Department of Computer Engineering
University of California at Santa Cruz

Camera Calibration

CMPE 264: Image Analysis and Computer Vision

Hai Tao

Department of Computer Engineering
University of California at Santa Cruz

The problem of camera calibration

- Estimation of the 3D geometry of the scene from images is an important task for a machine vision system.
- For a perspective camera, recall that the imaging process can be described as

$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} \cong K[R,\mathbf{T}]\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

  where K is the camera matrix and $\mathbf{T} = -RT$

$$K = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

  and $f_x = fs_x, f_y = fs_y$
- If the camera matrix K and the camera motion R and T are know in advance, then the scene geometry can be computed easily.
- The process of estimating K is called camera calibration

The problem of camera calibration

- Two categories of camera calibration algorithms
  - Calibration using calibration patterns – taking multiple images of a pattern from different viewpoints. Estimating camera matrix K using these images
  - Auto-calibration – estimating camera K directly from real image sequences
- Methods covered in this lecture are in the first category
  - Tsai's calibration algorithm – direct recovery of camera parameters
  - Estimating camera parameters from projection matrix
  - Zhenyou Zhang's calibration algorithm using a planar calibration object

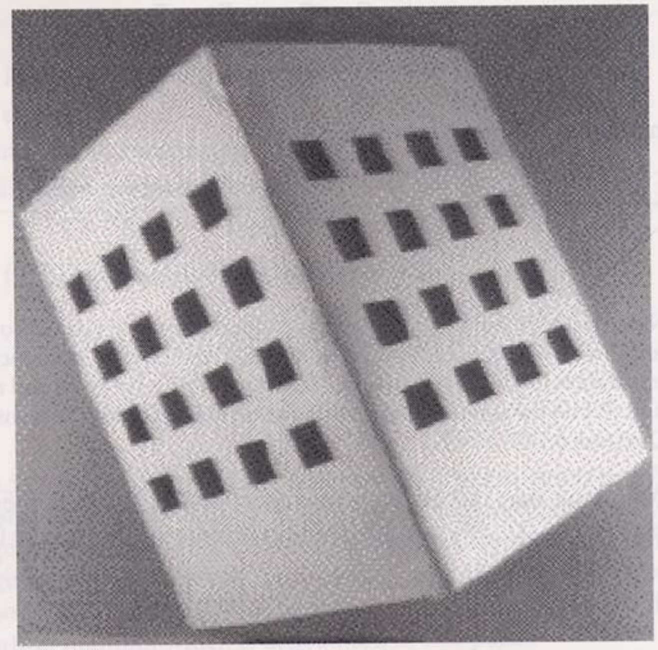


Figure 6.1 The typical calibration pattern used in this chapter.

Direct parameter calibration (Tsai 1987)

- Notation

$\boxed{P_i = [X_i^w, Y_i^w, Z_i^w]^T}$ - the known 3D position of the ith pattern point in the world coordinate system

$\boxed{p_i = [x_{i,im}, y_{i,im}]^T}$ - image coordinates of the ith point

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{21} & r_{32} & r_{33} \end{bmatrix}$$ - the rotation matrix

$$T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$ - the translation vector

$$K = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$ - the camera matrix, with four unknown parameters

Direct parameter calibration (Tsai 1987)

- The basic relationship for each pattern point

$$x_{im} - u_0 = f_x \frac{r_{11}X^w + r_{12}Y^w + r_{13}Z^w + T_x}{r_{31}X^w + r_{32}Y^w + r_{33}Z^w + T_z}$$
$$y_{im} - v_0 = f_y \frac{r_{21}X^w + r_{22}Y^w + r_{23}Z^w + T_y}{r_{31}X^w + r_{32}Y^w + r_{33}Z^w + T_z} \qquad (1)$$

Which variables are known ? Which need to be estimated ?

- We will assume that the principal point $[u_0, v_0]^T$ is known (usually as the center of the image). R, T, f_x, f_y need to be estimated. For simplicity, we denote

$$x_i = x_{i,im} - u_0$$
$$y_i = y_{im} - v_0$$

Direct parameter calibration (Tsai 1987)

From (1), we obtain the equation

$$x_i f_y (r_{21} X_i^w + r_{22} Y_i^w + r_{23} Z_i^w + T_y) = y_i f_x (r_{11} X_i^w + r_{12} Y_i^w + r_{13} Z_i^w + T_x) \qquad (2)$$

If we denote the aspect ration as $\alpha = f_x / f_y$ and

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

then (2) can be written as

$$x_i X_i^w v_1 + x_i Y_i^w v_2 + x_i Z_i^w v_3 + x_i v_4 - y_i X_i^w v_5 - y_i Y_i^w v_6 - y_i Z_i^w v_7 - y_i v_8 = 0$$

If we denote $\mathbf{v} = [v_1, \ldots, v_8]^T$ and

$$A = \begin{bmatrix} x_1 X_1^w & x_1 Y_1^w & x_1 Z_1^w & x_1 & -y_1 X_1^w & -y_1 Y_1^w & -y_1 Z_1^w & -y_1 \\ x_2 X_2^w & x_2 Y_2^w & x_2 Z_2^w & x_2 & -y_2 X_2^w & -y_2 Y_2^w & -y_2 Z_2^w & -y_2 \\ & & & & \cdots & & & \\ x_N X_N^w & x_N Y_N^w & x_N Z_N^w & x_N & -y_N X_N^w & -y_N Y_N^w & -y_N Z_N^w & -y_N \end{bmatrix} \qquad (3)$$

then $\boxed{A\mathbf{v} = 0}$

Direct parameter calibration (Tsai 1987)

- If $N \geq 7$, and the points are not coplanar, and the the rank of A is 7, then there is a nontrival solution, which is the eigenvector corresponding to the 0 eigenvalue of $A^T A$. In practice the solution is the eigenvector corresponding to the smallest eigenvalue of $A^T A$. Tsai has proved that the rank of A is 7 in the ideal case.
- Suppose the eigenvector is $\bar{\mathbf{v}}$, then $\bar{\mathbf{v}} = \gamma[r_{21}, r_{22}, r_{23}, T_y, \alpha r_{11}, \alpha r_{12}, \alpha r_{13}, \alpha T_x]^T$ The scale γ and α can be computed from

$$\gamma = \| [\bar{v}_1, \bar{v}_2, \bar{v}_3] \| \; or \; \gamma = -\| [\bar{v}_1, \bar{v}_2, \bar{v}_3] \|$$

$$\alpha = \| [\bar{v}_5, \bar{v}_6, \bar{v}_7] \| / \| [\bar{v}_1, \bar{v}_2, \bar{v}_3] \|$$

then

$$T_x = \frac{\bar{v}_8}{\gamma\alpha} \qquad [r_{21}, r_{22}, r_{23}] = [\bar{v}_1, \bar{v}_2, \bar{v}_3] / \gamma$$

$$[r_{11}, r_{12}, r_{13}] = [\bar{v}_5, \bar{v}_6, \bar{v}_7] / \gamma\alpha \qquad (4)$$

$$T_y = \bar{v}_4 / \gamma \qquad [r_{31}, r_{32}, r_{33}]^T = [r_{11}, r_{12}, r_{13}]^T \times [r_{21}, r_{22}, r_{23}]^T$$

Department of Computer Engineering
University of California at Santa Cruz

Direct parameter calibration (Tsai 1987)

- Make R orthonormal – in the previous computation, the first two rows of the rotation matrix may not be orthogonal to each other. To make R orthonormal, do the following:
  - Compute SVD of R as $R=UDV^T$, update R'= UIV^T
- Determine the sign of α by making sure x_i and $r_{11}X_i^w + r_{12}Y_i^w + r_{13}Z_i^w + T_x$ have the same sign
- Now, we have computed R, T_x, T_y, α and need to estimate T_z, f_x. To do this, recall (1)

$$x_i(r_{31}X_i^w + r_{32}Y_i^w + r_{33}Z_i^w + \boxed{T_z}) = \boxed{f_x}(r_{11}X_i^w + r_{12}Y_i^w + r_{13}Z_i^w + T_x)$$

For all the points this can be written as $F\begin{bmatrix} T_z \\ f_x \end{bmatrix} = \mathbf{b}$ where

$$F = \begin{bmatrix} x_1 & -r_{11}X_i^w + r_{12}Y_i^w + r_{13}Z_i^w + T_x \\ & \dots \\ x_N & -r_{11}X_N^w + r_{12}Y_N^w + r_{13}Z_N^w + T_x \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} -x_i(r_{31}X_i^w + r_{32}Y_i^w + r_{33}Z_i^w) \\ \dots \\ -x_N(r_{31}X_N^w + r_{32}Y_N^w + r_{33}Z_N^w) \end{bmatrix}$$

Solution

$$\boxed{\begin{bmatrix} T_z \\ f_x \end{bmatrix} = (F^T F)^{-1} F^T \mathbf{b}} \qquad (5)$$

Summary of Tsai's algorithm

- Construct the matrix A from 2D and 3D coordinates of feature points (in the 2D case, subtract the center) as in (3)
- Solve $A\mathbf{v} = 0$ by computing the eigenvector $\bar{\mathbf{v}}$ corresponding to the smallest eigenvalue of $A^T A$
- Solve R, T_x, T_y, α and the scale λ up to a sign using (4)
- Determine the sign of γ by making sure x_i and $r_{11}X_i^w + r_{12}Y_i^w + r_{13}Z_i^w + T_x$ have the same sign
- Construct the matrices F and $\mathbf{b}$ and solve T_z, f_x using (5)

Calibration via recovering the camera projection matrix

- Recall that

$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} \cong M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

This can be written as

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

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

The basic idea is to recover M first using linear a method, then recover all the other parameters

Department of Computer Engineering
University of California at Santa Cruz

Calibration via recovering the camera projection matrix

- Since
$$x_i(m_{31}X_i^w + m_{32}Y_i^w + m_{33}Z_i^w + m_{34}) = m_{11}X_i^w + m_{12}Y_i^w + m_{13}Z_i^w + m_{14}$$
$$y_i(m_{31}X_i^w + m_{32}Y_i^w + m_{33}Z_i^w + m_{34}) = m_{21}X_i^w + m_{22}Y_i^w + m_{23}Z_i^w + m_{24}$$

this can be written as a linear system

$$\boxed{A\mathbf{m} = 0}$$

where

$$A = \begin{bmatrix} X_1^w & Y_1^w & Z_1^w & 1 & 0 & 0 & 0 & 0 & -x_1X_1^w & -x_1Y_1^w & -x_1Z_1^w & -x_1 \\ 0 & 0 & 0 & 0 & X_1^w & Y_1^w & Z_1^w & 1 & -y_1X_1^w & -y_1Y_1^w & -y_1Z_1^w & -y_1 \\ & & & & & & & \dots & & & & \\ X_N^w & Y_N^w & Z_N^w & 1 & 0 & 0 & 0 & 0 & -x_NX_N^w & -x_NY_N^w & -x_NZ_N^w & -x_N \\ 0 & 0 & 0 & 0 & X_N^w & Y_N^w & Z_N^w & 1 & -y_NX_N^w & -y_NY_N^w & -y_NZ_N^w & -y_N \end{bmatrix}$$

$$\mathbf{m} = [m_{11} \quad m_{12} \quad m_{13} \quad m_{14} \quad m_{21} \quad m_{22} \quad m_{23} \quad m_{24} \quad m_{31} \quad m_{32} \quad m_{33} \quad m_{34}]^T$$

Calibration via recovering the camera projection matrix

- Since the rank of A is 11 (why ?), the solution is the column of V corresponding to the zero singular value of A, with $A = UDV^T$. In practice, the solution is the eigenvector corresponding to the smallest eigenvalue of $A^T A$.

  We denote the solution is $\hat{M} = \begin{bmatrix} \mathbf{q}_1^T & \\ \mathbf{q}_2^T & \mathbf{q}_4 \\ \mathbf{q}_3^T & \end{bmatrix} = \gamma M$

- Since

$$M = K[R, \mathbf{T}] = \begin{bmatrix} f_x r_{11} + u_0 r_{31} & f_x r_{12} + u_0 r_{32} & f_x r_{13} + u_0 r_{33} & f_x T_x + u_0 T_z \\ f_y r_{21} + v_0 r_{31} & f_y r_{22} + v_0 r_{32} & f_y r_{23} + v_0 r_{33} & f_y T_y + v_0 T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

  Therefore $\gamma = \pm\sqrt{\hat{m}_{31}^2 + \hat{m}_{32}^2 + \hat{m}_{33}^2}$. Using γ to normalize $\hat{M}$. The result is still denoted as $\hat{M}$, then

$$r_{3i} = \hat{m}_{3i} \qquad u_0 = \mathbf{q}_1^T \mathbf{q}_3 \qquad f_x = \sqrt{\mathbf{q}_1^T \mathbf{q}_1 - u_0^2} \qquad r_{1i} = (\hat{m}_{1i} - u_0 \hat{m}_{3i}) / f_x \qquad T_x = (\hat{m}_{14} - u_0 T_z) / f_x$$

$$T_z = \hat{m}_{34} \qquad v_0 = \mathbf{q}_2^T \mathbf{q}_3 \qquad f_y = \sqrt{\mathbf{q}_2^T \mathbf{q}_2 - v_0^2} \qquad r_{2i} = (\hat{m}_{2i} - v_0 \hat{m}_{3i}) / f_y \qquad T_y = (\hat{m}_{24} - v_0 T_z) / f_y$$

- Make R orthornormal. If the world reference frame is in front of the camera, choose sign so that $T_z > 0$

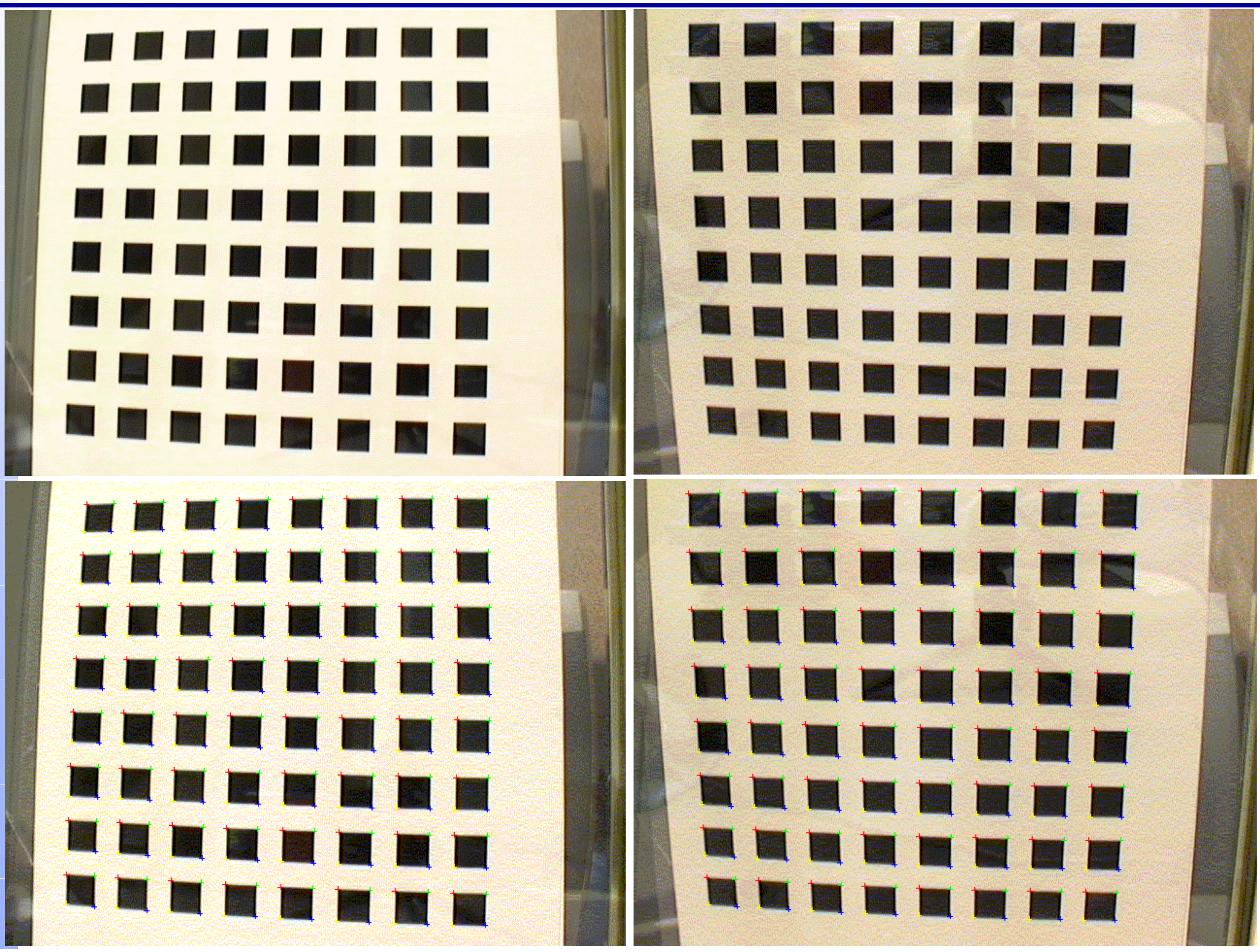
Camera calibration using planar patterns – Zhang 00

- Using known coordinates of many feature points on a planar pattern
- Capturing multiple images from different viewpoints (or equivalently, with the pattern at different positions and orientations)
- Automatically locating the feature points in the images
- Recovering homographies for each frames
- Recovering camera parameters from the homographies

- A full 5-parameter camera model is used

$$K = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Camera calibration using planar patterns – Zhang 00



Department of Computer Engineering
University of California at Santa Cruz

Calibration results

K=	832.5	0.204494	303.959
	0	832.53	206.585
	0	0	1

K1=-0.228601 k2=0.190353

R1= 0.992759 -0.026319 0.117201
0.0139247 0.994339 0.105341
-0.11931 -0.102947 0.987505

T1= -3.84019 3.65164 12.791

R2= 0.997397 -0.00482564 0.0719419
0.0175608 0.983971 -0.17746
-0.0699324 0.178262 0.981495

T2= -3.71693 3.76928 13.1974

R3= 0.915213 -0.0356648 0.401389
-0.00807547 0.994252 0.106756
-0.402889 -0.100946 0.909665

T3= -2.94409 3.77653 14.2456

…

Homogrpahy

- Without loss of generality, we assume the model plane is $Z_w = 0$, then

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong K[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{T}] \begin{bmatrix} X_w \\ Y_w \\ 0 \\ 1 \end{bmatrix} = K[\mathbf{r}_1, \mathbf{r}_2, \mathbf{T}] \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} = H \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}$$

- The 2D image coordinates and the corresponding 2D coordinates on the 3D image plane are related by a homography *H*. Homography is defined up to a scale factor

$$H = \lambda K[\mathbf{r}_1, \mathbf{r}_2, \mathbf{T}]$$

- Constraints on the intrinsic parameter

Suppose $H = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]$, then $[\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3] = \lambda K[\mathbf{r}_1, \mathbf{r}_2, \mathbf{T}]$. Since $\mathbf{r}_1$ and $\mathbf{r}_2$ are orthonormal, therefore

$$\mathbf{h}_1^T K^{-T} K^{-1} \mathbf{h}_2 = 0$$

$$\mathbf{h}_1^T K^{-T} K^{-1} \mathbf{h}_1 = \mathbf{h}_2^T K^{-T} K^{-1} \mathbf{h}_2$$

which gives two constraints for computing $B = K^{-T} K^{-1}$

Close-form solution

- We want to recover B first where

$$B = K^{-T}K^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{f_x^2} & -\frac{s}{f_x^2 f_y} & \frac{v_0 s - u_0 f_y}{f_x^2 f_y} \\ -\frac{s}{f_x^2 f_y} & \frac{s^2}{f_x^2 f_y^2} + \frac{1}{f_y^2} & \frac{s(v_0 s - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} \\ \frac{v_0 s - u_0 f_y}{f_x^2 f_y} & \frac{s(v_0 s - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} & \frac{(v_0 s - u_0 f_y)^2}{f_x^2 f_y^2} + \frac{v_0^2}{f_y^2} + 1 \end{bmatrix}$$

- If we denote

$$\mathbf{b} = [b_{11}, b_{12}, b_{22}, b_{13}, b_{23}, b_{33}] \qquad \mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$$

$$v_{ij} = [h_{i1}h_{j1},\ h_{i1}h_{j2} + h_{i2}h_{j1},\ h_{i2}h_{j2},\ h_{i3}h_{j1} + h_{i1}h_{j3},\ h_{i3}h_{j2} + h_{i2}h_{j3},\ h_{i3}h_{j3}]^T$$

then $\mathbf{h}_i^T B \mathbf{h}_j = v_{ij}^T \mathbf{b}$

To verify this, e.g. the coefficient of b_{12} is $h_{i1}h_{j2} + h_{i2}h_{j1}$

$$\mathbf{h}_i^T B \mathbf{h}_j = [h_{i1} \quad h_{i2} \quad h_{i3}] \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{bmatrix} \begin{bmatrix} h_{j1} \\ h_{j2} \\ h_{j3} \end{bmatrix}$$

Close-form solution

- The two constraints can be written as

$$\begin{bmatrix} v_{12}^T \\ v_{11}^T - v_{22}^T \end{bmatrix} \mathbf{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If *N* images of the model plane are observed, by stacking *N* such equations, we have

$$\boxed{V\mathbf{b} = 0}$$

If $N \geq 3$, we can solve *B*. If we have 2 images, 4 constraints are available, we can solve for 4 camera parameters. Usually, we will let $s = 0$. If we have only one image, we will further assume that u_0, v_0 are known and try to solve for f_x and f_y.

Department of Computer Engineering
University of California at Santa Cruz

Recovering camera parameters from B

- Once B is known, we can compute camera parameters from

$$B = \lambda \begin{bmatrix} \frac{1}{f_x^2} & -\frac{s}{f_x^2 f_y} & \frac{v_0 s - u_0 f_y}{f_x^2 f_y} \\ -\frac{s}{f_x^2 f_y} & \frac{s^2}{f_x^2 f_y^2} + \frac{1}{f_y^2} & \frac{s(v_0 s - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} \\ \frac{v_0 s - u_0 f_y}{f_x^2 f_y} & \frac{s(v_0 s - u_0 f_y)}{f_x^2 f_y^2} - \frac{v_0}{f_y^2} & \frac{(v_0 s - u_0 f_y)^2}{f_x^2 f_y^2} + \frac{v_0^2}{f_y^2} + 1 \end{bmatrix}$$

More specifically

$$v_0 = (b_{12} b_{13} - b_{11} b_{23}) / (b_{11} b_{22} - b_{12}^2)$$

$$\lambda = b_{33} - [b_{13}^2 + v_0 (b_{12} b_{13} - b_{11} b_{23})] / b_{11}$$

$$f_x = \sqrt{\lambda / b_{11}}$$

$$f_y = \sqrt{\lambda b_{11} / (b_{11} b_{22} - b_{12}^2)}$$

$$s = -b_{12} f_x^2 f_y / \lambda$$

$$u_0 = s v_0 / f_x - b_{13} f_x^2 / \lambda$$

Recover camera motion *R* and **T**

- Once camera matrix K is recovered, based on

$$H = K[\mathbf{r}_1, \mathbf{r}_2, \mathbf{T}]$$

we can recover the camera motion as follows

$$\mathbf{r}_1 = \lambda K^{-1}\mathbf{h}_1 \qquad \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$

$$\mathbf{r}_2 = \lambda K^{-1}\mathbf{h}_2 \qquad \mathbf{T} = \lambda K^{-1}\mathbf{h}_3$$

R matrix needs to be made orthonormal

- Gradient methods such as the Levenberg-Marquardt method is then used to refined the result by solving the following optimization problem

$$\min_{K, R_i, \mathbf{T}_i} \sum_{i=1}^{N} \sum_{j=1}^{m} \left\| m_{ij} - \hat{m}(K, R_i, \mathbf{T}_i, M_j) \right\|^2$$

where M_j is the 3D coordinates for the *j*th pattern point, m_{ij} is the projection of this point in the *i*th image

- For more details see

Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330-1334, 2000. Or http://www.research.microsoft.com/scripts/pubdb/pubsasp.asp?RecordID=212