = translation[0]
31     y = translation[1]
32     z = translation[2]
33
34     # Compute the rotation matrix
35     r00 = 2 * (q0 * q0 + q1 * q1) - 1
36     r01 = 2 * (q1 * q2 - q0 * q3)
37     r02 = 2 * (q1 * q3 + q0 * q2)
38
39     r10 = 2 * (q1 * q2 + q0 * q3)
40     r11 = 2 * (q0 * q0 + q2 * q2) - 1
41     r12 = 2 * (q2 * q3 - q0 * q1)
42
43     r20 = 2 * (q1 * q3 - q0 * q2)
44     r21 = 2 * (q2 * q3 + q0 * q1)
45     r22 = 2 * (q0 * q0 + q3 * q3) - 1
46
47     # 3x3 rotation matrix concatenated with the 3x1 translation vector
48     matrix = np.array([[r00, r01, r02, x],
49                       [r10, r11, r12, y],
50                       [r20, r21, r22, z]])
51
52     return matrix
53
54
55 if __name__ == "__main__":
56
57     # Check if the correct number of arguments is provided
58     if len(sys.argv) != 3:
59         print("Usage: python quaternion_to_matrix.py </path/to/your/text/file> </path/to/
60             the/pose/folder>")
61         sys.exit(1)
62
63     file_path = sys.argv[1]
64     pose_folder_path = sys.argv[2]
65     file_content = None
66
67     try:
68         with open(file_path, 'r') as file:
69             file_content = file.read()
70     except FileNotFoundError:
71         print(f"The file {file_path} was not found.")
72         sys.exit(1)
73     except Exception as e:
74         print(f"An error occurred: {e}")
75         sys.exit(1)
76
77     poses = file_content.split('\n')[:-1]
78
79     # Iterate through each pose, apply the transformation and save it
80     for pose in poses:
81         image_id, obj_id, q0, q1, q2, q3, x, y, z = pose.split(',')
82         Q = np.array([q0, q1, q2, q3], dtype=np.float32)
83         translation = np.array([x, y, z], dtype=np.float32)
84         matrix = quaternion_to_matrix(Q, translation)
85         np.save(pose_folder_path + '/pose' + str(int(image_id)), matrix)
86
87     print(f"Number of transformation processed: {len(poses)}.")

```

Listing A.3: Python script `quaternion_to_matrix.py` to transform a txt file with quaternions and the translation vector into multiple npy files containing the rotation matrix augmented with the translation vector.

```

1 """
2 Author:      Jeremy Chaverot
3 Date:        December 10, 2023
4 Description: Invert the masks from a given folder.
5 """

```

```

6
7 import cv2
8 import os
9 import sys
10
11
12 def inverse_masks_in_folder(folder_path):
13     # Iterate through the list of files at the specified path
14     for filename in os.listdir(folder_path):
15         # Filter to include only png image files and exclude MacOS temporary files
16         if filename.endswith(".png") and not filename.startswith('.'):
17             mask_path = os.path.join(folder_path, filename)
18             try:
19                 # Read the mask image
20                 mask = cv2.imread(mask_path, cv2.IMREAD_GRAYSCALE)
21                 if mask is None:
22                     print(f"Failed to read image: {mask_path}")
23                     continue
24
25                 # Invert the mask
26                 inverted_mask = cv2.bitwise_not(mask)
27
28                 # Save the inverted mask with a temporary name
29                 temp_path = os.path.join(folder_path, "temp_" + filename)
30                 cv2.imwrite(temp_path, inverted_mask)
31
32                 # Delete the original mask
33                 os.remove(mask_path)
34
35                 # Rename the inverted mask to the original filename
36                 os.rename(temp_path, mask_path)
37                 print(f"Inverted and replaced mask for: {mask_path}")
38             except Exception as e:
39                 print(f"Error processing {mask_path}: {e}")
40
41
42 if __name__ == "__main__":
43
44     # Check if the correct number of arguments is provided
45     if len(sys.argv) != 2:
46         print("Usage: python invert_mask.py <folder_path>")
47         sys.exit(1)
48
49     folder_path = sys.argv[1]
50     inverse_masks_in_folder(folder_path)

```

Listing A.4: Python script invert_mask.py to invert the masks from a specified folder. We aim to have a black object set against a white background.

```

1 """
2 Author:      Jeremy Chaverot
3 Date:        January 01, 2024
4 Description: Resize the images from a given folder.
5 """
6
7 import os
8 import sys
9 from PIL import Image
10
11
12 def resize_images(folder_path, resize_factor):
13     # Iterate through the list of files at the specified path
14     for filename in os.listdir(folder_path):
15         # Filter to include only png image files and exclude MacOS temporary files
16         if filename.endswith(".png") and not filename.startswith('.'):
17             img_path = os.path.join(folder_path, filename)
18             with Image.open(img_path) as img:
19                 # Calculate new size
20                 new_size = tuple([int(dim / resize_factor) for dim in img.size])
21                 # Resize the image

```



```
22         resized_img = img.resize(new_size, Image.ANTIALIAS)
23         # Save the resized image with a different name temporarily
24         temp_path = os.path.join(folder_path, "temp_" + filename)
25         resized_img.save(temp_path)
26
27         # Delete the original image
28         os.remove(img_path)
29
30         # Rename the resized image to the original filename
31         os.rename(temp_path, img_path)
32
33
34 if __name__ == "__main__":
35
36     # Check if the correct number of arguments is provided
37     if len(sys.argv) != 3:
38         print("Usage: resize.py <folder_path> <resize_factor>")
39         sys.exit(1)
40
41     folder_path = sys.argv[1]
42     factor = int(sys.argv[2])
43
44     resize_images(folder_path, factor)
```

Listing A.5: Python script `resize.py` designed to alter an image's size with respect to a specified resize factor.

B. Scitas Izar Setup Tutorial

Bibliography

- [1] Y. Liu, Y. Wen, S. Peng, C. Lin, X. Long, T. Komura, and W. Wang. *Gen6D: Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images*. 2023. arXiv: 2204.10776.
- [2] G. S. 3Blue1Brown. *Quaternions and 3d rotation, explained interactively*. Youtube. 2018. URL: <https://www.youtube.com/watch?v=zjMuIxRvygQ>.
- [3] W. R. Hamilton and W. E. Hamilton. *Elements of Quaternions*. London: Longmans, Green, & Company, 1866.
- [4] Y.-B. Jia. "Quaternions." In: (2022).
- [5] S. Hinterstoisser, V. Lepetit, S. Ilic, S. Holzer, G. Bradski, K. Konolige, and N. Navab. "Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes." In: *Computer Vision – ACCV 2012*. Springer Berlin Heidelberg, 2013, pp. 548–562.