Epipolar Geometry & PnP [Beta]

> 2-view Geometry: Fundamental Matrix F & Computation, RANSAC, Stereo Camera, Triangulation; and PnP.

💡 Upto date online link: https://www.notion.so/saishubodh/Epipolar-Geometry-PnP-Beta-de763584b7cb435cb63198bc997ba2a0

- Instructions to students
- 0. Revisiting single view geometry
  - 0.1 Difference between the ray and the image coordinates
- 1. Epipolar Geometry
  - 1.1 Introduction
  - 1.2 Motivation
  - 1.3 Intuition
  - 1.4 Some definitions
- 2. Fundamental matrix F
  - 2.1 Derivation of F
    - 2.1.1 Geometric Derivation
    - 2.1.2 Algebraic Derivation
  - 2.2 Computation of F
    - 2.2.1 Normalized 8 point algorithm
    - 2.2.2 Decomposition of E
    - 2.2.3 RANSAC: A quick overview
- 3. Stereo Camera
- 4. Triangulation
  - If they intersected
  - If they don't intersect
- 5. PnP
  - 5.0 Introduction
    - 5.0.1 What is the Perspective n Points (PnP) problem?

5.0.2 The P3P/Spatial Resection Problem
5.0.3 Difference between P3P and DLT
5.1 Solution to P3P
5.1.1 **Revisiting normalized coordinates**
5.1.2 **Two step process**
5.1.3 Length of projection rays
5.1.4 Transformation between camera frame and world frame
5.1.5 Additional Reading

> Page Author: Shubodh Sai

▼ Read this to know how to use this page the right way.

1. Firstly, enable dark mode.
2. Whenever you're studying this page, try to think critically and spend enough time wherever a question is asked before you "toggle" for the answer.
3. Any logged in Notion user can comment on this page. This is a wonderful way to collaborate. Whether you have doubts or a better way to explain a concept or even correct a typo, feel free to comment.
4. **DISCLAIMER:** This page is still in Beta stage and there are quite some non-reproducible elements. The final version (along with all the information on how to use it, say complementary Chrome extensions for example) will be updated by end of July. See next block for more info.

▼ For more information and/or if you're curious about this unorthodox way of teaching,

Follow this. This page is a part of a bigger project. By the end of July, the author will be making pages like this one for all of the MVG (transformations + projective + multi-view at the least) and will enable "Duplicate" option on Notion and make all the LaTeX code available on GitHub page for easier **reproducibility**. Once that is done, you can say goodbye to making your own notes and use these pages as starting points (Also, farewell to all those ridiculous course components like notes scribing!). According to a survey, the single biggest hurdle for learning is lack of good maintainable notes, this is a step towards overcoming that hurdle — even for a layman who doesn't want to use LaTeX (you can just drag and drop images of stuff you write on paper onto Notion). Since you bothered to read till here, let me shamelessly ask you to consider starring the repo if you find these ideas useful in your academic life — it'll put pressure on me to make it available sooner.😃

Instructions to students

1. Get a pen and paper: Let's get our hands dirty! (Can use the whiteboard on Teams for writing your answers during the class)
2. Participate! Don't feel shy.

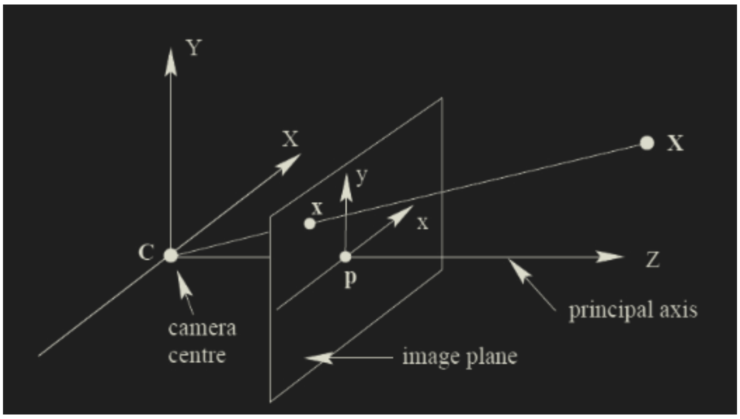
https://youtu.be/DgGV3l82NTk

Jokes aside, this song does explain a lot of things! It actually might have answer(s) to your assignment (TODO), so do watch the video at the end of my lecture again. 😉

Source: Cyrill Stachniss videos & Zisserman book unless explicitly mentioned.

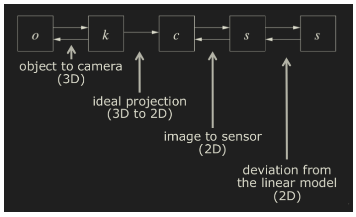
0. Revisiting single view geometry

0.1 Difference between the ray and the image coordinates



Pin hole camera

Few clarifications:



$$\begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

▼ Coordinates of image point in camera coordinate system? Let's call it **normalized image coordinates**.

Ans: $[x, y, f]^T$

▼ How to arrive at the above vector from homogenous **image coordinates** `[x,y,1]`? Forget about the scaling factor.
Clue ⇒

$$\lambda x = KX$$

Ans: K^{-1} . Consider $K^{-1} \times [x, y, 1]^T$ = ?

1. Epipolar Geometry

1.1 Introduction

1 view, 2 view and n view geometry.

- Intrinsic projective geometry between two views.

- Independent of scene structure.
- Dependent only on camera parameters and relative pose.

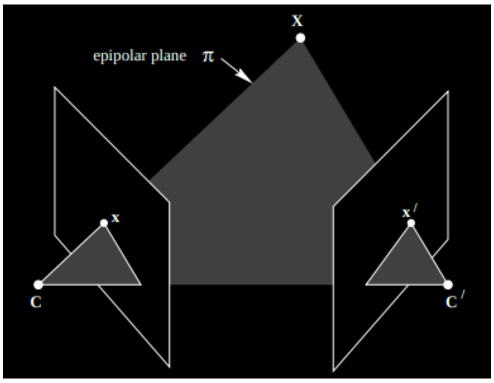
1.2 Motivation

All parameters: World points, image points, camera matrix, relative pose.

Think of a camera navigating. Typical robotics application.

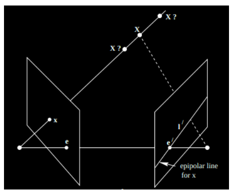
1. Restricts search for the corresponding points given F. (What is F?)
    - Knowns: Given P, R, t. Unknown: Corresponding point
2. F and its decomposition given enough corresponding points.
    - Knowns: Corresponding Points. Unknown: P, R, t

1.3 Intuition



The question: how is **x'** constrained? Two ways of looking at it:

▼ 1st way: figure



1. Knowns: Given P, R, t. Unknown: Corresponding point

▼ 1st way: Backproject **x**

Joining the 3D ray points to 2nd camera center, you end up with a "epipolar" line.

▼ 2nd way: 2 lines/rays make a plane

Intersection of the plane formed by baseline and first ray
with second image plane gives an "epipolar" line.

1.4 Some definitions

▼ Epipole: ***What is the image of C' in first view?***

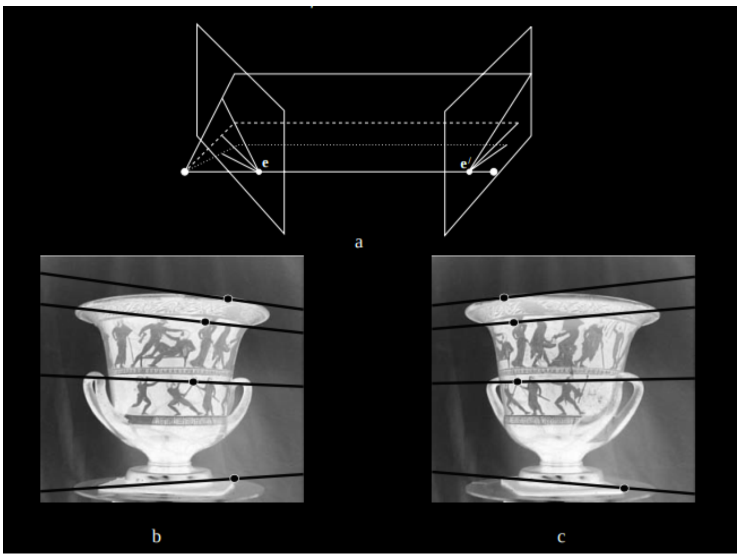
Also, point of intersection of the "baseline" (**CC'**) with the image planes.

▼ Epipolar plane: a plane containing baseline

▼ Epipolar line: ***Try using above "epi plane" definition.***

- Intersection of epipolar plane with the image plane.
- All epipolar lines intersect at?

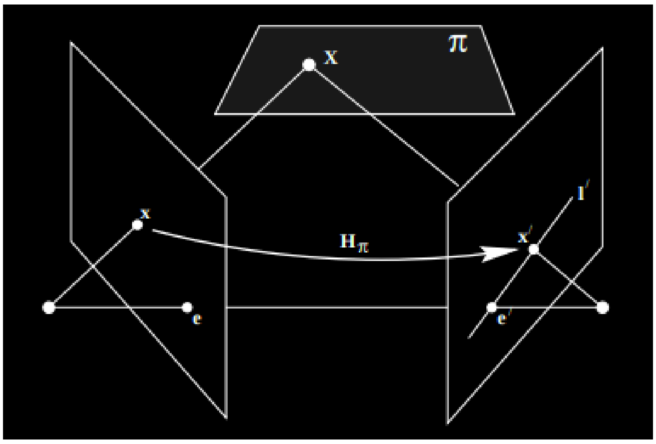
▼ Epipolar lines: Figure



2. Fundamental matrix F

2.1 Derivation of F

2.1.1 Geometric Derivation



▼ Clarification on notation

x' above is a **potential** corresponding point, it actually forms a family of points that is on the line l'. So for the first step of this derivation you can take any such point as you're just looking for a point on the epipolar line.

Two steps:

▼ 1. Point transfer via a **plane π**

▼ Revisiting projective geometry: Are lines mapped to lines?

Means it is a projectivity transform: x is projectively equivalent to X. So is x'.

Hence, projectivity theorem states that $\exists$ full-rank homography matrix $\mathbf{H}_\pi$ s.t.

$$\mathbf{x}^{'} = \mathbf{H}_\pi \mathbf{x}$$

▼ 2. Constructing epipolar line (Given two points, what is equation of line?)

$$\mathbf{l}^{'} = \mathbf{e}^{'} \times \mathbf{x}^{'} = [e^{'}]_\times \mathbf{x}^{'} = [e^{'}]_\times \mathbf{H}_\pi \mathbf{x} = \mathbf{F}\mathbf{x}$$

▼ F rank above?

Homography is a rank 3 invertible matrix, skew symmetric matrix is rank 2. So rank of F is 2.

▼ Given a point passing through a line, what's the equation?

$$\mathbf{x}^{'\mathrm{T}} \mathbf{F} \mathbf{x} = 0$$

Finally: Any point on the epipolar line would satisfy the above equation, so if some point in 2nd image y' satisfies the equation, that doesn't mean it is the corresponding point. BUT if you have an accurate corresponding point, then that MUST satisfy the condition. So the above is a necessary but not a sufficient condition for a point in 2nd image to be a corresponding point.

2.1.2 Algebraic Derivation

- Notation clarification: R_{21} is same as R_1^2 which is "1st to 2nd frame"

$$O_1 = \left[\ I_{3\times3} \quad | \quad 0_{3\times1}\ \right] \qquad\qquad O_2 = \left[\ R_{12} \quad | \quad t_{12}\ \right]_{3\times4}$$

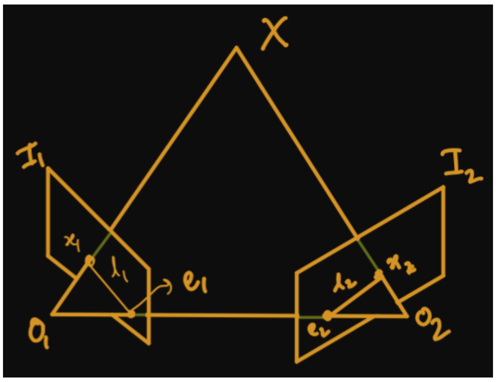
$$\lambda_1\vec{p_1} = K[I \quad 0]\vec{X}_{4\times1} \qquad\qquad \lambda_2\vec{p_2} = K[R_{21} \quad t_{21}]\vec{X}_{4\times1}$$

$$\lambda_1 K^{-1}\vec{p_1} = \vec{X} \qquad\qquad \lambda_2 K^{-1}\vec{p_2} = [R_{21} \quad t_{21}]\vec{X}$$

$$where \quad K^{-1}\vec{p_1} = \vec{x_1} \qquad\qquad where \quad K^{-1}\vec{p_2} = \vec{x_2}$$

💡 What really is $K^{-1}\vec{p}$?

O_1O_2P forms a triangle, the epipolar plane. So?



$$\implies \overrightarrow{O_1P} \cdot (\overrightarrow{O_1O_2} \times \overrightarrow{O_2P}) = 0$$
$$\implies \lambda_1\vec{x_1} \cdot \left(\vec{t_{12}} \times R_{12}\lambda_2\vec{x_2}\right) = 0$$
$$\implies \lambda_1\lambda_2\vec{x_1}\left(\vec{t_{12}} \times R_{12}\vec{x_2}\right) = 0$$
$$\implies \vec{x_1}^T\,[\mathrm{t}_{12}]_\times\ R_{12}\ \vec{x_2} = 0$$
$$\implies \vec{x_1}^T\ E_{12}\vec{x_2} = 0$$
$$\implies \vec{P_1}^T\ F_{12}\vec{P_2} = 0$$
$$where\ E_{12} = [\mathrm{t}_{12}]_\times\ R_{12}$$
$$\&\quad F_{12} = K^{-T}[\mathrm{t}_{12}]_\times\ R_{12}K^{-1}$$

> F_{12} is the **fundamental matrix** that related 2nd image to the 1st in terms of **pixel coordinates**. E_{12} is the **essential matrix** that related in terms of **normalized pixel coordinates.**

2.2 Computation of F

We have seen:

$$\mathrm{x}'^{\mathrm{T}}\mathrm{F}\mathrm{x} = 0$$

- Now, the natural question is: Can I find **F** given **enough** correspondences? **Enough?**

$$\mathbf{x} = (x, y, 1)^{\mathrm{T}} \text{ and } \mathbf{x}' = (x', y', 1)^{\mathbf{T}}$$

Expanding:

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1)\,\mathbf{f} = 0$$

From "n" points:

$$\mathrm{Af} = \begin{bmatrix} x_1'x_1 & x_1'y_1 & x_1' & y_1'x_1 & y_1'y_1 & y_1' & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x_n'x_n & x_n'y_n & x_n' & y_n'x_n & y_n'y_n & y_n' & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = \mathbf{0}$$

▼ Revisiting Linear Algebra `Ax=0` ...

1. If full rank (rectangular or square),
   1. If (n=9)>m, then you have these special solutions (as there is at least one free column. If rank less, even better). As the solution set is a linear combination of these solutions, it is infinite (at least a line).
   2. If m>(n=9), the below rows in the matrix can be made zero and you end up with only solution as zero.
   3. If m=n, full rank so inverstible only solution is zero vector.
2. If not full rank, you always have infinite solutions.

▼ **What rank must A have in our case for the solution to be (non-zero and) unique?**

8. However, with noisy data it's rare to be 8. It will be 9 (with n=9 or more points).

▼ What happens if less or **more**?

More: Least squares! — Just like in DLT - If $A = UDV^{\top}$, the LS solution is last column of V corresponding to the smallest singular value of A.

🔧 See this for understanding why the LS solution is the last column of V corresponding to the smallest singular value of A.

▼ `Homework: TODO1 - 10 Marks` What are the no of degrees of freedom of F? Elaborate.

Answer: (It is 7. Elaborate your understanding clearly.)

The solution vector $\mathbf{f}$ found in this way minimizes $\|\mathbf{Af}\|$ subject to the condition $\|\mathbf{f}\| = 1$.

2.2.1 Normalized 8 point algorithm

The above would've been fine F was a full rank matrix (as in DLT case), but it is not, so we need to enforce it in our solution.

Objective

> Given $n \geq 8$ image point correspondences $\{\mathbf{x_i} \leftrightarrow \mathbf{x_i}'\}$, determine the fundamental matrix F such that $\mathbf{x_i}'^{\top} F \mathbf{x_i} = 0$

Algorithm

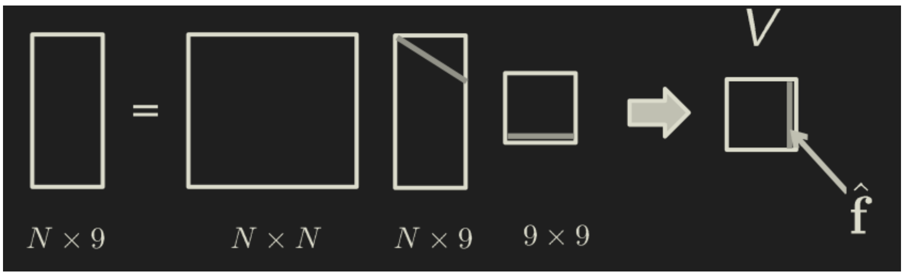
Algorithm EIGHT_POINT

The input is formed by n point correspondences, with $n \geq 8$.

1. Construct system (7.18) from n correspondences. Let A be the $n \times 9$ matrix of the coefficients of the system and $A = UDV^{\top}$ the SVD of A.
2. The entries of F (up to an unknown, signed scale factor) are the components of the column of V corresponding to the least singular value of A.
3. To enforce the singularity constraint, compute the singular value decomposition of F:
$$F = UDV^{\top}.$$
4. Set the smallest singular value in the diagonal of D equal to 0; let D' be the corrected matrix.
5. The corrected estimate of F, F', is finally given by
$$F' = UD'V^{\top}.$$

The output is the estimate of the fundamental matrix, F'.

Source: Trucco, Alessandro Verri. n.d. Introductory Techniques for 3-D Computer Vision-Prentice Hall (1998). Chapter 7, Stereopsis.



> **Normalization has to be done to avoid numerical instabilities! Important thing to remember during implementation!**

▼ `Homework: TODO2 - 5 Marks` - ***Singularity Cases in 8 point algorithm***

One singularity case when the above 8 point algorithm fails is under pure rotation (another case is if all the corresponding points are points on a plane (in the world)).

- ▼ 2a: Answer - Elaborate for pure rotation case.
- ▼ `BONUS 5 Marks` 2b: Answer - Elaborate for world points planar case.

2.2.2 Decomposition of E

Refer section 9.5, 9.6 of Zisserman for an elaborate discussion.

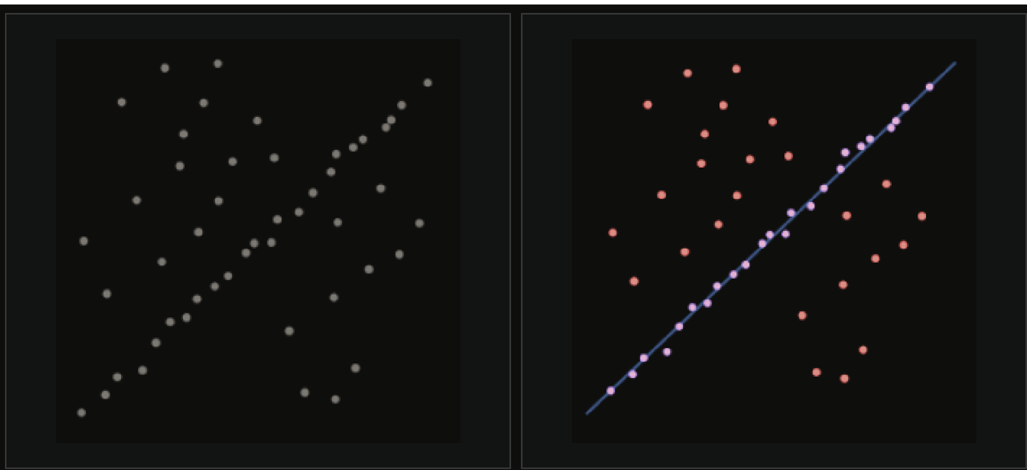
Suppose that the SVD of E is $\mathrm{U}\,\mathrm{diag}(1,1,0)\mathrm{V}^{\mathsf{T}}$. Two possible factorizations $\mathrm{E} = \mathrm{SR}$ of E are as follows:

$$\mathrm{S} = \mathrm{U}\mathrm{Z}\mathrm{U}^{\top} \quad \mathrm{R} = \mathrm{U}\mathrm{W}\mathrm{V}^{\top} \quad \text{or} \quad \mathrm{U}\mathrm{W}^{\mathsf{T}}\mathrm{V}^{\top}$$

Recollect

$$E_{12} = [\mathrm{t}_{12}]_x\, R_{12}$$

2.2.3 RANSAC: A quick overview



A data set with many outliers for which a line has to be fitted.

Fitted line with RANSAC; outliers have no influence on the result.

Algorithm 1 RANSAC

1: Select randomly the minimum number of points required to determine the model parameters.
2: Solve for the parameters of the model.
3: Determine how many points from the set of all points fit with a predefined tolerance ϵ.
4: If the fraction of the number of inliers over the total number points in the set exceeds a predefined threshold τ, re-estimate the model parameters using all the identified inliers and terminate.
5: Otherwise, repeat steps 1 through 4 (maximum of N times).

Source: Konstantinos G. Derpanis

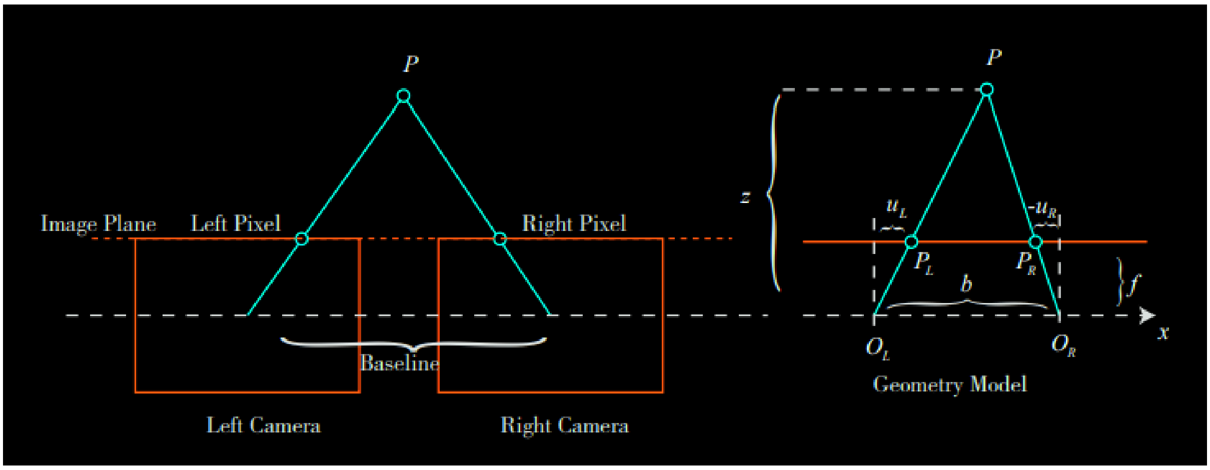
- ▼ `Homework: TODO3 - 5 Marks` - ***RANSAC***
  - ▼ Explain why RANSAC is necessary for accurate F estimation.

Answer:

▼ Explain briefly how you'd apply RANSAC algorithm to computation of F (using normalized 8 point algorithm)?

Answer:

3. Stereo Camera



Finding the world point!

$$\frac{z-f}{z} = \frac{b-u_L+u_R}{b} \qquad \implies z = \frac{fb}{d}, \quad d \triangleq u_L - u_R$$

$$(Clue \rightarrow \triangle PP_LP_R \text{ and } \triangle PO_LO_R) \qquad where \quad d \rightarrow disparity/parallax$$

▼ `Homework: TODO4 - 15 Marks`

1. Deriving depth equation

   ▼ a. Derive the depth equation step by step with proper explanation.

   Answer:

   ▼ b. Can you give a real world example where "d" is zero?

   Answer:

2. There are many ways we humans reason about depth. One most common cue seems to be the same principle used in stereo.

   ▼ a. Elaborate on this statement.

   Answer:

   ▼ b. Also, if that is really the case, why do you think we do pretty good in sensing depth with 1 eye closed (or people born with 1 eye)?

   Answer:

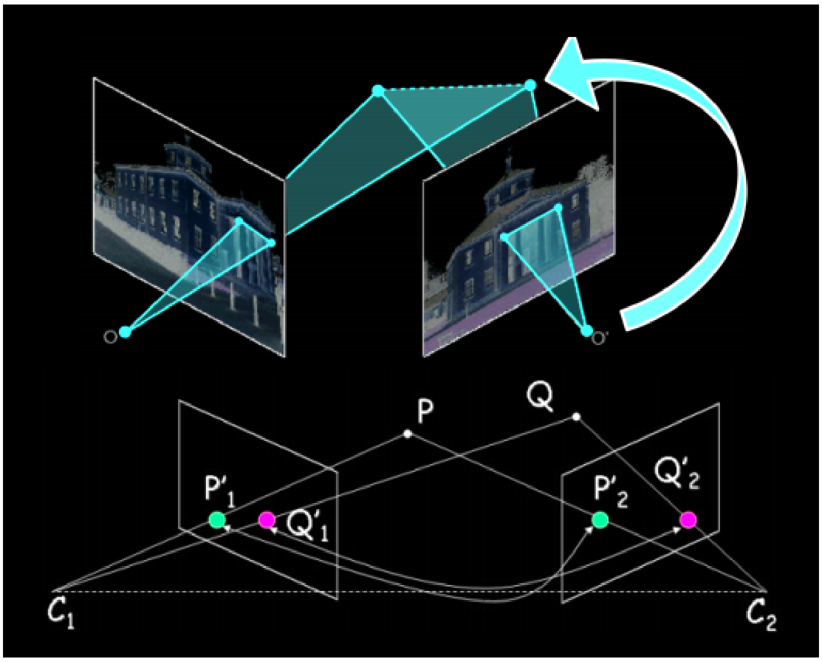
3. What happens if the left and right camera in the above image aren't facing in the same direction (but with some common overlap region of the world)? In other words, they have a rotational transformation in addition to existing translational transformation. Put your answer under "Answer: 3a" section. After writing down your answer, google "stereo rectification" and explain what you understand under "Answer: 3b" section.

    ▼ Answer: 3a

    ▼ Answer: 3b

4. Triangulation

> How to compute the position of a point in 3-space given its image in two views and the camera matrices of those views?



Rays from the camera to the 3D point in the world:

$$
\begin{aligned}
\boldsymbol{r} &= K_1^{-1}\boldsymbol{x_1} \\
\boldsymbol{s} &= R_2^1 K_2^{-1}\boldsymbol{x_2}
\end{aligned}
$$

- $\boldsymbol{x_1}, \boldsymbol{x_2}$: image pixel coordinate (homogeneous) in camera 1 & 2.
- K_1 and K_2 are intrinsic matrices for camera 1 and 2.
- R_2^1 is the relative orientation of camera 2 with respect to camera 1.

Equations of Lines:

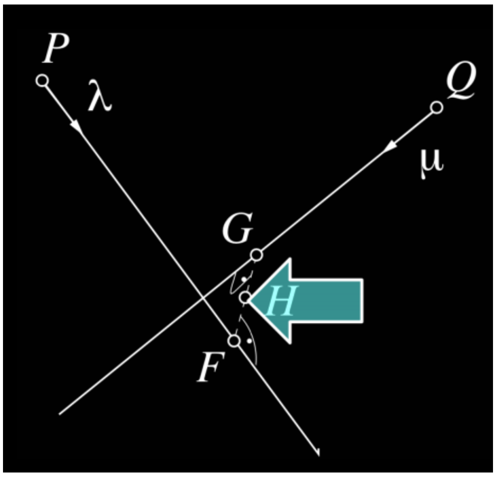
$$\boldsymbol{f} = \boldsymbol{P} + \lambda \boldsymbol{r}$$
$$\boldsymbol{g} = \boldsymbol{Q} + \mu \boldsymbol{s}$$

Knowns:

- $\boldsymbol{P}$ and $\boldsymbol{Q}$ are centers of camera 1 and 2 in the world respectively.
- The world points are F and G and $\boldsymbol{r}, \boldsymbol{s}$ are direction vectors.

Unknown:

- scalars μ and λ .



If they intersected

$$1. \quad \vec{f} = \vec{g} \iff \|f\| = \|g\|$$
$$2. \quad \frac{\vec{f} \cdot \vec{g}}{\|f\|\|g\|} = 1$$

If they don't intersect

- Ensure distance is minimum.
- Line $(\mathrm{F} - \mathrm{G})$ perpendicular to both lines r and s .

$$(\boldsymbol{F} - \boldsymbol{G}) \cdot \boldsymbol{r} = 0$$
$$(\boldsymbol{F} - \boldsymbol{G}) \cdot \boldsymbol{s} = 0$$

$$\boldsymbol{f} = \boldsymbol{P} + \lambda \boldsymbol{r}$$
$$\boldsymbol{g} = \boldsymbol{Q} + \mu \boldsymbol{s}$$

We have two equations and two unknowns λ and μ .

$$(\boldsymbol{P} + \lambda \boldsymbol{r} - (\boldsymbol{Q} + \mu \boldsymbol{s})) \cdot \boldsymbol{r} = 0$$
$$(\boldsymbol{P} + \lambda \boldsymbol{r} - (\boldsymbol{Q} + \mu \boldsymbol{s})) \cdot \boldsymbol{s} = 0$$

$$\begin{bmatrix} r \cdot r - s \cdot r \\ r \cdot s - s \cdot s \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} (\boldsymbol{Q} - \boldsymbol{P}) \cdot r \\ (\boldsymbol{Q} - \boldsymbol{P}) \cdot s \end{bmatrix}$$

$$\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} r \cdot r - s \cdot r \\ r \cdot s - s \cdot s \end{bmatrix}^{-1} \begin{bmatrix} (\boldsymbol{Q} - \boldsymbol{P}) \cdot r \\ (\boldsymbol{Q} - \boldsymbol{P}) \cdot s \end{bmatrix}$$

- λ and μ found.
- Obtain $\boldsymbol{F}$ and $\boldsymbol{G}$ from right equation.
- Mid-point of this line segment $\boldsymbol{H} = \boldsymbol{F} - \boldsymbol{G}$ is the final estimate for the $3\mathrm{D}$ triangulated world point.

$$\vec{h} = \vec{f} + \frac{\|FG\|}{2} \widehat{fg}$$

$$\boldsymbol{f} = \boldsymbol{P} + \lambda \boldsymbol{r}$$
$$\boldsymbol{g} = \boldsymbol{Q} + \mu \boldsymbol{s}$$

5. PnP

5.0 Introduction

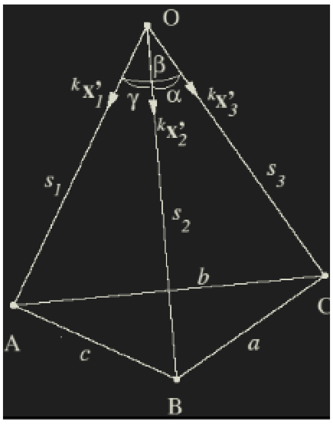
5.0.1 What is the Perspective n Points (PnP) problem?

Given: known 3D landmarks positions in the **world frame** and given their 2D image correspondences in the **camera frame**.

Determine: 6DOF pose of the camera (or camera motion) in the world frame (including the intrinsic parameters if uncalibrated).

- The 2D−2D epipolar geometry method
  - 8 or more point pairs (example: the eight-point method).
  - Problems with initialization, pure rotation, and scaling.
- However, if the 3D position of one of the two feature points is known, then at least 3 point pairs (and at least one additional point verification result) are needed to estimate camera motion. (This is P3P)

5.0.2 The P3P/Spatial Resection Problem



Given:

- $3D$ coordinates of object points X_i
- 2D image coordinates x_i of corresponding object points

Find:

- Extrinsic parameters R, X_O of the **calibrated** camera (unlike DLT)

5.0.3 Difference between P3P and DLT

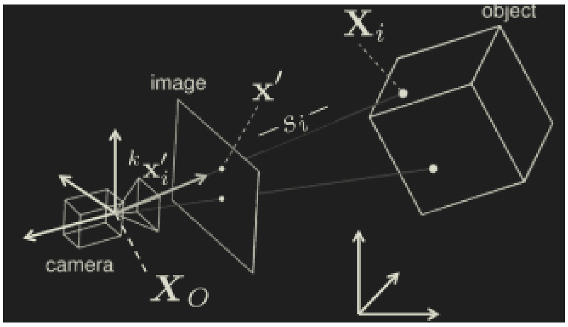
- P3P/Spatial Resection for calibrated cameras
  - 6 unknowns, so at least 3 points are needed
- DLT for uncalibrated cameras (seen)
  - 11 unknowns, so at least 6 points are needed

5.1 Solution to P3P

5.1.1 Revisiting normalized coordinates

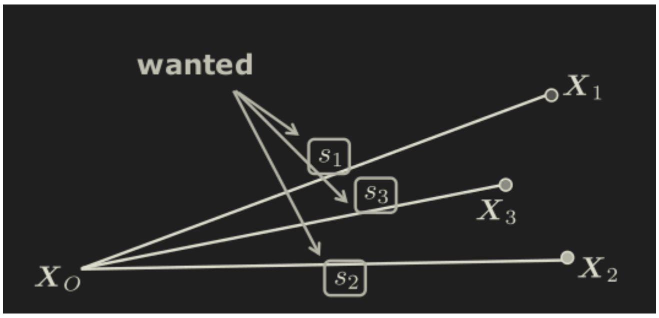
$$\mathrm{x} = \mathrm{K}R\,[I_3| - X_O]\,\mathrm{X}$$

$${}^k\mathbf{x}_i' = \mathrm{K}^{-1}\mathbf{x}_i'$$



5.1.2 Two step process

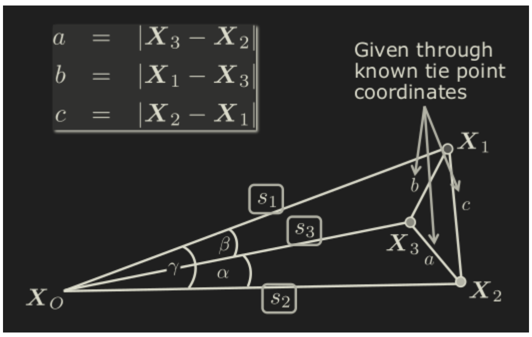
1. Length of projection rays
2. Orientation



💡 Clarity about camera frame and world frame: angles and distances between points

5.1.3 Length of projection rays

1. **Do we know a, b, c?**

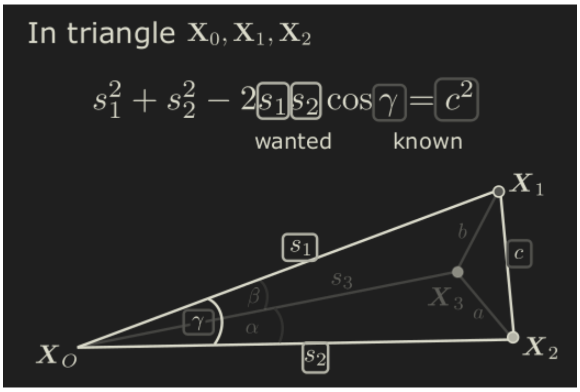


2. Do we know angles?

$$\cos\gamma = \frac{(X_1 - X_0)\cdot(X_2 - X_0)}{\|X_1 - X_0\|\,\|X_2 - X_0\|}$$

Clue: Normalized Coords

Cosine rule:



$$
\begin{aligned}
a^2 &= s_2^2 + s_3^2 - 2 s_2 s_3 \cos\alpha \\
b^2 &= s_1^2 + s_3^2 - 2 s_1 s_3 \cos\beta \\
c^2 &= s_1^2 + s_2^2 - 2 s_1 s_2 \cos\gamma
\end{aligned}
$$

$$\text{We have:} \quad a^2 = s_2^2 + s_3^2 - 2 s_2 s_3 \cos\alpha$$

$$\text{Define:} \quad u = \tfrac{s_2}{s_1} \quad v = \tfrac{s_3}{s_1}$$

$$\implies a^2 = s_1^2\left(u^2 + v^2 - 2uv\cos\alpha\right)$$

$$
\begin{aligned}
s_1^2 &= \frac{a^2}{u^2 + v^2 - 2uv\cos\alpha} \\
&= \frac{b^2}{1 + v^2 - 2v\cos\beta} \\
&= \frac{c^2}{1 + u^2 - 2u\cos\gamma}
\end{aligned}
$$

$$
\begin{aligned}
b^2 &= s_1^2 + s_3^2 - 2 s_1 s_3 \cos\beta \\
c^2 &= s_1^2 + s_2^2 - 2 s_1 s_2 \cos\gamma
\end{aligned}
$$

Substitute ***u*** in other equation — **4th degree polynomial:**

$$A_4 v^4 + A_3 v^3 + A_2 v^2 + A_1 v + A_0 = 0$$

$$
\begin{aligned}
A_4 &= \left(\tfrac{a^2 - c^2}{b^2} - 1\right)^2 - \tfrac{4c^2}{b^2}\cos^2\alpha \\
A_3 &= 4\left[\frac{a^2 - c^2}{b^2}\left(1 - \frac{a^2 - c^2}{b^2}\right)\cos\beta \right. \\
&\quad \left. - \left(1 - \frac{a^2 + c^2}{b^2}\right)\cos\alpha\cos\gamma + 2\frac{c^2}{b^2}\cos^2\alpha\cos\beta\right]
\end{aligned}
$$

$$
\begin{aligned}
A_2 &= 2\left[\left(\frac{a^2 - c^2}{b^2}\right)^2 - 1 + 2\left(\frac{a^2 - c^2}{b^2}\right)^2 \right. \\
&\quad + 2\left(\frac{b^2 - c^2}{b^2}\right)\cos^2\alpha \\
&\quad - 4\left(\frac{a^2 + c^2}{b^2}\right)\cos\alpha\cos\beta\cos\gamma \\
&\quad \left. + 2\left(\frac{b^2 - a^2}{b^2}\right)\cos^2\gamma\right]
\end{aligned}
$$

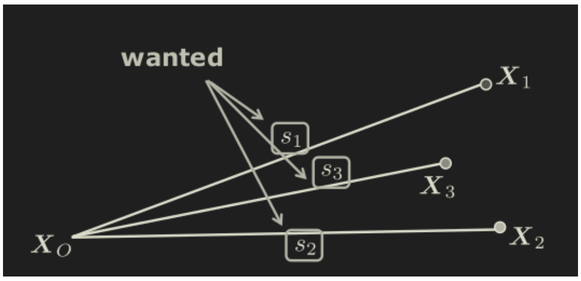
$$A_1 = 4\left[-\left(\frac{a^2-c^2}{b^2}\right)\left(1+\frac{a^2-c^2}{b^2}\right)\cos\beta + \frac{2a^2}{b^2}\cos^2\gamma\cos\beta - \left(1-\left(\frac{a^2+c^2}{b^2}\right)\right)\cos\alpha\cos\gamma\right]$$

$$A_0 = \left(1+\frac{a^2-c^2}{b^2}\right)^2 - \frac{4a^2}{b^2}\cos^2\gamma$$

But upto 4 possible solutions possible. So we consider 4th point to confirm the right solution.

5.1.4 Transformation between camera frame and world frame

$$\begin{aligned} {}^cX_1 &= s_1\, {}^c\hat{X}_1 \\ {}^cX_2 &= s_2\, {}^c\hat{X}_2 \\ {}^cX_3 &= s_3\, {}^c\hat{X}_3 \end{aligned}$$



$$P = {}^wX - {}^w\bar{X}, \quad Q = {}^cX - {}^c\bar{X}$$

Let

$$\begin{gathered} \text{Covariance Matrix: } S = PQ^T \\ \mathbf{SVD}(S) = UDV^T \\ \text{Rotation : } R = VU^T \\ \text{Translation: } t = \bar{x} - R\bar{X} \end{gathered}$$

5.1.5 Additional Reading

To be added on this page.

> We will revisit these concepts at a high level and understand the big picture of MVG in SLAM Lectures!

> Contact: Shubodh Sai