ECE 661
PURDUE UNIV.
AVI KAK
Lecture 19

Some Cool Properties of The Camera Projection Matrix

Reference: Multiple View Geometry in Computer Vision by Hartley and Zisserman

- The 3x4 camera projection matrix P that you learned in the previous lecture is a pretty wonderous thing. No matter how you slice and dice it, every part you get has a useful interpretation.
- Terminology: From this point on, I'll refer to the Center of Projection as the Camera Center. Additionally, I may refer to the Camera Projection Matrix P as just the Projection Matrix or as just the camera itself.
- Let's start by determining the location of the camera center for a given 3x4 matrix P.

Given a P, Where is the Camera Center?

- The 3x4 matrix P has exactly one null vector and that vector is the homogeneous representation of the camera center of the camera to which P corresponds.
- To prove the above assertion, we first note that P maps a four-dimensional space to a three-dimensional space. Therefore, P has at least a one-dimensional null space. Therefore, there must exist a vector $\vec{C}$ such that $P\vec{C} = \vec{O}$. We next prove that such a vector $\vec{C}$ must be the camera center for the given P. Consider a line that passes through such a point $\vec{C}$ and another point $\vec{A}$ that can be anywhere. The points on this 3D line are expressed as a vector span of $\vec{C}$ and $\vec{A}$: $\vec{X} = \lambda_1\vec{C} + \lambda_2\vec{A} \equiv \vec{C} + \lambda\vec{A}$. Let $\vec{x}(\lambda)$ be the pixel that corresponds to $\vec{X}(\lambda)$. By definition, $\boxed{\vec{x}(\lambda) = P\vec{X}(\lambda) = \lambda P\vec{A}}$ since $P\vec{C} = 0$. However, our use of homogeneous coordinates implies that $\lambda P\vec{A} \equiv P\vec{A}$. This means, all points on the line passing through $\vec{C}$ and $\vec{A}$ are mapped to exactly the same point in the image plane. From the geometry of camera imaging, we know that this can happen only if $\vec{C}$ is at the CoP.
- Armed with the fact that the null vector of P is the camera center, I'll now derive formulas for the camera centers for three different kinds of cameras: (1) pinhole cameras, (2) Finite Projective cameras, and (3) General Projective cameras.
- For the pinhole camera, let's consider the camera $P = KR[I|-\tilde{C}]$ where $\tilde{C} = \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix}$ is the translation vector from the world frame to the camera frame. It is trivial to show that $\vec{C} = (c_x, c_y, c_z, 1)$ is a null vector of P:

$$P\vec{C} = KR\begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 1 & -c_z \end{bmatrix}\begin{pmatrix} c_x \\ c_y \\ c_z \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This is just as well since we know already from a visual examination of P that its camera center is at $\begin{pmatrix} c_x \\ c_y \\ c_z \\ 1 \end{pmatrix}$ in homogeneous coordinates.

- Let's next consider the more general case of a finite projective camera that is only constrained by the property that the first three columns of P must form a non-singular submatrix. As was shown in Lecture 18, the projection matrix P of such cameras is expressed as $P = [M | \vec{p}_4]$ where M is a 3x3 nonsingular submatrix. It can be shown straightforwardly that the null vector of such a P is

$$\vec{C} = \begin{pmatrix} -M^{-1}\vec{p}_4 \\ 1 \end{pmatrix}$$

To verify:

$$[M|\vec{p}_4]\begin{pmatrix} -M^{-1}\vec{p}_4 \\ 1 \end{pmatrix} = [I|\vec{p}_4]\begin{pmatrix} -MM^{-1}\vec{p}_4 \\ 1 \end{pmatrix} = [I|\vec{p}_4]\begin{pmatrix} -\vec{p}_4 \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & p_{4,x} \\ 0 & 1 & 0 & p_{4,y} \\ 0 & 0 & 1 & p_{4,z} \end{bmatrix}\begin{pmatrix} -p_{4,x} \\ -p_{4,y} \\ -p_{4,z} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

- Finally, let's consider a general projective camera that is NOT a finite projective camera. We can again write $P = [M|\vec{p}_4]$ but now we assume that M is singular. (while P still has rank 3, as mentioned in Lecture 18). Since M is singular, it cannot be of full rank. That is, M has at least a one dimensional null space. Therefore, there must exist a 3-vector $\vec{d}$ such that $M\vec{d} = \vec{0}$. Given such a $\vec{d}$, it is easy to show that

$$\vec{C} = \begin{pmatrix} \vec{d} \\ 0 \end{pmatrix}$$

is the null vector of P and therefore its camera center. In fact,

$$[M|\vec{p}_4]\begin{pmatrix} \vec{d} \\ 0 \end{pmatrix} = M\vec{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Note the very important conclusion: The center of projection of a camera that is truly general projective is located at infinity.

- One last thing before I end this section: How are we to interpret $P\vec{C} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ when $\vec{C}$ is the camera center? The homogeneous vector $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ represents an undefined point in the image plane. The camera center is one unique point in the world frame for which an image point is <u>not</u> defined.

What do the Columns of P Stand For?

- Let's express the 3x4 matrix P as $P = [\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4]$

- We now show that the first three colums, $\vec{p}_1$, $\vec{p}_2$, and $\vec{p}_3$, are the vanishing points of the ideal points along the three cardinal directions in the world coordinate frame. That is, $\vec{p}_1$ is the image pixel in homogeneous coordinates of the point at infinity along the world X-axis. Similarly, $\vec{p}_2$ is the image pixel for the point at infinity along the world Y-axis. The same goes for $\vec{p}_3$ vis-a-vis the world-Z axis.

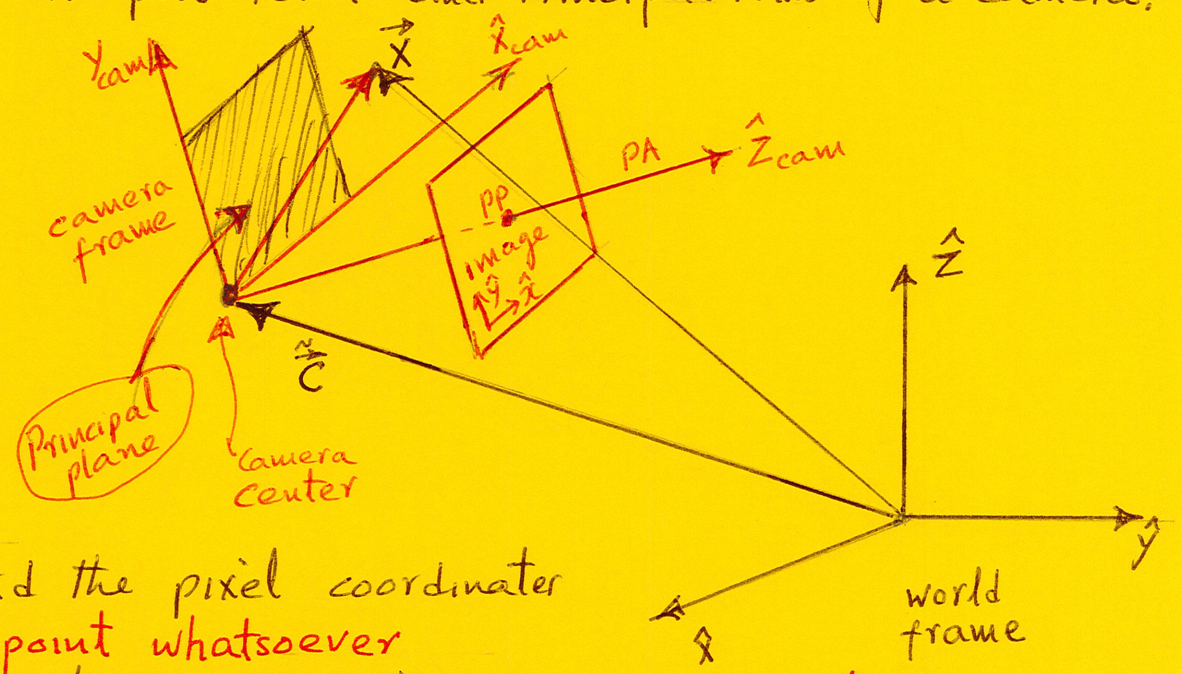
- In order to prove what $\vec{p}_1$ stands for, note that the ideal point along the world X-axis is given by $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$. Since $\boxed{P\cdot\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \vec{p}_1}$ we have that the image of this ideal point is formed at the pixel whose homogeneous representation is $\vec{p}_1$. By the same reasoning, $\vec{p}_2$ is the image pixel for the ideal point along the world Y-axis, and $\vec{p}_3$ is the image pixel for the ideal point along the world Z-axis.

- That brings us to the last column of P. The last column, $\vec{p}_4$, is the image of the world origin. The homogeneous representation of the world origin is $\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$. Its image is $\boxed{P\cdot\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = \vec{p}_4}$.

What do the Rows of P Stand For?

- This section shows that each row of P is the homogeneous representation of a plane and the fact that the three planes corresponding to the three rows of P intersect at the camera center. But first we need to define the notion of the Principal Plane of a camera. You already know from Lecture 16 what we mean by the Principal Point and Principal Axis of a camera.

- As shown in the figure at right, the Principal Plane is the plane that is parallel to the image plane and that passes through the camera center.



- The camera matrix P is supposed to allow us to find the pixel coordinates for the image of any world point whatsoever (except for just one world point – the CoP of the camera). So let's see what image pixels correspond to the world points on the Principal Plane. Let's consider the point $\vec{X}$ shown. In general, we obtain the image pixel of a world point $\vec{X}$ by finding the point where the ray from the camera center to $\vec{X}$ pierces the image plane. When we do the same for the ~~pou~~ world point $\vec{X}$ shown, we get a pixel located on l_∞ in the image plane. The pixel coordinates of this image point will by $\begin{pmatrix}x\\y\\0\end{pmatrix}$ where the ratio of x and y would be dictated by the orientation of the line joining the camera center with $\vec{X}$.

- Let's now rewrite P in the following form:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} \vec{p}^{1T} \\ \vec{p}^{2T} \\ \vec{p}^{3T} \end{bmatrix}$$

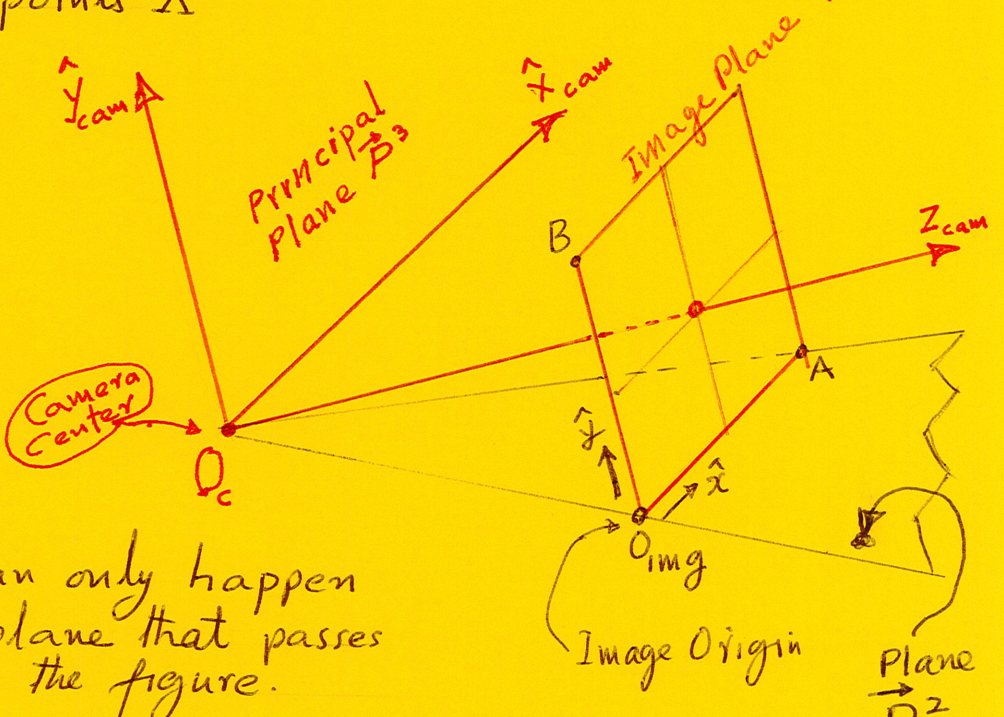
$\vec{p}^1$, $\vec{p}^2$, and $\vec{p}^3$ are the individual row vectors of P.

$$\vec{p}^1 = \begin{pmatrix}p_{11}\\p_{12}\\p_{13}\\p_{14}\end{pmatrix} \quad \ldots$$

- Obviously, for any world point $\vec{X}$, $P\vec{X} = \begin{pmatrix}\vec{p}^{1T}\vec{X}\\ \vec{p}^{2T}\vec{X}\\ \vec{p}^{3T}\vec{X}\end{pmatrix}$

- We know that when $\vec{X}$ is on the Principal Plane, $P \cdot \vec{X} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ for some (x, y). This implies $\vec{P}^3 \cdot \vec{X} = 0$ for all $\vec{X}$ on the Principal Plane. That means $\vec{P}^3$ must be the homogeneous representation of the Principal Plane.

- Let's now try to identify the world plane that is represented by the first row of P. Whatever this plane is, any world point $\vec{X}$ on that plane must satisfy $\boxed{\vec{P}^{1T} \cdot X = 0}$. That implies that for all such world points, the image pixels will have homogeneous coordinates of the form $\begin{pmatrix} 0 \\ y \\ w \end{pmatrix}$. This would be the case for all world points $\vec{X}$ that are on the plane that passes through the points O_c, O_{img}, and B on the right.



- It is easier to visualize the plane represented by the second row of P. For all such world points $\vec{P}^{2T} \cdot \vec{X} = 0$, implying that the pixel coordinates for all such world points must be of the form $\begin{pmatrix} x \\ 0 \\ w \end{pmatrix}$. This can only happen for the world points that are on the plane that passes through the points O_c, O_{img}, and A in the figure.

- The ~~pl~~ world planes $\vec{P}^1$ and $\vec{P}^2$ are called the Axis Planes of a camera.

Given a P, Where is the Principal Point in the Image?

- As you know, the Principal Point is where the Principal Axis pierces the image plane.

- Even if you have no need to find the image coords of the Principal Point of a camera, the result shown below is utterly "beautiful" — it shows the submatrix M multiplying its own last row to yield something that has practical significance. Remember, we sometimes write $P = [M | \vec{p}_4]$

- We will find the Principal Point by projecting into the image plane the ideal point along the Principal Axis. The Principal Axis is perpendicular to the Principal Plane whose homogeneous representation is $\begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{pmatrix}$.

- From Lecture 6, you already know that if the homogeneous representation of a plane is $(p_{31}, p_{32}, p_{33}, p_{34})^T$, then the normal to the plane is encoded in the physical 3-vector $\begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \end{pmatrix}$. From Lecture 6, you should also know that the ideal point along this normal would be given by $\begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \\ 0 \end{pmatrix}$.

- The pixel coords of the image of this ideal point are given by:

The last row of M is made up of the first 3 elements of the last row of P.

$$P \cdot \begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \\ 0 \end{pmatrix} = [M | \vec{p}_4] \begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \\ 0 \end{pmatrix} = M \cdot \begin{pmatrix} p_{31} \\ p_{32} \\ p_{33} \end{pmatrix} = M \cdot \begin{pmatrix} m_{31} \\ m_{32} \\ m_{33} \end{pmatrix}$$

which is a homogeneous representation of pixel coords of the Principal Point

Backprojecting an Image Pixel into World Frame

- Backprojecting image pixels into the world frame has useful applications in multiview scene reconstruction.
- For each pixel $\vec{x} = \begin{pmatrix} x \\ y \\ w \end{pmatrix}$, we define a ray as the line that passes through the camera center and the pixel. This ray is our backprojection of the pixel.
- If, besides the camera center, we had one additional world point on the ray, we could then define a vector span from the two points for a homogeneous representation of the ray. [See Lecture 6 for how a vector span is used to represent a line in world 3D.]
- We now claim that, for a given pixel $\vec{x} = \begin{pmatrix} x \\ y \\ w \end{pmatrix}$, there exists a world point $P^+\vec{x}$ on the corresponding ray where $\boxed{P^+ = P^T(PP^T)^{-1}}$ is the pseudoinverse of P. This claim is based on the observation that the location of the image of this world point is the same as $\vec{x}$:

$$P\cdot(P^+\vec{x}) = P\cdot\left(P^T(PP^T)^{-1}\vec{x}\right) = PP^T(PP^T)^{-1}\vec{x} = \vec{x}$$

Since P is of rank 3, the 3x3 matrix PP^T is of full rank. The inverse $(PP^T)^{-1}$ is therefore guaranteed to exist.

- The third bullet above mentioned using a vector span of the camera center and the world point $P^+\vec{x}$ for a homogeneous reprentation of the backprojected ray in world 3D. By a vector span formed by two given points, we mean a linear combination of the two points. This means that a world point on the ray can be expressed as $\vec{X} = \lambda_1 C + \lambda_2 P^+\vec{x}$ for any arbitrary values of the coefficients λ_1 and λ_2. Since $\vec{X}$ is homogeneous, we can also write $\vec{X} = \lambda_1\left(\vec{C} + \frac{\lambda_2}{\lambda_1} P^+\vec{x}\right)$ $\equiv \vec{C} + \lambda P^+\vec{x}$ where $\vec{C}$ is the 4-vector homogeneous representation of the camera center. So we can say $\vec{X}(\lambda) = \vec{C} + \lambda P^+\vec{x}$ is the parameterized representation of the world points on the backprojected ray through a given pixel $\vec{x}$.

Cameras at Infinity

- Despite the sense conveyed by "infinity" in the title of this section, these cameras are of great practical importance when images are formed under orthographic and weak perspective conditions.
- The second bullet on page 19-2 shows that a general projective camera (that is not finite projective) has the following property: The submatrix M of $P = [M|\vec{p}_4]$ is singular and that the null vector $\vec{d}$ of M gives us the camera center $\vec{C} = \begin{pmatrix} \vec{d} \\ 0 \end{pmatrix}$ at infinity for such a camera.

- We will now consider a sub-category of such general projective cameras — the sub-category is known as the Affine Camera.
- Under Affine Cameras, we will consider: ① Orthographic Projection Camera, ② Scaled Orthographic Projection Camera, and ③ Weak Perspective Projection Camera.
- In general, a camera is affine if the last row of the 3x4 P is $(0,0,0,1)$. Obviously, with $P=[M|\vec{p}_4]$, M for an affine camera is given by $M=\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 0 \end{bmatrix}$.
- Orthographic Projection Camera: In its generic form, this affine camera is defined by $P=\begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{bmatrix}$. Since $\begin{pmatrix}0\\0\\1\end{pmatrix}$ is the null vector of M in this case, the CoP of this camera is at $\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$, which is a point at ∞ along the world-Z axis. This camera gives us an orthographic projection of the object onto the xy-plane. (By orthographic we mean projection with a bundle of parallel rays. Obviously, only those object points would be projected that are visible to the camera.) For other poses of the camera, we introduce a rotation R and a translation $\vec{t}$ from the world frame to the camera frame. Following the same logic as on page 18-3 of Lecture 16, we can write $\vec{x}=\begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{bmatrix}\begin{bmatrix} R & \vec{t} \\ \vec{0}^T & 1 \end{bmatrix}\vec{X} = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{bmatrix}\begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{r}^{3T} & t_3 \\ \vec{0}^T & 1 \end{bmatrix}\vec{X} = \begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{0}^T & 1 \end{bmatrix}\vec{X}$ world point where $\vec{r}^1$, $\vec{r}^2$, and $\vec{r}^3$ are the three row vectors of R, and where t_1, t_2, t_3 are the three components of $\vec{t}$. Therefore, the projection matrix P of an ortho projection camera at pose $(R, \vec{t})$ with respect to the world frame is given by $P=\begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{0}^T & 1 \end{bmatrix}$. By world frame here we mean the coord frame in which the generic config is defined
- Scaled Orthographic Projection Camera: A scaled orthographic is a regular orthographic projection camera as described above that is followed by isotropic scaling of the image. What that means is that whatever image is produced by the P matrix of the previous bullet needs to be scaled by a factor k. Therefore, the new P would be given by $P=\begin{bmatrix} k&0&0 \\ 0&k&0 \\ 0&0&1 \end{bmatrix}\cdot\begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{0}^T & 1 \end{bmatrix} \equiv \begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{0}^T & 1/k \end{bmatrix}$. For both the P here and in the previous case, the first two rows of M submatrix of P must be orthogonal.
- Weak Perspective Projection Camera: This camera differs from the one in the previous case in only one sense: we now use two different scale factors for the $\hat{x}$ and the $\hat{y}$ directions in the image plane. We now write for the projection matrix:

$$P=\begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \vec{r}^{1T} & t_1 \\ \vec{r}^{2T} & t_2 \\ \vec{0}^T & 1 \end{bmatrix}$$

The first two rows of the M submatrix of P would still be orthogonal.

- IMPORTANT: The last two affine camera models — Scaled and Weak-Perspective Cameras — are often used for images of distant objects recorded with long focal-length lenses and when the objects subtend small angles at the camera center.