TTK4255 Robotic Vision

Assignment 6

Written Spring 2022 by A. Fanebust and Ø. Solbø

Part 1: Various theory questions

Task 1.1)

From equations (7) to (9) in the assignment text, it is assumed that θ represent the angle between the Z-axes and the vector to said point. This is shown in figure 1, and implies that $\theta = 0$, means that a point is straight in front of the camera.

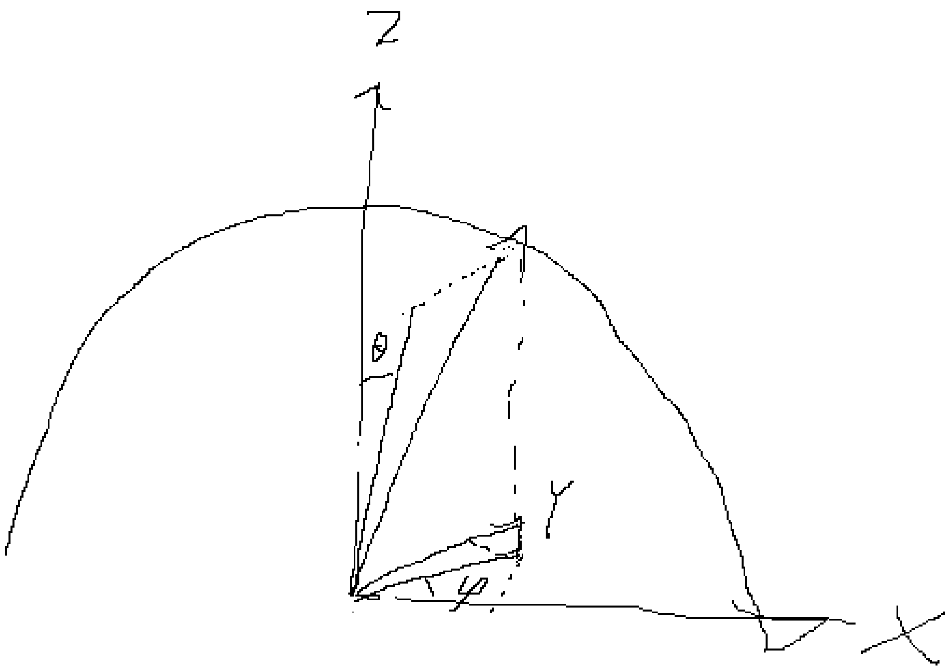

Figure 1: Angles and axes for task 1.1).

With a fish-eye lens with a FOV of π rad, it is therefore expected that the angle θ to be in the interval $\theta \in [-\pi/2, \pi/2]$. Equation (5) and (6) is developed from $\frac{X}{Z}$ and $\frac{Y}{Z}$ respectively, giving $tan(\theta)$ in the equation. This is ill-defined for $\theta = \pm\frac{\pi}{2}$, and undistorting the images could therefore prove difficult.

Task 1.2)

a)

c_x, c_y, fs_x, fs_y will be halved when downscaling the image. Since s_x and s_y corresponds to the pixel width and height, these values will be maintained. This implies that f must be halved when downscaling. k_i, p_j, where $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$, must be maintained, as the distortion is unaffected by a downscaling.

b)
Assuming that a calibrated image is cropped, the pixel density will remain identical. Due to cropping, c_x will be shifted to $c_x = \frac{W-(l+r)}{2}$. As the observed distortion is relative to the original center of the image, cropping will result in having a different set of distortion parameters. For example, the radius r might appear smaller for some points. This requires larger parameters to maintain the radial distortion. A larger radius on the other hand, would imply the usage of smaller parameters. A more general solution could be to shift the x used in the distortion algorithms back to the original coordinates.

Task 1.3)

Left image uses $k_1 > 0$ and right image uses $k_1 < 0$. Using equations (3) and (4) in the assignment text, the radial distortion in x is given as $k_1 r^2$. Simplifying the equation for radius by assuming $y = 0$, one gets that $u = c_x + f_x(x + k_1 x^2)$. Since the value for x is relative to c_x, using $k_1 > 0$ will push the horizontal pixels outwards, as observed in the left image. For $k_1 < 0$, the horizontal pixels are pushed inwards, as observed in the right image.

Task 1.4)

It is assumed the pose of each image must be estimated, where each image contains 6 DOF. As there assumed to be 9 intrinsic parameters to be estimated, $\{c_x, c_y, s_x f_x, s_y f_y, k_1, k_2, k_3, p_1, p_2\}$, one gets a total number of estimated parameters as $6N + 9$.

Task 1.5)

When estimating the intrinsic parameters, it is sufficient to use an arbiter scale of the checkerboards. Given that enough images from enough views are used, the true scale of the checkerboards are redundant. The true scale is only necessary when calculating the camera position, due to the relationship between distance and observed image scale compared to true scale.

Using an arbiter scale for the checkerboard, the differences in image scale could be observed over the checkerboard. Thus, observing the changes to horizontal and vertical scales, as well as the image skew. By constraining over multiple locations, one could separate the effects from tangential and radial distortions. Thus, the intrinsic parameters can be estimated from an image on the computer.

In addition, the square size is not mentioned as a necessary input to the calibration algorithm in Zhang's paper [1]. Instead, only the need for unique camera orientations are highlighted in the paper.

Task 1.6)

It is theorized that the calibration will be dependent on where the points are located. For example, if the points are all located near the center, all close together, it may be difficult to extract the distortion parameters. The estimates may therefore produce a small reprojection error for the observed points, but be totally useless for any point further away from the center.

Another theorized scenario, is that a low resolution calibration image is used. Due to the lack of resolution, it may be difficult to estimate the distortion errors that occur.

Task 1.7)

It is assumed the camera moves orthogonal to the image frame, constantly having the checkerboard in front. This results in $\theta = 0$, as described in section . This means the image point will always have $X = 0 = Y$, resulting in $x = \frac{X}{Z} = \frac{0}{\lambda} = 0 = \frac{0}{\lambda} = \frac{Y}{Z} = y$. Under the assumption that there is no distortion, the following equations are achieved:

$$
\begin{aligned}
u &= c_x + f_x(x + \delta x) = c_x \\
v &= c_y + f_y(y + \delta y) = c_y
\end{aligned}
$$

This means the focal length does not contain enough constraints for evaluating it, making it impossible to separate from the translation between images.

Zhang justifies the same based on the conditions the originate from the two images. If the two images have the same orientation, the camera plane intersects the checkerboard plane at the same point. The images do therefore not introduce more than one unique condition based on the orientation.

Part 2: Analysis of an example calibration

Task 2.1)

Using equation (1) in the assignment text, with the assumption of zero distortion, the equation for the horizontal pixel-index, u, is given as $u = c_x + f_x x$. Reforming the equation and solving for x, given $u = W$ and assuming no errors, gives

$$
\begin{aligned}
x &= \frac{W - c_x}{f_x} \\
&= \frac{2816 - 1370.0592}{2359.40946} \\
&= 0.612840291.
\end{aligned}
$$

The maximum obtainable horizontal index is therefore given as

$$
\begin{aligned}
u_{max} &= c_x + (f_x + f_{x95\%})x \\
&= 1370.05852 + (2359.40946 + 1.96 \cdot 0.842) \cdot 0.612840291 \\
&= 2817.010883.
\end{aligned}
$$

The maximum pixel error is therefore given as $2817.010883 - 2816 = 1.010883$.

Task 2.2)

With $p_1 = 0$ and $p_2 = 0$ the task22 script calculated the following width and height as 2829.38740098 and 2116.25613737 respectively. As the errors have a magnitude far greater than 1 pixel, it is obvious that the p-values cannot be omitted.

References

[1] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, December 2000, mSR-TR-98-71, Updated March 25, 1999. [Online]. Available: https://www.microsoft.com/en-us/research/publication/a-flexible-new-technique-for-camera-calibration/