Geometric Fundamentals in Robotics
Homogeneous Coordinates

Basilio Bona

DAUIN-Politecnico di Torino

July 2009
Homogeneous coordinates are an augmented representation of points and lines in $\mathbb{R}^n$ spaces, embedding them in $\mathbb{R}^{n+1}$, hence using $n+1$ parameters. This representation is useful in dealing with perspective and projective transformation (computer graphics, etc.) and for rigid displacement representation.

We start introducing the concept on 2D spaces.

Given a point $\mathbf{p}$ in $\mathbb{R}^2$, represented as $P = (p_1, p_2)$, i.e., the vector $\mathbf{p} = \begin{bmatrix} p_1 & p_2 \end{bmatrix}^T$ its homogeneous representation (using homogeneous coordinates) is

$$\tilde{\mathbf{p}} = \begin{bmatrix} \tilde{p}_1 & \tilde{p}_2 & \tilde{p}_3 \end{bmatrix}^T; \quad \text{with } [0 \ 0 \ 0]^T \text{ not allowed}$$

The vector representation is obtained dividing the first $n$ homogeneous components by the $(n+1)$-th, that is often called scale.

$$p_1 = \tilde{p}_1/\tilde{p}_3; \quad p_2 = \tilde{p}_2/\tilde{p}_3$$
A geometric interpretation of the homogeneous coordinates is given in the following Figure.

\[ \pi : \{ z = 1 \} \]

\[
\begin{bmatrix}
  x \\
  y \\
  z
\end{bmatrix} \rightarrow \begin{bmatrix}
  x/z \\
  y/z
\end{bmatrix}
\]

**Figure:** Geometric interpretation of homogeneous coordinates.
We embed the plane $\pi : z = 1$ in a 3D space and we consider the planar coordinates as the intersection of a projection line that goes through the origin of the RF $\{i, j, k\}$.

As $z$ goes to $\infty$ we can represent the origin, while as $z$ goes to 0, we can represent points at infinity, along a well defined direction. The set of all points at infinity represents the line at infinity.

**Figure:** Points at $\infty$. 

```
\begin{itemize}
  \item We embed the plane $\pi : z = 1$ in a 3D space and we consider the planar coordinates as the intersection of a projection line that goes through the origin of the RF $\{i, j, k\}$.
  \item As $z$ goes to $\infty$ we can represent the origin, while as $z$ goes to 0, we can represent points at infinity, along a well defined direction. The set of all points at infinity represents the line at infinity.
\end{itemize}
```
The same approach is valid for 3D homogeneous coordinates: given a point \( \mathbf{p} \) in \( \mathbb{R}^3 \), represented as \( P = (p_1, p_2, p_3) \), or else by a vector

\[
\mathbf{p} = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}^T
\]

its homogeneous representation is

\[
\tilde{\mathbf{p}} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix}^T
\]

Hence

\[
p_1 = \tilde{p}_1/\tilde{p}_4; \quad p_2 = \tilde{p}_2/\tilde{p}_4; \quad p_3 = \tilde{p}_3/\tilde{p}_4
\]

In robotics it is common to set \( \tilde{p}_4 = 1 \). We will adhere to this convention.
Some author treats differently the homogeneous representation of a point (geometric vector $v_p$) by that of an oriented segment (physical vector $v_{ab}$). In the first case

$$\tilde{v}_p = \begin{bmatrix} v_p \\ 1 \end{bmatrix}$$

while in the second case

$$\tilde{v}_{ab} = \begin{bmatrix} v_b \\ 1 \end{bmatrix} - \begin{bmatrix} v_a \\ 1 \end{bmatrix} = \begin{bmatrix} v_{ab} \\ 0 \end{bmatrix}$$

Not all textbooks adhere to this convention, since in this case the sum of two geometric vectors produces

$$\tilde{v}_p + \tilde{v}_q = \begin{bmatrix} v_p \\ 1 \end{bmatrix} + \begin{bmatrix} v_q \\ 1 \end{bmatrix} = \begin{bmatrix} v_p + v_q \\ 2 \end{bmatrix}$$

This sum, once transformed back in 3D space, becomes not the usual parallelogram sum of segments, but the midpoint of the parallelogram (see also Affine Spaces).
Figure: Homogeneous sum of two geometric vectors.
Briefly, affine spaces are spaces where the origin is not considered to be a special point, and homogeneous coordinates are used to characterize the affine elements (points and planes).

Figure: Linear and affine subspaces.
Rigid displacements and homogeneous coordinates

Using homogeneous representation, the generic displacement \((R, t)\) is given by the following **homogeneous** transformation matrix

\[
H \equiv H(R, t) = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}
\]

since

\[
H\tilde{v} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} v \\ 1 \end{bmatrix} = Rv + t
\]

Hence, the \(SE(3)\) group transformation \(g = (R, t) \in SE(3)\), is represented by the \(4 \times 4\) homogeneous matrix \(H\) through the use of homogeneous coordinates:

\[
p_a = g_{ab} p_b \leftrightarrow \tilde{p}_a = H_{ab}(R_{ab}, t_{ab})\tilde{p}_b
\]
Transformation composition $g_{ab}g_{bc}$ is obtained by the homogeneous matrix product

$$H_{ab}(R_{ab}, t_{ab})H_{bc}(R_{bc}, t_{bc})$$

i.e.,

$$\begin{bmatrix} R_{ab} & t_{ab} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} R_{bc} & t_{bc} \\ 0^T & 1 \end{bmatrix} = \begin{bmatrix} R_{ab}R_{bc} & R_{ab}t_{bc} + t_{ab} \\ 0^T & 1 \end{bmatrix}$$

Often it is convenient to write the time derivative of a homogeneous vector as

$$\dot{\tilde{p}} = \begin{bmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \dot{p}_3 \\ 0 \end{bmatrix}$$

This notation will prove useful when dealing with twists.
Introduction to projective/perspective geometry

Homogeneous coordinates \( \tilde{p} \) are used for describing point in projective spaces \( \mathbb{P}^n \) of dimension \( n \).

Two points \( \tilde{a} \) and \( \tilde{b} \) are “equal” (\( \tilde{a} \sim \tilde{b} \)) if there exist a real \( \lambda \neq 0 \) such that

\[
\tilde{a} = \lambda \tilde{b} \quad \rightarrow \quad \tilde{b} \sim \lambda \tilde{b}
\]

collinear points are equal

this also means that vectors are equal up to a scale factor.

We say that \( \tilde{a} \) and \( \tilde{b} \) belong to the same line, i.e., to the same affine subspace of \( \mathbb{R}^{n+1} \).

Consider the planar (2D) projection through a pinhole camera of a 3D point \( P, p = [x \ y \ z]^T \). The ray of light intercepts the image plane in the 2D image point \( P_i \), whose homogeneous coordinates on the image plane are

\[
\tilde{p}_i = \begin{bmatrix} x \\ y \\ f \end{bmatrix} \quad \rightarrow \quad p_i = \begin{bmatrix} x/f \\ y/f \end{bmatrix}^T = \frac{1}{f} \begin{bmatrix} x \\ y \end{bmatrix}^T
\]

where \( f \) is the focal distance.
Figure: Example of projection through a pinhole camera.
Considering the similar triangles in Figure, we have:

\[
\frac{x}{f} = \frac{X}{Z}, \quad \frac{y}{f} = \frac{Y}{Z}
\]

If we set \( \lambda = Z/f \), where \( f \) is known and the distance \( Z \) from \( \pi \) of the point \( p \) is unknown, we have

\[
X = \lambda x, \quad Y = \lambda y, \quad Z = \lambda f
\]

Hence, if we compute the cartesian coordinates of \( p_i \) on \( \pi \) from homogeneous coordinates, we have

\[
p_i = \begin{bmatrix} x/f \\ y/f \end{bmatrix} \quad \leftrightarrow \quad \tilde{p}_i = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} \sim \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \sim \begin{bmatrix} x \\ y \\ f \end{bmatrix}
\]

and every 3D point with homogeneous coordinates \( \lambda [x \ y \ f]^T \) is projected onto the same point \( p_i \) of the image plane \( \pi \).

Notice that the above transformation is a nonlinear one.
The nonlinear mapping $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ can be expressed as a (linear) matrix transformation between the homogeneous spaces $\mathbb{R}^4 \rightarrow \mathbb{R}^3$ as

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} \sim \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \equiv \lambda \tilde{p}_i = P \begin{bmatrix} p \\ 1 \end{bmatrix}$$

Nonsingular (i.e., invertible) mappings between projective spaces are described by products of full rank square matrices $T$ by homogeneous coordinates, $\tilde{p}_a = T_{ab} \tilde{p}_b$

An invertible mapping from $\mathcal{P}^n$ onto itself preserves collinearities, i.e., if two points belong to a line, also the transformed points will belong to a line.

Affine transformations preserve not only collinearities, but also parallelism between lines and ratios of distances.

Similarity transformations preserves also angles, but not scales, while Euclidean (rigid) transformations preserve also scales.
Figure: Examples of different mappings.
A hyperplane in $\mathbb{R}^n$ is defined by the equation between a set of points $p_i$ and a set of (real) coefficients $\ell_i$

$$\ell_1 p_1 + \ell_2 p_2 + \cdots + \ell_n p_n + \ell_{n+1} = 0$$

Using the homogeneous coordinates we obtain an homogeneous linear equation in $\mathcal{P}^n$

$$\tilde{\ell}^T \tilde{p} = 0$$

where

$$\tilde{\ell} = \begin{bmatrix} \ell_1 \\ \vdots \\ \ell_n \\ \ell_{n+1} \end{bmatrix}, \quad \tilde{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_n \\ 1 \end{bmatrix}$$

The two homogeneous vectors can exchange their roles; we say that they are dual.
Remember that, given a vector $p \in \mathcal{V}(\mathcal{F})$, the set of all linear transformations $f(p) : \mathcal{V} \mapsto \mathbb{R}^1$ form another vector space $\mathcal{V}^*$, and $f \in \mathcal{V}^*$ are called dual vectors or 1–forms.

For example, the total time derivative of a scalar function $f(p)$ on $\mathbb{R}^3$ is

$$\frac{df(p(t))}{dt} = \nabla_p \phi(p) \dot{p}(t)$$

is a dual vector wrt $\dot{p}$. Another dual vector is the divergence

$$\nabla \cdot p = \left[ \frac{\partial}{\partial p_1} \cdots \frac{\partial}{\partial p_n} \right]^T \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}$$

or the Laplacian (gradient of the divergence)

$$\nabla^2 f = \nabla \cdot \nabla f = \left[ \frac{\partial}{\partial p_1} \cdots \frac{\partial}{\partial p_n} \right]^T \begin{bmatrix} \frac{\partial f}{\partial p_1} \\ \vdots \\ \frac{\partial f}{\partial p_n} \end{bmatrix}$$
Duality in projective spaces

The space of all hyperplanes in $\mathbb{R}^n$ defines another perspective space, called the dual of the original $\mathbb{P}^n$.

**Duality principle**: for all projective results established using points and planes, a symmetrical result holds interchanging the role of planes and points.

Example:

two points define a line $\leftrightarrow$ two lines define a point

Notice that the projective plane $\mathbb{P}^2$ whose points are defined by homogeneous coordinates, has more points that the usual $\mathbb{R}^2$ plane, since it includes also points at $\infty$. 
A collineation given by a full rank matrix $\mathbf{T}$ transforms points as

$$\tilde{\mathbf{p}}_a = \mathbf{T}_{ab} \tilde{\mathbf{p}}_b$$

A hyperplane $\tilde{\ell}_b \tilde{\mathbf{p}}_b = 0$ is transformed as

$$\tilde{\ell}_b \tilde{\mathbf{p}}_b = 0 \iff \tilde{\ell}_a \tilde{\mathbf{p}}_a = 0$$

hence

$$\tilde{\ell}_a \mathbf{T}_{ab} \tilde{\mathbf{p}}_b \Rightarrow \tilde{\ell}^T = \tilde{\ell}_a \mathbf{T}_{ab}$$

yielding

$$\tilde{\ell}_a = \mathbf{T}_{ab}^{-T} \tilde{\ell}_b$$

Matrix $\mathbf{T}_{ab}^{-T}$ is called the dual of $\mathbf{T}_{ab}$

Comparing this last equation with the first one, we conclude that points and lines can be transformed, but using different (dual) collineation matrices $\mathbf{T}_{ab}$ and $\mathbf{T}_{ab}^{-T}$. 

Basilio Bona (DAUIN-Politecnico di Torino)
Rigid motions
July 2009 19 / 32
We restrict our analysis on the 2D projective space $\mathcal{P}^2$; for the sake of clarity, we drop the $\tilde{}$ sign on top of the homogeneous vectors, i.e., $\tilde{a}$ is now indicated as $a$.

A line $\ell = [\ell_1 \ \ell_2 \ \ell_3]$ can be characterized by three parameters:

1. the slope $-\ell_1/\ell_2$
2. the $x$-intercept $-\ell_3/\ell_1$
3. the $y$-intercept $-\ell_3/\ell_2$

Points $p$ and lines $\ell$ are related by

$$\ell^T p = 0$$

in the sense that a point is on a line or a line include a point iff the previous relation holds.
Duality

- A line \( \ell \) passing for two points \( p_1 \) and \( p_2 \) is defined as

\[
\ell = p_1 \times p_2 = S(p_1)p_2
\]

- A point \( p \) belonging to two lines (intercept) \( \ell_1 \) and \( \ell_2 \) is defined as

\[
p = \ell_1 \times \ell_2 = S(\ell_1)\ell_2
\]

Points at infinity or **ideal points** have the following representation

\[
p_{\infty} = [p_1 \ p_2 \ 0]^T
\]

All ideal points lie on the **ideal line**, represented by

\[
\ell_{\infty} = [0 \ 0 \ 1]^T
\]
Three points $p_1, p_2, p_3$ lie on the same line (they are \textit{collinear}) iff

$$\det \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix} = 0$$

Since the duality principle holds, three lines $\ell_1, \ell_2, \ell_3$ meet at the same point (they are \textit{concurrent}) iff

$$\det \begin{bmatrix} \ell_1 & \ell_2 & \ell_3 \end{bmatrix} = 0$$
Parallelism

Two lines

\[
\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}; \quad \mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}
\]

are parallel when the intersection point is at infinity, i.e.,

\[
\mathbf{p} = \mathbf{m} \times \mathbf{n} = \begin{bmatrix} p_1 \\ p_2 \\ 0 \end{bmatrix}
\]

This reduces to the conditions

\[
\frac{m_1}{m_2} = \frac{n_1}{n_2} \quad \text{or} \quad \det \begin{bmatrix} m_1 & n_1 \\ m_2 & n_2 \end{bmatrix} = 0
\]

The direction \( k \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \) points toward the point at infinity.
<table>
<thead>
<tr>
<th>Point</th>
<th>$p = [p_1 \ p_2 \ p_3]^T$</th>
<th>Line</th>
<th>$\ell = [\ell_1 \ \ell_2 \ \ell_3]^T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence</td>
<td>$p^T\ell = 0$</td>
<td>Incidence</td>
<td>$\ell^T p = 0$</td>
</tr>
<tr>
<td>Line</td>
<td>$\ell = p_1 \times p_2$</td>
<td>Point</td>
<td>$p = \ell_1 \times \ell_2$</td>
</tr>
<tr>
<td>Collinearity</td>
<td>$\det [p_1 \ p_2 \ p_3] = 0$</td>
<td>Collinearity</td>
<td>$\det [\ell_1 \ \ell_2 \ \ell_3] = 0$</td>
</tr>
<tr>
<td>$\infty$ Point</td>
<td>$[p_1 \ p_2 \ 0]^T$</td>
<td>$\infty$ Line</td>
<td>$[0 \ 0 \ \ell_3]^T$</td>
</tr>
</tbody>
</table>
The generic **projective** transformation (also called **homographies**) in \( \mathcal{P}^2 \) is represented by a non singular \( 3 \times 3 \) matrix (acting on homogeneous vectors), whose elements are

\[
T = \begin{bmatrix}
t_{11} & t_{12} & t_{13} \\
t_{21} & t_{22} & t_{23} \\
t_{31} & t_{32} & t_{33}
\end{bmatrix}
\]

Since the transformations are equivalent up to a scale factor, the number of dof in \( T \) is 8.

According to structural constraint, we have a hierarchy of transformations, from the most general to the least one; in particular

(\textbf{Transformations in } \mathcal{P}^2)

(most general) projective \( \rightarrow \) affine \( \rightarrow \) similarity \( \rightarrow \) Euclidean (rigid)
The affine transformation must preserve the ideal line and the ideal points (as well as parallelism); hence, for any arbitrary scaling factor $\lambda$

$$\lambda \begin{bmatrix} p_1 \\ p_2 \\ 0 \end{bmatrix} = T \begin{bmatrix} p_1 \\ p_2 \\ 0 \end{bmatrix}$$

that implies

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix} \rightarrow \text{dof}=6$$
The similarity transformation must preserve angles and ratios of lengths; that implies

\[
T = \begin{bmatrix}
\cos \theta & -\sin \theta & t_{13} \\
\sin \theta & \cos \theta & t_{23} \\
0 & 0 & t_{33}
\end{bmatrix} \rightarrow \text{dof}=3
\]

with \( t_{33} \neq 1 \).

The effects of the last column in \( T \) is to produce a shear i.e., a scale variation that is non uniform in all directions.
The Euclidean (rigid) transformation must preserve also lengths; that implies

$$ T = \begin{bmatrix} \cos \theta & -\sin \theta & t_{13} \\ \sin \theta & \cos \theta & t_{23} \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \text{dof}=3 $$
Homogeneous coordinates were introduced in 1827 by Möbius in the context of barycentric coordinates. Given a triangle, the barycentric coordinates of a point $P$ are the weights required at the triangle vertices such that $P$ become the center of gravity of the triangle.

The point is computed as $p = \begin{bmatrix} w_a x & w_b y & w_c z \end{bmatrix}^T$, with

$$w_a + w_b + w_c = 1$$

Figure: Example of barycentric coordinates.
There exist a link between homogeneous coordinates and quaternions that goes through the Möbius transformation. Given two complex numbers \( x = x_1 + jx_2, \ y = y_1 + jy_2 \in \mathbb{C} \), the ordered pair of complex numbers 

\[
\tilde{z} = \begin{bmatrix} x \\ y \end{bmatrix}^T \in \mathbb{C}^2, \quad \tilde{z} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}^T
\]

are the homogeneous coordinates of \( z = z_1 + jz_2 \)

\[
z = \frac{x}{y} = z_1 + jz_2 = \frac{x_1 + jx_2}{y_1 + jy_2}
\]

To every pair of complex numbers \((x, y)\) a unique complex number \(z\) corresponds, while for every complex number \(z\) there is an infinite number of complex pairs (projective transformation).
Transformation between homogeneous representations

\[ \tilde{z} = \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \tilde{w} = \begin{bmatrix} u \\ v \end{bmatrix} \text{ where } \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \]

where \( a, b, c, d \) are complex numbers.

We can compute the transformed complex number as

\[ w = \frac{u}{v} = \frac{ax + by}{cx + dy} = \frac{a \frac{x}{y} + b}{c \frac{x}{y} + d} = \frac{az + b}{cz + d} \]

This is nothing else that the Möbius transformation.

Möbius transformation

\[ w = \frac{az + b}{cz + d}; \quad a, b, c, d, w, z \in \mathbb{C} \]
If we take a particular form of the Möbius transformation, due to Gauss:

\[ w = \frac{az + b}{-bz + a} \]

we know that every rotation of the sphere can be expressed by a complex function as the one above, and we notice that it can be represented by the matrix of its coefficients

\[
\begin{bmatrix}
a & b \\
-b^* & a^*
\end{bmatrix} = \begin{bmatrix}
\alpha + j\beta & \gamma + j\delta \\
-\gamma + j\beta & \alpha - j\beta
\end{bmatrix}
\]

\[
= \alpha \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} + \beta \begin{bmatrix}j & 0 \\ 0 & j\end{bmatrix} + \gamma \begin{bmatrix}0 & 1 \\ -1 & 0\end{bmatrix} + \delta \begin{bmatrix}0 & j \\ j & 0\end{bmatrix}
\]

\[
= \alpha 1 + \beta i + \gamma j + \delta k
\]

\[
= \text{a quaternion}
\]

For further details see: T. Needham, Visual Complex Analysis, OUP, 1997.