Lab III Computer Vision: Inferring 3D from 2D

Amit Bharti
University of Groningen
Groningen, The Netherlands
a.bharti.1@student.rug.nl

Pedro Rodriguez de Ledesma Jimenez
University of Groningen
Groningen, The Netherlands
p.rodriguez.de.ledesma.jimenez@student.rug.nl

I. Introduction

When a screen is used to form a planar image of the 3D object, the image obtained won't be clear since any point on the screen will have multiple rays of light being received from the object. So to constrain that, pinholes and lens are used to form a sharp planar image of the 3D object on a 2D plane. Lens need to be calibrated and therefore is necessary to obtain the internal characteristics. In the Figure 1 we can see how a point P is projected on the image plane at P_I which is intersection of the image plane and the line PO_c. Let x_w, y_w, z_w be the co-ordinate of P and u, v be the co-ordinate of the projected point P_I on the image plane.

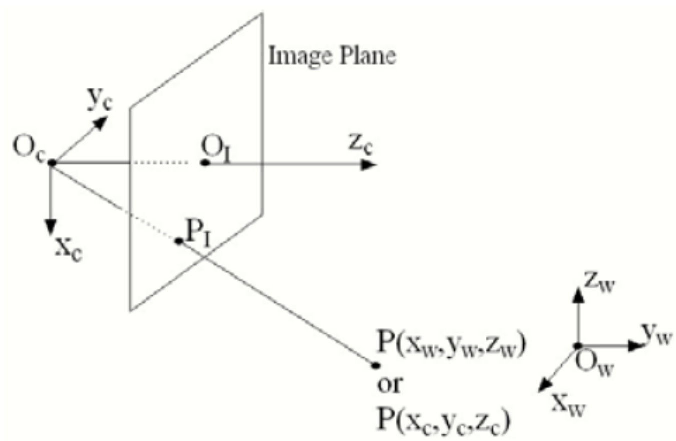

Fig. 1: Projection of point P from world to planer surface
Source : Determination of Food Portion Size by Image Processing [1]

From similar triangle $O_cP_IO_I$ and O_cPZ_c as seen in the Figure 1, we could write the following relationships.

$$\frac{u}{f} = \frac{x_w}{z_w} \tag{1}$$

$$u = f * \frac{x_w}{z_w} \tag{2}$$

Similarly:

$$v = f * \frac{y_w}{z_w} \tag{3}$$

where *f*, the focal length of the camera is the distance between O_c and O_I as seen in Figure 1.
Furthermore, the equations of the line in 3D are given by following equations.

$$L = \vec{p} + \lambda\vec{w} \tag{4}$$

$$L = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + \lambda \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

where w is the direction vector, and is normalized: $\|w\|^2$ = 1. If instead of a point two parallel lines with non-zero slope with Z_c lanes are projected, they will meet in the projection image in a *vanishing point*. This point can be found as the intersection with the image plane of a line through the projection center parallel to the given 3D lines.

Referring to the equation 2 and 12, equation for the perspective projection of the line in the image plane will be as below.

$$u = f * \frac{p_1 + \lambda w_1}{p_3 + \lambda w_3} \tag{5}$$

$$v = f * \frac{p_2 + \lambda w_2}{p_3 + \lambda w_3} \tag{6}$$

All the points of this lines which are infinitely far removed from the projection center will have coordinates as below:

$$u_\infty = \lim_{\lambda\to\infty} f * \frac{p_1 + \lambda w_1}{p_3 + \lambda w_3} \tag{7}$$

$$v_\infty = \lim_{\lambda\to\infty} f * \frac{p_2 + \lambda w_2}{p_3 + \lambda w_3} \tag{8}$$

In the next section four exercise are presented, where we start identifying camera constant parameters like the focal length from an image with concentric rectangles. Afterwards, the direction of the edges are computed and the normal of the plane where they are is calculated. We will explain in detail the process and a Jupyter notebook is attached with the process.

II. Exercises

A. Exercise 1

To compute the focal length f we used two different approaches. But first, the vanishing points in which the projections of the parallel lines converges were identified by extending sets of parallel lines identified. The lines are the edges that are identified by applying the Canny edge detector and the Hough transform. In the case of a rectangle, which has parallel lines in 2 different directions, we get two vanishing points as shown in the Figure 2. In this representation, ABCD is the projection of a 3D rectangle in which AB and CD line segment that are parallel (in 3D to each other) end in V_1 as vanishing point 1 while AD and BC line segments are parallel so giving V_2 as another vanishing point.

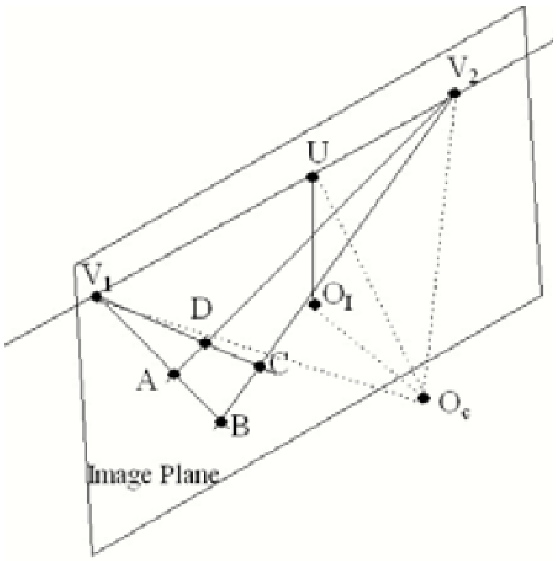

Fig. 2: Vanishing points obtained from a rectangle.
Source : Determination of Food Portion Size by Image Processing [1]

$$\begin{pmatrix} w1 \\ w2 \\ w3 \end{pmatrix} = \frac{1}{\sqrt{u_\infty^2 + v_\infty^2 + f^2}} \begin{pmatrix} u_\infty \\ v_\infty \\ f \end{pmatrix}$$

In the first approach, the focal length was computed from the above equation, where the direction vector of the two parallel lines projection is defined by the vanishing point of them and the focal length. Since we have two sets of parallel lines, we have two vanishing points and two direction vectors, although the focal length has the same value because it is the same image. However, the direction vectors and the focal length are unknown. Therefore, another constraint is used to solve the equation. Since we know that rectangle edges are orthogonal to each other, the dot product between the directions vector of the orthogonal edges will be zero and leads to below equation:

$$w_1 * w_1\prime + w_2 * w_2\prime + w_3 * w_3\prime = 0 \tag{9}$$

where: w_1, w_2 and w_3 are the direction vectors of 1st set of parallel lines while $w_1\prime$, $w_2\prime$ and $w_3\prime$ are the direction vectors of the 2nd set of parallel lines. In the below equation, we substituted the value of direction vectors in the equation 9 to get the dot product between them.

$$\frac{1}{\sqrt{u_\infty^2 + v_\infty^2 + f^2}} * u_\infty . \frac{1}{\sqrt{u\prime_\infty^2 + v\prime_\infty^2 + f^2}} * u\prime_\infty + \frac{1}{\sqrt{u_\infty^2 + v_\infty^2 + f^2}} * v_\infty . \frac{1}{\sqrt{u\prime_\infty^2 + v\prime_\infty^2 + f^2}} * v\prime_\infty + \frac{1}{\sqrt{u_\infty^2 + v_\infty^2 + f^2}} * f . \frac{1}{\sqrt{u\prime_\infty^2 + v\prime_\infty^2 + f^2}} * f = 0$$

$$u_\infty . u\prime_\infty + v_\infty . v\prime_\infty + f^2 = 0 \tag{10}$$

where u_∞ and v_∞ are the vanishing point coordinates of the 1st set of parallel lines, while $u\prime_\infty$ and $v\prime_\infty$ are the coordinates of the 2nd set of parallel lines.
Since we know the value of all the parameters in the equation 10 other than focal length, the focal length of the camera is obtained. Once the focal length is determined, the direction vectors of the rectangle edges are computed. The normal of the planar patch containing the rectangles is calculated by taking the cross product between both the direction vectors as given by the below equation.

$$Normal\ to\ planar\ patch\ = (w_1, w_2, w_3) * (w_1\prime, w_2\prime, w_3\prime) \tag{11}$$

The second approach acomplish the task geometricaly [2].

B. Exercise 2

A concentric rectangle picture with parallel sides was processed to get two sets of parallel lines that are orthogonal to each other. The Canny edge detector was used to identify the points of the edges and the Hough transform was used to identify lines from the points, all this using Python and not Matlab. Once the lines were found, they were extended until the intersection to obtain the vanishing point of each direction. Later and using the vanishing points, the focal length of the camera was calculated to get the intrinsic parameters of the camera for its calibration. The image used for drawing the parallel lines from the rectangle is shown in Figure 3 where the lines identified by the Canny edge detector and the Hough transform are displayed. For more information of the process check-up the corresponding Jupyter notebook of the Lab.

C. Exercise 3

The function *parlines* computes the equation $A\vec{w} = \vec{0}$, where A is the matrix with the direction vector coefficients obtained by linking the equations of the 3D line with the their projection by the relation between u and v (coordinates of the projection) with their 3D coordinates (4). Furthermore, a change of variable has been done to take out of the matrix the variable f. In the first part of the function the direction vector of the projection is computed (h_1 and g_1). The last column of the matrix is computed in the next lanes.

Once the matrix is calculated, the system is solved by performing a singular decomposition of the matrix A to find a solution that minimized $\|A\vec{w}\|^2$ and satisfies that the direction vector is normalized. Once the function computes the SVD, we select the singular vector with the smallest singular value.

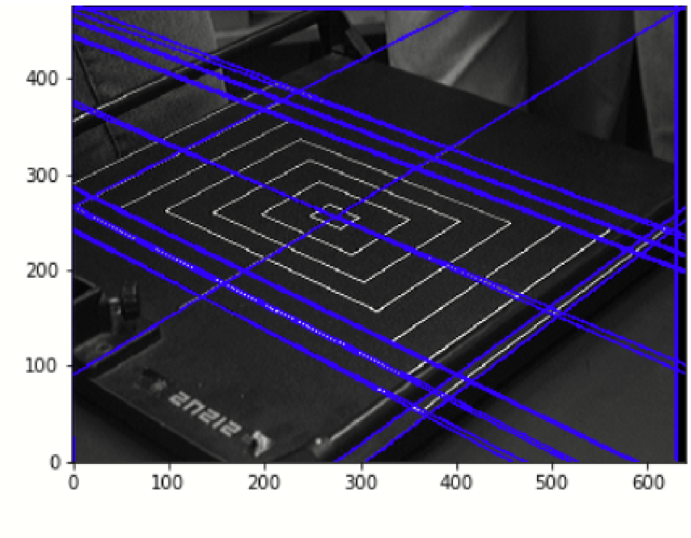

Fig. 3: Construction of parallel lines for getting the perspective projection

Once we have the best solution of the direction vector that meets the requirements previously mentioned, the change of variable accomplished before gets undone by dividing w_1 and w_2 by the focal length. At the last step, the vector obtained is normalized, a requirement of the directional vector.

$$\mathbf{A}'\,\vec{x} = \begin{pmatrix} h_1 & -g_1 & d_1g_1 - c_1h_1 \\ . & . & . \\ . & . & . \\ h_N & -g_N & d_Ng_N - c_Nh_N \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ . \\ . \\ 0 \end{pmatrix}$$

Fig. 4: The equation $A\vec{x} = \vec{0}$, where $x_1 = w_1 * f$, $x_2 = w_2 * f$ and $x_3 = w_3$

D. Exercise 4

The image *rectangle.tif* was read and the edges of the rectangles were identify by applying the Canny edge detector to obtain more accurate vanishing points coordinates. The camera constant f or focus length was computed by two different methods, the first one explained in previous sections in this documents and the second by geometrical relationships. The projection center is in the hemisphere that passes through the two vanishing points. Furthermore as another constrain, the projection center is in the center of the image. With this two constrains we compute the projections and afterwards the focal length. The focal length obtained is $f = 1337.3$. Once the focal length of the camera is known, the direction vector of the rectangle edges are computed.

$$\vec{w}_1 = [-0.66734558, 0.45638976, 0.5885221] \qquad (12)$$

$$\vec{w}_2 = [0.74210182, 0.4740619, 0.47386728] \qquad (13)$$

The normal of the plane is computed by the vectorial multiplication of the directional vectors.

$$\vec{w}_n = [-0.06272773, 0.75297656, -0.65505079] \qquad (14)$$

REFERENCES

[1] M. Sun, Q. Liu, K. Schmidt, J. Yang, N. Yao, J. D. Fernstrom, M. H. Fernstrom, J. P. DeLany, and R. J. Sclabassi, "Determination of food portion size by image processing," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2008, pp. 871–874.

[2] D. H. Lee, K. Jang, and S. Jung, "Intrinsic camera calibration based on radical center estimation." 01 2004, pp. 7–13.