Camera Calibration

Dr. Gerhard Roth

Winter 2011

Finding camera parameters (intrinsic)

- Can use the EXIF tag for any digital image
  - Has focal length f in millimeters but not the pixel size
  - But you can get the pixel size from the camera manual
  - There are only a finite number of different pixels sizes because number of sensing element sizes is limited
  - If there is not a lot of image distortion due to optics then this approach is sufficient (only linear calibration)
- Can perform explicit camera calibration
  - Put a calibration pattern in front of the camera
  - Take a number of different pictures of this pattern
  - Now run the calibration algorithm (different types)
  - Result is intrinsic camera parameters (linear and non-linear) and the extrinsic camera parameters of all the images

Explicit camera calibration

- Use a calibration pattern with known geometry
  - In Opencv use a checkerboard
  - Other systems use special targets with known 3d geometry
- Write equations linking co-ordinates of the projected points, and the camera parameters
- From images of the calibration target
  - Intrinsic camera parameters
    - (depend only on camera characteristics)
  - Extrinsic camera parameters
    - (depend only on position camera)
  - In OpenCV the calibration process finds fx, fy, ox, oy, along with the distortion parameters
  - We study a method that does not find the distortion parameters

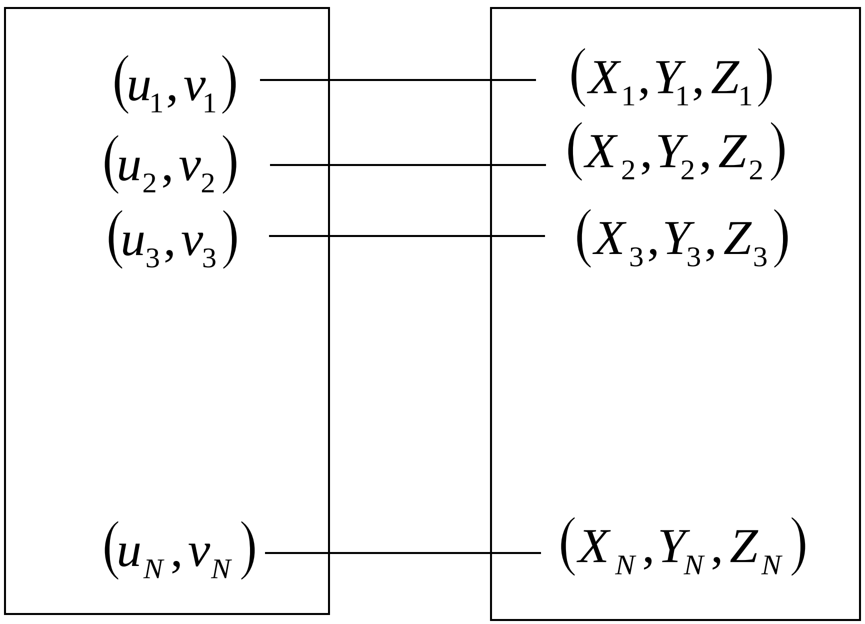
Calibration using known 3d geometry

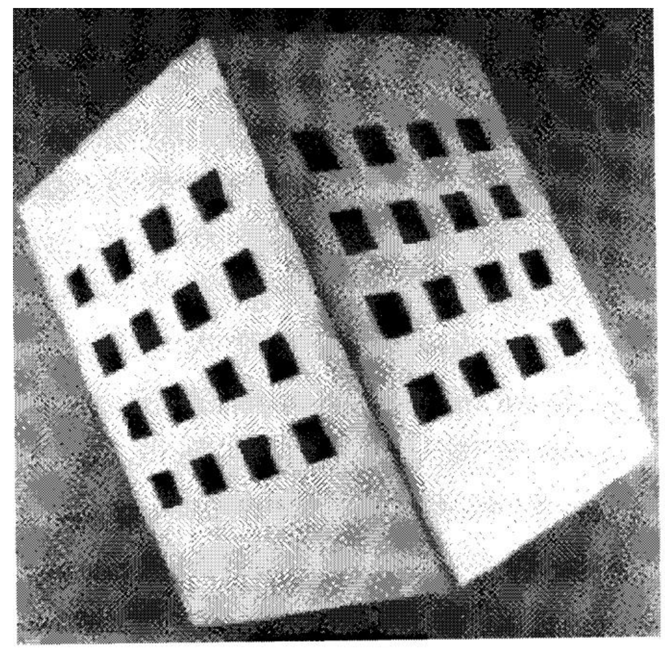
- Use a calibration pattern with known 3d geometry (often a box, not planar)
- Write projection equations linking known co-ordinates of a set of 3-D points and their projections and solve for camera parameters
- Given a set of one or more images of the calibration pattern estimate
  - Intrinsic camera parameters
    - (depend only on camera characteristics)
  - Extrinsic camera parameters
    - (depend only on position camera)
- We do not estimate distortion parameters

Estimating camera parameters

- Projection matrix

Calibration pattern





Camera parameters

- Intrinsic parameters (K matrix)
  - There are 5 intrinsic parameters
  - Focal length f
  - Pixel size in x and y directions, sx and sy
  - Principal point ox, oy
- But they are not independent
  - Focal length fx = f / sx and fy = f / sy
  - Principal point ox, oy
  - This makes four intrinsic parameters
- Extrinsic parameters [R| T]
  - Rotation matrix and translation vector of camera
  - Relations camera position to a known frame
  - [R|T] are the intrinsic parameters
- Projection matrix
  - 3 by 4 matrix P =K [R | T] is called projection matrix

Projection Equations

Projective Space

- Add fourth coordinate
  - $P_w = (X_w, Y_w, Z_w, 1)^T$
- Define $(u,v,w)^T$ such that
  - $u/w = x_{im}$, $v/w = y_{im}$

$$\begin{pmatrix} x_{im} \\ y_{im} \end{pmatrix} = \begin{pmatrix} u/w \\ v/w \end{pmatrix} \Leftarrow \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \mathbf{M_{int}M_{ext}} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

3x4 Matrix E_{ext}

- Only extrinsic parameters
- World to camera

$$\mathbf{M}_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T & T_x \\ \mathbf{R}_2^T & T_y \\ \mathbf{R}_3^T & T_z \end{bmatrix}$$

3x3 Matrix E_{int}

- Only intrinsic parameters
- Camera to frame

$$\mathbf{M}_{int} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

Simple Matrix Product! Projective Matrix

$M = M_{int}M_{ext}$

- $(X_w, Y_w, Z_w)^T \rightarrow (x_{im}, y_{im})^T$
- Linear Transform from projective space to projective plane
- M defined up to a scale factor – 11 independent entries

Two different calibration methods

- Both use a set of 3d points and 2d projections
- Direct approach (called Tsai method)
  - Write projection equations in terms of all the parameters
    - That is all the unknown intrinsic and extrinsic parameters
  - Solve for these parameters using non-linear equations
- Projection matrix approach
  - Compute the projection matrix (the 3x4 matrix M)

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}$$

  - Compute camera parameters as closed-form functions of M

•Two different calibration methods

- Both approaches work with same data
  - Projection matrix approach is simpler to explain than the direct approach
- Direct approach requires an extra step
  - There are also other calibration methods
- But all calibration methods
  - Use patterns with know geometry or shape
  - Take multiple views of theses patterns
  - Match the information across the different views
- Perform some mathematics to calculate the intrinsic and extrinsic camera parameters
- We look at simplified case of only one view!

Estimating the projection matrix

World – Frame Transform

- Drop “im” and “w”
- N pairs (xi,yi) <-> (Xi,Yi,Zi)

$$x_i = \frac{u_i}{w_i} = \frac{m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

$$y_i = \frac{u_i}{w_i} = \frac{m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

Linear equations of m

- 2N equations, 11 independent variables
- N >=6 , SVD => m up to a unknown scale

$$\mathbf{Am} = \mathbf{0}$$

$$\mathbf{A} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1X_1 & -x_1Y_1 & -x_1Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1X_1 & -y_1Y_1 & -y_1Y_1 & -y_1 \\ . & . & . & . & . & . & & & & & & \end{bmatrix}$$

$$\mathbf{m} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}^T$$

Homogeneous System

- M linear equations of form A**x** = 0
- If we have a given solution x1, s.t. Ax1 = 0 then c * x1 is also a solution A(c* x1) = 0
- Need to add a constraint on **x**,
  - Basically make **x** a unit vector $x^T x = 1$
- Can prove that the solution is the eigenvector corresponding to the single zero eigenvalue of that matrix $A^T A$
  - This can be computed using eigenvector of SVD routine
  - Then finding the zero eigenvalue (actually smallest)
  - Returning the associated eigenvector

Decompose projection matrix

- 3x4 Projection Matrix M computed previously
  - Both intrinsic (4) and extrinsic (6) – 10 parameters

$$\mathbf{M} = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

From M^ to parameters (p134-135)

- Find scale $|\gamma|$ by using unit vector R_3^T
- Determine T_z and sign of γ from m_{34} (i.e. q_{43})
- Obtain R_3^T
- Find (Ox, Oy) by dot products of Rows q1. q3, q2.q3, using the orthogonal constraints of R
- Determine fx and fy from q1 and q2 All the rests: R_1^T, R_2^T, Tx, Ty
- Enforce orthogonality on R by using SVD

Rotation Matrices

- Are special 3 by matrices R

$$RR^T = R^T R = I \qquad \det(R) = 1$$

- Rows and columns of R are mutually orthogonal unit vectors
- Each row and column has unit norm
- Dot product of each row with others is zero
- Dot product of each column with others is zero

Ensuring the orthogonality of R

- The computation of R does not take into account explicitly the orthogonality constraints on R
- The estimate $\hat{R}$ of R cannot be expected to be orthogonal:
- We can "enforce" the orthogonality on $\hat{R}$ using SVD:

$$\hat{R}\hat{R}^T = I \qquad \hat{R} = UDV^T$$

- Replace *D (singular values)* with Identity Matrix *I*:

$$\hat{R}' = UIV^T \ (\hat{R}'\hat{R}'^T = I)$$

Direct Approach

- Main steps

(1) Assuming that o_x and o_y are known, estimate all the remaining parameters.

(2) Estimate o_x and o_y

Projection Equations (book notation)

- Projection equations (camera and world)

$$x = f\,\frac{X_C}{Z_C} \qquad y = f\,\frac{Y_C}{Z_C}$$

- Calibration equations

$$x = -(x_{im} - o_x)s_x \qquad y = -(y_{im} - o_x)s_y$$

- Using the projection equations and camera calibration information we get:

$$x_{im} = -\frac{f}{s_x}\frac{X^C}{Z^C} + o_x \qquad y_{im} = -\frac{f}{s_y}\frac{Y^C}{Z^C} + o_y$$

Projection Equations (book notation)

- 3D point P in world reference frame is $\left[X^w, Y^w, Z^w\right]$
- $\left[X^c, Y^c, Z^c\right]$ are the co-ordinates of P in camera frame
- Extrinsic parameters, apply rotation then translation

$$\begin{bmatrix} X^c \\ Y^c \\ Z^c \end{bmatrix} = R \begin{bmatrix} X^w \\ Y^w \\ Z^w \end{bmatrix} + T$$

$$X^c = r_{11} X^w + r_{12} Y^w + r_{13} Z^w + T_x$$

$$Y^c = r_{21} X^w + r_{22} Y^w + r_{23} Z^w + T_y$$

$$Z^c = r_{31} X^w + r_{32} Y^w + r_{33} Z^w + T_z$$

Projection Equations (book notation)

- Now let $f_x = f / s_x,\ f_y = f / s_y, \alpha = s_y / s_x$

$$x - o_x = -f_x \frac{r_{11}X^{(W)} + r_{12}Y^{(W)} + r_{13}Z^{(W)} + T_X}{r_{31}X^{(W)} + r_{32}Y^{(W)} + r_{33}Z^{(W)} + T_Z}$$

$$y - o_y = -f_y \frac{r_{21}X^{(W)} + r_{22}Y^{(W)} + r_{23}Z^{(W)} + T_Y}{r_{31}X^{(W)} + r_{32}Y^{(W)} + r_{33}Z^{(W)} + T_Z}$$

- Now assume that Ox, and Oy are zero then

$$x_i f_y \left(r_{21}X_i + r_{22}Y_i^{(W)} + r_{23}Z_i^{(W)} + T_Y \right) = y_i f_x \left(r_{11}X_i^{(W)} + r_{12}Y_i^{(W)} + r_{13}Z_i^{(W)} + T_X \right)$$

$$x_i X_i^{(W)} r_{21} + x_i Y_i^{(W)} r_{22} + x_i Z_i^{(W)} r_{23} + x_i T_Y - y_i X_i^{(W)} \alpha r_{11} - y_i Y_i^{(W)} \alpha r_{12} - y_i Z_i^{(W)} \alpha r_{13} - y_i \alpha T_X = 0$$

Remember that x, y are now pixel (not camera) co-ordinates

Projection Equations (book notation)

For each corresponding pair (u_i, v_i) and (X_i, Y_i, Z_i)

$$x_i\ X_i^{(W)} v_1 + x_i\ Y_i^{(W)} v_2 + x_i Z_i^{(W)} v_3 + x_i v_4 - y_i\ X_i^{(W)} v_5 - y_i\ Y_i^{(W)} v_6 + y_i\ Z_i^{(W)} v_7 + y_i v_8 = 0$$

$$v = [v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5 \quad v_6 \quad v_7 \quad v_8]^T \qquad Av = 0$$

$$v_1 = r_{21} \quad v_5 = \alpha r_{11}$$
$$v_2 = r_{22} \quad v_6 = \alpha r_{12}$$
$$v_3 = r_{23} \quad v_7 = \alpha r_{13}$$
$$v_4 = T_Y \quad v_8 = \alpha T_X$$

$$A = \begin{bmatrix} x_1 X_1^w & x_1 Y_1^w & x_1 Z_1^w & x_1 & -y_1 X_1^w & -y_1 Y_1^w & -y_1 Z_1^w & -y_1 \\ x_2 X_2^w & x_2 Y_2^w & x_2 Z_2^w & x_2 & -y_2 X_2^w & -y_2 Y_2^w & -y_2 Z_2^w & -y_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_N X_N^w & x_N Y_N^w & x_N Z_N^w & x_N & -y_N X_N^w & -y_N Y_N^w & -y_N Z_N^w & -y_N \end{bmatrix}$$

- If there are N points, A is an N by 8 matrix
  - Need to solve A v = 0 but require an extra constraints
  - This is a homogenerous system with constraints

Estimating Camera Parameters

Let $\bar{v}$ be the obtained solution vector, then

$$\bar{v} = \gamma\left(r_{21} \quad r_{22} \quad r_{23} \quad T_Y \quad \alpha r_{11} \quad \alpha r_{12} \quad \alpha r_{13} \quad \alpha T_X\right)^T$$

Since $r_{21}^2 + r_{22}^2 + r_{23}^2 = 1$, we have

$$\sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{\gamma^2\left(r_{21}^2 + r_{22}^2 + r_{23}^2\right)} = |\gamma|$$

- Using magnitude of γ we can determine α

Since $r_{21}^2 + r_{22}^2 + r_{23}^2 = 1$, we have

$$\sqrt{v_5^2 + v_6^2 + v_7^2} = \sqrt{\gamma^2\alpha^2\left(r_{21}^2 + r_{22}^2 + r_{23}^2\right)} = \alpha|\gamma|$$

Estimating Camera Parameters

Determine r_{31}, r_{32}, r_{33}

- Can be estimated as the cross product of R_1 and R_2:

$$R_3 = R_1 \ x \ R_2$$

- The sign of R_3 is already fixed (the entries of R_3 remain unchanged if the signs of all the entries of R_1 and R_2 are reversed).

Estimating Camera Parameters

Determine the sign of γ

- Consider the following equations again:

$$x = -f/s_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z} = -f/s_x \frac{X_c}{Z_c}$$

$$y = -f/s_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}r_w + r_{33}Z_w + T_z} = -f/s_y \frac{Y_c}{Z_c}$$

- If $Z_c > 0$, then x and $r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x$ must have opposite signs (it is sufficient to check the sign for one of the points).

if $x(r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x) > 0$, then
  reverse the signs of r_{11}, r_{12}, r_{13}, and T_x
else
  no further action is required

Estimating Camera Parameters

- Similarly, if $Z_c > 0$, then y and $r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_x$ must have opposite signs (it is sufficient to check the sign for one of the points).

if $y(r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) > 0$, then
    reverse the signs of r_{21}, r_{22}, r_{23}, and T_y
else
    no further action is required

Must do more work to compute Tz and fx, and also to find Ox, Oy

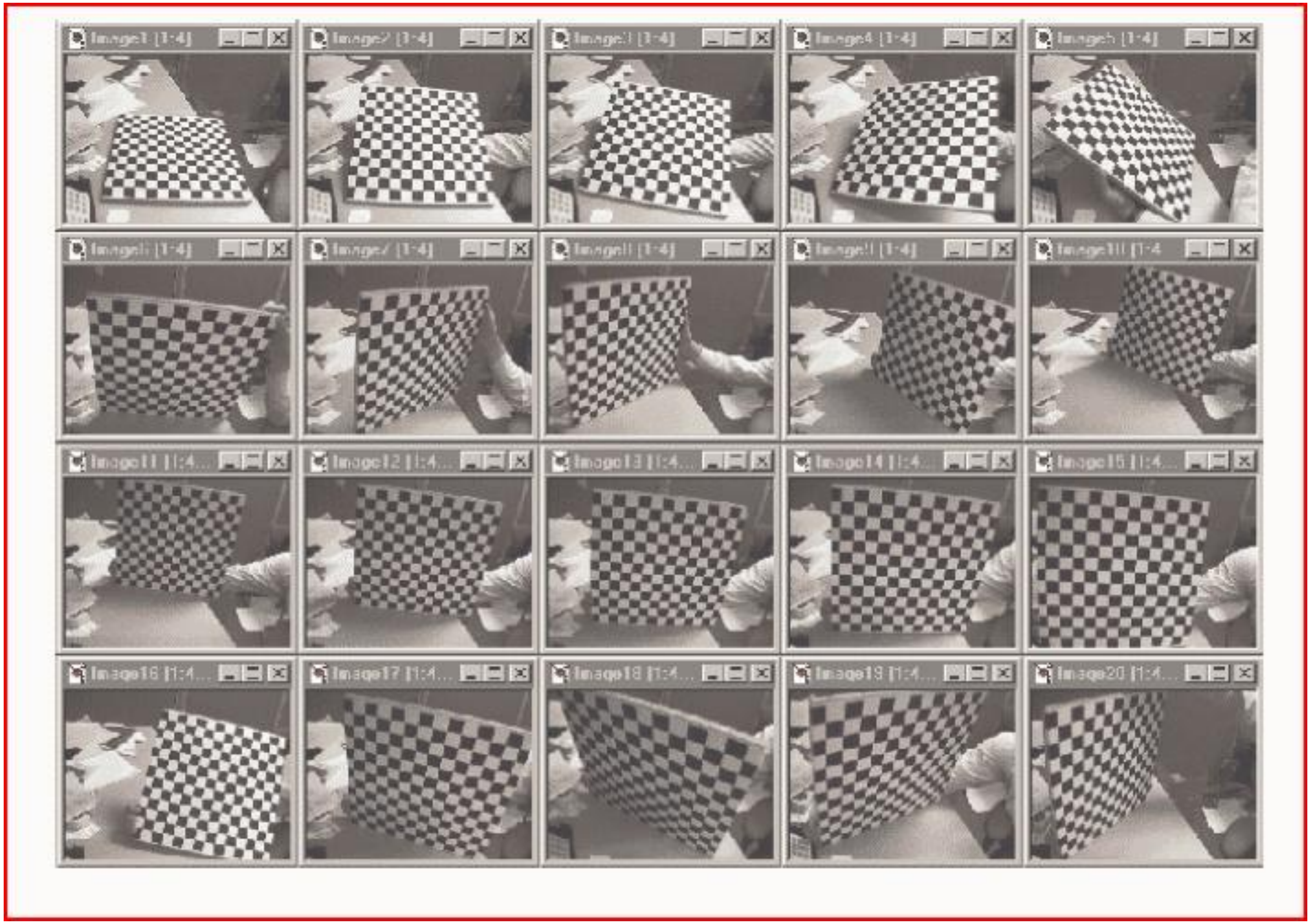
Calibration Summary

- Comparison of methods
  - Direct approach requires extra step to find Ox, Oy
  - Projection approach finds Ox, and Oy at same time
    - Is simpler mathematically than the direct approach
  - Both methods require a refit to find a “valid” R matrix
- There are other calibration methods
  - Zhang approach uses flat plane (implemented in OpenCV)
  - Plane must be flat, but do not need 3D co-ordinates
- But all calibration methods
  - Have some known targets with known 3D geometry or shape
  - Take a number of images of these targets
  - From these measurements calculate the camera partakers
  - Are essential for further processing like reconstruction

Multiple View/Camera Calibration

- Previous math describes the calibration process for a single image
  - We usually take multiple images of the same calibration target (from a variety of different views)
  - Simultaneously find all extrinsic parameters and all the intrinsic parameters of the single camera
- Also calibrate radial distortion using fact that there are straight lines in the pattern
- OpenCV code can do this using a checkerboard pattern
- Zhang’s algorithm is used most in practice

Input set of 2d Calibration Patterns



Final Camera positions and the pattern