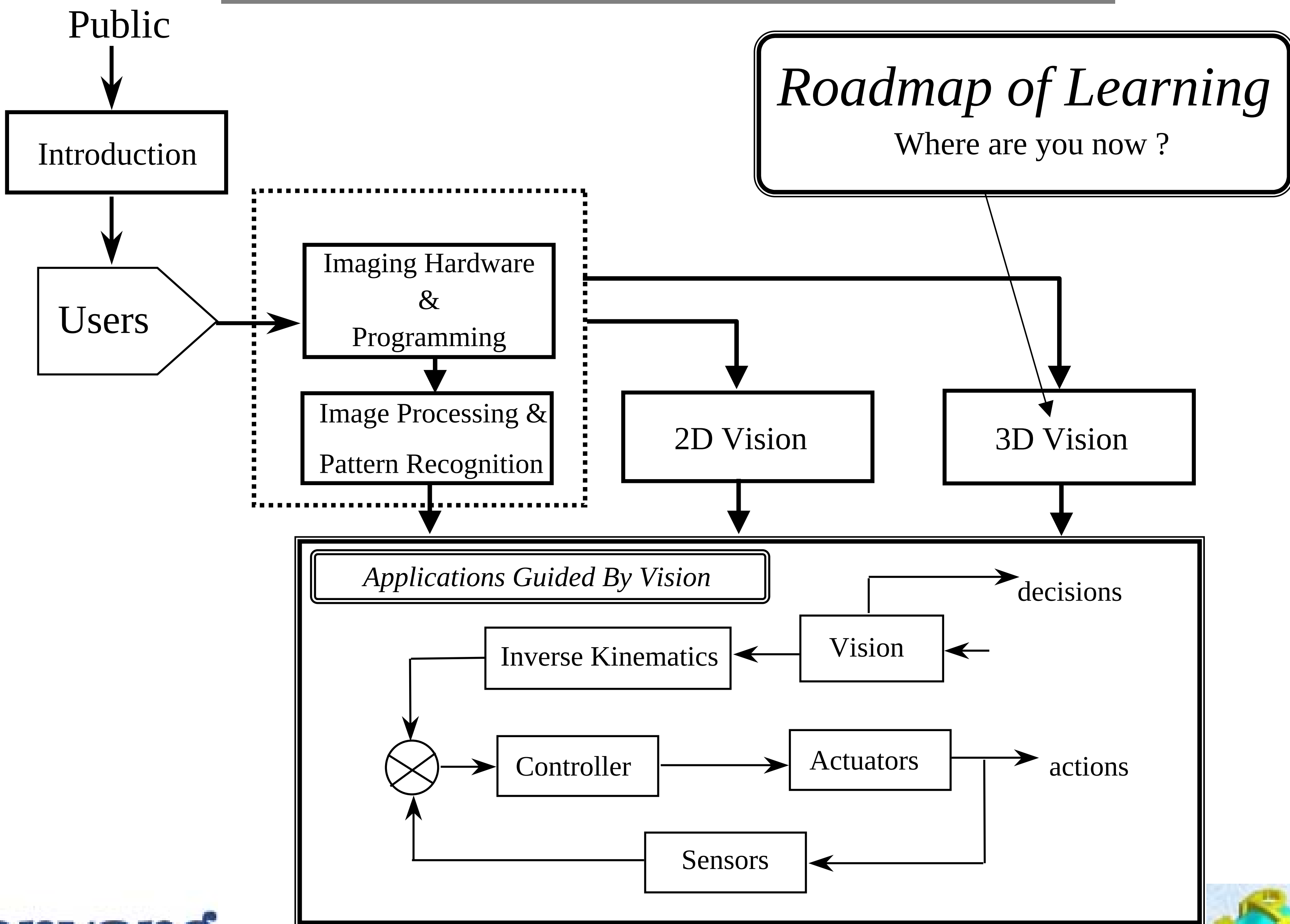
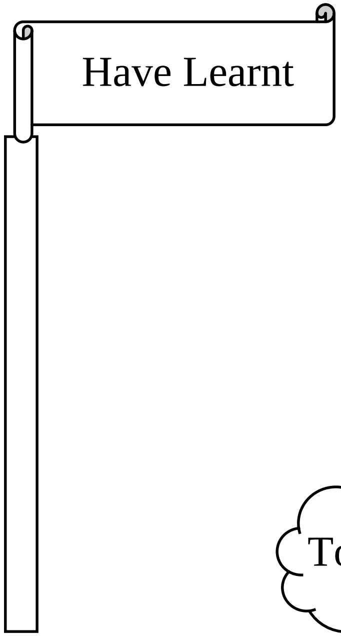
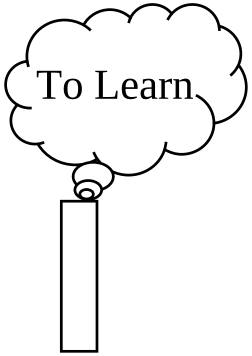


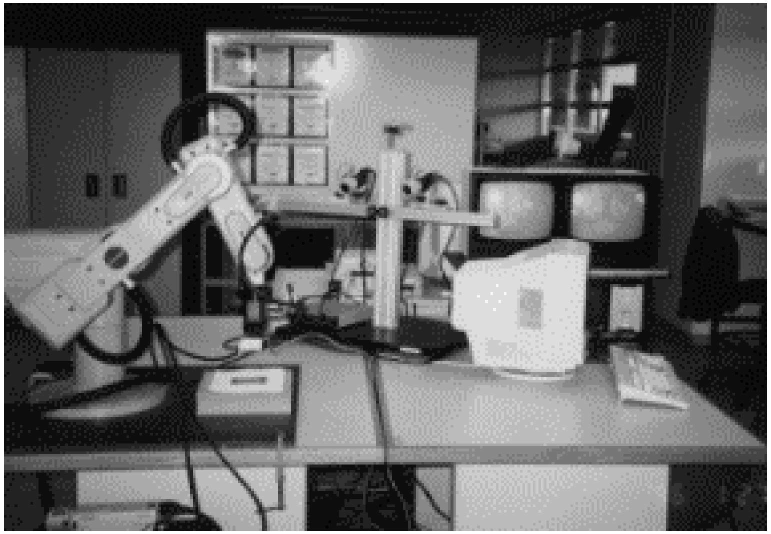
Nanyang Technological University

CONTENT

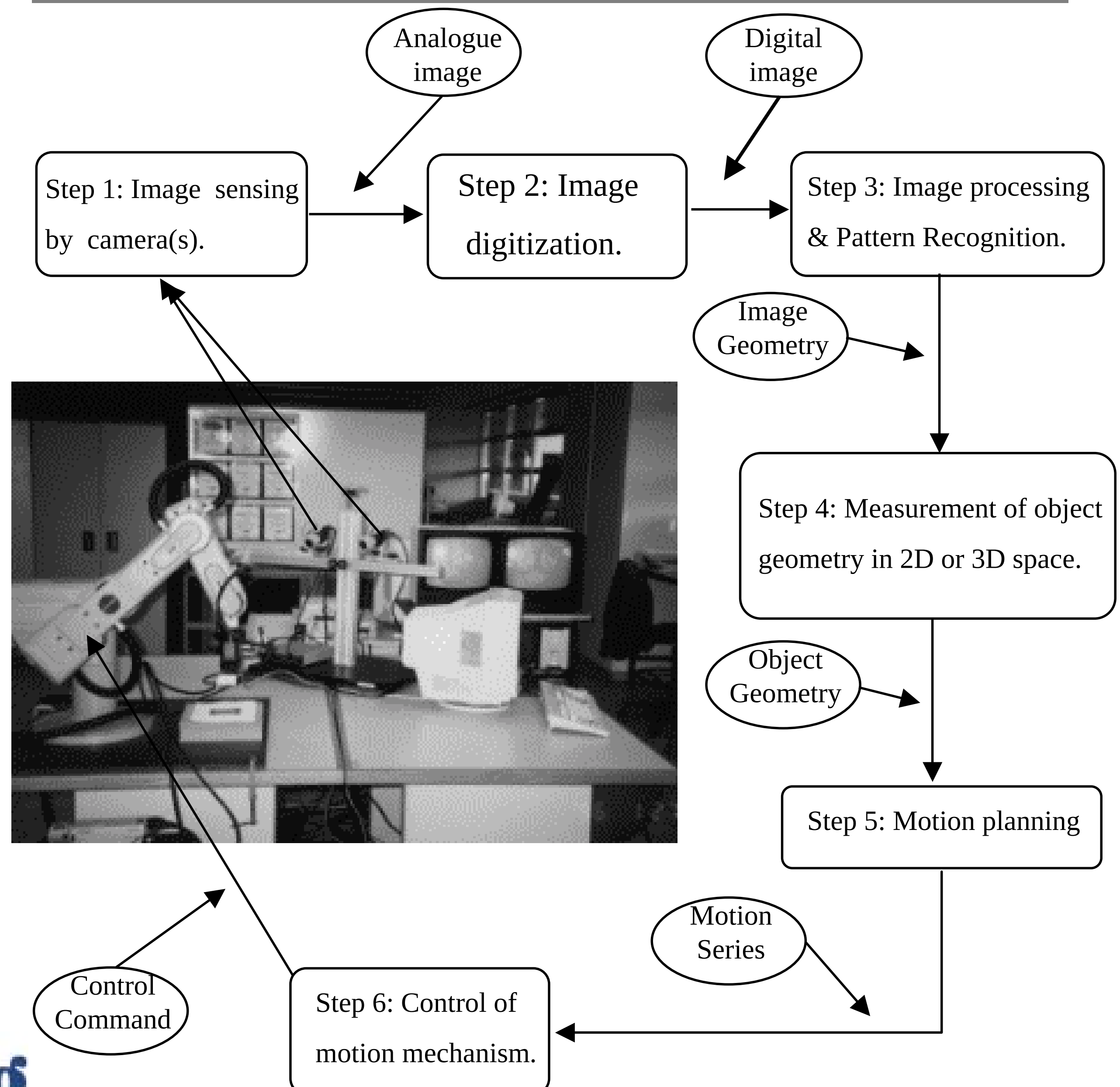




How does a machine vision system work for the task of visual guidance or measurement ?



Answer:

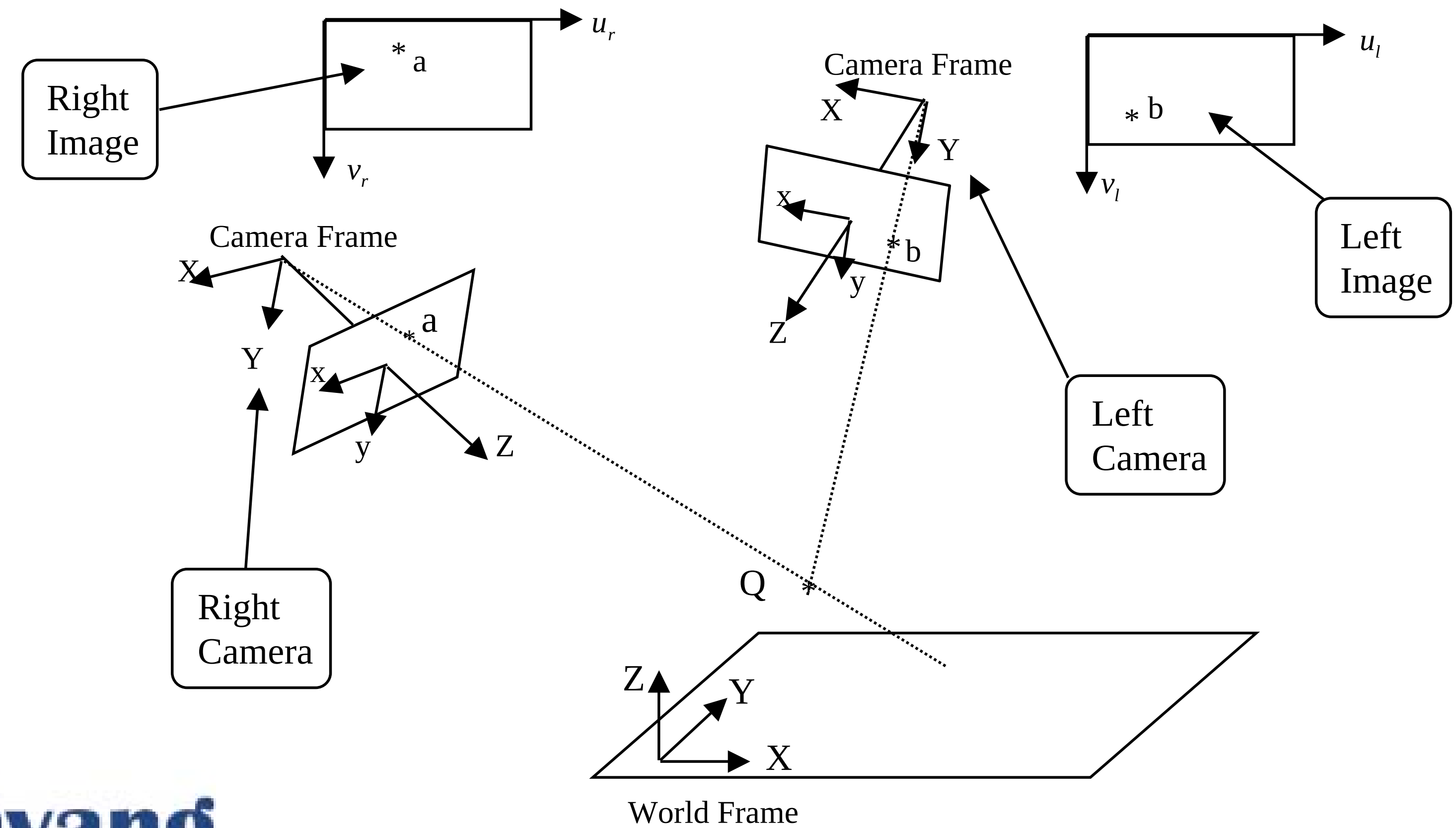


Nanyang Technological University

Computation of 3D Coordinates by Using Stereo Vision (A Review):

By using two cameras, the 3D coordinates of an object point can be computed through a simple triangulation:

$$\begin{pmatrix} {}^{r}X \\ {}^{r}Y \\ {}^{r}Z \end{pmatrix} = (A^t \bullet A)^{-1} \bullet (A^t \bullet B).$$



$$A = \begin{pmatrix} (m_{l1} - m_{l9} \bullet u_l) & (m_{l2} - m_{l10} \bullet u_l) & (m_{l3} - m_{l11} \bullet u_l) \\ (m_{l5} - m_{l9} \bullet v_l) & (m_{l6} - m_{l10} \bullet v_l) & (m_{l7} - m_{l11} \bullet v_l) \\ (m_{r1} - m_{r9} \bullet u_r) & (m_{r2} - m_{r10} \bullet u_r) & (m_{r3} - m_{r11} \bullet u_r) \\ (m_{r5} - m_{r9} \bullet v_r) & (m_{r6} - m_{r10} \bullet v_r) & (m_{r7} - m_{r11} \bullet v_r) \end{pmatrix} \qquad B = \begin{pmatrix} u_l - m_{l4} \\ v_l - m_{l8} \\ u_r - m_{r4} \\ v_r - m_{r8} \end{pmatrix}$$

Necessary Conditions:

But, there are two necessary conditions imposed to stereo vision:

Condition 1: The 3D calibration matrices are known.

Condition 2: The stereo correspondence between the image coordinates in left camera and the image coordinates in right camera is known.

Nanyang Technological University

Questions :

1. How to determine the coefficients of the calibrations matrices:

$* \quad (m_{l1}, m_{l2}, ..., m_{l11})$

$* \quad (m_{r1}, m_{r2}, ..., m_{r11})$

?

2. How to automatically determine that the image coordinates

$* \quad (u_l, v_l)$

$* \quad (u_r, v_r)$

correspond to the projection of a same 3D object point ?

How to determine the coefficients inside a 3D calibration matrix ?

ANSWER:

1. Problem Analysis:

From the proof in Lecture 7-1, we know that the relationship between 3D coordinates and its 2D image coordinates is:

$$\begin{bmatrix} \rho \bullet u \\ \rho \bullet v \\ \rho \end{bmatrix} = \begin{bmatrix} 1/D_x & 0 & u_0 \\ 0 & 1/D_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \bullet \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} {}^rX \\ {}^rY \\ {}^rZ \\ 1 \end{bmatrix}$$

1. The focal length of the camera is “f”.
2. Sampling steps in image plane are: (D_x, D_y).
3. The coordinates of the origin of real image plane with respect to the digital image plane are: (u_0, v_0).
4. The motion transformation between the reference frame and the camera frame is:

$$ {}^cM_r = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Intrinsic Parameters

Extrinsic Parameters

How to determine the coefficients inside a 3D calibration matrix ?

ANSWER:

2. Solutions:

Solution 1:

If we know the parameters of a camera in 3D space:

1. The focal length of the camera: "f",
2. Sampling steps in image plane: (D_x, D_y),
3. The coordinates of the origin of real image plane with respect to the digital image plane: (u_0, v_0),
4. The motion transformation between the world frame and the camera frame,

then, the 3D calibration matrix can be computed.

Question:

How to know the above parameters in practice ?

Solution 2:

To do calibration because there are only 11 linearly-related unknowns inside the 3D calibration matrix.

$$\begin{pmatrix} s \bullet u_l \\ s \bullet v_l \\ s \end{pmatrix} = \begin{bmatrix} m_{l1} & m_{l2} & m_{l3} & m_{l4} \\ m_{l5} & m_{l6} & m_{l7} & m_{l8} \\ m_{l9} & m_{l10} & m_{l11} & 1 \end{bmatrix} \bullet \begin{pmatrix} {}^{r}X \\ {}^{r}Y \\ {}^{r}Z \\ 1 \end{pmatrix}$$

Advantages:

1. Easy to implement in practice.

2. Precise.

3. Solution to derive the parameters of a camera in 3D space (how ?)

How to determine the coefficients inside a 3D calibration matrix ?

ANSWER:

3. Principle of 3D calibration:

Step 1: Use one camera as example. The 3D calibration matrix relates 3D coordinates to image coordinates in the following way:

$$\begin{pmatrix} s \bullet u \\ s \bullet v \\ s \end{pmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ m_5 & m_6 & m_7 & m_8 \\ m_9 & m_{10} & m_{11} & 1 \end{bmatrix} \bullet \begin{pmatrix} {}^rX \\ {}^rY \\ {}^rZ \\ 1 \end{pmatrix} \tag{1}$$

Step 2: Write out the equations in algebraic form:

$$\left| \begin{aligned} u &= \frac{m_1 \bullet {}^rX + m_2 \bullet {}^rY + m_3 \bullet {}^rZ + m_4}{m_9 \bullet {}^rX + m_{10} \bullet {}^rY + m_{11} \bullet {}^rZ + 1} \\ v &= \frac{m_5 \bullet {}^rX + m_6 \bullet {}^rY + m_7 \bullet {}^rZ + m_8}{m_9 \bullet {}^rX + m_{10} \bullet {}^rY + m_{11} \bullet {}^rZ + 1} \end{aligned} \right. \tag{2}$$

Step 3: Assuming that (u,v) and (X,Y, Z) are known. We have:

$$\begin{cases} ({}^rX)\bullet m_1 + ({}^rY)\bullet m_2 + ({}^rZ)\bullet m_3 + m_4 - ({}^rX\bullet u)\bullet m_9 - ({}^rY\bullet u)\bullet m_{10} - ({}^rZ\bullet u)\bullet m_{11} = u \\ ({}^rX)\bullet m_5 + ({}^rY)\bullet m_6 + ({}^rZ)\bullet m_7 + m_8 - ({}^rX\bullet v)\bullet m_9 - ({}^rY\bullet v)\bullet m_{10} - ({}^rZ\bullet v)\bullet m_{11} = v \end{cases} \quad (3)$$

Step 4: One pair of {(u,v), (X,Y,Z)} gives two linear equations.

Since there are eleven unknowns, we need at least eleven equations. Therefore, the necessary and sufficient condition to estimate the 3D calibration matrix is to have six pairs of {(u,v}, (X,Y,Z)}.

Step 5: Assume that we have “n” pairs (n>5) of {(u,v), (X,Y,Z)}.

We have the following equations:

$$\begin{cases} {}^rX_i\bullet m_1 + {}^rY_i\bullet m_2 + {}^rZ_i\bullet m_3 + m_4 - {}^rX_i\bullet u_i\bullet m_9 - {}^rY_i\bullet u_i\bullet m_{10} - {}^rZ_i\bullet u_i\bullet m_{11} = u_i \\ {}^rX_i\bullet m_5 + {}^rY_i\bullet m_6 + {}^rZ_i\bullet m_7 + m_8 - {}^rX_i\bullet v_i\bullet m_9 - {}^rY_i\bullet v_i\bullet m_{10} - {}^rZ_i\bullet v_i\bullet m_{11} = v_i \end{cases} \quad (4)$$

$$i = 1,2,\ldots,n$$

Step 6: Write the equation system (4) in matrix form:

$$A \bullet V = B \qquad (5)$$

$$A = \begin{bmatrix} {}^rX_1 & {}^rY_1 & {}^rZ_1 & 1 & 0 & 0 & 0 & 0 & -{}^rX_1 \bullet u_1 & -{}^rY_1 \bullet u_1 & -{}^rZ_1 \bullet u_1 \\ 0 & 0 & 0 & 0 & {}^rX_1 & {}^rY_1 & {}^rZ_1 & 1 & -{}^rX_1 \bullet v_1 & -{}^rY_1 \bullet v_1 & -{}^rZ_1 \bullet v_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ {}^rX_n & {}^rY_n & {}^rZ_n & 1 & 0 & 0 & 0 & 0 & -{}^rX_n \bullet u_n & -{}^rY_n \bullet u_n & -{}^rZ_n \bullet u_n \\ 0 & 0 & 0 & 0 & {}^rX_n & {}^rY_n & {}^rZ_n & 1 & -{}^rX_n \bullet v_n & -{}^rY_n \bullet v_n & -{}^rZ_n \bullet v_n \end{bmatrix} \quad V = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ m_5 \\ m_6 \\ m_7 \\ m_8 \\ m_9 \\ m_{10} \\ m_{11} \end{bmatrix} \quad B = \begin{bmatrix} u_1 \\ v_1 \\ \dots \\ u_n \\ v_n \end{bmatrix}$$

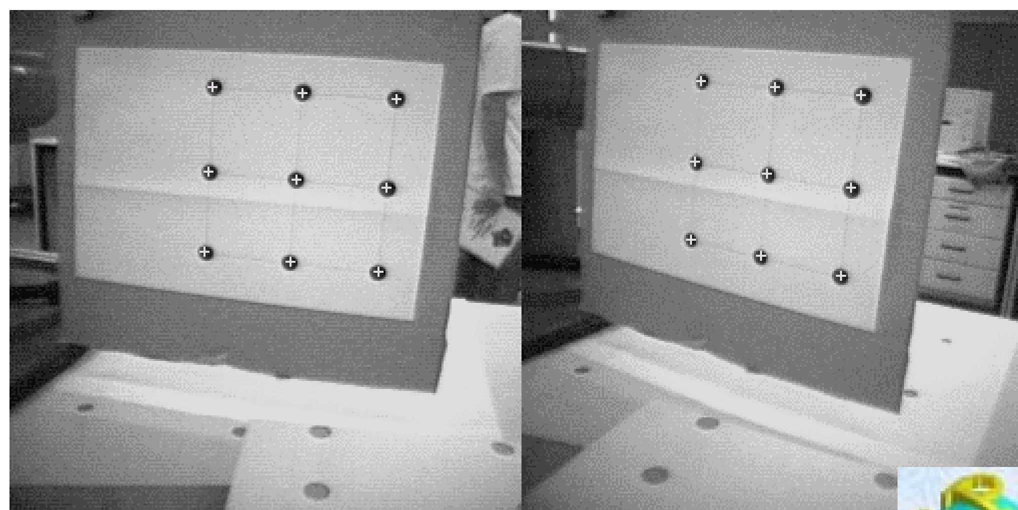
Step 7: The coefficients can be computed by a Least-Square estimation:

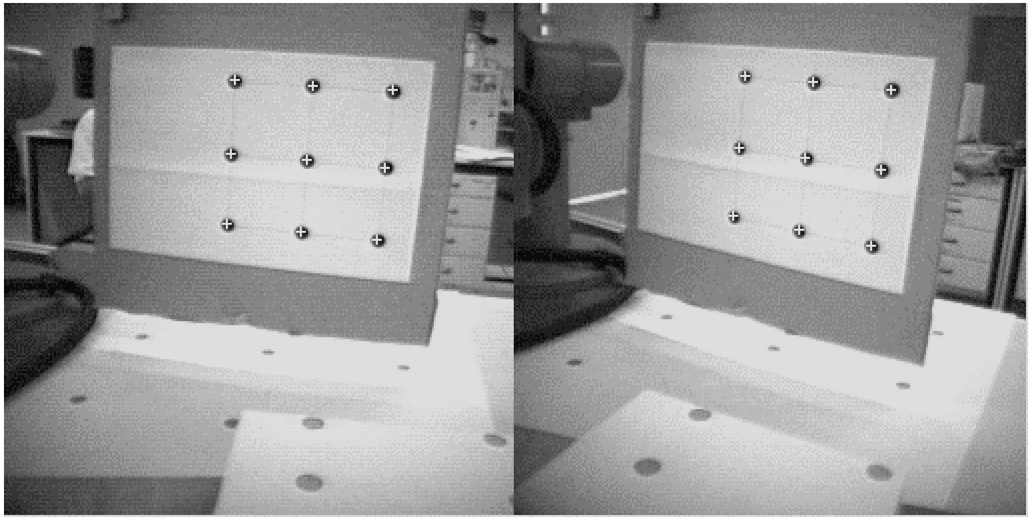
$$V = (A^t \bullet A)^{-1} \bullet (A^t \bullet B) \qquad (6)$$

Example

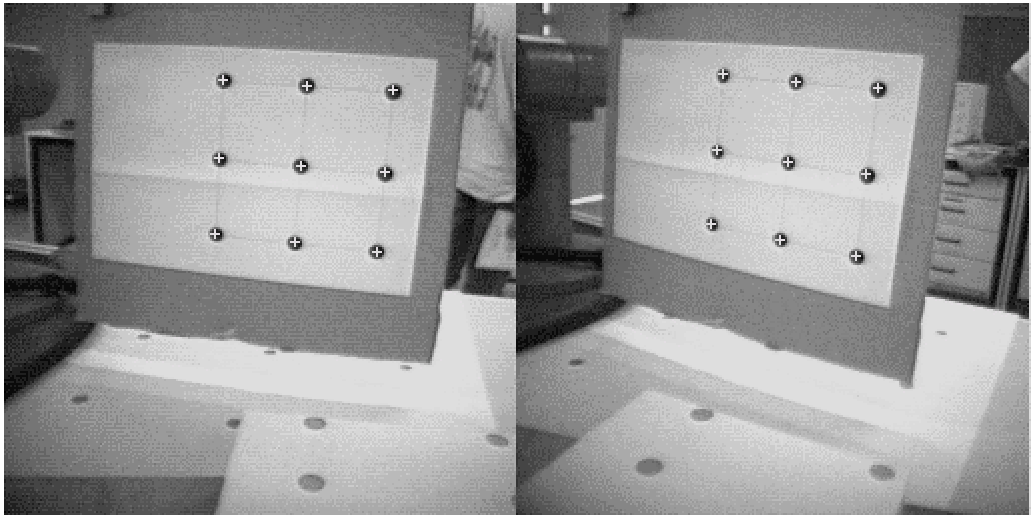
1. System Setup:

A panel with 3x3 dot pattern is hold by a robot arm, and placed in front of a stereo vision system. The horizontal and vertical distances between the dots are (100, 100) (mm). The initial position of the panel defines the XY plane of a world frame. Then, the robot arm moves the panel along Z axis to the locations (0,0,200) (mm) and (0,0,400) (mm), sequentially.
Three pairs of stereo images are captured by the stereo cameras at three different location along Z axis: Z=0, Z=200 (mm) and Z=400 (mm).

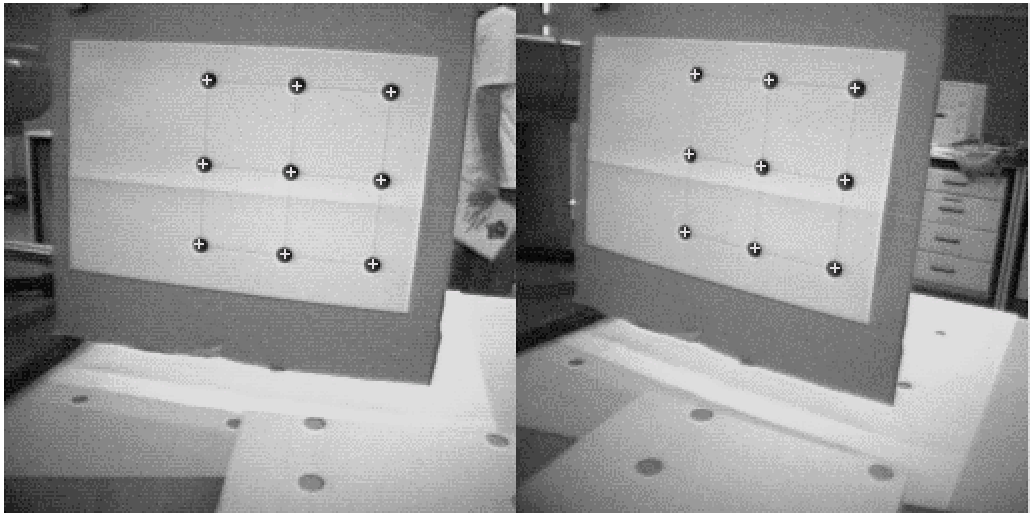




Z=0 (mm)



Z=200 (mm)



Z=400 (mm)

Example

2. Calibration Points Detection:

The dot pattern at each location is detected by a simple algorithm that involves two operations:

a) Dot detection by thresholding.

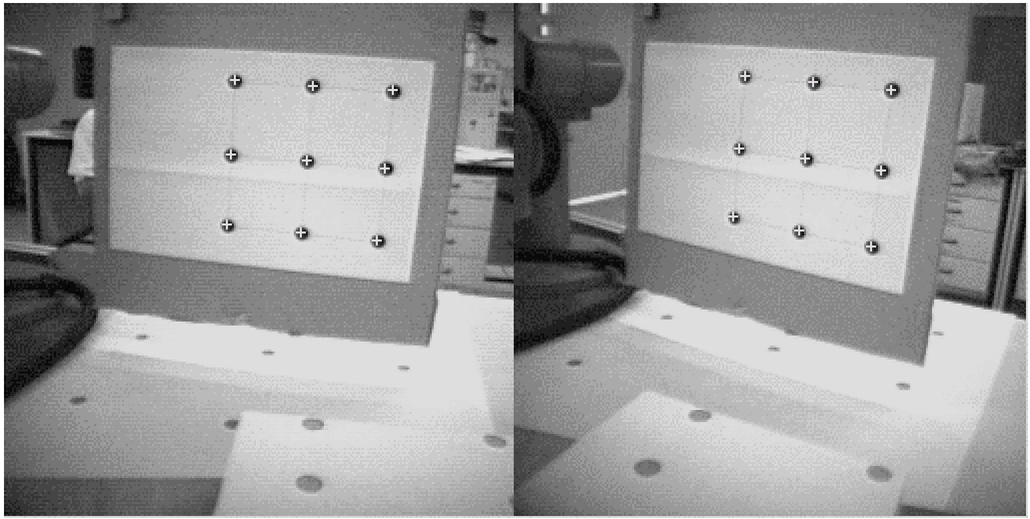
b) Estimation of dot's image coordinates by first grouping the pixels belonging to a dot and then computing the centre of gravity of the grouped pixels.

3. Calibration Data for The Left Camera:

There are 3x9=27 pairs of {(u,v), (X,Y,Z)}. The tables show the values:

Z=0.0 (mm)

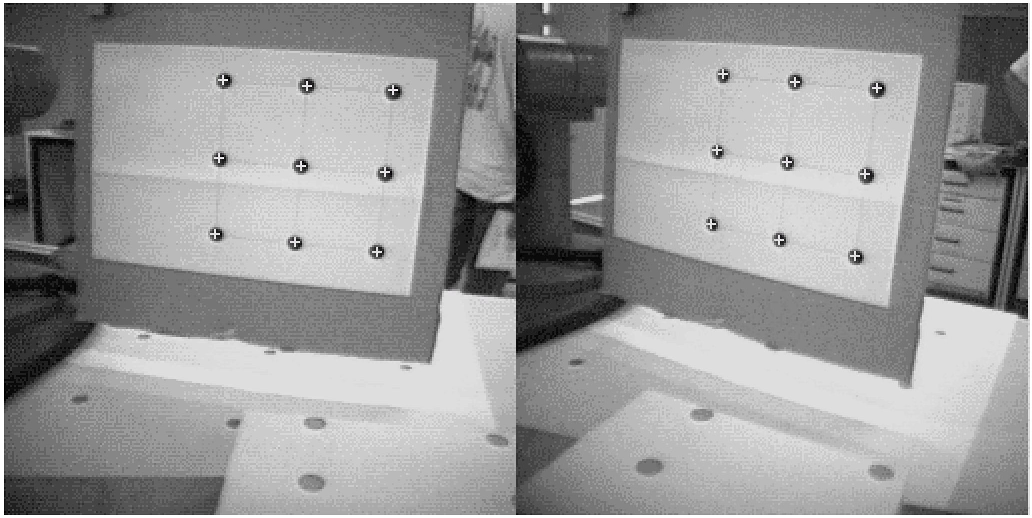
u	v	X	Y	Z
116.210	38.842	0.0	200.0	0.0
155.200	41.075	100.0	200.0	0.0
195.667	44.000	200.0	200.0	0.0
114.500	75.500	0.0	100.0	0.0
152.294	78.500	100.0	100.0	0.0
191.891	82.027	200.0	100.0	0.0
112.451	110.580	0.0	0.0	0.0
149.500	114.500	100.0	0.0	0.0
187.852	118.382	200.0	0.0	0.0



Z=0 (mm)

Z=200.0 (mm)

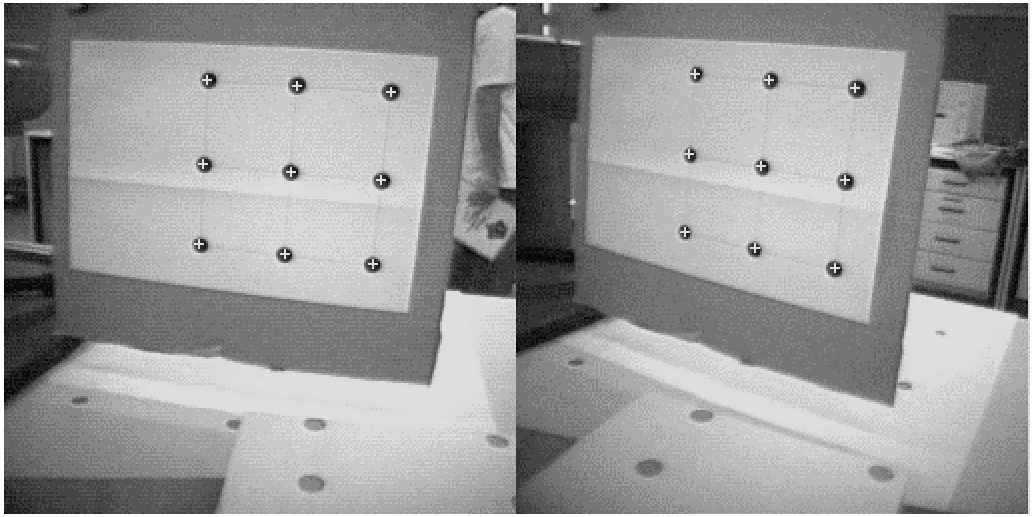
u	v	X	Y	Z
109.814	38.512	0.0	200.0	200.0
151.087	41.000	100.0	200.0	200.0
194.400	44.140	200.0	200.0	200.0
107.789	77.842	0.0	100.0	200.0
148.128	81.128	100.0	100.0	200.0
190.444	84.933	200.0	100.0	200.0
105.806	115.055	0.0	0.0	200.0
145.102	119.641	100.0	0.0	200.0
186.075	123.950	200.0	0.0	200.0



Z=200 (mm)

Z=400.0 (mm)

u	v	X	Y	Z
102.160	38.260	0.0	200.0	400.0
146.358	40.679	100.0	200.0	400.0
193.175	44.175	200.0	200.0	400.0
99.808	80.362	0.0	100.0	400.0
143.081	84.102	100.0	100.0	400.0
188.500	88.500	200.0	100.0	400.0
97.928	120.690	0.0	0.0	400.0
139.837	125.674	100.0	0.0	400.0
184.146	130.583	200.0	0.0	400.0



Z=400 (mm)

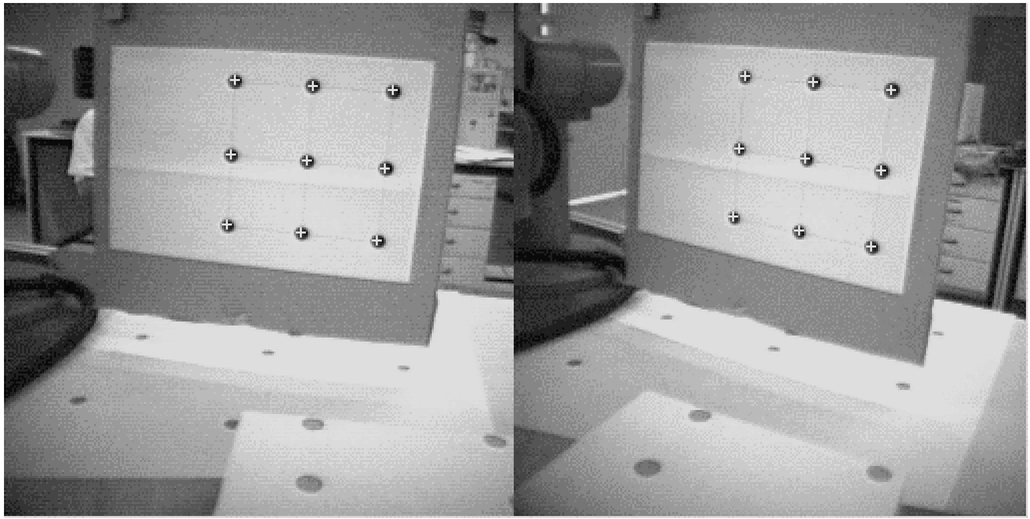
Nanyang Technological University

4. Calibration Data for The Right Camera:

There are 3x9=27 pairs of {(u,v), (X,Y,Z)}. The tables show the values:

Z=0.0 (mm)

u	v	X	Y	Z
115.074	36.407	0.0	200.0	0.0
149.571	39.628	100.0	200.0	0.0
188.659	43.182	200.0	200.0	0.0
112.414	72.448	0.0	100.0	0.0
146.097	77.484	100.0	100.0	0.0
183.684	83.553	200.0	100.0	0.0
109.869	106.782	0.0	0.0	0.0
142.276	113.655	100.0	0.0	0.0
178.576	121.576	200.0	0.0	0.0

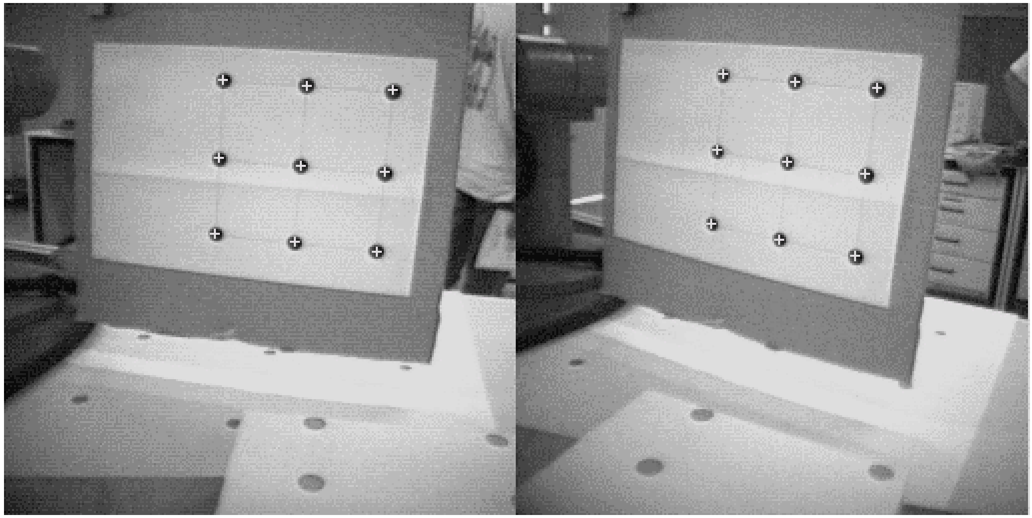


Z=0 (mm)

Nanyang Technological University

Z=200.0 (mm)

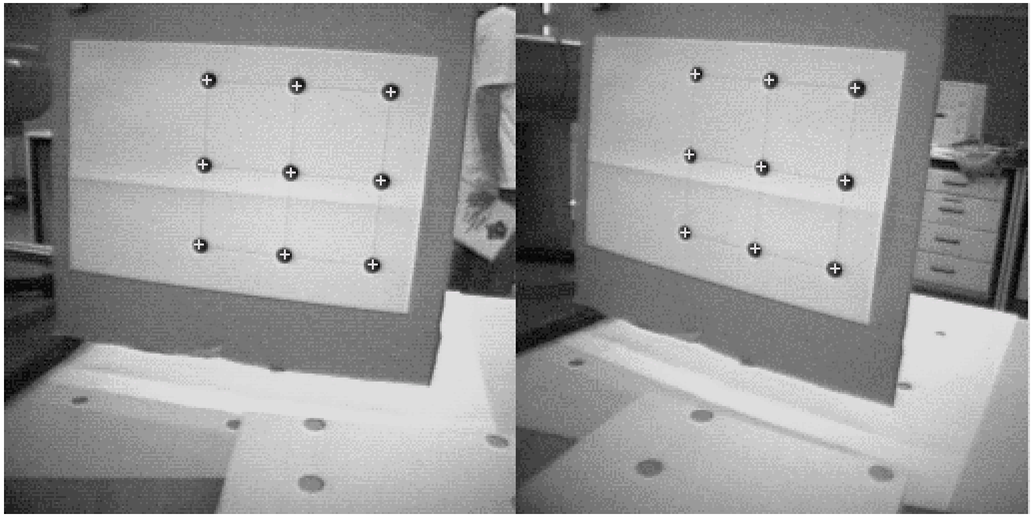
u	v	X	Y	Z
103.314	35.914	0.0	200.0	200.0
139.325	38.975	100.0	200.0	200.0
180.081	42.857	200.0	200.0	200.0
100.375	73.844	0.0	100.0	200.0
135.412	79.529	100.0	100.0	200.0
175.048	85.833	200.0	100.0	200.0
97.724	110.345	0.0	0.0	200.0
131.500	118.000	100.0	0.0	200.0
169.692	126.615	200.0	0.0	200.0



Z=200 (mm)

Z=400.0 (mm)

u	v	X	Y	Z
89.925	35.275	0.0	200.0	400.0
127.109	38.369	100.0	200.0	400.0
169.852	42.500	200.0	200.0	400.0
86.686	75.828	0.0	100.0	400.0
122.925	81.800	100.0	100.0	400.0
164.563	88.963	200.0	100.0	400.0
84.125	114.219	0.0	0.0	400.0
118.846	122.974	100.0	0.0	400.0
158.936	132.638	200.0	0.0	400.0



Z=400 (mm)

5. Estimated 3D Calibrations Matrices:

Left Camera:

$$M_l = \begin{bmatrix} 0.337599 & -0.004174 & -0.068044 & 112.500411 \\ 0.015303 & -0.367294 & -0.013599 & 110.527718 \\ -0.000209 & -0.000212 & -0.000322 & 1.0 \end{bmatrix}$$

Right Camera:

$$M_r = \begin{bmatrix} 0.245747 & 0.004458 & -0.087352 & 109.666495 \\ 0.007531 & -0.358938 & -0.012992 & 106.705739 \\ -0.000554 & -0.000216 & -0.000280 & 1.0 \end{bmatrix}$$

Quiz: *How to verify that the above 3D calibration matrices are correct ?*

Self-Study

Question:

How to derive the intrinsic and extrinsic parameters of a camera in 3D space ?

Hints:

Step 1: Define R1, R2, and R3 to be the row vectors of the rotation matrix:

$$R_1 = \begin{bmatrix} r_{11} & r_{12} & r_{13} \end{bmatrix}$$
$$R_2 = \begin{bmatrix} r_{21} & r_{22} & r_{23} \end{bmatrix}$$
$$R_3 = \begin{bmatrix} r_{31} & r_{32} & r_{33} \end{bmatrix}$$

Step 2: From the equation in Slide 6, we can have the following relationship:

$$\begin{bmatrix} \rho \bullet u \\ \rho \bullet v \\ \rho \end{bmatrix} = t_z \bullet \begin{bmatrix} (\frac{f}{D_x} \bullet \frac{R_1}{t_z} + u_0 \bullet \frac{R_3}{t_z}) & (\frac{f}{D_x} \bullet \frac{t_x}{t_z} + u_0) \\ (\frac{f}{D_y} \bullet \frac{R_2}{t_z} + v_0 \bullet \frac{R_3}{t_z}) & (\frac{f}{D_y} \bullet \frac{t_y}{t_z} + v_0) \\ \frac{R_3}{t_z} & 1 \end{bmatrix} \bullet \begin{bmatrix} {}^rX \\ {}^rY \\ {}^rZ \\ 1 \end{bmatrix}$$

Step 3: If the 3D calibration matrix is known:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & 1 \end{bmatrix} \qquad \begin{bmatrix} (\frac{f}{D_x} \bullet \frac{R_1}{t_z} + u_0 \bullet \frac{R_3}{t_z}) & (\frac{f}{D_x} \bullet \frac{t_x}{t_z} + u_0) \\ (\frac{f}{D_y} \bullet \frac{R_2}{t_z} + v_0 \bullet \frac{R_3}{t_z}) & (\frac{f}{D_y} \bullet \frac{t_y}{t_z} + v_0) \\ \frac{R_3}{t_z} & 1 \end{bmatrix} = M$$

Step 4: The above equality gives the following equations:

$$\frac{f}{D_x} \bullet \frac{R_1}{t_z} + u_0 \bullet \frac{R_3}{t_z} = [m_{11} \quad m_{12} \quad m_{13}] \qquad \frac{f}{D_x} \bullet \frac{t_x}{t_z} + u_0 = m_{14}$$

$$\frac{f}{D_y} \bullet \frac{R_2}{t_z} + v_0 \bullet \frac{R_3}{t_z} = [m_{21} \quad m_{22} \quad m_{23}] \qquad \frac{f}{D_y} \bullet \frac{t_y}{t_z} + v_0 = m_{24}$$

$$\frac{R_3}{t_z} = [m_{31} \quad m_{32} \quad m_{33}]$$

Step 5: From the above equations and the fact that the rotation matrix is orthonormal, you can derive the solution.

SUMMARY

1. For a camera in 3D space, the "intrinsic" parameters include:
   a. Focal length: f.
   b. The location of the origin of the real image frame with respect to the digital image frame: (u_0, v_0). (this is also called "optical center").
   c. The sampling steps of digital image: (D_x, D_y).
2. For a camera in 3D space, the "extrinsic" parameters include:
   a. The rotation matrix from the world frame to the camera frame: R.
   b. The translation vector from the world frame to the camera frame: T.
3. Knowing the intrinsic and extrinsic parameters of a camera in 3D space, the 3D calibration matrix can be computed.
4. Knowing six pairs of {(u,v), (X,Y,Z)}, the 3D calibration matrix for a camera in 3D space can be estimated.
5. Knowing the 3D calibration matrix, the intrinsic and extrinsic parameters of the corresponding camera in 3D space can be derived.

Nanyang Technological University