

# Cameras in Homogeneous Coordinates<sup>1</sup>

## Lecture 15

See Sections 6.2.2 and 6.3.1 in  
Reinhard Klette: Concise Computer Vision  
Springer-Verlag, London, 2014

`ccv.wordpress.fos.auckland.ac.nz`

---

<sup>1</sup>See last slide for copyright information.

# Agenda

① Homogeneous Coordinates

② Camera Calibration

# Homogeneous Coordinates in the Plane

In homogeneous coordinates, subsequent matrix multiplication and vector addition reduce to just one matrix multiplication - just to state one benefit

## Homogeneous Coordinates in the Plane:

For coordinates  $x$  and  $y$ , add a third coordinate  $w$

$(x', y', w)$  are the *homogeneous coordinates* for a 2D point  $(x'/w, y'/w)$

$w \neq 0$ , then  $(x', y', w)$  represents  $(x'/w, y'/w)$  in 2D inhomogeneous coordinates; scale of  $w$  is unimportant; with  $w = 1$  use  $x = x'$  and  $y = y'$

$w = 0$ , then  $(x, y, 0)$  defines a *point at infinity*

## Lines in the Plane

Straight line  $\gamma = (a, b, c)$  in homogeneous coordinates in the plane:

$$a \cdot x + b \cdot y + 1 \cdot c = [a, b, c]^T \cdot [x, y, 1] = 0$$

Two straight lines  $\gamma_1 = (a_1, b_1, c_1)$  and  $\gamma_2 = (a_2, b_2, c_2)$  intersect at

$$\gamma_1 \times \gamma_2 = (b_1 c_2 - b_2 c_1, a_2 c_1 - a_1 c_2, a_1 b_2 - a_2 b_1)$$

This point is the *cross product* of two vectors

Two lines  $\gamma_1$  and  $\gamma_2$  parallel iff  $a_1 b_2 = a_2 b_1$

Thus: Parallel lines intersect at a point at infinity

Calculus using homogeneous coordinates applies uniformly for existing points as well as for points at infinity

## Two Points

Consider two different points  $p_1 = (x_1, y_1, w_1)$  and  $p_2 = (x_2, y_2, w_2)$

They define (i.e. are incident with) the line  $p_1 \times p_2$

**Example:** One point at infinity, say  $p_1 = (x_1, y_1, 0)$

Then:  $p_1 \times p_2 = (y_1 w_2, x_1 w_2, x_1 y_2 - x_2 y_1)$  is an existing straight line

Point  $p_1 = (x_1, y_1, 0)$  is at infinity in direction  $[x_1, y_1]^T$

**Example:** Both points at infinity, i.e.  $w_1 = w_2 = 0$

Then: straight line  $p_1 \times p_2 = (0, 0, x_1 y_2 - x_2 y_1)$  is at infinity

Note:  $x_1 y_2 \neq x_2 y_1$  for  $p_1 \neq p_2$

# Translation

Point  $p = (x, y)$  and translation  $\mathbf{t} = [t_1, t_2]^T$   
in inhomogeneous 2D coordinates

Multiplication in homogeneous coordinates

“translation times point”

$$\begin{bmatrix} 1 & 0 & t_1 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot [x, y, 1]^T = [x + t_1, y + t_2, 1]^T$$

Result: defines point  $(x + t_1, y + t_2)$  in inhomogeneous coordinates

This way we can also translate a point at infinity

# Observation

Homogeneous coordinates allow us to perform uniquely defined calculations in the plane also covering the cases that we were not able to express before in our calculus when using only inhomogeneous  $xy$  coordinates

# Homogeneous Coordinates in 3D Space

Point  $(X, Y, Z) \in \mathbb{R}^3$  represented by  $(X', Y', Z', w)$  in homogeneous coordinates, with  $(X, Y, Z) = (X'/w, Y'/w, Z'/w)$

Affine transforms can be represented by  $4 \times 4$  matrix multiplications

## Translation Example

Point  $P = (X, Y, Z)$  and translation  $\mathbf{t} = [t_1, t_2, t_3]^\top$   
in inhomogeneous 3D coordinates

Multiplication

$$\begin{bmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot [X, Y, Z, 1]^\top = [X + t_1, Y + t_2, Z + t_3, 1]^\top$$

results in point  $(X + t_1, Y + t_2, Z + t_3)$  in inhomogeneous coordinates



# Rotation and Translation Example

Consider an affine transform defined by rotation and translation

$4 \times 4$  matrix multiplication

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot [X, Y, Z, 1]^T = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \cdot [X, Y, Z, 1]^T \\ = [X_s, Y_s, Z_s, 1]^T$$

results in point  $(X_s, Y_s, Z_s)$  in inhomogeneous coordinates

This introduces a notation of  $4 \times 4$  matrices by means of a

$3 \times 3$  submatrix  $\mathbf{R}$ , a column 3-vector  $\mathbf{t}$ , and a row 3-vector  $\mathbf{0}^T$

# Agenda

① Homogeneous Coordinates

② Camera Calibration

# Intrinsic and Extrinsic Parameters

Camera calibration specifies *intrinsic* (i.e. camera-specific) and *extrinsic* parameters of a given one- or multi-camera configuration

## **Intrinsic or internal parameters:**

- 1 the (effective) focal length,
- 2 dimensions of the sensor matrix,
- 3 sensor cell size or aspect ratio of sensor height to width,
- 4 radial distortion parameters,
- 5 coordinates of the principal point, or
- 6 the scaling factor

**Extrinsic parameters:** specification of affine transforms for identifying *poses* (i.e. location and direction) of cameras in a world coordinate system

Here:

Overview on calibration for understanding what is happening *in principle*

# A User's Perspective on Camera Calibration

Use **geometric patterns** on 2D or 3D surfaces that we are able to measure very accurately (e.g. calibration rig)

Geometric patterns are **recorded, localized** in the resulting images, and their appearance in the image grid is compared with measurements about their geometry in the real world using a formal **camera model**

---

Calibration possibly for **one camera only** (e.g. of a multi-camera system) assuming that cameras are static or that we only calibrate internal parameters

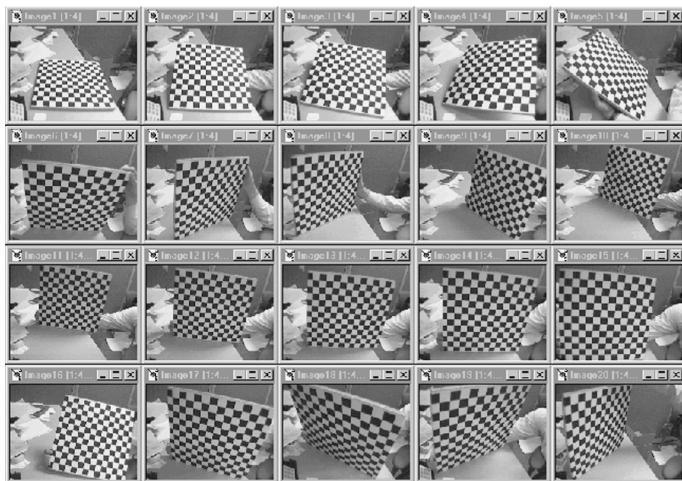
---

Typically: a movable **multi-camera system** and a multi-camera approach for calibration, aiming at calibrating internal and external parameters

All cameras need to be exactly **time-synchronized**, especially if the calibration rig moves during the procedure

# Calibration Software

C sources by J.-Y. Bouquet: calibration rig recorded under various poses



## Designing a Calibration Method

Define parameters to be calibrated and a corresponding camera model

**Example:** Radial distortion parameters  $\kappa_1$  and  $\kappa_2$  and model includes

$$\begin{aligned}x_u &= c_x + (x_d - c_x)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + e_x) \\y_u &= c_y + (y_d - c_y)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + e_y)\end{aligned}$$

Radial distortion parameters already known, then model possibly defined by

$$\begin{aligned}\begin{bmatrix} x - c_x \\ y - c_y \\ f \end{bmatrix} &= \begin{bmatrix} x_u \\ y_u \\ f \end{bmatrix} = f \begin{bmatrix} X_s/Z_s \\ Y_s/Z_s \\ 1 \end{bmatrix} \\ &= f \begin{bmatrix} \frac{r_{11}(X_w+t_1)+r_{12}(Y_w+t_2)+r_{13}(Z_w+t_3)}{r_{31}(X_w+t_1)+r_{32}(Y_w+t_2)+r_{33}(Z_w+t_3)} \\ \frac{r_{21}(X_w+t_1)+r_{22}(Y_w+t_2)+r_{23}(Z_w+t_3)}{r_{31}(X_w+t_1)+r_{32}(Y_w+t_2)+r_{33}(Z_w+t_3)} \\ 1 \end{bmatrix}\end{aligned}$$

## Measured Points into Subpixel Accurate Image Points

Points  $(X_w, Y_w, Z_w)$  on calibration rig or *calibration marks* are known by their physically measured world coordinates

For  $(X_w, Y_w, Z_w)$ , identify corresponding point  $(x, y)$  that is the projection of  $(X_w, Y_w, Z_w)$  in the image plane

---

**Example:** 100 different points  $(X_w, Y_w, Z_w)$  define 100 equations in the form as given on Page 14, where only  $c_x, c_y, f, r_{11}$  to  $r_{33}, t_1, t_2,$  and  $t_3$  appear as unknowns

We have an overdetermined equational system and need to find a “clever” optimization procedure for solving it for those few unknowns

# Refined Model

## Examples

- (1) focal length  $f_x$  in  $x$ -direction and  $f_y$  in  $y$ -direction
  - (2) edge length  $e_x$  and  $e_y$  of sensor cells in sensor matrix of camera
  - (3) transition from camera coordinates in world units to homogeneous camera coordinates in pixel units
  - (4) shearing factor  $s$  for evaluating the orthogonality of recorded images
- 

Accordingly

resulting equational systems becomes more complex with more unknowns



# General Summary of a Calibration Procedure

- 1 Known positions ( $X_w, Y_w, Z_w$ ) in the world are related to identifiable locations ( $x, y$ ) in recorded images
- 2 Equations defining our camera model then contain  $X_w, Y_w, Z_w, x,$  and  $y$  as known values and intrinsic or extrinsic camera parameters as unknowns
- 3 Resulting equational system (necessarily nonlinear due to central projection or radial distortion) needs to be solved for the specified unknowns
- 4 Over-determined situations provide for stability of a used numeric solution scheme

# Manufacturing a Calibration Board

A rigid board wearing a black and white checkerboard pattern is common  
It should have  $7 \times 7$  squares at least

The squares need to be large enough such that their minimum size, when recorded on the image plane during calibration, is  $10 \times 10$  pixels at least

**Example:** Camera with effective focal length  $f$  allows us to estimate the visible size of  $a \times a$  cm for each square assuming a distance of  $b$  m between camera and board

## Localizing Corners in the Checkerboard

For the checkerboard, *calibration marks* are the corners of the squares, and those can be identified by approximating intersection points of grid lines, thus defining the corners of the squares potentially with subpixel accuracy

**Example:** Assume 10 vertical and 10 horizontal grid lines on a checkerboard

This should result in  $10 + 10$  peaks in the  $d\alpha$  Hough space for detecting line segments

Each peak defines a detected grid line

Intersection points of those define the corners of the checkerboard in the recorded image

Method requires that lens distortion has been removed from the recorded images prior to applying the Hough-space method

## Localizing Calibration Marks

A calibration pattern can also be defined by marks such as circular or square dots

---

**Example:** Cameras calibrated in the same location, with calibration marks permanently painted on walls or other static surfaces

---

We identify an image region  $S$  of pixels as the area that shows a calibration mark, say, in grey levels

Position of the calibration mark can then be identified at subpixel accuracy by calculating the centroid of this region

# Copyright Information

This slide show was prepared by Reinhard Klette with kind permission from Springer Science+Business Media B.V.

The slide show can be used freely for presentations. However, *all the material* is copyrighted.

R. Klette. Concise Computer Vision.  
©Springer-Verlag, London, 2014.

In case of citation: just cite the book, that's fine.