

Exercises in 3D Computer Vision

Exercise 1 Introduction to numerical computing with Matlab (Exemplary Solution)

- a) Once you have added out directory to your working directory it is as easy as this:
`I=imread('lena.gif'); image(I);`
- b) `colormap(gray(256))`
- c) `image(I+50);`
- d) `I=double(I);`
- e) `I=I*2-50`
- f) `I(35,:)=zeros(1,256); I(:,75:80)=255*ones(256,6);`

Exercise 2 Homogeneous Representation of Points and Lines in Projective Space (Exemplary Solution)

- a)
- In \mathbb{R}^2 a point is represented by (2×1) vector $\mathbf{x} = (x_1, x_2)^\top$. Lines can be represented in vector-form or parametric form. The vector form of a line \mathbf{l} consists of a point on the line and a direction vector, i.e. $\mathbf{l}(\lambda) = \mathbf{p} + \lambda \mathbf{v}$. The parametric form is $\mathbf{l} : ax + by + c = 0$, where $(a, b)^\top$ is the normal vector to \mathbf{l} and, if $(a, b)^\top$ is a unit vector, $|c|$ is the distance between the line and center of the coordinate system. This notation of points and lines is called *inhomogeneous* notation.
 - In the projective space \mathbb{P}^2 , points and lines are written in *homogeneous* coordinates. Here, a point is a (3×1) vector $\mathbf{x} = (x_1, x_2, x_3)^\top$ and a line is also a (3×1) vector that contains the three entries of the parametric form in \mathbb{R}^2 , i.e., $\mathbf{l} = (a, b, c)^\top$.
 - A plane π in \mathbb{R}^3 can be represented in vector-form $\pi(\lambda, \mu) = \mathbf{p} + \lambda \mathbf{v}_1 + \mu \mathbf{v}_2$, where \mathbf{p} is a point on the plane, \mathbf{v}_1 and \mathbf{v}_2 are two (3×1) vectors parallel to the plane. It can also be represented in parametric form $\pi : n_1x + n_2y + n_3z + d = 0$, where $\mathbf{n} = (n_1, n_2, n_3)^\top$ is a vector perpendicular to π and, if \mathbf{n} is a unit vector, $|d|$ is the distance between the plane and the center of the coordinate system.
- b) For explaining the geometric meaning of homogeneous representation of points and lines, one can consider the 3-vectors to be in \mathbb{R}^3 (note, however, that here both, points and lines are just vectors). Being in the projective space means that these vectors are projected onto a subspace of \mathbb{R}^3 without losing the information of the third coordinate. For a projection, one always needs a center of projection (also for orthographic projections, where the center would be at infinity) and an image plane. Without loss of generality, one can assume the projection center to be in the origin in \mathbb{R}^3 .

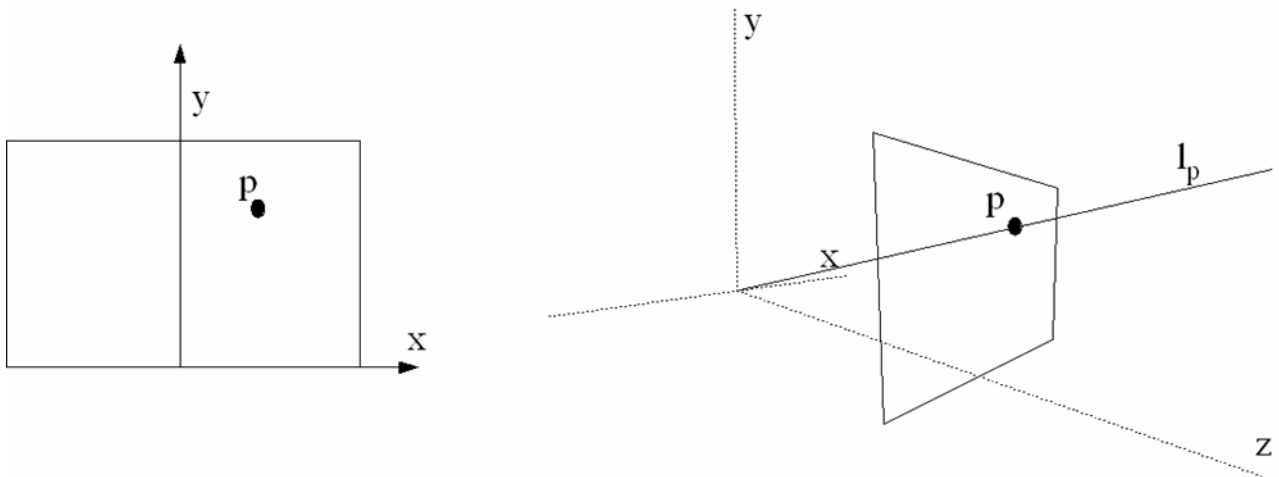


Figure 1: A point \mathbf{p} lies on a ray through image plane and projection center

- Projecting a point from 3D to 2D means to determine the intersection of the image plane with the ray connecting the projection center with the 3D point (see figure 1). Thus, a homogeneous 3-vector describing a point can be thought of as the direction vector of the ray projecting a 3D point onto a 2D image plane (Note that the direction vector is enough to describe the ray since the origin (projection center) lies on it).
- Think of a plane π_1 in 3D including the projection center. Now, define an arbitrary plane π_2 in 3D that does not include the projection center as image plane. If π_1 and π_2 are not parallel they will intersect in a line \mathbf{g} (see figure 2). The line \mathbf{g} can now be represented with the parameters of the plane π_1 , where the parametric form is $\pi_1 : ax + by + cz = 0$ (note again, that the origin lies on the plane, i.e. $d = 0$). That is, the homogeneous (3×1) vector representing a line can be thought of as the normal of the plane π_1 going through the projection center and the line.

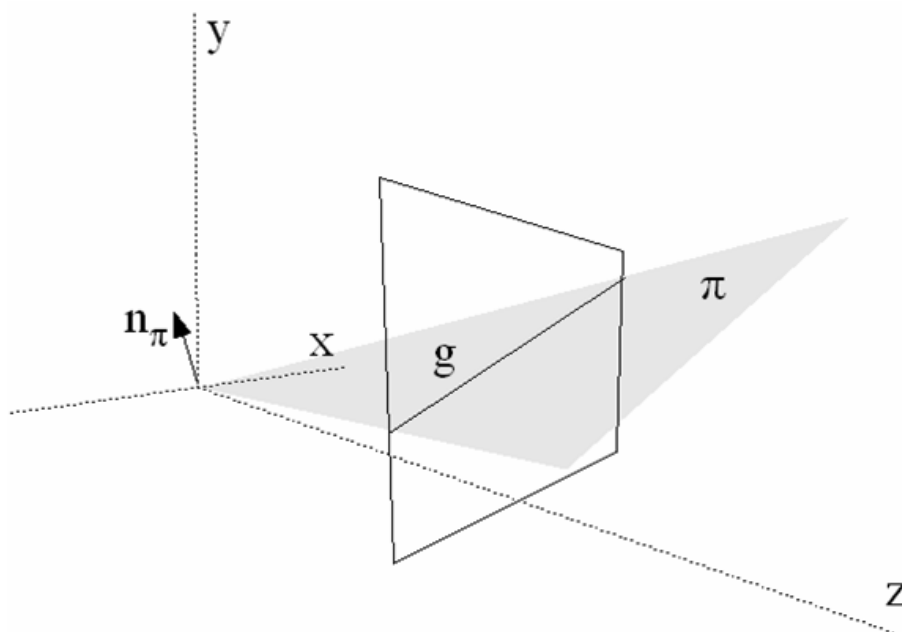


Figure 2: A line \mathbf{g} lies on a plane π through image plane and projection center

You might argue now that it is not enough to just describe points with the direction of a ray and lines with the normal of the plane since in both cases the image plane can be chosen arbitrarily. That is when one has to think of homogeneous coordinates as

equivalence classes, i.e. a class of points that all lie on this ray (see figure 3). Just

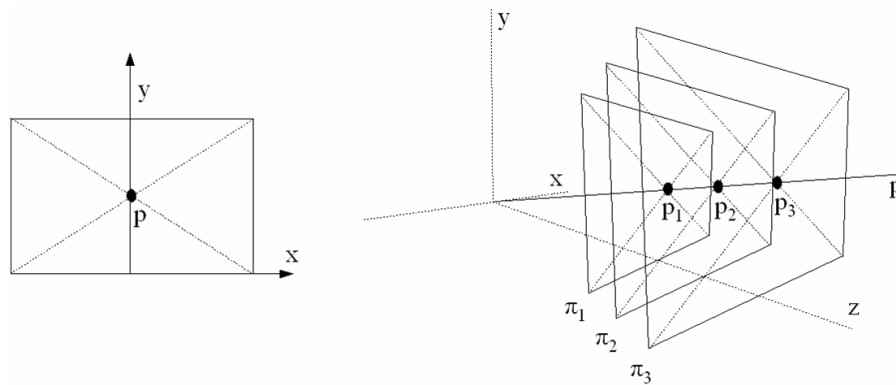


Figure 3: The ray intersects multiple image planes and thus the point \mathbf{p} represents an equivalence class, i.e. $p_1 = p_2 = p_3$

when the image plane is defined, i.e. the distance from image plane to projection center, the homogeneous representation is unique. For mathematical computations, however, it is not necessary to constrain the image plane to a certain distance from the projection center. Thus, homogeneous coordinates are always up to scale, e.g. the 2D point with homogeneous coordinates $(1, 2, 1)^\top$ is equivalent to $(2, 4, 2)^\top$ or, in general, $k(1, 2, 1)^\top$.

- c) Since we are in 2D space, the normal of a vector $(a, b)^\top$ is simply $(b, -a)$. In inhomogeneous coordinates, the direction of a line is either given (vector-form) or perpendicular to the normal described in the parametric form $\mathbf{l} : ax + by + c = 0$, i.e. $(b, -a)^\top$. In homogeneous coordinates $\mathbf{l} = (a, b, c)^\top$, as they express the parametric form implicitly, the direction vector is again by $(b, -a)$.

There are several ways to compute a line's distance to the origin. Given the vector-form, one can build the scalar product of the line's normal \mathbf{n} and a point on the line. If the normal has been normalized before (i.e. $|\mathbf{n}| = 1$, this gives the distance to the origin. In parametric form as well as in homogeneous coordinates, the distance to the origin is inherently given by the third parameter/coordinate c . This is the unnormalized distance to the origin, i.e. if the normal vector $(a, b)^\top$ is normalized, c is the distance, if not, $c/\sqrt{a^2 + b^2}$ is the distance.

- d) Taking the line \mathbf{l} again as the projection of a plane from 3D to 2D, \mathbf{p} lies on that line if it also lies on the plane π . Let $\pi : ax + by + cz = 0$. Then a point $\mathbf{p} = (p_1, p_2, p_3)^\top$ lies on this plane if $ap_1 + bp_2 + cp_3 = 0$. Written in vector-notation, it must satisfy $\mathbf{p}^\top \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0$, which is exactly the condition for a point lying on a line described by the plane's normal.

- e) Intersecting two lines with inhomogeneous notation is simple: given two lines \mathbf{l} and \mathbf{g} in the parametric form, setting $\mathbf{l}^\top \mathbf{x} = 0$ and $\mathbf{g}^\top \mathbf{x} = 0$ and solving for x_1 and x_2 . The same thing can be done in the vector form in order to find λ and μ . This will give the intersection point. It is easier, however, when dealing with homogeneous coordinates. Think again of the two lines in \mathbb{P}^2 as two planes in 3D space \mathbb{R}^3 . Building the cross-product of their normals \mathbf{n}_1 and \mathbf{n}_2 gives us a unique 3-vector \mathbf{x} that is perpendicular to both normals. This is exactly the direction of the ray representing the intersection point, i.e. $\mathbf{x} = \mathbf{n}_1 \times \mathbf{n}_2$.

The same applies when building the cross product of two homogeneous points. A 3D vector \mathbf{v} that is perpendicular to both rays is the result. \mathbf{v} describes a line where both points lie on.

- f) Two parallel lines in the Cartesian space never intersect, or intersect very very far "at the infinity". However, their images, obtained by a projective transformation intersect. The intersection point is called vanishing point, and it is the image of a point at the infinity. If we take the image of two other parallel lines (parallel in the Cartesian space), they will intersect in the image in another vanishing point. The line that passes by these two vanishing point is called a vanishing line, it is the image of the line at infinity \mathbf{l}_∞ .

Exercise 3 Vector Notation (Exemplary Solution)

a)

$$\mathbf{R}^{-1} = \mathbf{R}^\top = \begin{bmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \\ \mathbf{r}_3^\top \end{bmatrix}$$

b)

$$\mathbf{T}\mathbf{v} = \begin{bmatrix} \mathbf{R}\mathbf{x} + \mathbf{t} \\ 1 \end{bmatrix}$$

c)

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{R}^\top & -\mathbf{R}^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^\top & -\mathbf{r}_1^\top \mathbf{t} \\ \mathbf{r}_2^\top & -\mathbf{r}_2^\top \mathbf{t} \\ \mathbf{r}_3^\top & -\mathbf{r}_3^\top \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}$$

d)

$$\mathbf{T}^{-1}\mathbf{v} = \begin{bmatrix} \mathbf{r}_1^\top (\mathbf{x} - \mathbf{t}) \\ \mathbf{r}_2^\top (\mathbf{x} - \mathbf{t}) \\ \mathbf{r}_3^\top (\mathbf{x} - \mathbf{t}) \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^\top & 0 \\ \mathbf{r}_2^\top & 0 \\ \mathbf{r}_3^\top & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} - \mathbf{t} \\ 1 \end{bmatrix}$$