NTUST course: Computer Vision and Applications (CI5336701, 2024 Spring)

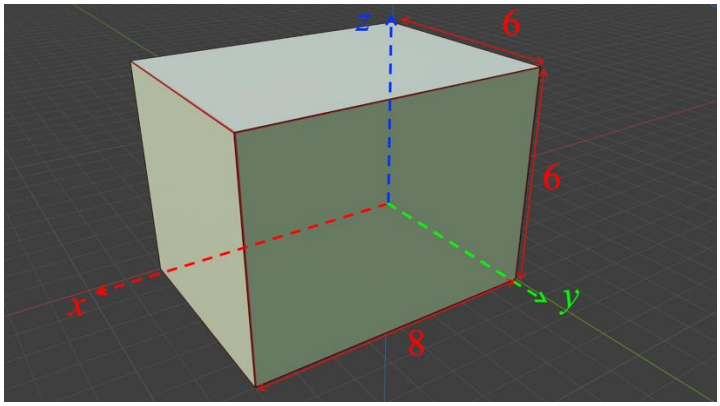
Homework#4: Calibrate the camera from single image

Date Due：2024. May. 20th, PM11:59。

Description：

---

1. Write a program for determining the camera position from ONE given image (enclosed file: imgToCalib.png) and known structures. You may need to calculate the intrinsic and extrinsic parameters of the camera. No lens distortion on this image. (choose your tools, ex. python, C++/C, openCV, Matlab).
2. In this image, there is only one box whose size is 8 x 6 x 6 m^3(in length, width and height). The world coordinate is on a corner and its *x*, *y* and *z* axes are on the edges as shown in the figure (the given image has no notation).



3. Please manually select 3 rectangles on the surface of the box, then determine their homography from them (real-dimension to image domain). Based on Zhang's method (or IAC), please calculate the intrinsic parameter of the camera.
4. In this assignment, there is NO need to read this image or detect features in your program. You can use an image editor tool (ex. photoshop, XnView) for picking out the pixel coordinate then copy text into your program.
5. Deliverable: the data you should provide:
   1) Source code in python, C++/C or Matlab (or similar program with simple comment), being able to print out 3 information: **a) intrinsic parameter K**, **b) extrinsic parameter RT** and **c) how far the camera to the origin**.
   2) (Optional) Execution file (.exe), if appliable.

   Please zip all your files, then upload to moodle2 (https://moodle2.ntust.edu.tw/) by due date.
6. Score evaluation:
   1. 75%: for determining **intrinsic parameter K** correctly.
   2. 25%: **extrinsic parameter RT**, and +10% (bonus) for calculating the camera position

(relative to world coordinate).

Hint: Camera Position and Ground Truth (for your reference):

The following data are not used in your homework#4. These are ONLY for verifying your result. Please do NOT use any of these data in your program.

Ground truth (generated by an ideal camera):

K matrix

[1866.7, 0.00, 960.0

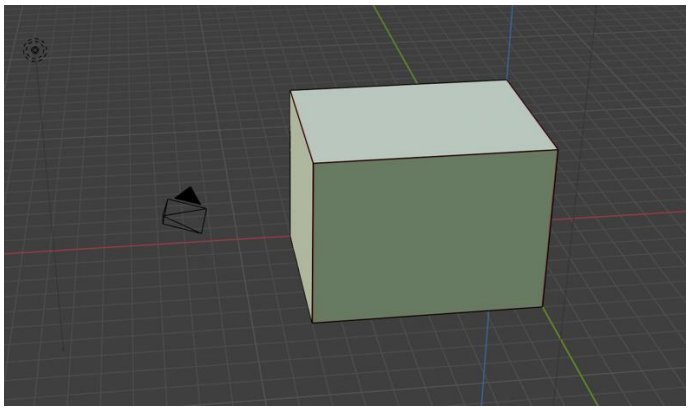
0.00, 1866.7, 540.0

0.00, 0.00, 1.0]

RT matrix (updated @2024/05/07)

[-0.8309382	0.5563648	0.00000	0.9749568
0.2477334	0.369993	-0.8953956	0.1147192
-0.4981667	-0.7440184	-0.4452714	18.974983]

The blender file is given, if you need to view the structure.



(blank below this line)