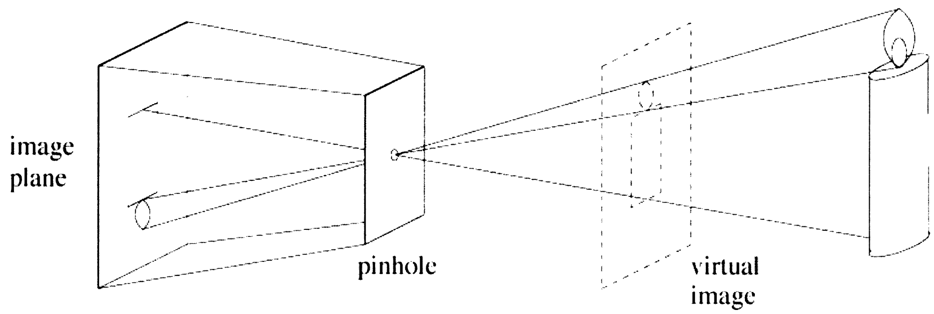
Cameras

- First photograph due to Niepce
- First on record shown in the book - 1822
- Basic abstraction is the pinhole camera
  - lenses required to ensure image is not too dark
  - various other abstractions can be applied

Pinhole cameras

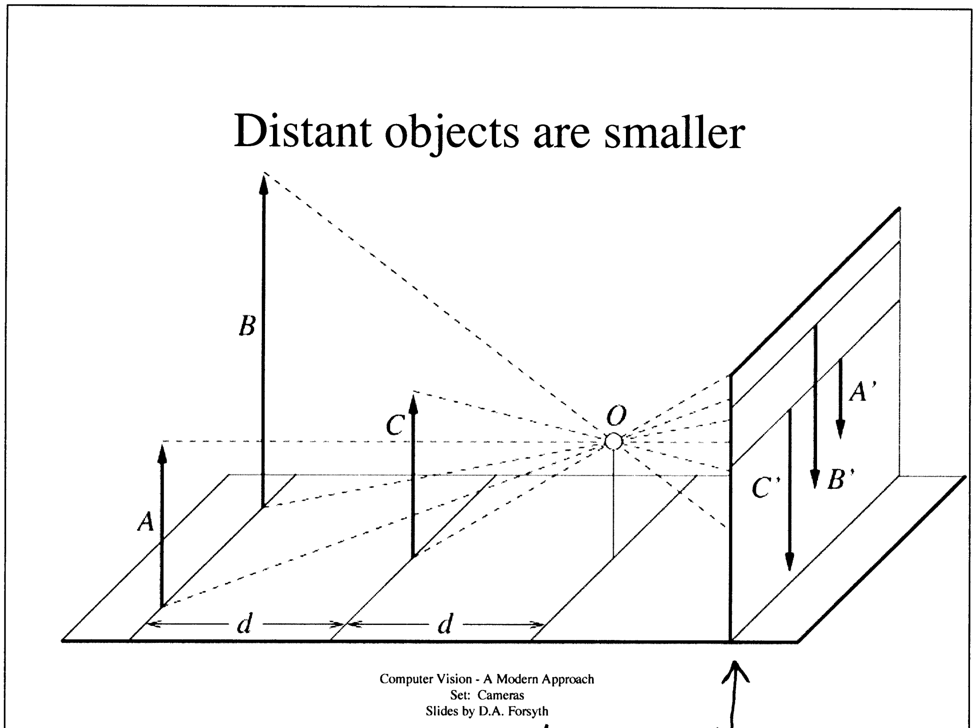
- Abstract camera model - box with a small hole in it
- Pinhole cameras work in practice



Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

The point to make here is that each point on the image plane sees light from only

one direction, the one that passes through the pinhole.

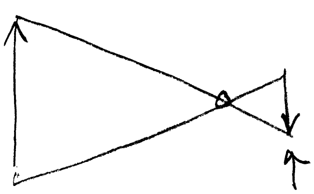


≈ f

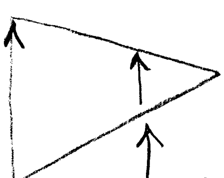
common to put virtual image here

image plane

NOTE: Due to camera scaling B' and C' are shown as the same height although they are different.



real image



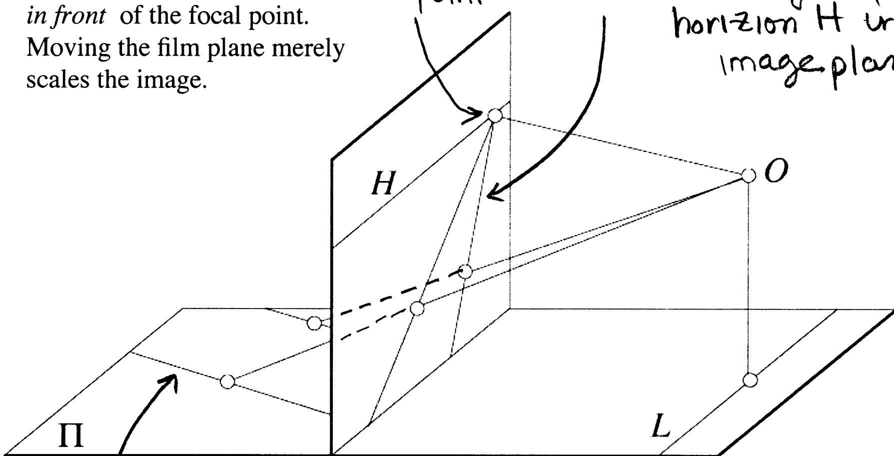
virtual image

Parallel lines meet

Common to draw film plane *in front* of the focal point. Moving the film plane merely scales the image.

Vanishing point

lines converge to a point at horizion H in virtual image plane



Slides by D.A. Forsyth

lines parallel in real world

virtual image plane

focal plane

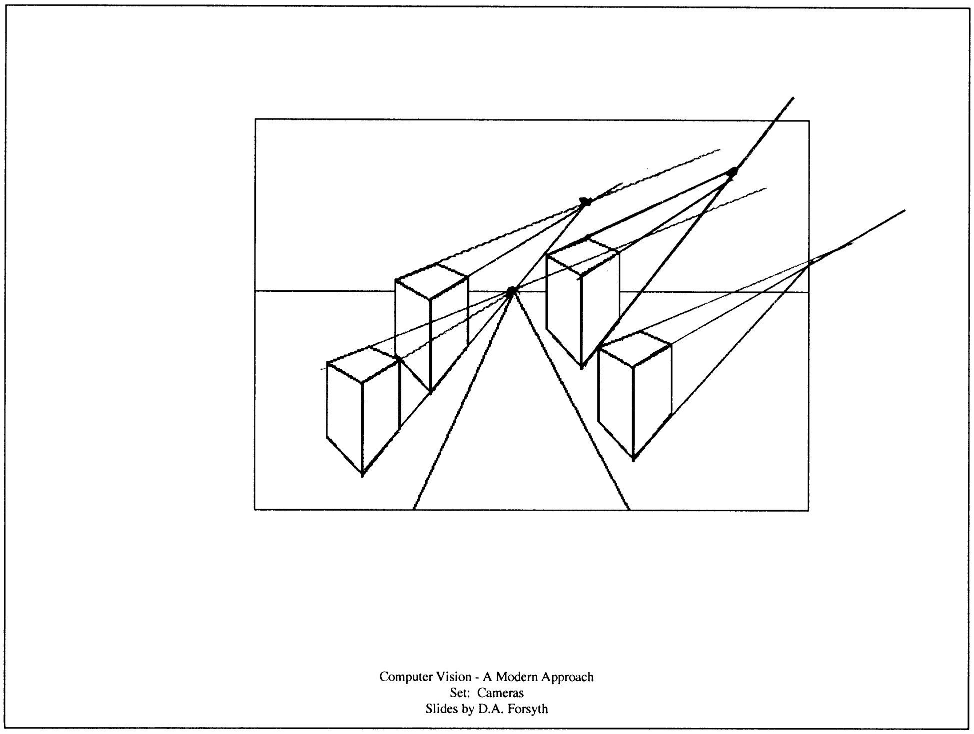
Vanishing points

- each set of parallel lines (=direction) meets at a different point
  - The *vanishing point* for this direction
- Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  - The line is called the *horizon* for that plane

- Good ways to spot faked images
  - scale and perspective don't work
  - vanishing points behave badly
  - supermarket tabloids are a great source.

Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

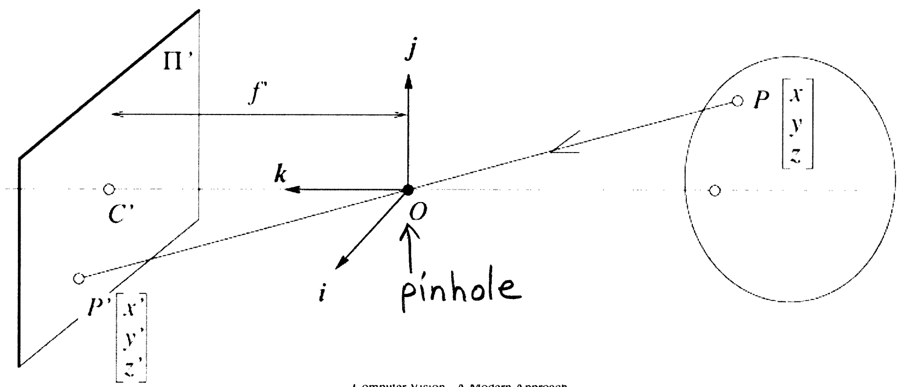
all vanishing points are on same horizion



Is this a perspective image of four identical buildings? why not?

perspective is bad. All vanishing points should be on horizon
same.

The equation of projection



Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

since $\overrightarrow{OP}$ and $\overrightarrow{OP'}$ are collinear.

$$x' = \lambda x$$
$$y' = \lambda y \quad \Rightarrow \quad \lambda = \frac{x'}{x} = \frac{y'}{y} = \frac{f'}{z} \leftarrow \text{in place of } z'$$
$$f' = \lambda z$$

therefore $x' = f' \dfrac{x}{z}$

$$y' = f' \frac{y}{z}$$

The equation of projection

- Cartesian coordinates:
  - We have, by similar triangles, that (x, y, z) -> (f x/z, f y/z, -f)
  - Ignore the third coordinate, and get

$$(x, y, z) \rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

Homogenous coordinates

- Add an extra coordinate and use an equivalence relation
- for 2D
  - equivalence relation k*(X,Y,Z) is the same as (X,Y,Z)
- for 3D
  - equivalence relation k*(X,Y,Z,T) is the same as (X,Y,Z,T)

- Basic notion
  - Possible to represent points "at infinity"
    - Where parallel lines intersect
    - Where parallel planes intersect
  - Possible to write the action of a perspective camera as a matrix

The camera matrix

- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

2-D image

3-D world coordinates

Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

$U = X$

$V = Y$

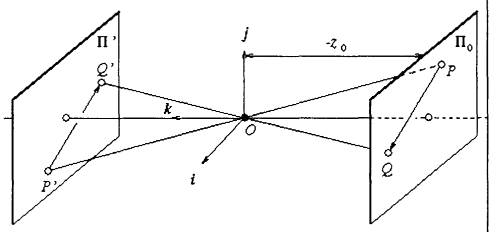
$W = \frac{Z}{f}$ ← scale factor

scaling $U = \frac{X}{\frac{Z}{f}} = f\frac{X}{Z}$

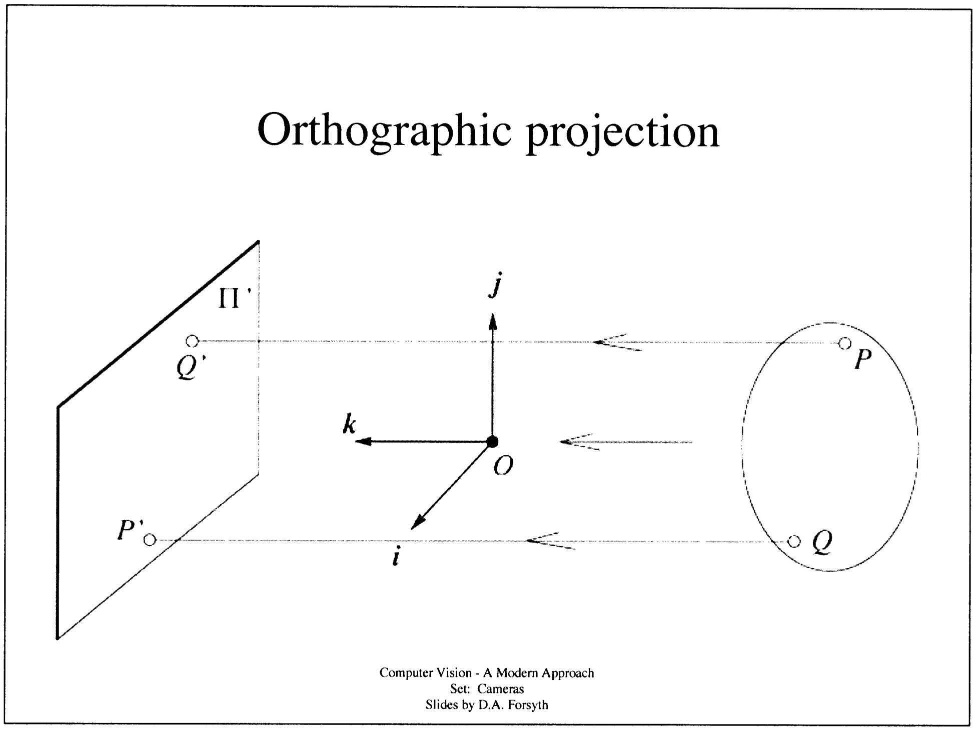
$V = \frac{Y}{\frac{Z}{f}} = f\frac{Y}{Z}$

Weak perspective

- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group
  - Adv: easy
  - Disadv: wrong



If depth of field of objects in a group is similar, assume magnification for this group is constant.



Extreme case of weak perspective
From previous drawing we normalize image coordinates to have magnification $m = -1$ which looks wrong but is correct.

The projection matrix for orthographic projection

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

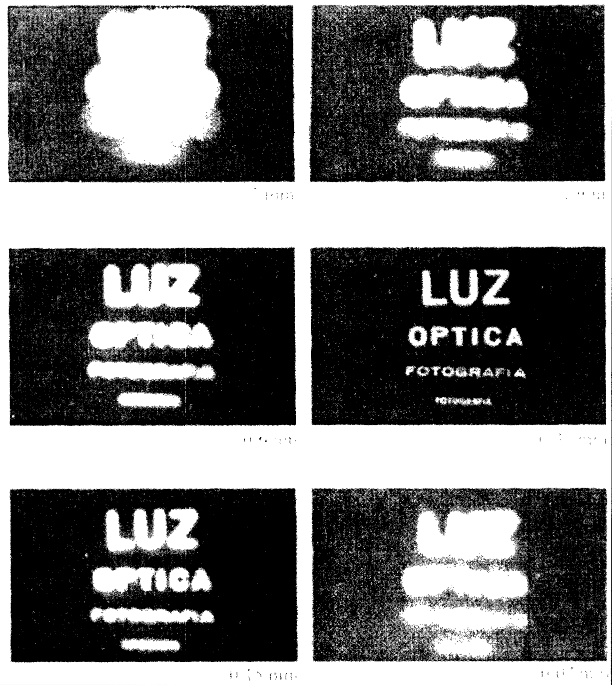
Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

This is a simplification of previous projection matrix in the limit that $f \to \infty$

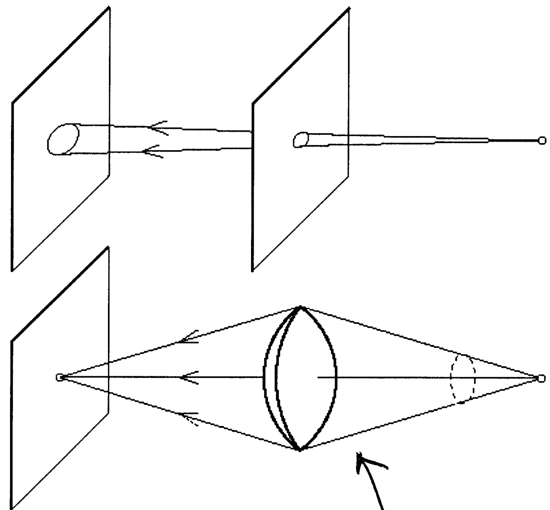
Pinhole too big -
many directions are averaged, blurring the image

Pinhole too small-
diffraction effects blur the image

Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.



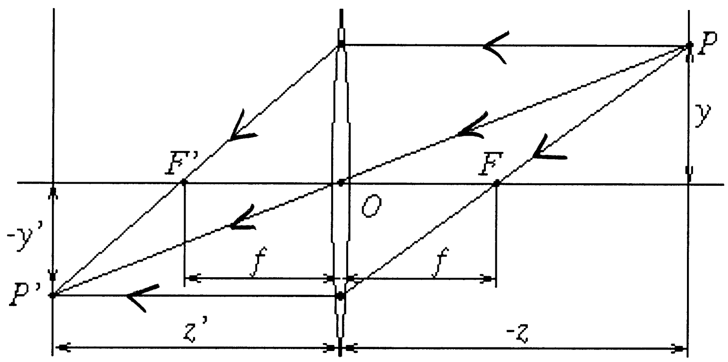
The reason for lenses



Slides by D.A. Forsyth

gather many more rays of light

The thin lens



$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$ ← lens law

image distance   object distance

Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

geometric optics for a thin lens

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2$$

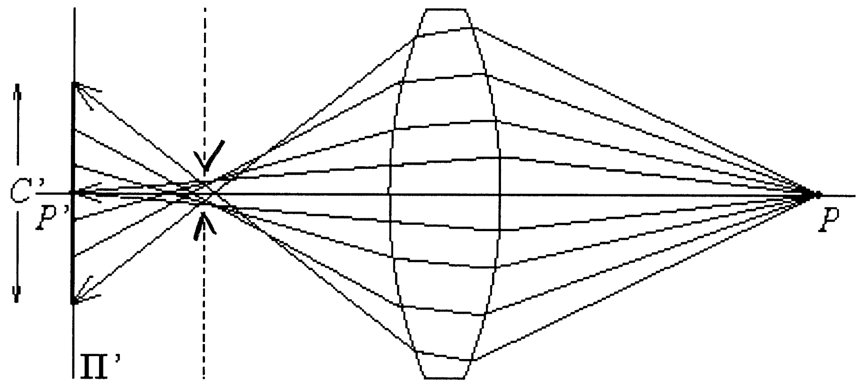
Assumes y and $-y'$ are close to z-axis

ray PP' is not refracted $\alpha_1 = \alpha_2$
but other rays are refracted.

f is the point at which entering rays parallel to the z-axis are focussed, i.e., cross the z-axis

Spherical aberration

circle of least confusion

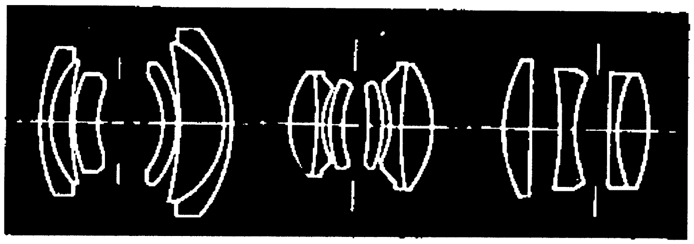


Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

due to assuming $\sin\alpha_1 \approx \alpha_1$
$\sin\alpha_2 \approx \alpha_2$
for paraxial rays

focal point not at same z-coordinate

Lens systems



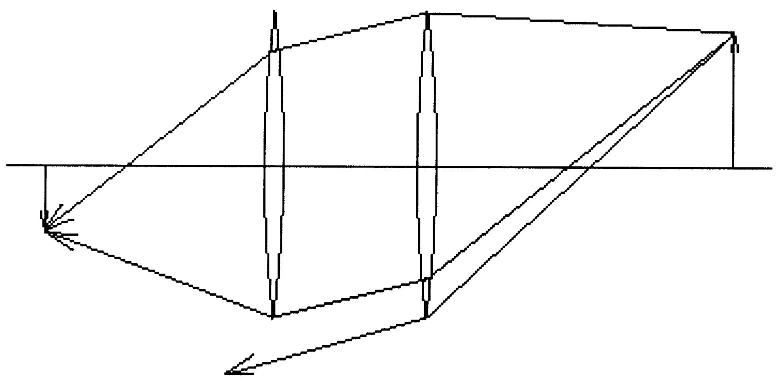
Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

Complicated len systems to minimize aberrations, etc.

Common for modern lenses (except front and rear) to be plastic

- light weight
- aspheric

Vignetting



Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

light follows a path which prevents it from reaching image plane.
Can be accidental or deliberate

Other (possibly annoying) phenomena

- Chromatic aberration
  - Light at different wavelengths follows different paths; hence, some wavelengths are defocussed
  - Machines: coat the lens
  - Humans: live with it
- Scattering at the lens surface
  - Some light entering the lens system is reflected off each surface it encounters (Fresnel's law gives details)
  - Machines: coat the lens, interior
  - Humans: live with it (various scattering phenomena are visible in the human eye)
- Geometric phenomena (Barrel distortion, etc.)

Camera parameters

- Issue
  - camera may not be at the origin, looking down the z-axis
    - extrinsic parameters
  - one unit in camera coordinates may not be the same as one unit in world coordinates
    - intrinsic parameters - focal length, principal point, aspect ratio, angle between axes, etc.

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

location of camera origin relative to world coordinates

focal length
principal points (thick lens)
etc.

Camera calibration

- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object
  - identify image points
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix
- Error minimization:
  - Linear least squares
    - easy problem numerically
    - solution can be rather bad
  - Minimize image distance
    - more difficult numerical problem
    - solution usually rather good,
    - start with linear least squares
  - Numerical scaling is an issue

* due to Tsai

Computer Vision - A Modern Approach
Set: Cameras
Slides by D.A. Forsyth

* assume radial alignment constraint

i.e. spherical distortion changes radial <u>scale</u> but not direction

Allows LSE to find all parameters <u>but</u>

- t_z z translation of camera
- magnification
- distortion parameters

$$\lambda = 1 + \underline{\kappa_1 d^2} + K_2 d^4 + K_3 d^6$$

usually this is all we use

d = distance between image point and image center.