Image formation

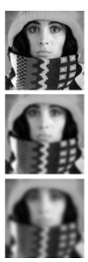
Monday March 21

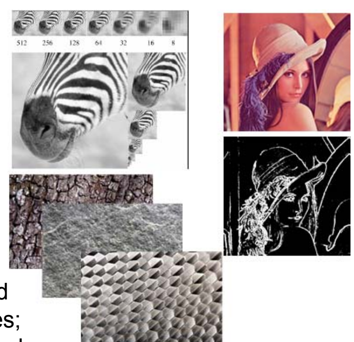
Prof. Kristen Grauman
UT-Austin

Announcements

- Reminder: Pset 3 due March 30
- Midterms: pick up today

Recap: Features and filters

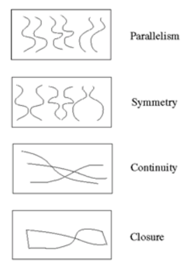




Transforming and describing images; textures, colors, edges

Kristen Grauman

Recap: Grouping & fitting

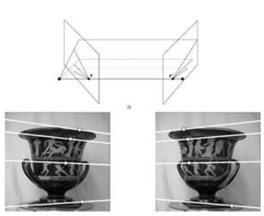


[fig from Shi et al]

Clustering, segmentation, fitting; what parts belong together?

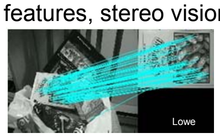
Kristen Grauman

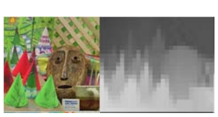
Multiple views



Hartley and Zisserman

Multi-view geometry, matching, invariant features, stereo vision





Kristen Grauman

Plan

- Today:
  - Local feature matching btwn views (wrap-up)
  - Image formation, geometry of a single view
- Wednesday: Multiple views and epipolar geometry
- Monday: Approaches for stereo correspondence

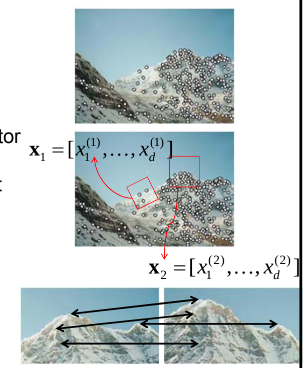
Previously

Local invariant features

- Detection of interest points
  - Harris corner detection
  - Scale invariant blob detection: LoG
- Description of local patches
  - SIFT : Histograms of oriented gradients
- Matching descriptors

Local features: main components

1) Detection: Identify the interest points
2) Description: Extract vector feature descriptor surrounding each interest point.
3) Matching: Determine correspondence between descriptors in two views

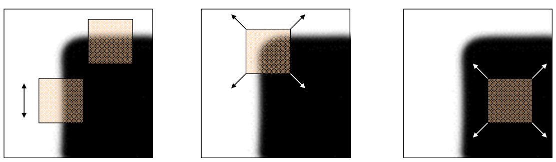


$\mathbf{x}_1 = [x_1^{(1)}, \ldots, x_d^{(1)}]$

$\mathbf{x}_2 = [x_1^{(2)}, \ldots, x_d^{(2)}]$

Kristen Grauman

Recall: Corners as distinctive interest points



"edge": $\lambda_1 >> \lambda_2$, $\lambda_2 >> \lambda_1$

"corner": λ_1 and λ_2 are large, $\lambda_1 \sim \lambda_2$;

"flat" region: λ_1 and λ_2 are small;

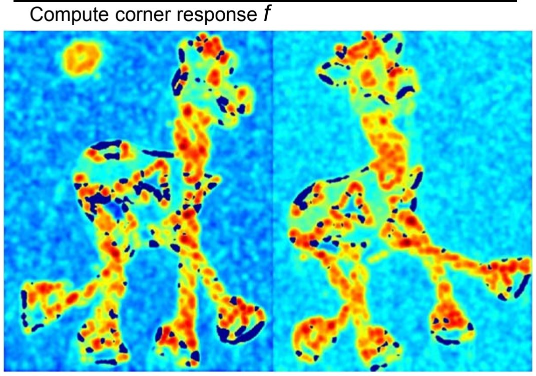
One way to score the cornerness:

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

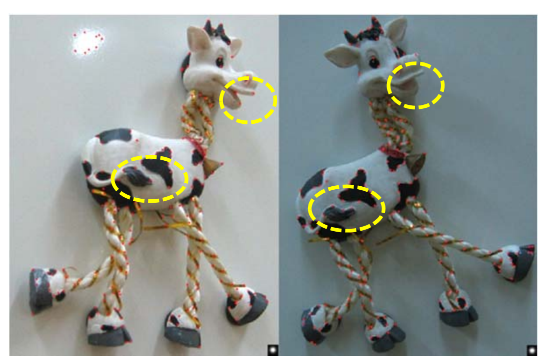
Harris Detector: Steps

Harris Detector: Steps

Compute corner response *f*

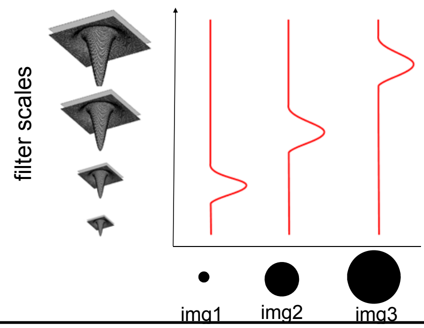


Harris Detector: Steps



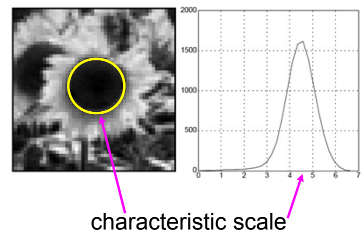
Blob detection in 2D: scale selection

Laplacian-of-Gaussian = "blob" detector $\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$



Blob detection in 2D

We define the *characteristic scale* as the scale that produces peak of Laplacian response



Slide credit: Lana Lazebni

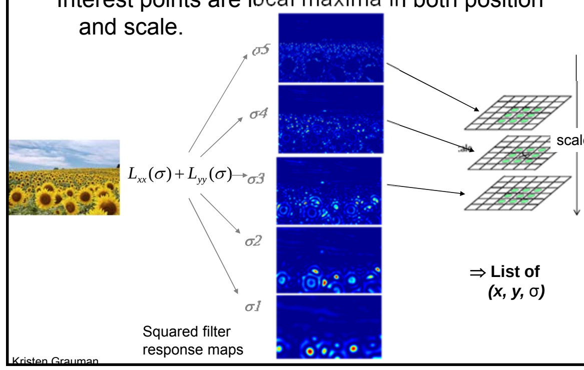
Example

Original image at ¾ the size

Kristen Grauman

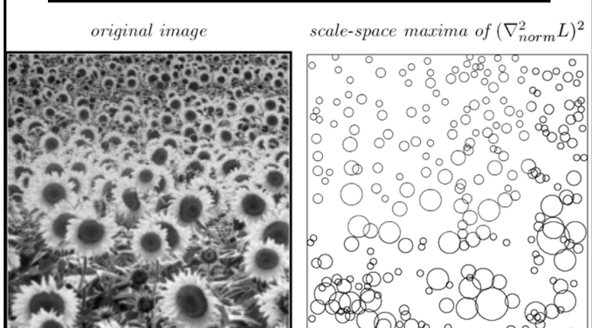
Scale invariant interest points

Interest points are local maxima in both position and scale.



Kristen Grauman

Scale-space blob detector: Example

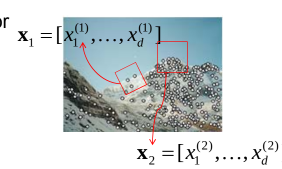


T. Lindeberg. Feature detection with automatic scale selection. IJCV 1998.

Local features: main components

1) Detection: Identify the interest points

2) Description: Extract vector feature descriptor surrounding each interest point.

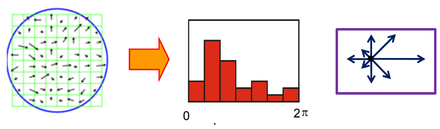


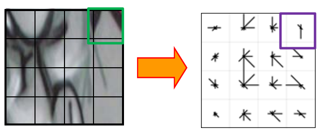
3) Matching: Determine correspondence between descriptors in two views

Kristen Grauman

SIFT descriptor [Lowe 2004]

- Use histograms to bin pixels within sub-patches according to their orientation.



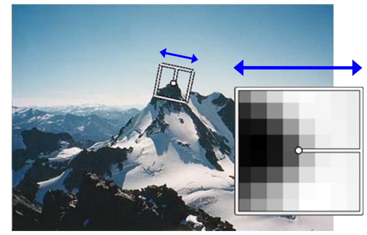


Why subpatches?

Why does SIFT have some illumination invariance?

Kristen Grauman

Making descriptor rotation invariant

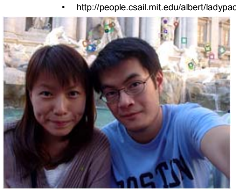


- Rotate patch according to its dominant gradient orientation
- This puts the patches into a canonical orientation.

Image from Matthew Brown

SIFT descriptor [Lowe 2004]

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint
    - Up to about 60 degree out of plane rotation
  - Can handle significant changes in illumination
    - Sometimes even day vs. night (below)
  - Fast and efficient—can run in real time
  - Lots of code available
    - http://people.csail.mit.edu/albert/ladypack/wiki/index.php/Known_implementations_of_SIFT



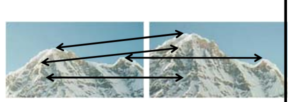
Steve Seitz

SIFT properties

- Invariant to
  - Scale
  - Rotation
- Partially invariant to
  - Illumination changes
  - Camera viewpoint
  - Occlusion, clutter

Local features: main components

1) Detection: Identify the interest points
2) Description: Extract vector feature descriptor surrounding each interest point.
3) Matching: Determine correspondence between descriptors in two views



Kristen Grauman

Matching local features

Kristen Grauman

Matching local features

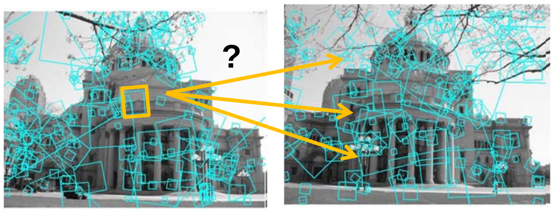


Image 1 Image 2

To generate **candidate matches**, find patches that have the most similar appearance (e.g., lowest SSD)

Simplest approach: compare them all, take the closest (or closest k, or within a thresholded distance)

Kristen Grauman

Ambiguous matches

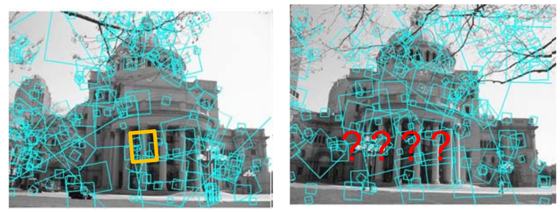


Image 1 Image 2

At what SSD value do we have a good match?

To add robustness to matching, can consider **ratio** : distance to best match / distance to second best match

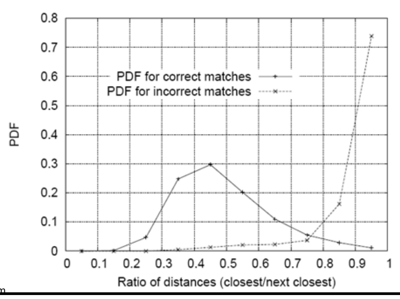
If low, first match looks good.

If high, could be ambiguous match.

Kristen Grauman

Matching SIFT Descriptors

- Nearest neighbor (Euclidean distance)
- Threshold ratio of nearest to 2nd nearest descriptor



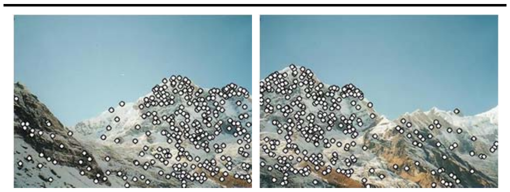
Lowe IJCV 2004

Derek Hoiem

Recap: robust feature-based alignment

Source: L. Lazebnik

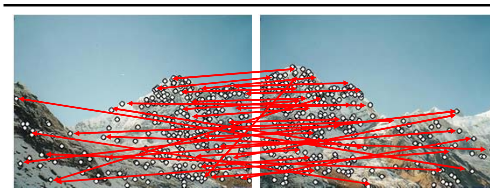
Recap: robust feature-based alignment



- Extract features

Source: L. Lazebnik

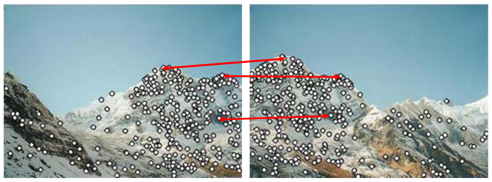
Recap: robust feature-based alignment



- Extract features
- Compute *putative matches*

Source: L. Lazebnik

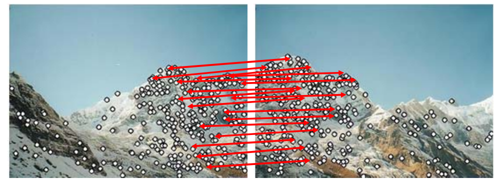
Recap: robust feature-based alignment



- Extract features
- Compute *putative matches*
- Loop:
  - *Hypothesize* transformation *T* (small group of putative matches that are related by *T*)

Source: L. Lazebnik

Recap: robust feature-based alignment



- Extract features
- Compute *putative matches*
- Loop:
  - *Hypothesize* transformation *T* (small group of putative matches that are related by *T*)
  - *Verify* transformation (search for other matches consistent with *T*)

Source: L. Lazebnik

Recap: robust feature-based alignment

- Extract features
- Compute *putative matches*
- Loop:
  - *Hypothesize* transformation *T* (small group of putative matches that are related by *T*)
  - *Verify* transformation (search for other matches consistent with *T*)

Source: L. Lazebnik

Applications of local invariant features

- Wide baseline stereo
- Motion tracking
- Panoramas
- Mobile robot navigation
- 3D reconstruction
- Recognition
- …

Automatic mosaicing

http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html

Wide baseline stereo

[Image from T. Tuytelaars ECCV 2006 tutorial]

Recognition of specific objects, scenes

Schmid and Mohr 1997

Sivic and Zisserman, 2003

Rothganger et al. 2003

Lowe 2002

Kristen Grauman

Plan

- Today:
  - Local feature matching btwn views (wrap-up)
  - Image formation, geometry of a single view
- Wednesday: Multiple views and epipolar geometry
- Monday: Approaches for stereo correspondence

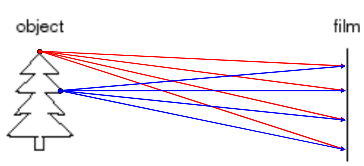
Image formation

- How are objects in the world captured in an image?

Physical parameters of image formation

- Geometric
  - Type of projection
  - Camera pose
- Optical
  - Sensor's lens type
  - focal length, field of view, aperture
- Photometric
  - Type, direction, intensity of light reaching sensor
  - Surfaces' reflectance properties

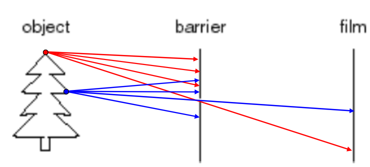
Image formation



- Let's design a camera
  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Slide by Steve Seitz

Pinhole camera

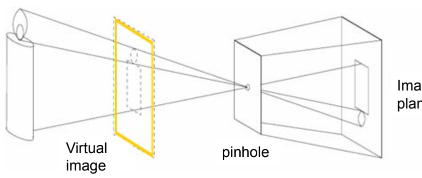


- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**
  - How does this transform the image?

Slide by Steve Seitz

Pinhole camera

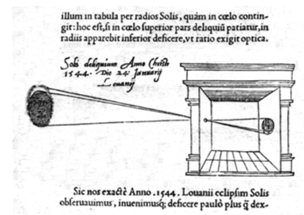
- Pinhole camera is a simple model to approximate imaging process, perspective **projection**.



If we treat pinhole as a point, only one ray from any given point can enter the camera.

Fig from Forsyth and Ponce

Camera obscura



In Latin, means 'dark room'

"**Reinerus Gemma-Frisius**, observed an eclipse of the sun at Louvain on January 24, 1544, and later he used this illustration of the event in his book De Radio Astronomica et Geometrica, 1545. It is thought to be the first published illustration of a camera obscura..."
Hammond, John H., The Camera Obscura, A Chronicle

http://www.acmi.net.au/AIC/CAMERA_OBSCURA.html

Camera obscura

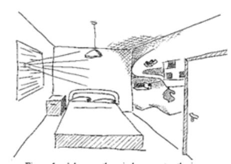

Jetty at Margate England, 1898.


Around 1870s

An attraction in the late 19^{th} century

http://brightbytes.com/cosite/collection2.html
Adapted from R. Duraiswami

Camera obscura at home



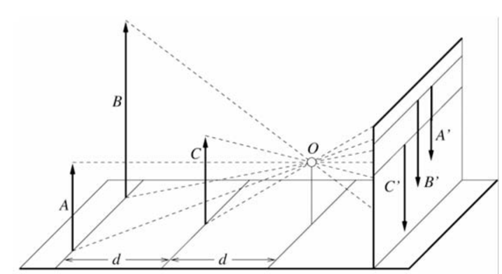
Sketch from http://www.funsci.com/fun3_en/sky/sky.htm

http://blog.makezine.com/archive/2006/02/how_to_room_sized_camera_obscu.html

Perspective effects

Perspective effects

- Far away objects appear smaller

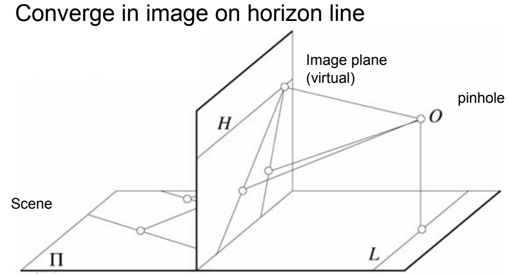


Forsyth and Ponce

Perspective effects

Perspective effects

- Parallel lines in the scene intersect in the image
- Converge in image on horizon line



Projection properties

- Many-to-one: any points along same ray map to same point in image
- Points → points
- Lines → lines (collinearity preserved)
- Distances and angles are **not** preserved

- Degenerate cases:
  – Line through focal point projects to a point.
  – Plane through focal point projects to line
  – Plane perpendicular to image plane projects to part of the image.

Perspective and art

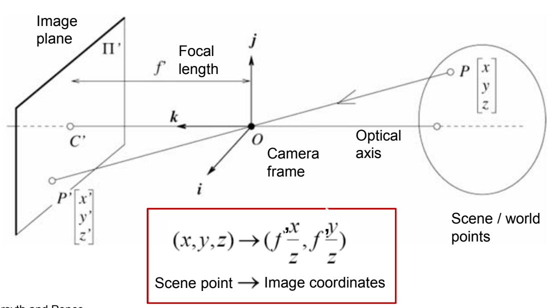
- Use of correct perspective projection indicated in 1st century B.C. frescoes
- Skill resurfaces in Renaissance: artists develop systematic methods to determine perspective projection (around 1480-1515)

Raphael

Durer, 1525

Perspective projection equations

- 3d world mapped to 2d projection in image plane



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

Scene point → Image coordinates

Forsyth and Ponce

Homogeneous coordinates

Is this a linear transformation?

- no—division by z is nonlinear

Trick: add one more coordinate:

$$(x,y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (x,y,z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous image coordinates

homogeneous scene coordinates

Converting *from* homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w) \qquad \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Slide by Steve Seitz

Perspective Projection Matrix

- Projection is a matrix multiplication using homogeneous coordinates:

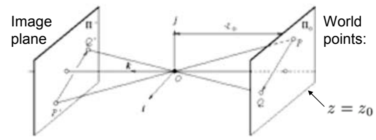
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f' \end{bmatrix} \Rightarrow (f'\frac{x}{z}, f'\frac{y}{z})$$

divide by the third coordinate to convert back to non-homogeneous coordinates

Slide by Steve Seitz

Weak perspective

- Approximation: treat magnification as constant
- Assumes scene depth << average distance to camera

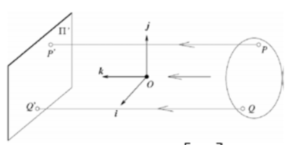


$$x' = f\frac{x}{z} \approx \frac{f}{z_0}x$$

$$y' = f\frac{y}{z} \approx \frac{f}{z_0}y$$

Orthographic projection

- Given camera at **constant** distance from scene
- World points projected along rays parallel to optical access



$$x' = x$$

$$y' = y$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow (x, y)$$

Physical parameters of image formation

- Geometric
  - Type of projection
  - Camera pose
- Optical
  - Sensor's lens type
  - focal length, field of view, aperture
- Photometric
  - Type, direction, intensity of light reaching sensor
  - Surfaces' reflectance properties

Pinhole size / aperture

How does the size of the aperture affect the image we'd get?

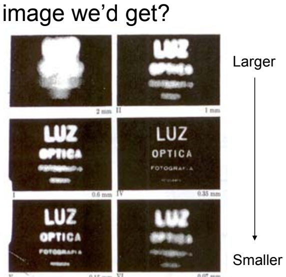
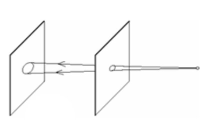
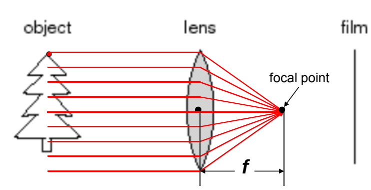


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]



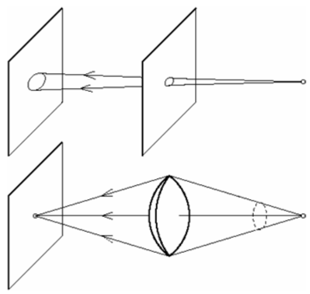
Adding a lens



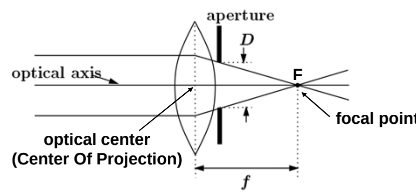
- A lens focuses light onto the film
  - Rays passing through the center are not deviated
  - All parallel rays converge to one point on a plane located at the *focal length f*

Slide by Steve Seitz

Pinhole vs. lens

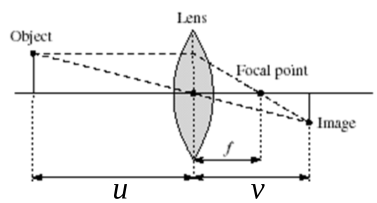


Cameras with lenses



- A lens focuses parallel rays onto a single focal point
- Gather more light, while keeping focus; make pinhole perspective projection practical

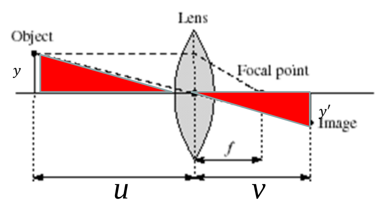
Thin lens equation



- How to relate distance of object from optical center (u) to the distance at which it will be in focus (v), given focal length f?

Thin lens equation

$$\frac{y'}{y} = \frac{v}{u}$$

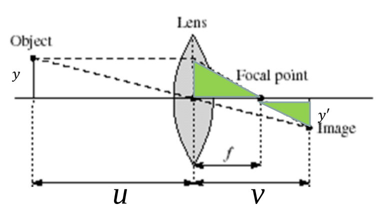


- How to relate distance of object from optical center (u) to the distance at which it will be in focus (v), given focal length f?

Thin lens equation

$$\frac{y'}{y} = \frac{v}{u}$$

$$\frac{y'}{y} = \frac{(v-f)}{f}$$

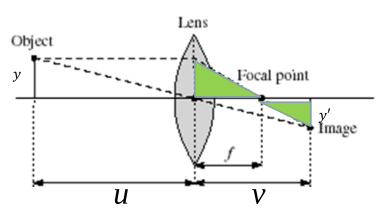


- How to relate distance of object from optical center (u) to the distance at which it will be in focus (v), given focal length f?

Thin lens equation

$$\frac{y'}{y} = \frac{v}{u}$$

$$\frac{y'}{y} = \frac{(v-f)}{f}$$



$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

- Any object point satisfying this equation is in focus

Focus and depth of field

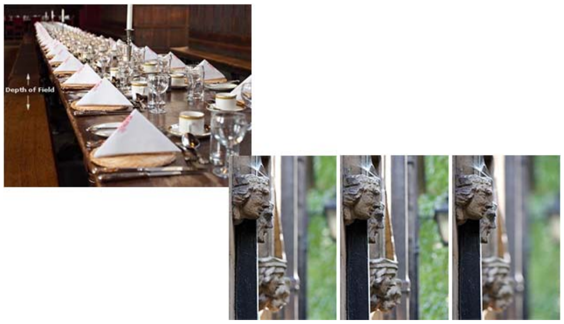
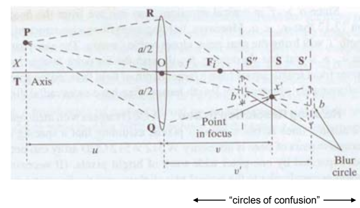


Image credit: cambridgeincolour.com

Focus and depth of field

- Depth of field: distance between image planes where blur is tolerable



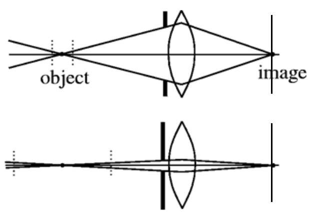
Thin lens: scene points at distinct depths come in focus at different image planes.

(Real camera lens systems have greater depth of field.)

Shapiro and Stockman

Focus and depth of field

- How does the aperture affect the depth of field?



- A smaller aperture increases the range in which the object is approximately in focus

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field Slide from S. Seitz

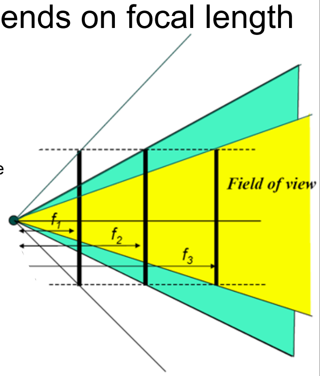
Field of view

- Angular measure of portion of 3d space seen by the camera

Images from http://en.wikipedia.org/wiki/Angle_of_view

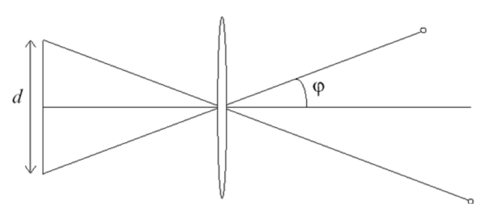
Field of view depends on focal length

- As ***f*** gets smaller, image becomes more *wide angle*
  - more world points project onto the finite image plane
- As ***f*** gets larger, image becomes more *telescopic*
  - smaller part of the world projects onto the finite image plane



from R. Duraiswami

Field of view depends on focal length



Size of field of view governed by size of the camera retina:

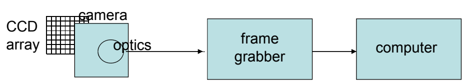
$$\varphi = \tan^{-1}\left(\frac{d}{2f}\right)$$

Smaller FOV = larger Focal Length

Slide by A. Efros

Digital cameras

- Film → sensor array
- Often an array of charge coupled devices
- Each CCD is light sensitive diode that converts photons (light energy) to electrons



Summary

- Image formation affected by geometry, photometry, and optics.
- Projection equations express how world points mapped to 2d image.
  – Homogenous coordinates allow linear system for projection equations.
- Lenses make pinhole model practical.
- Parameters (focal length, aperture, lens diameter,…) affect image obtained.

Next time

- Geometry of two views