Real Cameras and their Calibration

COMPSCI 527 — Computer Vision
Outline

1 Real Cameras
   Depth of Field
   Distortion

2 Camera Calibration
   A Camera Model
   Parameter Optimization
   Lab Setup and Imaging
Depth of Field

- Focal length: focal distance when an object at $\infty$ is in focus
- Focal length is a lens property
- Focal distance can be changed by rotating the focusing ring
- Nothing to do with zoom, which changes focal length
- Alas, $f$ is often used for either focal distance or focal length
Changing Depth of Field

- **Aperture**: diameter of the hole in front of the lens
- Measured in *stops*, or *f-numbers* \( n = \frac{f}{a} \)
  
  \( (a \) is aperture diameter, \( f \) is focal length)  
- Area (light flux) is proportional to *square* of diameter  
- Small aperture (big *f*-number) \( \Rightarrow \) great depth of field  
- A shallow depth of field is sometimes desirable
Depth of Field and Exposure

http://www.boostyourphotography.com/2014/10/depth-of-field.html
Lens Distortion

pincushion

barrel
not distortion:

Real Cameras

Distortion

85mm @ 200cm
35mm @ 85cm
16mm @ 40cm
12mm @ 30cm
8mm @ 20cm
Camera Calibration

- Cameras have *intrinsic parameters*: focal distance, pixel size, principal point, lens distortion parameters
- ... and *extrinsic parameters*: Rotation, translation relative to some world reference system
- *Camera calibration* is a combination of lab measurements and algorithms aimed at determining both types of parameters
- We do not calibrate for finite depth of field (stop down the aperture, flood the scene with light)
Calibration as Learning

- Many variants, the general idea is the same
- Looks very much like machine learning:
  1. Make a parametric model of what a camera does: Inputs are world points \( W \) in world coordinates, outputs are image points \( \xi \) in image pixel coordinates (“predictor architecture”)
  2. Collect a sufficiently large set \( T \) of input-output pairs \((W_n, \xi_n)\) (“training set”)
  3. “Loss function” measures discrepancy between \( \hat{\xi}_n \) predicted by model and \( \xi_n \) measured in image
  4. Fit the parameters to \( T \) by numerical optimization (“training”)
- We even have generalization requirements: The parameters should be correct for pairs \((W, \xi)\) not in \( T \)
- We already know how to do 3, 4. Need to figure out 1, 2.
Camera Model

\[ \mathbf{X} = R (\mathbf{W} - \mathbf{t}) \]
\[ \mathbf{x} = \rho(\mathbf{X}) = \frac{1}{X_3} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \]
\[ \mathbf{y} = d(\mathbf{x}) \quad \text{(lens distortion)} \]
\[ \xi = S\mathbf{y} + \pi \]
\[ S = f \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \]

Can only determine products \( fs_x \) and \( fs_y \), not \( f, s_x, s_y \) individually.
Lens Distortion Model

- Distortion is radial around the principal point:
  \[ y = d(x) = \delta(r) x \quad \text{where} \quad r = ||x|| \]
- Radial distortion function \( \delta(\cdot) \) is nonlinear
- Must be analytical everywhere (Maxwell). Implication:
  - Restrict to \( x \) axis: \( \delta(r(x)) = \delta(|x|) \)
  - Odd powers of \( |x| \) have a cusp at the origin
  - Therefore, \( \delta(r) = 1 + k_1 r^2 + k_2 r^4 + \ldots \quad (k_0 \neq 1 \text{ can be folded into } f) \)
- Large powers of \( r \) only affect peripheral areas and cannot be determined well
- Typically, \( \delta(r) = 1 + k_1 r^2 + k_2 r^4 \)
Camera Parameters

\[
X = R(W - t)
\]

\[
x = p(X) = \frac{1}{X_3} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}
\]

\[
y = x(1 + k_1 \|x\|^2 + k_2 \|x\|^4)
\]

\[
\xi = Sy + \pi
\]

- Extrinsic parameters: \( R, t \) (6 degrees of freedom)
- Intrinsic parameters: \( \pi, f_{sx}, f_{sy}, k_1, k_2 \) (6 numbers)

\[\xi = c(W; p) \text{ where } p \in \mathbb{R}^{12}\]
Data Fitting

- Collect input-output pairs \((W_n, \xi_n)\) for \(n = 1, \ldots, N\)
- \(\hat{\xi} = c(W; p)\) where \(p \in \mathbb{R}^{12}\)
- \(p^* = \arg \min_p e(p)\) where \(e(p) = \frac{1}{N} \sum_{n=1}^{N} \|\xi_n - c(W_n; p)\|^2\)
- \(e\) is nonlinear
- To initialize: clamp \(k_1 = k_2 = 0\), solve a linear system
- Approximate because of clamping and because the residual is different from \(e(p)\)
- Use any optimization algorithm to refine
Calibration Target

http://www.mdpi.com/1424-8220/9/6/4572/htm

Duke Computer Vision Lab
Circles are Problematic
Calibration Protocol Summary

- Place calibration target in front of camera (fill the image)
- Measure image coordinates (with software help?)
- Make a file with \((W_n, \xi_n)\) pairs
- Fit parameters by numerical optimization
- Redo if you touch the lens or the camera!
An Example for Distortion Only