



Pose Tracking I

CS 6334 Virtual Reality

Professor Yu Xiang

The University of Texas at Dallas

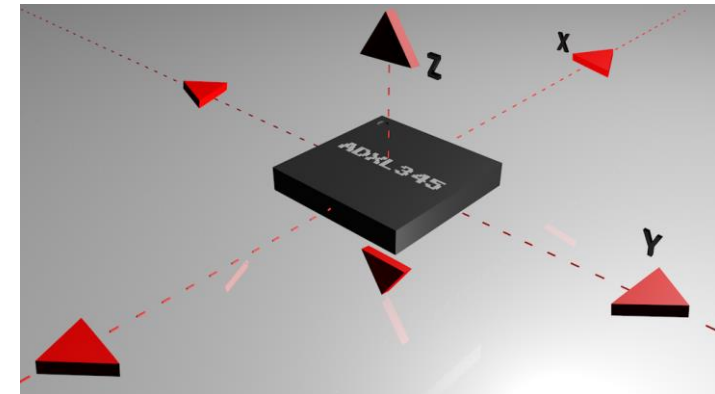
Tracking in VR

- Tracking the user's sense organs
 - E.g., Head and eye
 - Render stimulus accordingly
- Tracking user's other body parts
 - E.g., human body and hands
 - Locomotion and manipulation
- Tracking the rest of the environment
 - Augmented reality
 - Obstacle avoidance in the real world



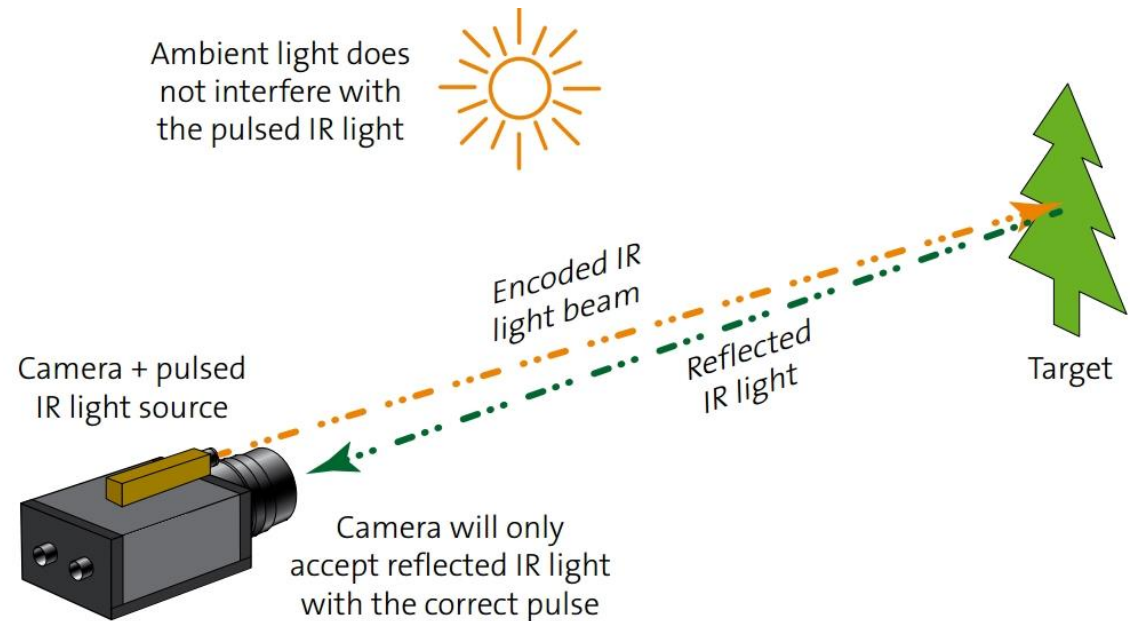
Tracking 3D Orientation and 3D Position

- Why not just integrate the accelerometer?
 - Drift error, quadratically growing with calibration error (difference between true value and the measured value)
 - The true body acceleration cannot be accurately extracted, especially when the body quickly rotates
 - IMU itself cannot help
 - Cannot distinguish motions at constant velocity



Wave-based Tracking

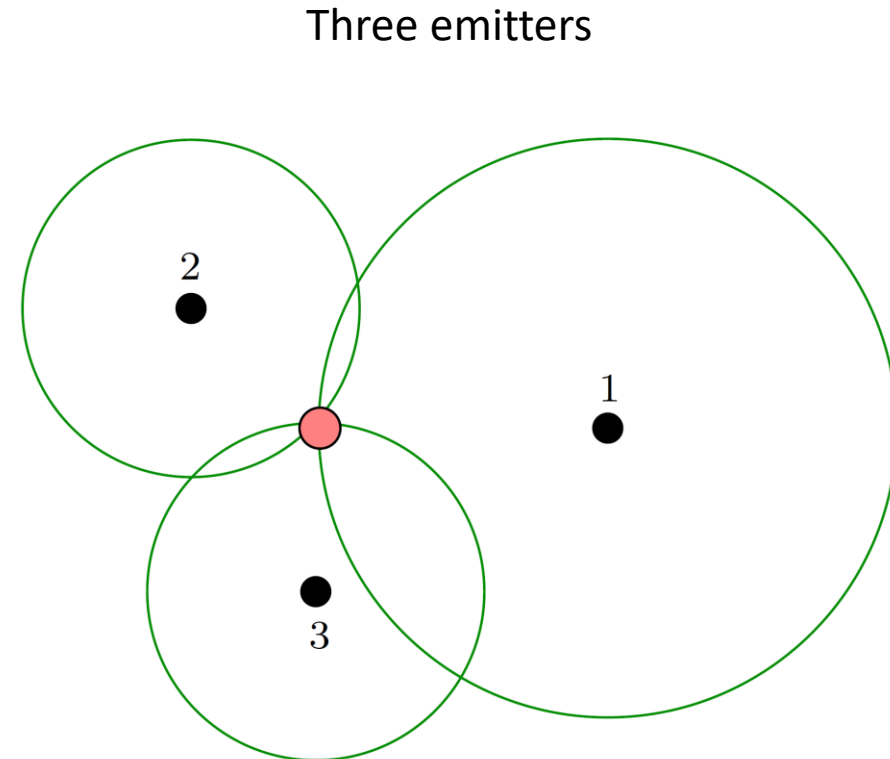
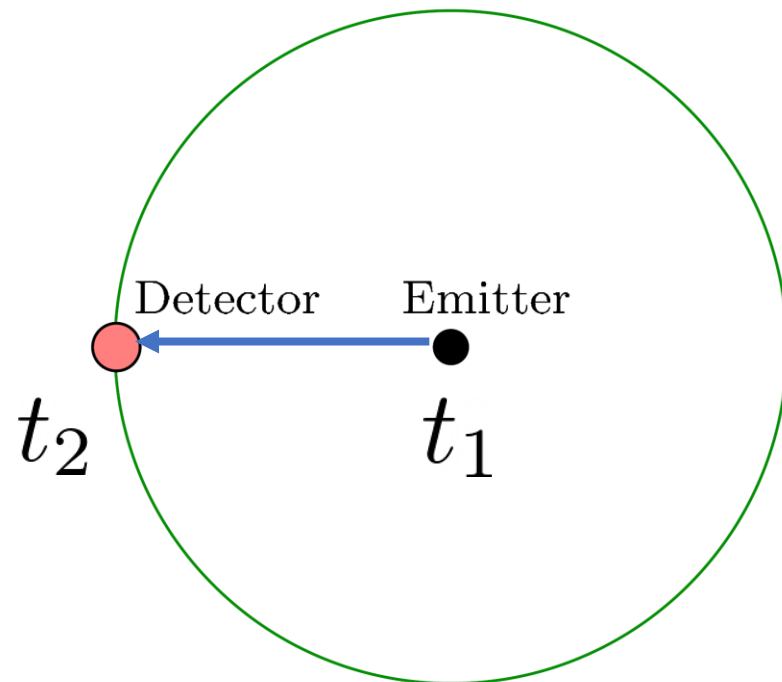
- Active tracking systems
 - Infrared, ultrasound, and electromagnetic fields
- Emitter-detector pair
 - Speaker-microphone
 - Time of flight to estimate distance



<https://www.stemmer-imaging.com/en-ie/knowledge-base/cameras-3d-time-of-flight-cameras/>

Wave-based Tracking

- Trilateration



Time measurement is synchronized between source and destination
Use time of arrival (time of flight) to compute distance

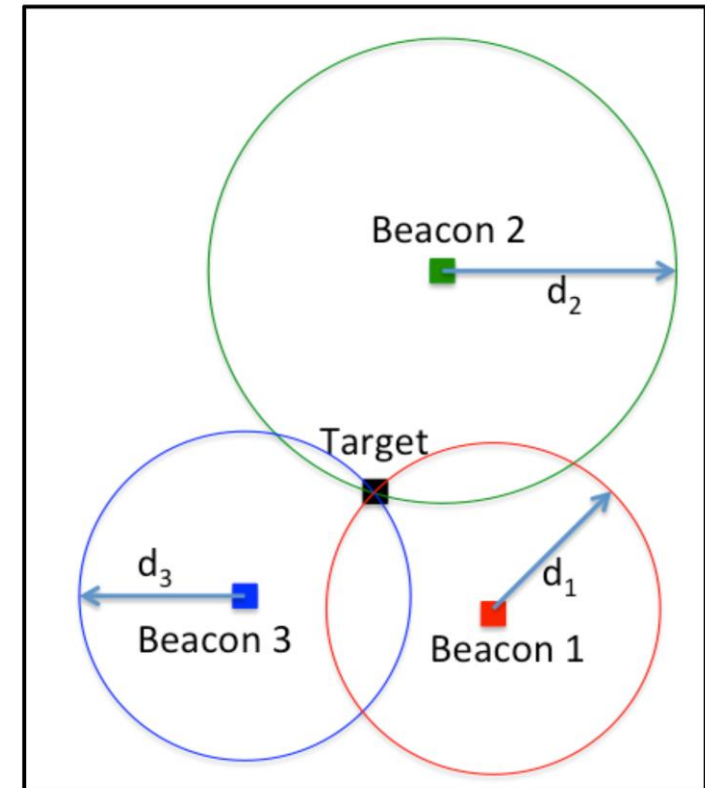
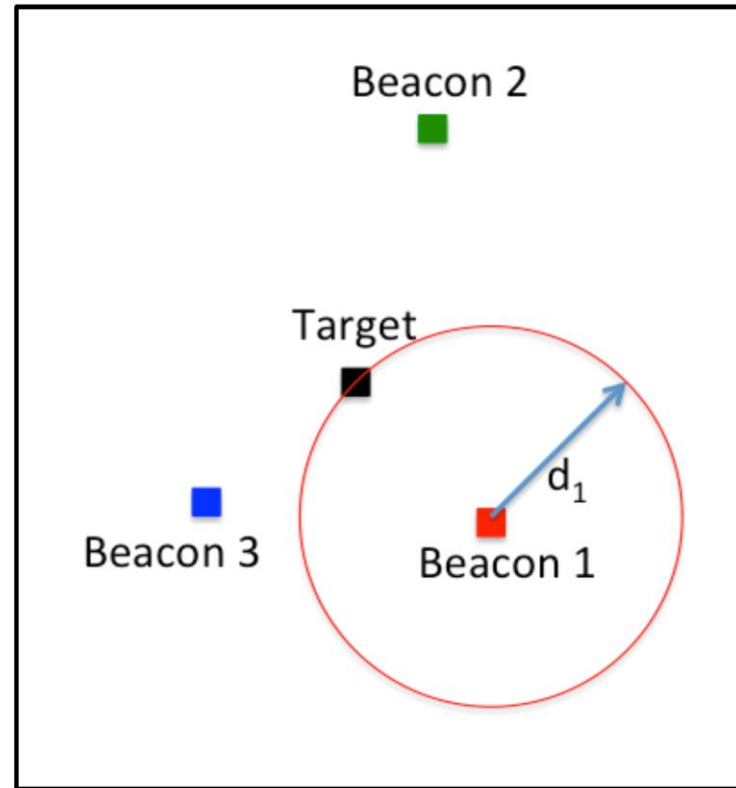
- 330 m/s for ultrasound

Time of Arrival (ToA)

- Signals sent from a beacon to the target
- Known sending time and arrival time

$$d = c * (t_{arrival} - t_{sent})$$

$$d = \sqrt{(x_{ref} - x)^2 + (y_{ref} - y)^2}$$



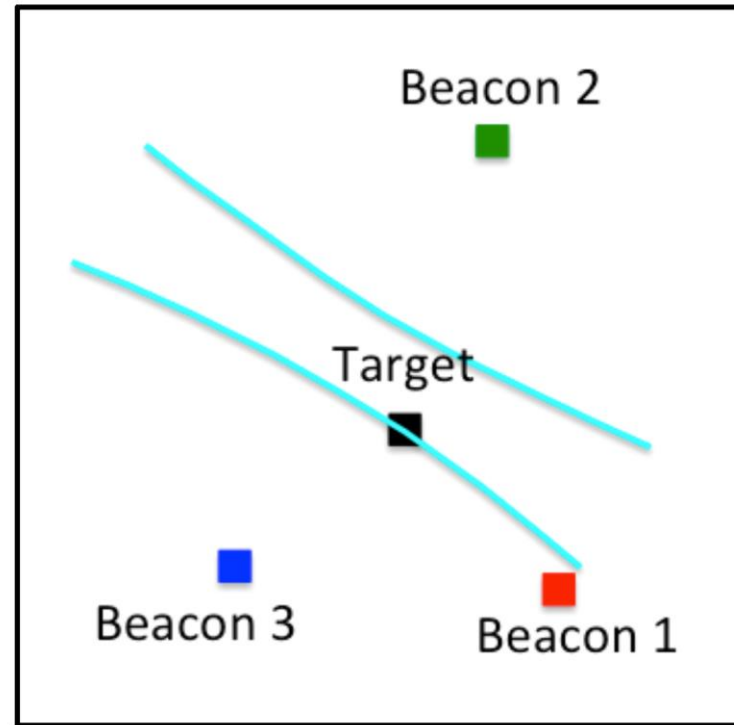
https://sites.tufts.edu/eesenior/designhandbook/files/2017/05/FireBrick_OKeefe_F1.pdf

Time Difference of Arrival (TDoA)

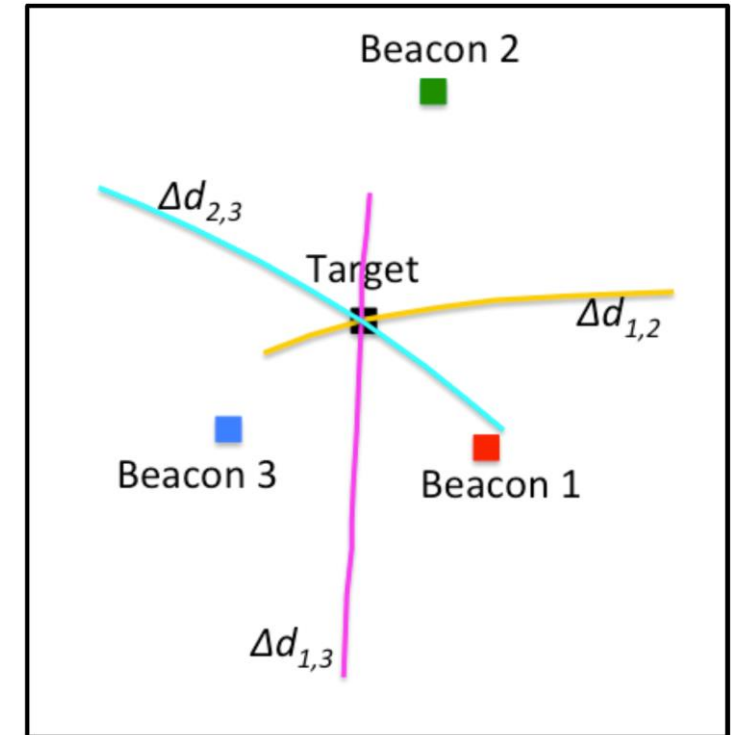
- Do not know sending time
- Signals sent from the target to beacons

$$\Delta d = c * (\Delta t)$$

$$\Delta d = \sqrt{(x_2 - x)^2 - (y_2 - y)^2} - \sqrt{(x_1 - x)^2 - (y_1 - y)^2}$$



hyperbola



https://sites.tufts.edu/eesenior/designhandbook/files/2017/05/FireBrick_OKeefe_F1.pdf

Wave-based Tracking

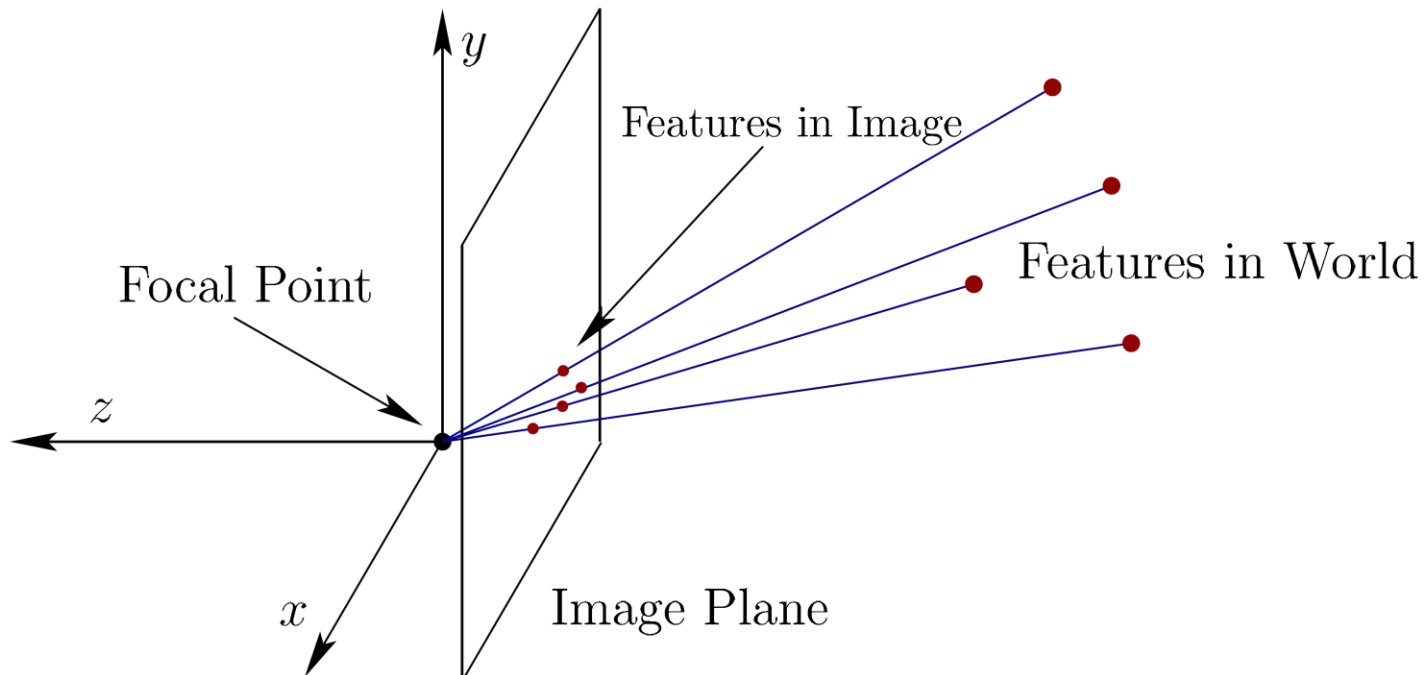
- Magnetic fields
 - The field should vary greatly across different locations
 - Drawback: the field may become unpredictably warped in each environment, causing straight-line motions to be estimated as curved



Figure 9.10: (a) A magnetic dipole offers a field that varies its magnitude and direction as the position changes (figure by Wikipedia user Geek3). (b) The Razer Hydra, a game controller system that generates a weak magnetic field using a base station, enabling it to track the controller positions (figure by <http://www.GamingShogun.com/>).

Feature-based Tracking

- Using visibility of features in the real world



- Natural Features
 - No setup cost
 - A difficult computer vision problem
- Artificial features
 - Print a special tag



QR code

QR Code (Quick Response Code)

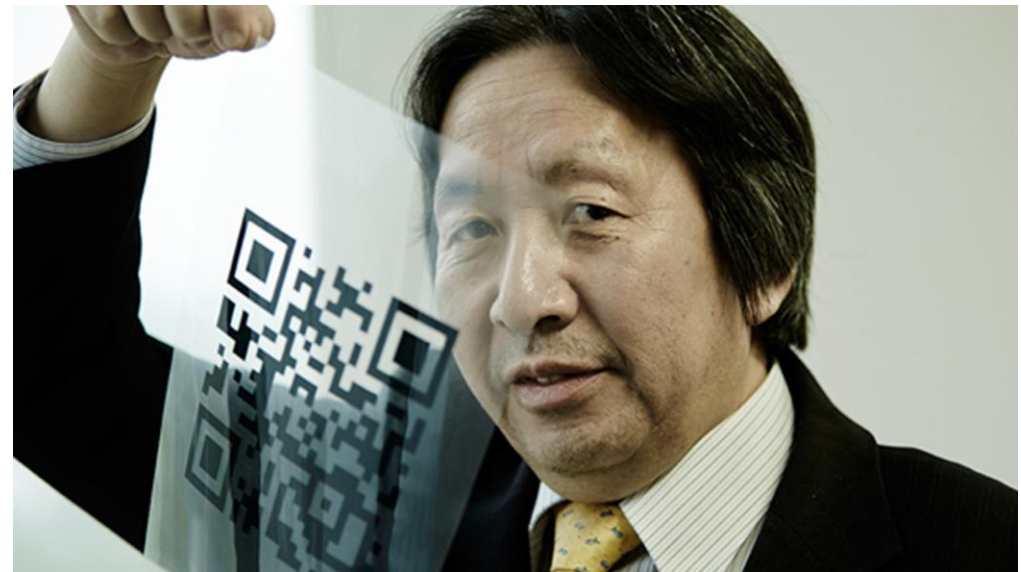
- QR codes contains more information than barcodes



7,089 numeric characters or 2,953 alphanumeric characters

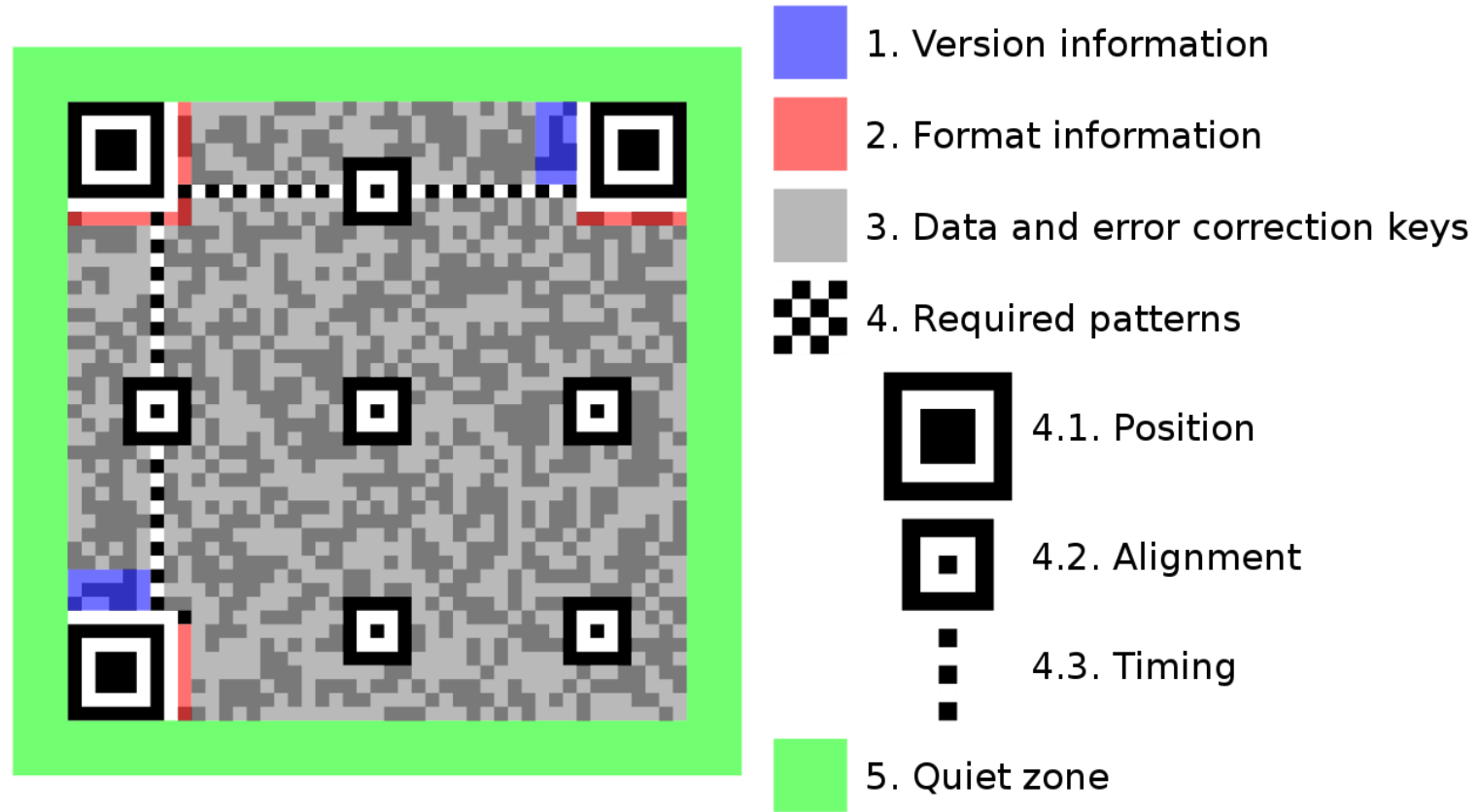


20 to 100 characters



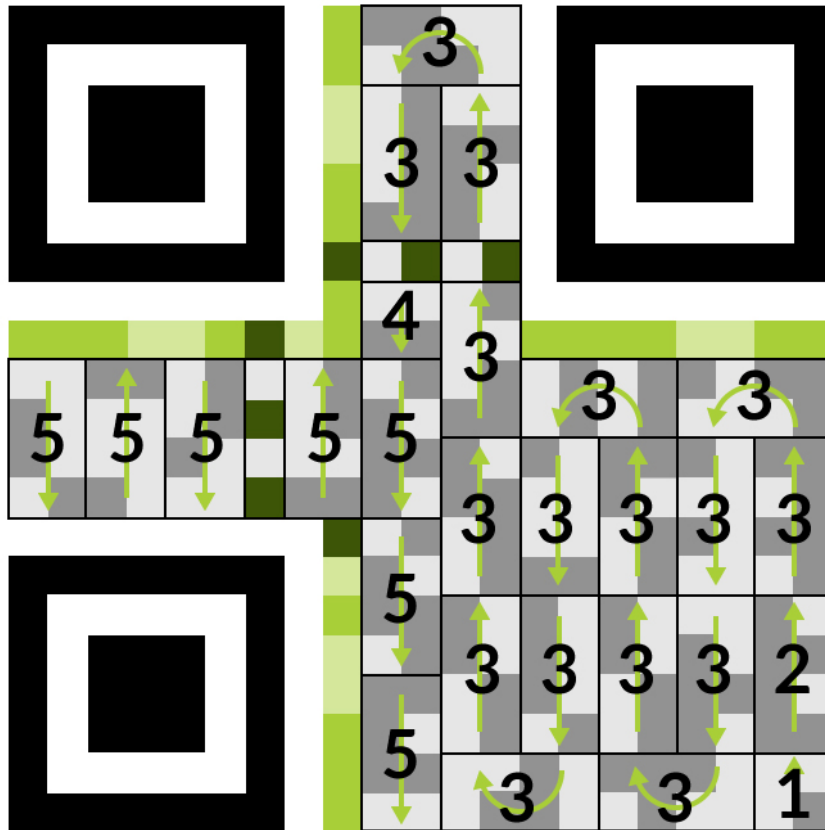
*Photo of QR Code inventor Masahiro Hara. Source: qrcode.com
DENSO WAVE*

QR Code



Structure of a QR code (version 7), highlighting functional elements (Wikipedia)

QR Code



- 1 Mode indicator
- 2 Character count indicator
- 3 Standard data modules
- 4 Stop indicator
- 5 Error correction modules

- Fixed patterns
- Format information

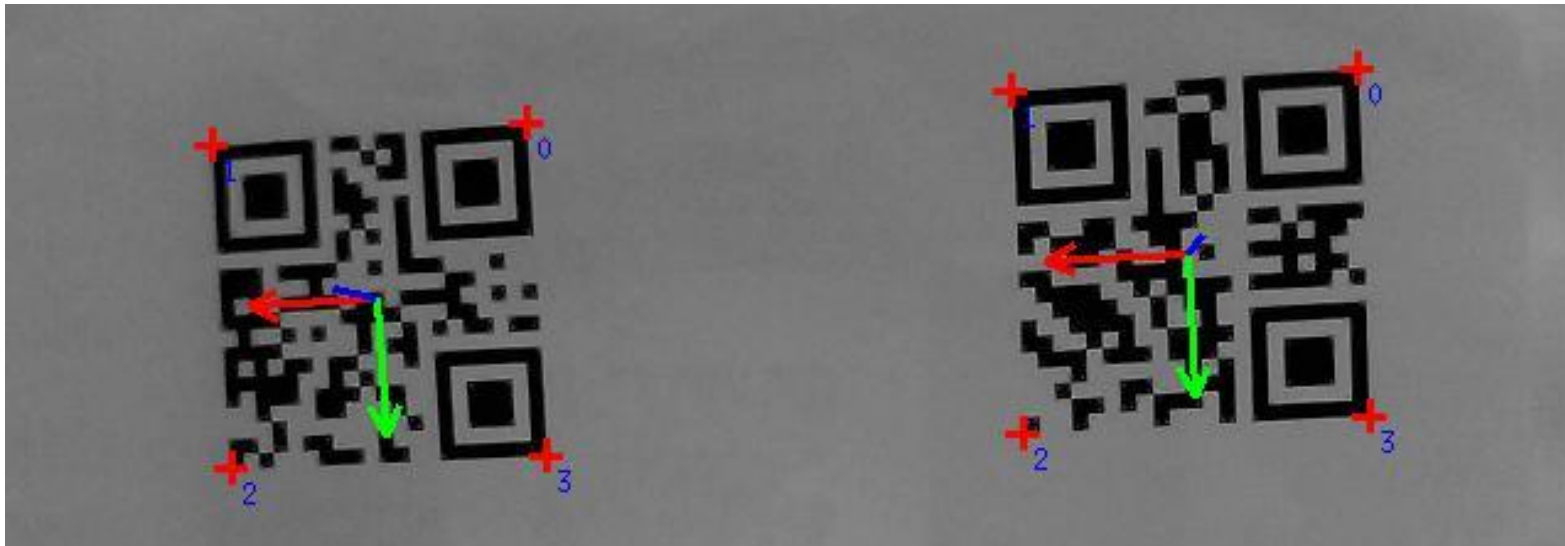


Mode type: numeric, alphanumeric, byte, or kanji (Japanese)

Data is stored in bits
• e.g., 01011011

QR Code for Pose Estimation

- Using the 4 corners of a QR code as features

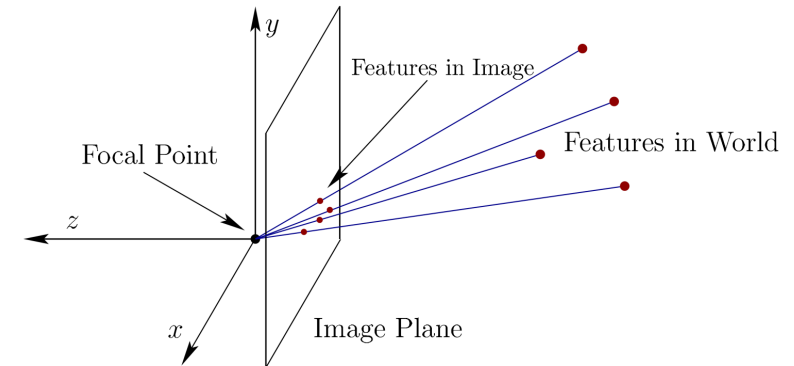


<https://visp-doc.inria.fr/doxygen/visp-daily/tutorial-pose-estimation-qr.html>

The Perspective-n-Point (PnP) Problem

- Given/known variables

- A set of n 3D points in the world coordinates p_w
- Their projections (2D coordinates) on an image p_c
- Camera intrinsics K



- Unknown variables

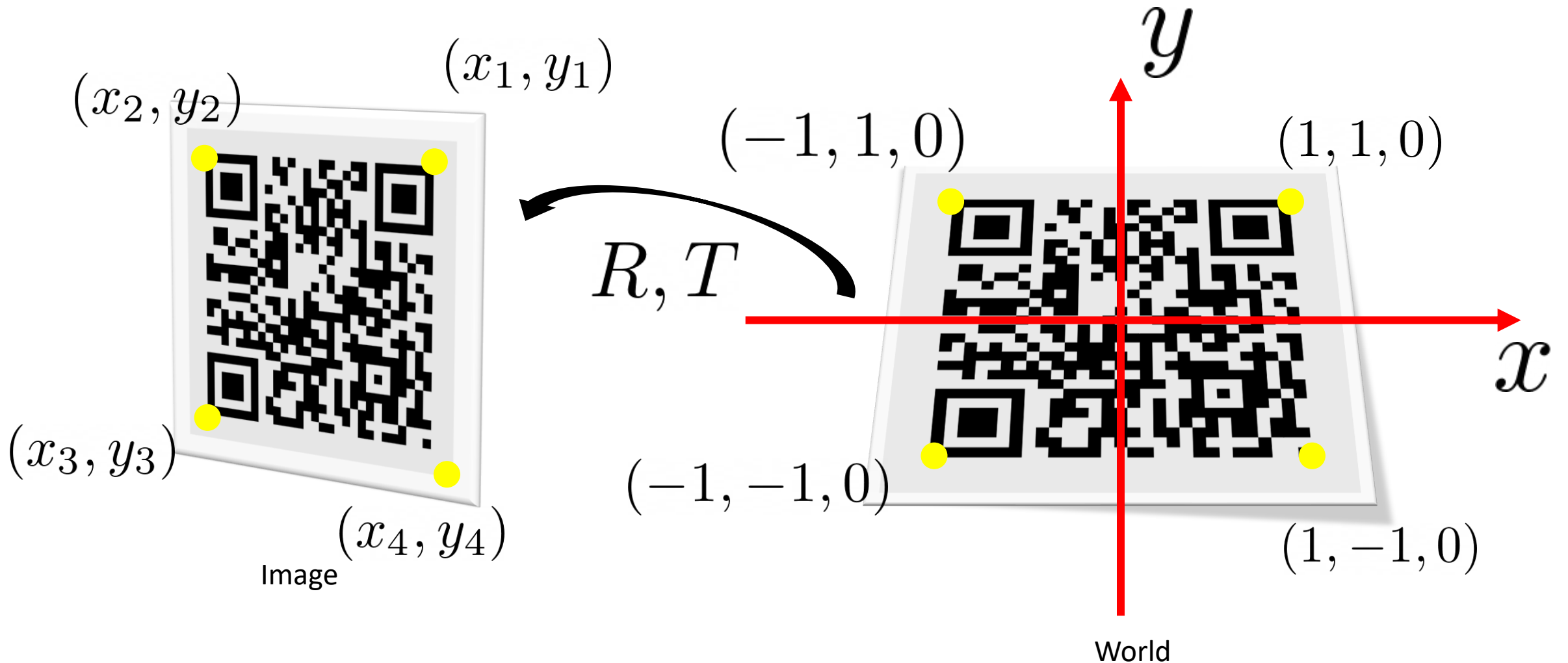
- 3D rotation of the camera with respect to the world coordinates R
- 3D translation of the camera T

$$s p_c = K [R | T] p_w.$$

↙ ↘ ↗
 Unknown

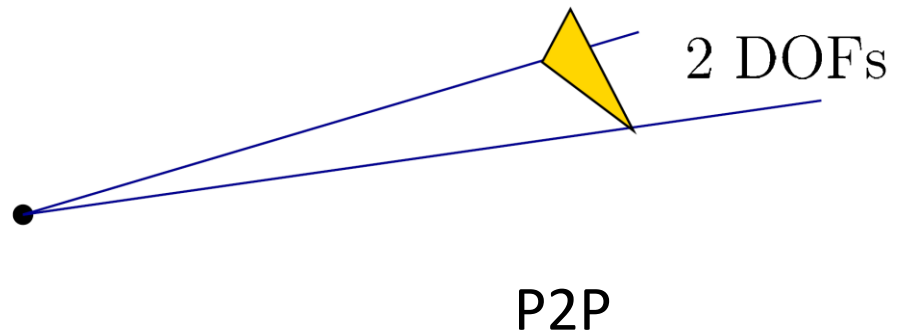
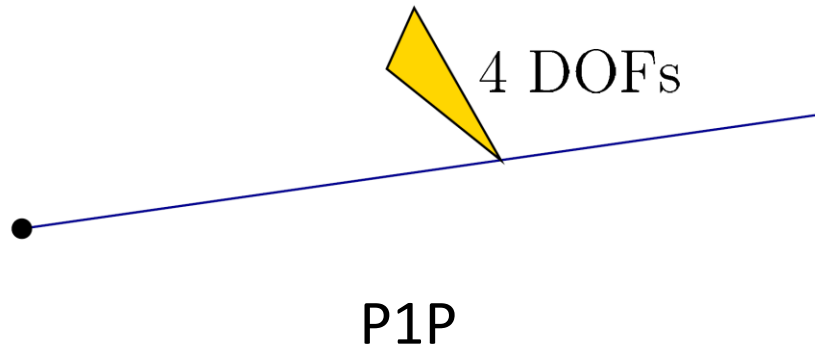
$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

The PnP Problem with QR Code

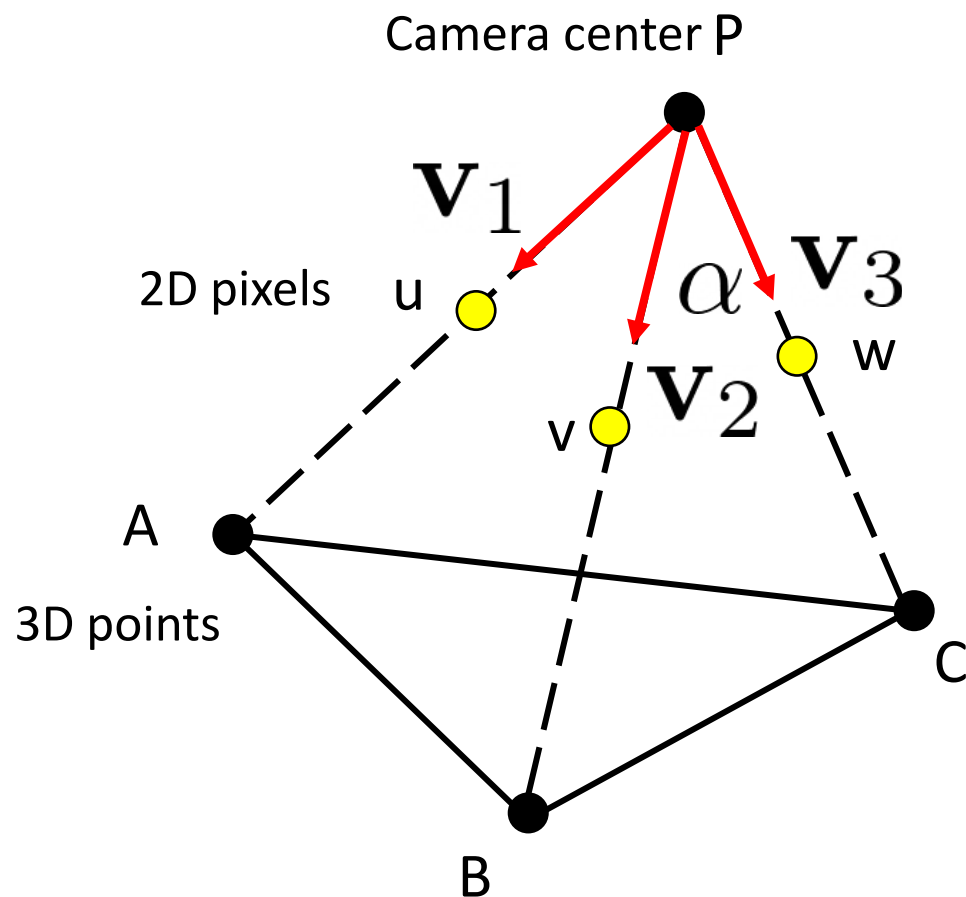


The Perspective-n-Point (PnP) Problem

- 6 degrees of freedom (DOFs)
 - 3 DOF rotation, 3 DOF translation
- Each feature that is visible eliminates 2 DOFs



P3P



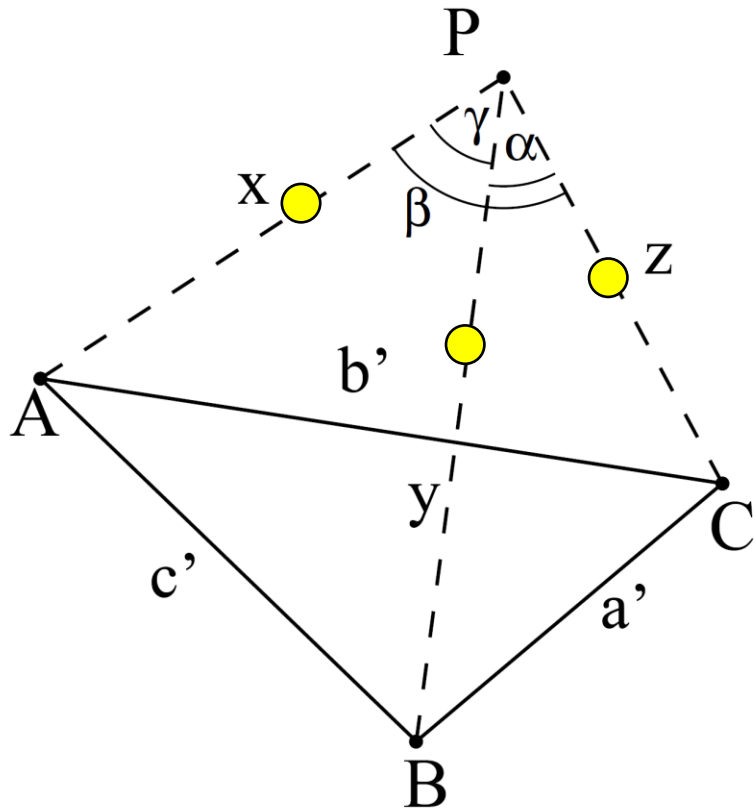
$$\mathbf{v}_1 = K^{-1} \mathbf{u}$$

$$\mathbf{v}_2 = K^{-1} \mathbf{v}$$

$$\mathbf{v}_3 = K^{-1} \mathbf{w}$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = \|\mathbf{v}_2\| \|\mathbf{v}_3\| \cos \alpha$$

P3P



$$X = |PA| \quad Y = |PB| \quad Z = |PC|$$

$$p = 2 \cos \alpha \quad a' = |BC|$$

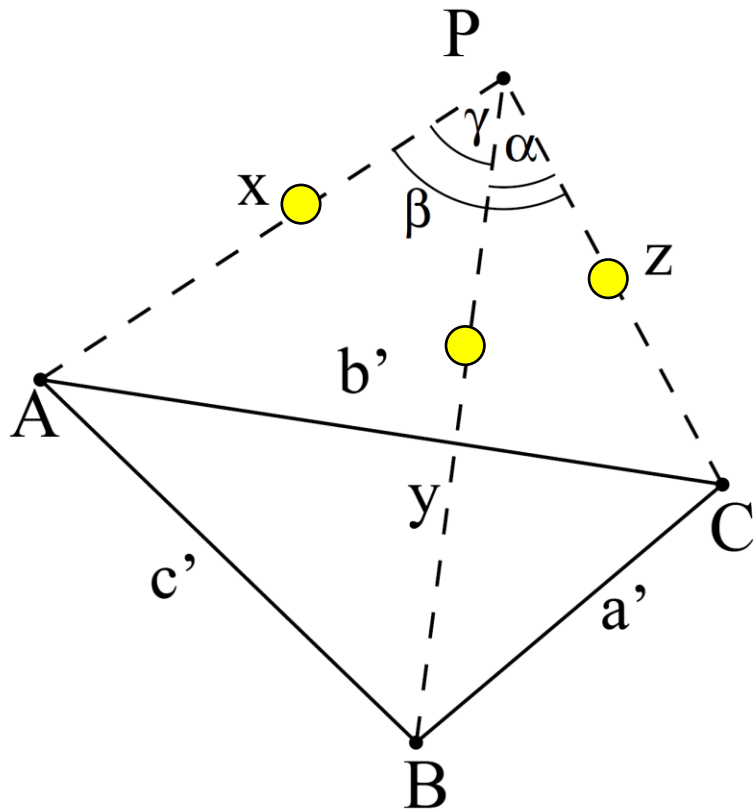
$$q = 2 \cos \beta \quad b' = |AC|$$

$$r = 2 \cos \gamma \quad c' = |AB|$$

$$\text{law of cosines} \left\{ \begin{array}{l} Y^2 + Z^2 - YZp - a'^2 = 0 \\ Z^2 + X^2 - XZq - b'^2 = 0 \\ X^2 + Y^2 - XYr - c'^2 = 0. \end{array} \right.$$

X, Y, Z are the unknowns

P3P



- Find the solutions for X , Y , Z (depth of the 3 pixels)
- Obtain the coordinates of A , B , C in camera frame
- Compute R and T using the coordinates of A , B , C in camera frame and in world frame

6D Pose Between Two Sets of 3D Points

- A well-studied problem
- Least-squares Estimation

$$\Sigma^2 = \sum_{i=1}^N \| p'_i - (R p_i + T) \|^2.$$

K.S. Arun, T.S. Huang, and S.D. Blostein. Least-Squares Fitting of Two 3-D Points Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.

S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4), April 1991

Further Reading

- Section 9.3, Virtual Reality, Steven LaValle