CS 6320, 3D Computer Vision
Spring 2012, Prof. Guido Gerig
Solutions Assignment 1

Out: Mon Jan 30 2012
In: Mon Feb 13 2012

Summary Grading HW1

Theory:	10
Practical (using existing calibration SW):	18
Total	**28**
Record Bonus separately	

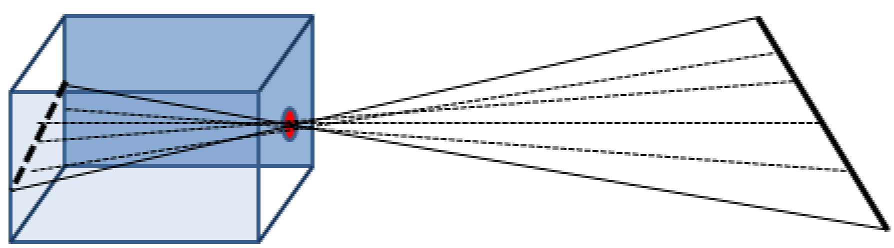
Grading Theoretical Part: Total 8 + 1*

Material: "Computer Vision" by Forsyth & Ponce, Chapters 1, 2 and 3

1. Pinhole camera [1+1 = 2 total]

a) *"A straight line in the world space is projected onto a straight line at the image plane". Prove by geometric consideration.*

This important feature can be proven by a simple geometric consideration: All light rays emitted from a straight line pass through the pinhole. Consequently they all lie on a plane which is spanned by the straight line and the pinhole. This plane intersects with the image plane in a straight line.



b) *Show that, in the pinhole camera model, three collinear points in 3-D space are imaged into three collinear points on the image plane.*

Solution 1a does not make any restriction on the choice of points. As we already know that points on a straight line are imaged to a straight line, the same is true for collinear points as these are all imaged to a line, where by definition points are collinear.

Formally, we can also choose two 3D scene points, project them into the image plane, then choose a third scene point that lies on the line defined by the 2 points. It can be easily shown that the image of this 3rd point indeed is part of a line through the images of points 1 and 2.

Or, coming up with a line equation for a line in 3D space with parameter t, we can substitute with the equations for the pinhole camera. It can be shown that the resulting line equation is again linear in t.

straight line in scene:	$\underline{X}(t) = \underline{X}_0 + t\underline{U}$, where $\underline{U}$ is $(\underline{X}_2-\underline{X}_1)$
pinhole camera:	$x = -f(X/Z)$, similarly for y
image plane: substitution:	$x = -f(X_0 + tU)/(Z_0 + tW)$, similarly for y

More elegant solution:

3 collinear points in world coordinates $\underline{X_1, X_2, X_3}$ form a matrix whose determinant is 0. We can apply the perspective projection to each point to get the three points $\underline{x_1, x_2, x_3}$ and their determinant in image space:

World coordinates: $$Det\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix} = 0$$, via collinearity.

Image coordinates: $$Det\begin{vmatrix} \frac{X_1}{Z_1} & \frac{Y_1}{Z_1} & 1 \\ \frac{X_2}{Z_2} & \frac{Y_2}{Z_2} & 1 \\ \frac{X_3}{Z_3} & \frac{Y_3}{Z_3} & 1 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 .$$

For points on a line: $\bar{x}_3 = (k)\bar{x}_2 + (1-k)\bar{x}_1$ and the same for all components. Substituting this into the equation above shows that the Det is zero and that the projected points are collinear.

---

2. Perspective Projection [1+2 =3 total]

- *2.a)* Prove geometrically that the projections of two parallel lines lying in some plane Π appear to converge on a horizon line H formed by the intersection of the image plane with the plane parallel to Π and passing through the pinhole.

(see diagram slides for discussion of the geometry of vanishing points)

- *2.b)* Prove the same result algebraically using the perspective projection equation. You can assume for simplicity that the plane Q is orthogonal to the image plane (as you might see in an image of railway tracks, e.g.).

Prove the same result algebraically using the perspective projection Eq. (1.1). You can assume for simplicity that the plane Π is orthogonal to the image plane.

Solution Let us define the plane Π by $y = c$ and consider a line Δ in this plane with equation $ax + bz = d$. According to Eq. (1.1), a point on this line projects onto the image point defined by

$$\begin{cases} x' = f'\dfrac{x}{z} = f'\dfrac{d - bz}{az}, \\ y' = f'\dfrac{y}{z} = f'\dfrac{c}{z}. \end{cases}$$

This is a parametric representation of the image δ of the line Δ with z as the parameter. This image is in fact only a half-line since when $z \to -\infty$, it stops at the point $(x', y') = (-f'b/a, 0)$ on the x' axis of the image plane. This is the vanishing point associated with all parallel lines with slope $-b/a$ in the plane Π. All vanishing points lie on the x' axis, which is the horizon line in this case.

---

3. **Depth of Field** **[1+1+1 = 3 total]**

A) An interesting and desirable property of the pinhole camera is the infinite depth of focus. Give an intuitive explanation.

Each scene point emits one ray that passes through the hole onto the image plane. There is only one ray per point P that creates p, and therefore every P independent of the distance z creates a sharp picture in the image plane. The depth of field therefore is infinite.

B) Consider a camera equipped with a thin lens, with its image plane at position $z^{'}$ and the plane of scene points in focus at position z. What is the size of the blur circle obtained by imaging a point located at position $z + \delta z$ on the optical axis?

1.5. Consider a camera equipped with a thin lens, with its image plane at position z' and the plane of scene points in focus at position z. Now suppose that the image plane is moved to $\hat{z}'$. Show that the diameter of the corresponding blur circle is

$$d\frac{|z' - \hat{z}'|}{z'},$$

where d is the lens diameter. Use this result to show that the depth of field (i.e., the distance between the near and far planes that will keep the diameter of the blur circles below some threshold ε) is given by

$$D = 2\varepsilon f z(z+f)\frac{d}{f^2d^2 - \varepsilon^2 z^2},$$

and conclude that, for a *fixed* focal length, the depth of field increases as the lens diameter decreases, and thus the f number increases.
Hint: Solve for the depth $\hat{z}$ of a point whose image is focused on the image plane at position $\hat{z}'$, considering both the case where $\hat{z}'$ is larger than z' and the case where it is smaller.

Solution If ε denotes the diameter of the blur circle, using similar triangles immediately shows that

$$\varepsilon = d\frac{|z' - \hat{z}'|}{z'}.$$

Now let us assume that $z' > \hat{z}'$. Using the thin lens equation to solve for the depth $\hat{z}$ of a point focused on the plane $\hat{z}'$ yields

$$\hat{z} = fz\frac{d - \varepsilon}{df + \varepsilon z}.$$

By the same token, taking $z' < \hat{z}'$ yields

$$\hat{z} = fz\frac{d + \varepsilon}{df - \varepsilon z}.$$

Finally taking D to be the difference of these two depths yields

$$D = \frac{2\varepsilon d f z(z+f)}{f^2 d^2 - \varepsilon^2 z^2}.$$

Now we can write $D = kd/(f^2d^2 - \varepsilon^2 z^2)$, where $k > 0$ since $z' = fz/(f+z) > 0$. Differentiating D with respect to d for a fixed depth z and focal length f yields

$$\frac{\partial D}{\partial d} = -k\frac{f^2d^2 + \varepsilon^2 z^2}{(f^2d^2 - \varepsilon^2 z^2)^2} < 0,$$

which immediately shows that D decreases when d increases, or equivalently, that D increases with the f number of the lens.

C) Use this result to derive the equation for $\triangle Z_{0}^{+}$ similarly to the way $\triangle Z_{0}^{-}$ is derived in the chapter 1 slides, using the thin lens assumption.

Formal definition of depth of field: Using the lens equation for Z_0 and Z_{0+}, and similar triangles, we obtain the following result (see slides to course lecture):

$$\Delta Z_0^+ = Z_0 - Z_0^+ = \frac{Z_0(f - Z_0)}{f(1 + d/b) - Z_0}$$

- Depth of field decreases with d (smaller aperture) and increases with Z_0 (close objects have very narrow field of view, often seen in macro pictures of small objects, far objects are always sharp).
- Photographer: To strike a balance between incoming light and sharp depth range (smaller aperture means less light and thus much longer exposure times).

4. 3D Rotation **[1+1 = 2 total]**

a) Show that 3D Rotation expressed as $R_x.R_y.R_z$ is not commutative.

b) Given very small angles of rotation which is often found as misalignment, where cos(angle) $\rightarrow$1 and sin(angle) $\rightarrow\varepsilon$. Would that have an effect on the commutativeness?

a) Matrix multiplications in general are not commutative ($|A|.|B| \neq |B|.|A|$). This can be easily shown for the rotation matrices by demonstrating that Rx.Ry $\neq$ Ry.Rx, for example. You can also take a cube, mark one corner, rotate around two axis and then observe the position of the corner. Permuting the rotations gives a different result.

b) Calculating Rx.Ry and Ry.Rx, for example, shows that if we set α =very small or 0, the matrices only contain rotations around β and are therefore very similar and nearly commutative. Thus, if rotations around one axis are nearly 0, it becomes commutative when we combined it with rotation around another axis. We can also replace cos(angle) by 1 and sin(angle) by ε and solve Rx.Ry.Rz and other combinations. This shows that due to $\varepsilon^3 << \varepsilon^2 << \varepsilon$, the resulting matrices are becoming practically the same, so that the transformations are close to commutative.

Practical Assignment

Will be discussed in class.

Grading Practical Part: Total 18

Record Bonus separately

Criteria:

“Completed” Experiment (images, coordinates, setup of Matlab solution, obtain results).
6
Success: Did they get reasonable results? 4
Does report cover everything from design, implementation, discussion of results.3
Quality of report: Style, clarity, organization. 3
Additional evaluation of measured versus calculated locations? 2