Image Mosaics

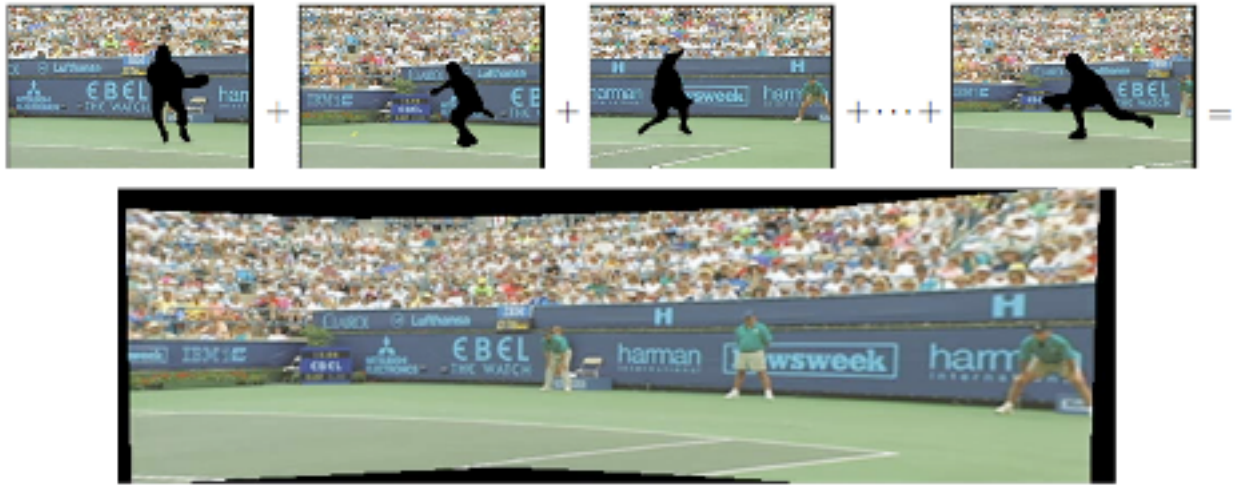
with some slides from R. Szeliski, S. Seitz, D. Lowe, A. Efros,

Image Stitching

- Aligning multiple images together to form a composite
- The type of alignment depends on the motion between images

Applications

- Video stitching of background scene



Rotation Only - Calibrated Case

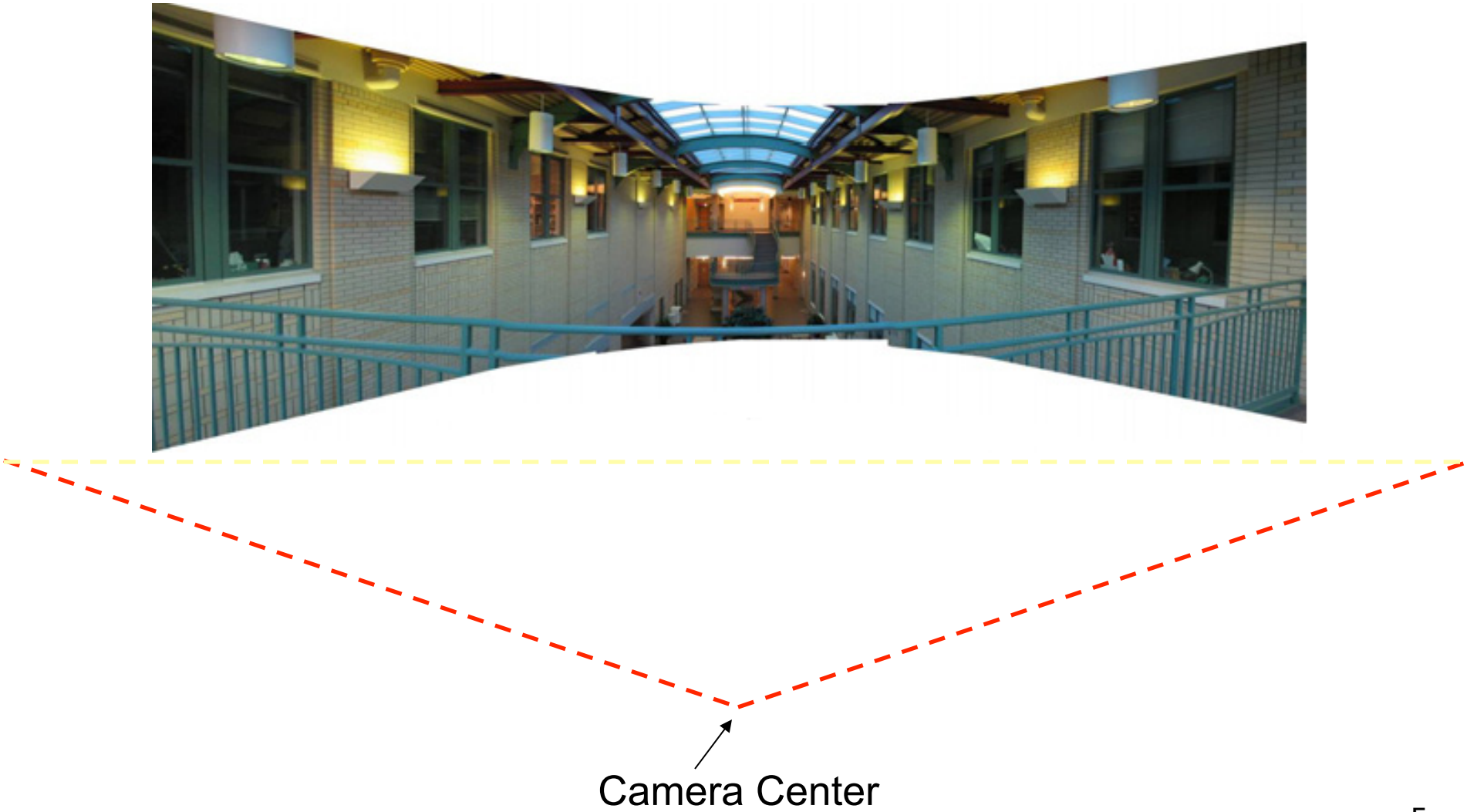
- Calibrated Two views related by rotation only

$$\lambda_2 \mathbf{x}_2 = R\lambda_1 \mathbf{x}_1 \qquad \widehat{\mathbf{x}_2} R \mathbf{x}_1 = 0$$

- Mapping to a reference view – rotation can be estimated

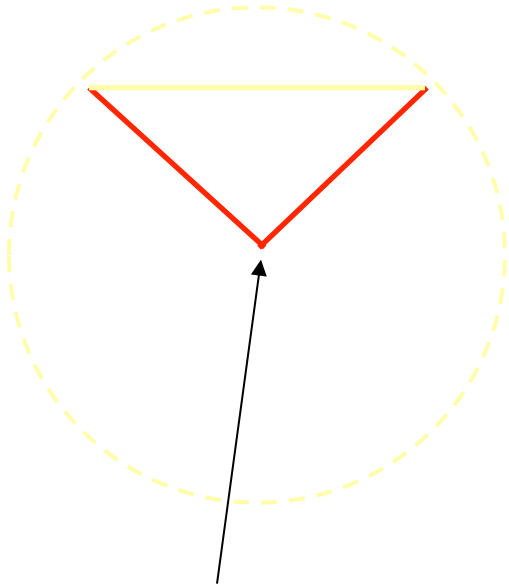
- Mapping to a cylindrical surface

Do we have to project on to a plane ?



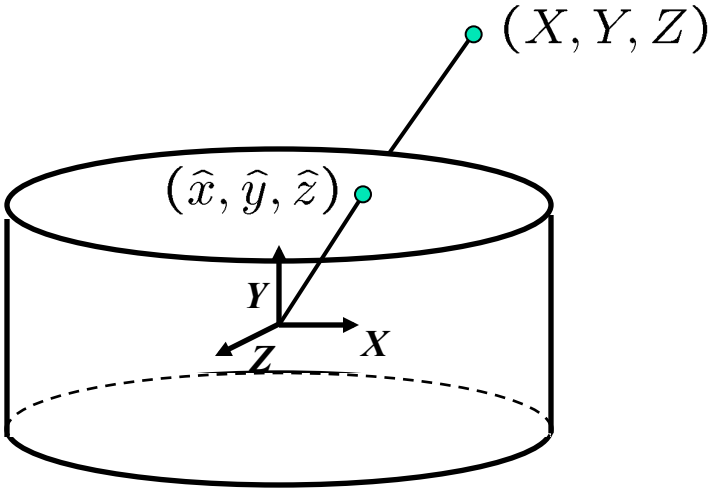
Cylindrical Projection

360˚ Panorama
[Szeliski & Shum 97]



Camera Center

Cylindrical projection



unit cylinder

- Map 3D point (X,Y,Z) onto cylinder

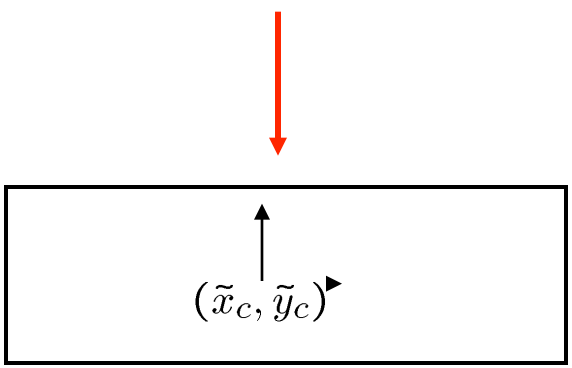
$$(\hat{x}, \hat{y}, \hat{z}) = \frac{1}{\sqrt{X^2+Z^2}}(X, Y, Z)$$

- Convert to cylindrical coordinates

$$(sin\theta, h, cos\theta) = (\hat{x}, \hat{y}, \hat{z})$$

- Convert to cylindrical image coordinates

$$(\tilde{x}, \tilde{y}) = (f\theta, fh) + (\tilde{x}_c, \tilde{y}_c)$$

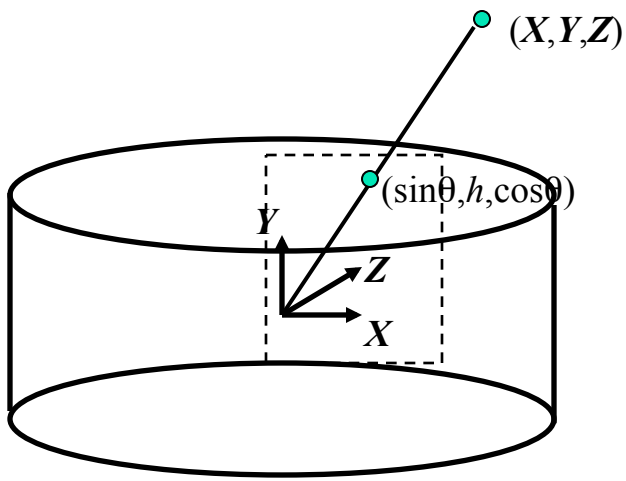


unwrapped cylinder

cylindrical image

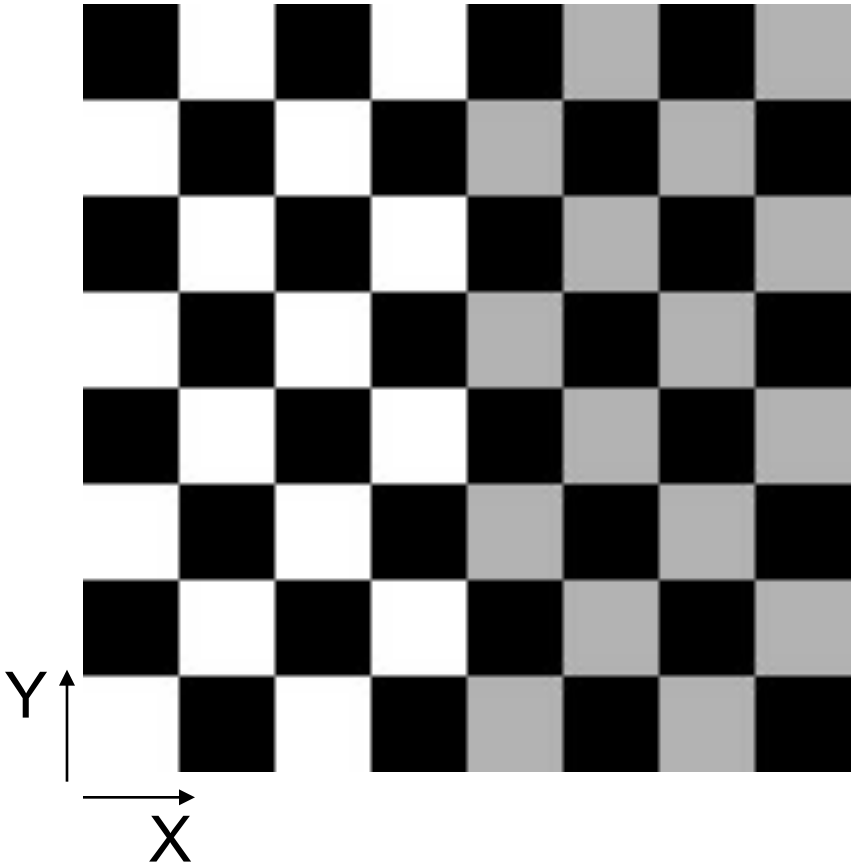
Inverse Cylindrical projection

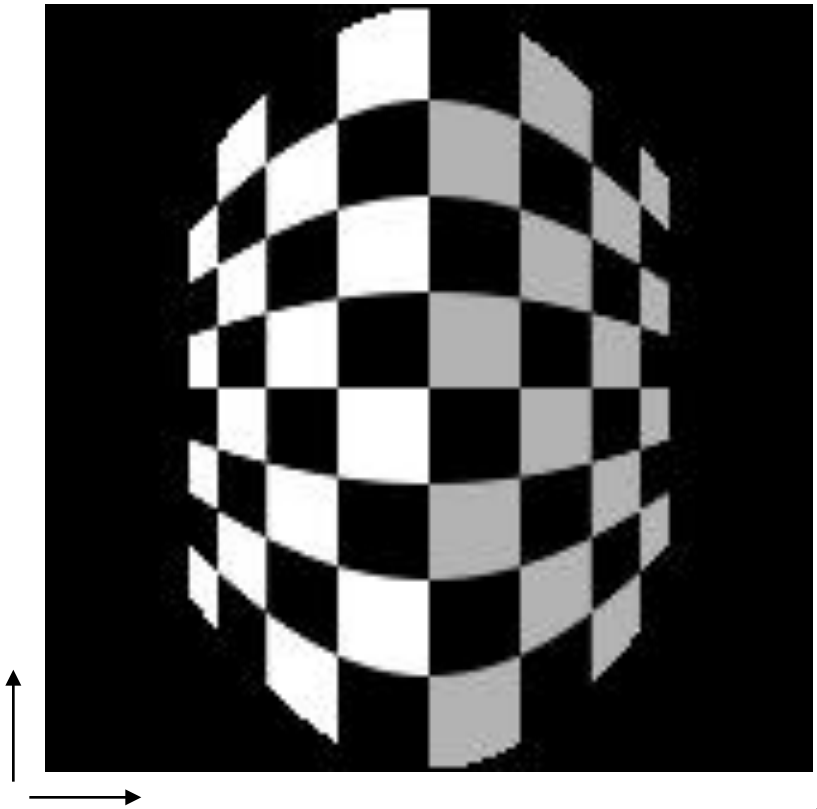
$$(\tilde{x}, \tilde{y}) = (f\theta, fh) + (\tilde{x}_c, \tilde{y}_c)$$



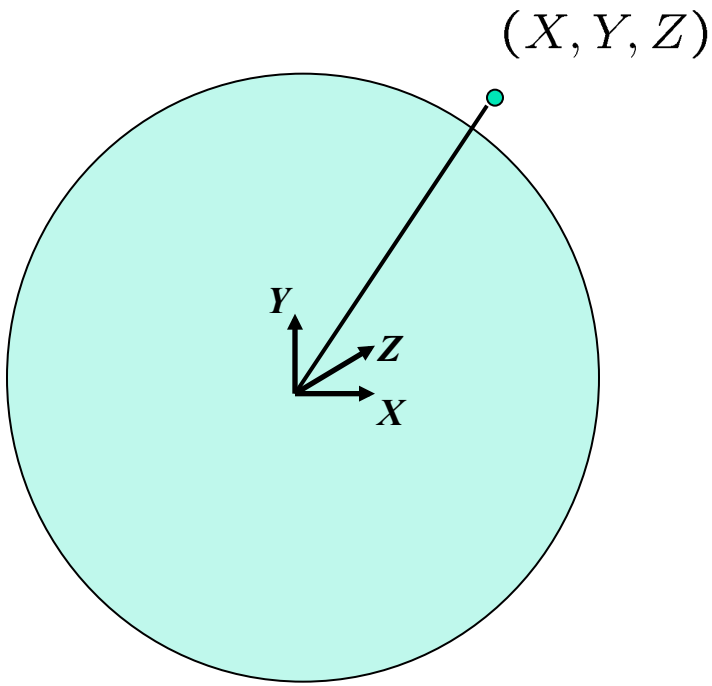
$$\begin{aligned}
\theta &= (x_{cyl} - x_c)/f \\
h &= (y_{cyl} - y_c)/f \\
\hat{x} &= \sin\theta \\
\hat{y} &= h \\
\hat{z} &= \cos\theta \\
x &= f\hat{x}/\hat{z} + x_c \\
y &= f\hat{y}/\hat{z} + y_c
\end{aligned}$$

Cylindrical Projection





Spherical projection



- Map 3D point (X,Y,Z) onto sphere

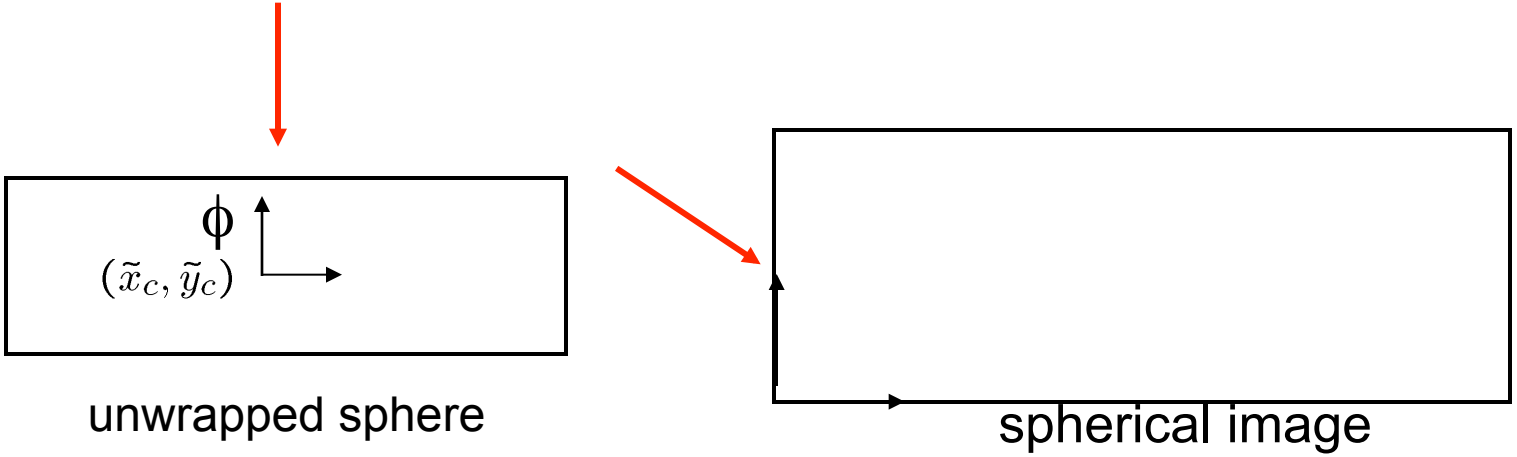
$$(\hat{x}, \hat{y}, \hat{z}) = \frac{1}{\sqrt{X^2 + Y^2 + Z^2}}(X, Y, Z)$$

- Convert to spherical coordinates

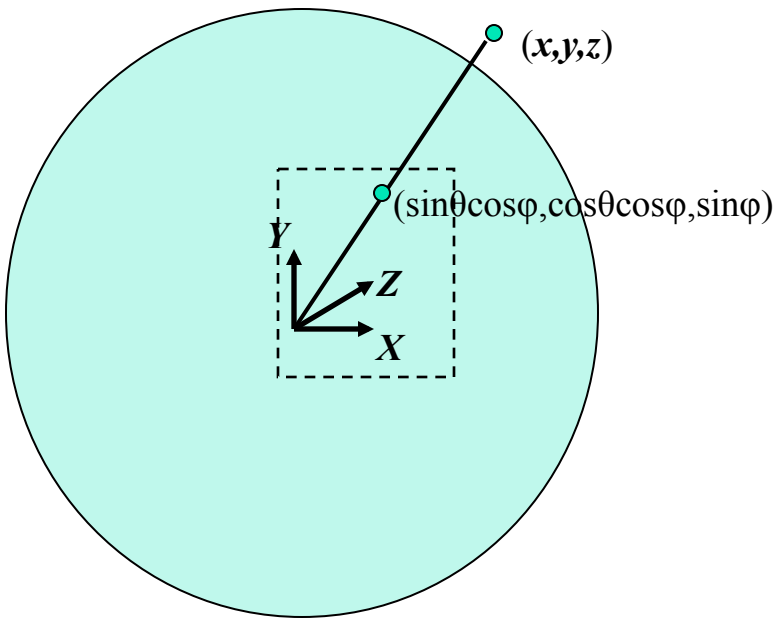
$$(\sin\theta \cos\phi, \sin\phi, \cos\theta \cos\phi) = (\hat{x}, \hat{y}, \hat{z})$$

- Convert to spherical image coordinates

$$(\tilde{x}, \tilde{y}) = (f\theta, fh) + (\tilde{x}_c, \tilde{y}_c)$$



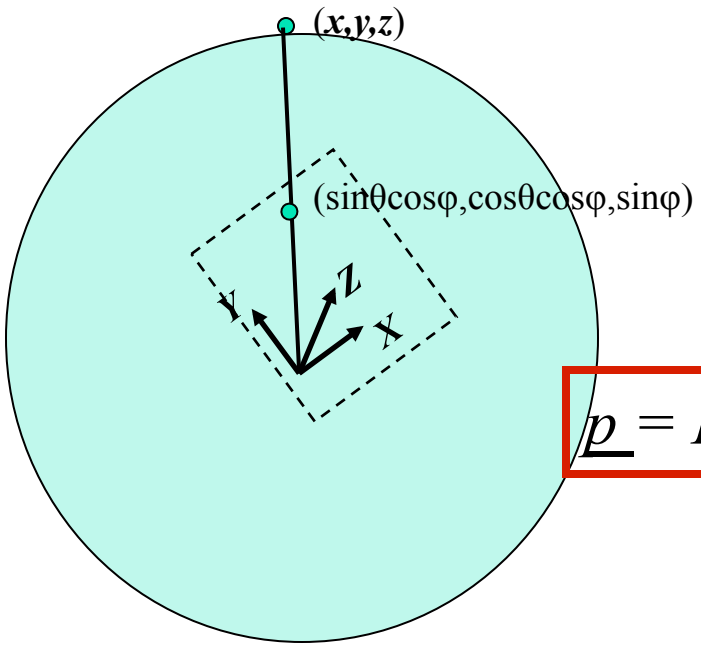
Inverse Spherical projection



$$
\begin{aligned}
\theta &= (x_{sph} - x_c)/f \\
h &= (y_{sph} - y_c)/f \\
\hat{x} &= \sin\theta \ \cos \\
\hat{y} &= \sin\varphi \\
\hat{z} &= \cos\theta \cos\varphi \\
x &= f\hat{x}/\hat{z} + x_c \\
y &= f\hat{y}/\hat{z} + y_c
\end{aligned}
$$

3D rotation

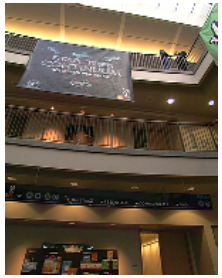
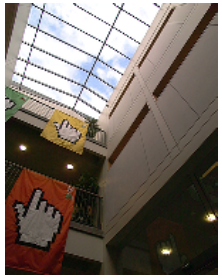
- Rotate image before placing on unrolled sphere

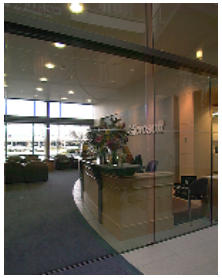


$$\underline{p} = \boldsymbol{R}\, p$$

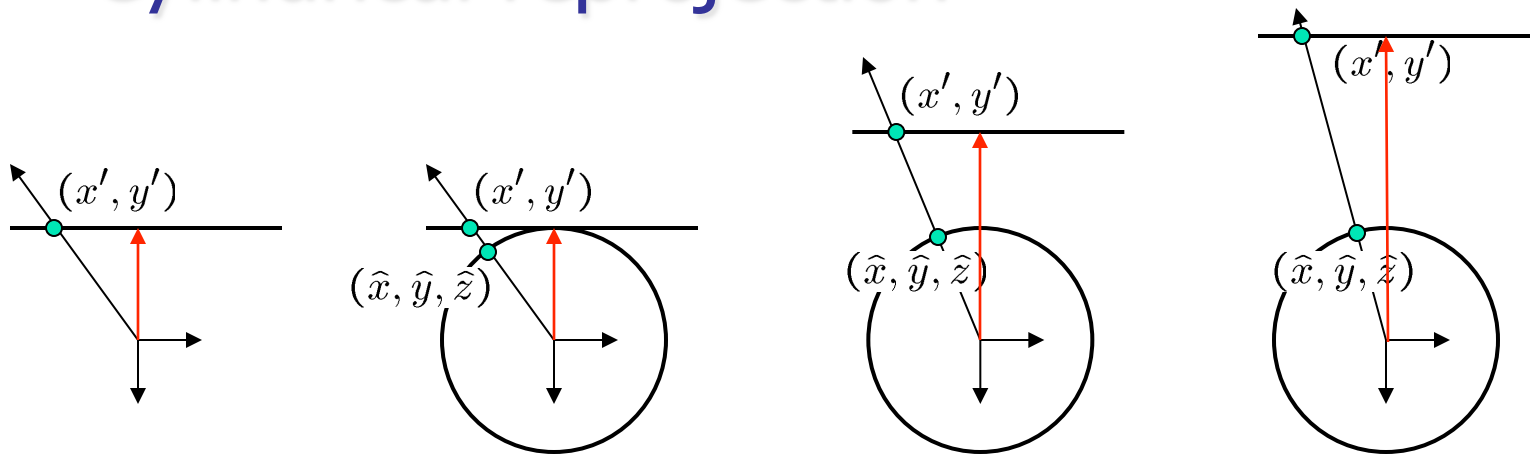
$$\begin{aligned}
\theta &= (x_{sph} - x_c)/f \\
h &= (y_{sph} - y_c)/f \\
\hat{x} &= \sin\theta \ \cos \\
\hat{y} &= \sin\varphi \\
\hat{z} &= \cos\theta \cos\varphi \\
x &= f\hat{x}/\hat{z} + x_c \\
y &= f\hat{y}/\hat{z} + y_c
\end{aligned}$$

Full-view Panorama


+

+

+

+


Cylindrical reprojection

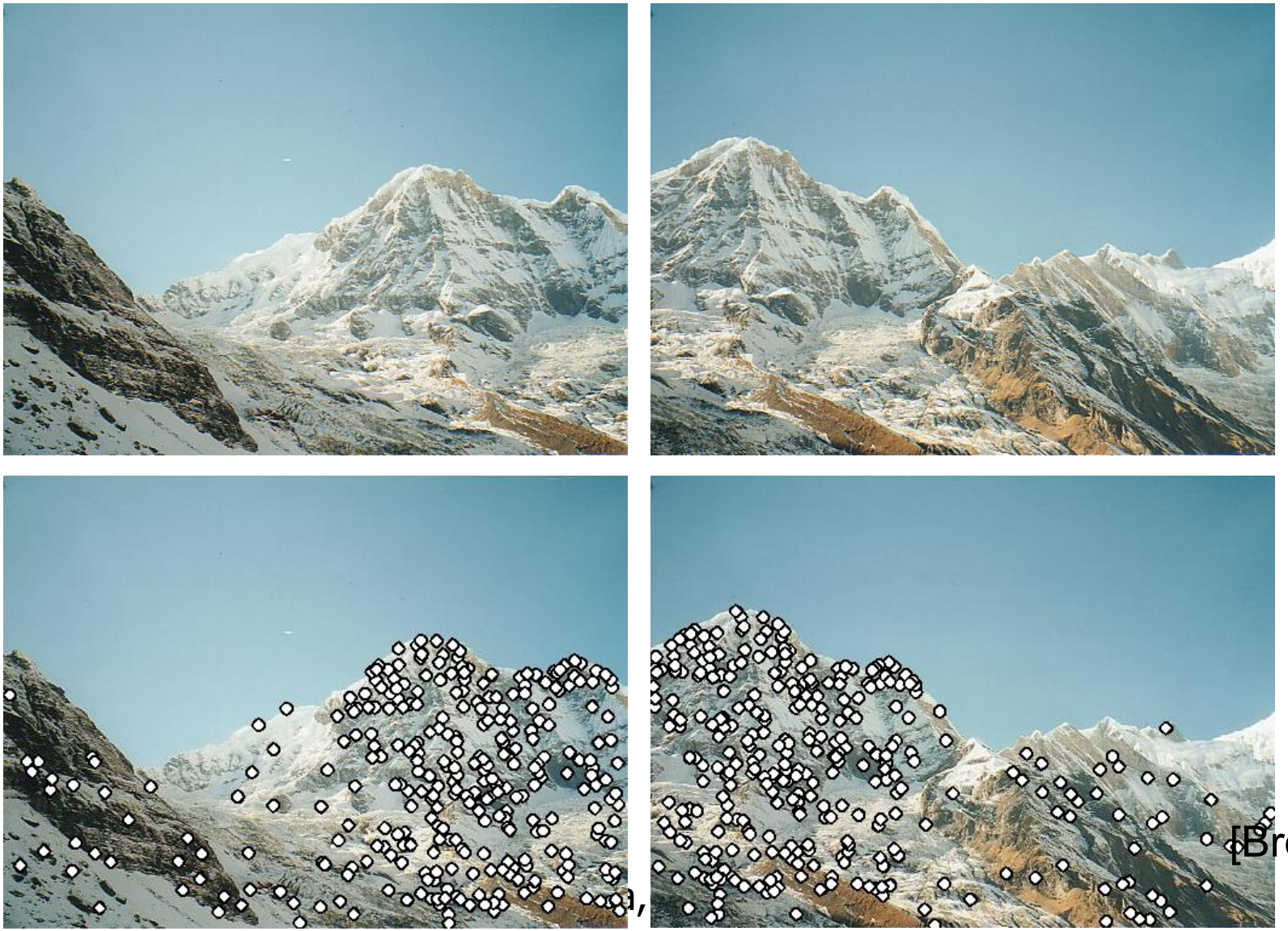


top-down view

Focal length – the dirty secret…

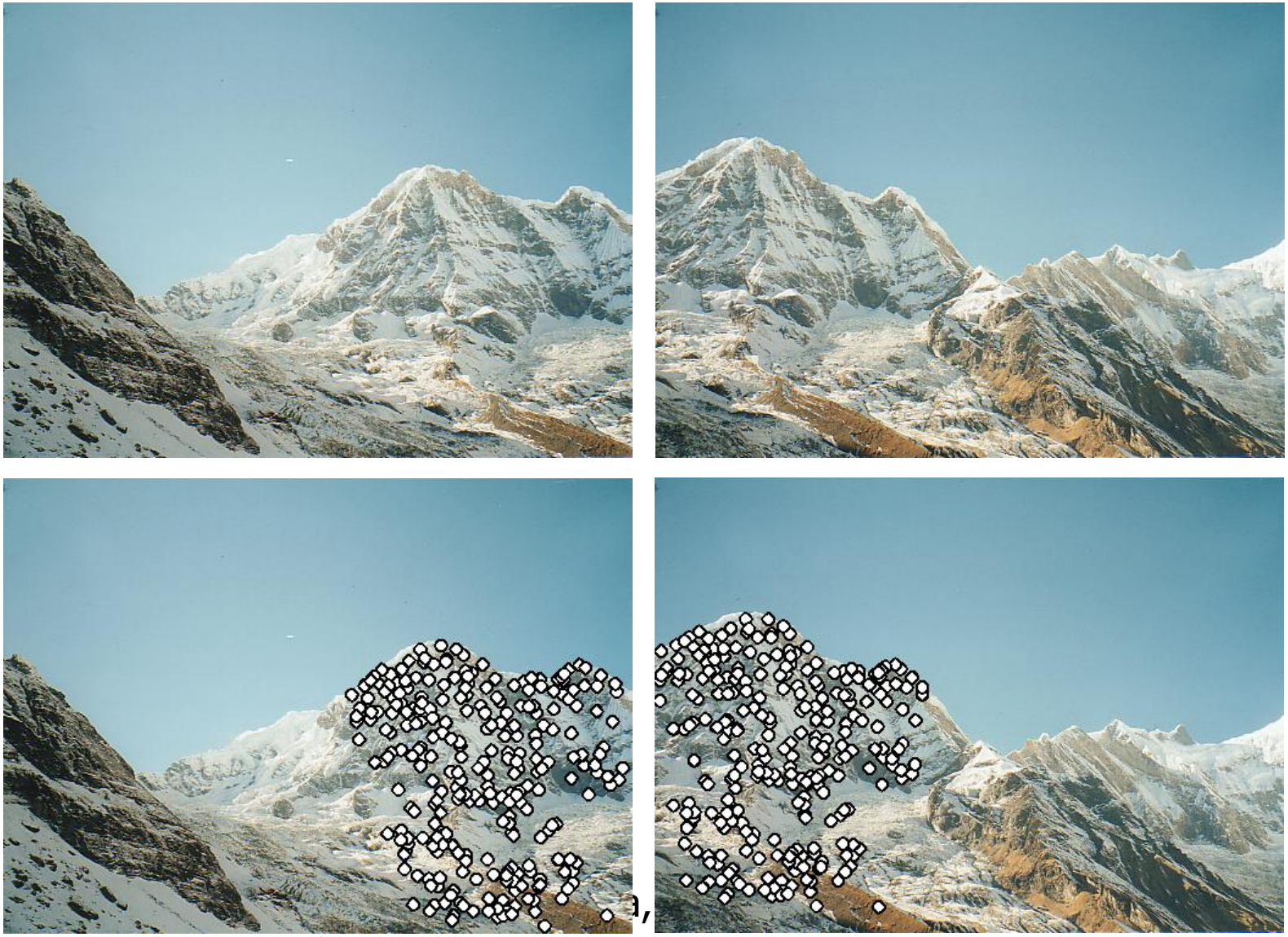
Image 384x300 **f = 180 (pixels)** **f = 280** **f = 380**

Matching SIFT Features



[Brown 2003]

Reject Outliers using RANSAC



Stitching Images

Removing Seams

- Vignetting – reduction of image brightness at the periphery – caused in digital cameras by angle dependence of the light incidence on the sensor
- Small misalignments

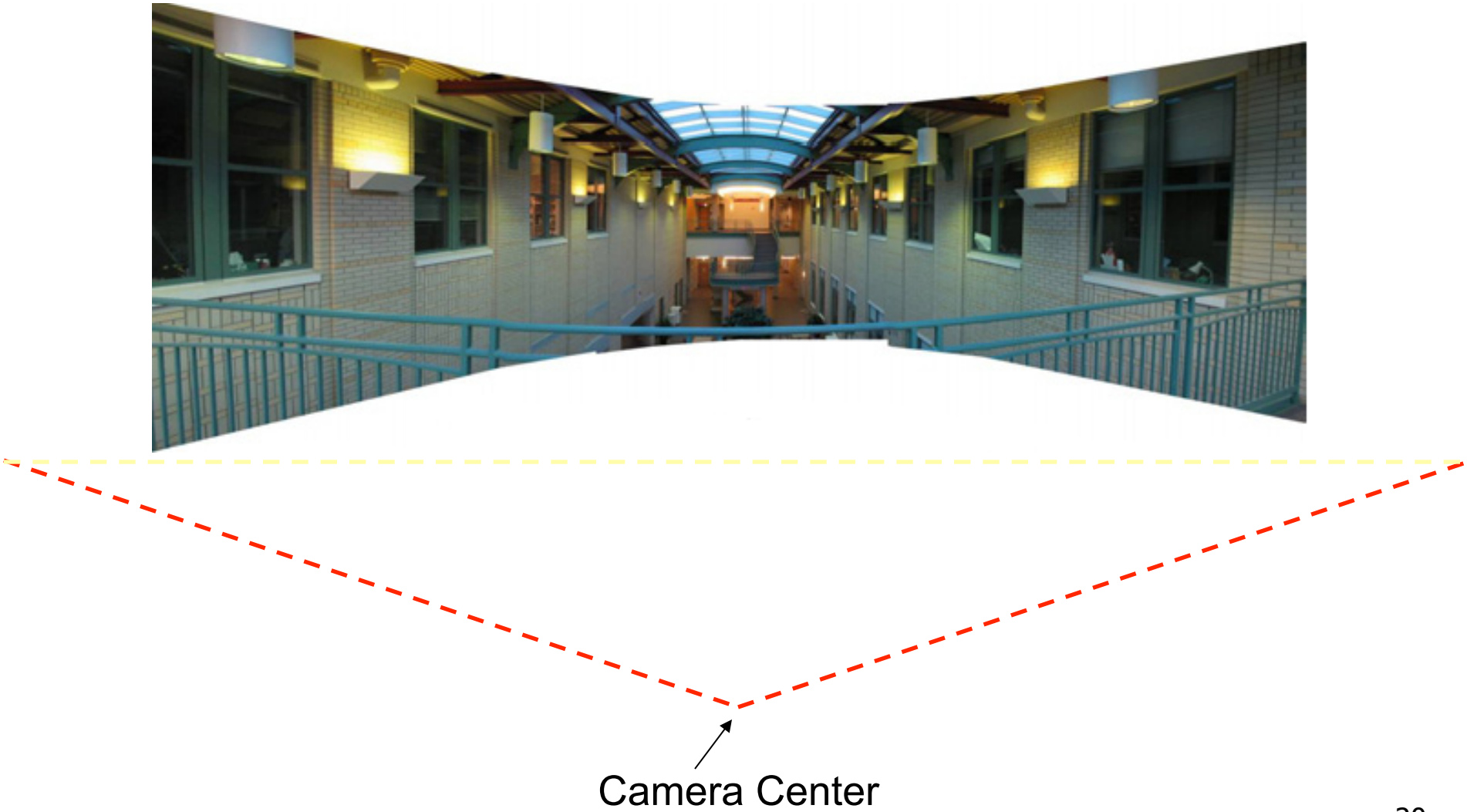
[Brown 2003]

Multi-band Blending

- Multi-resolution technique using image pyramid
- Hides seams but preserves sharp detail

[Brown 2003]

Do we have to project on to a plane ?



How to do it?

- Basic Procedure
  - Take a sequence of images from the same position
    - Rotate the camera about its optical center
  - Compute transformation between second image and first
  - Transform the second image to overlap with the first
  - Blend the two together to create a mosaic
  - If there are more images, repeat

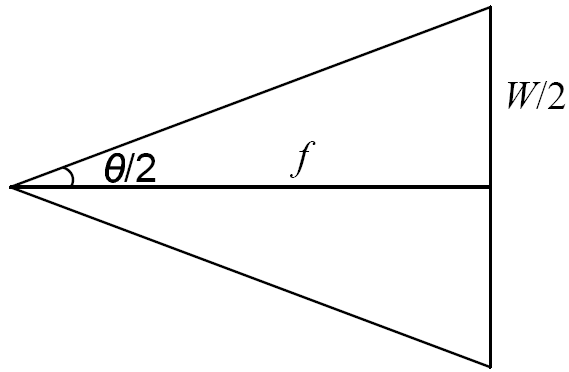
- …but **wait**, why should this work at all?
  - What about the 3D geometry of the scene?
  - Why aren't we using it?

Mosaics: stitching images together

virtual wide-angle camera

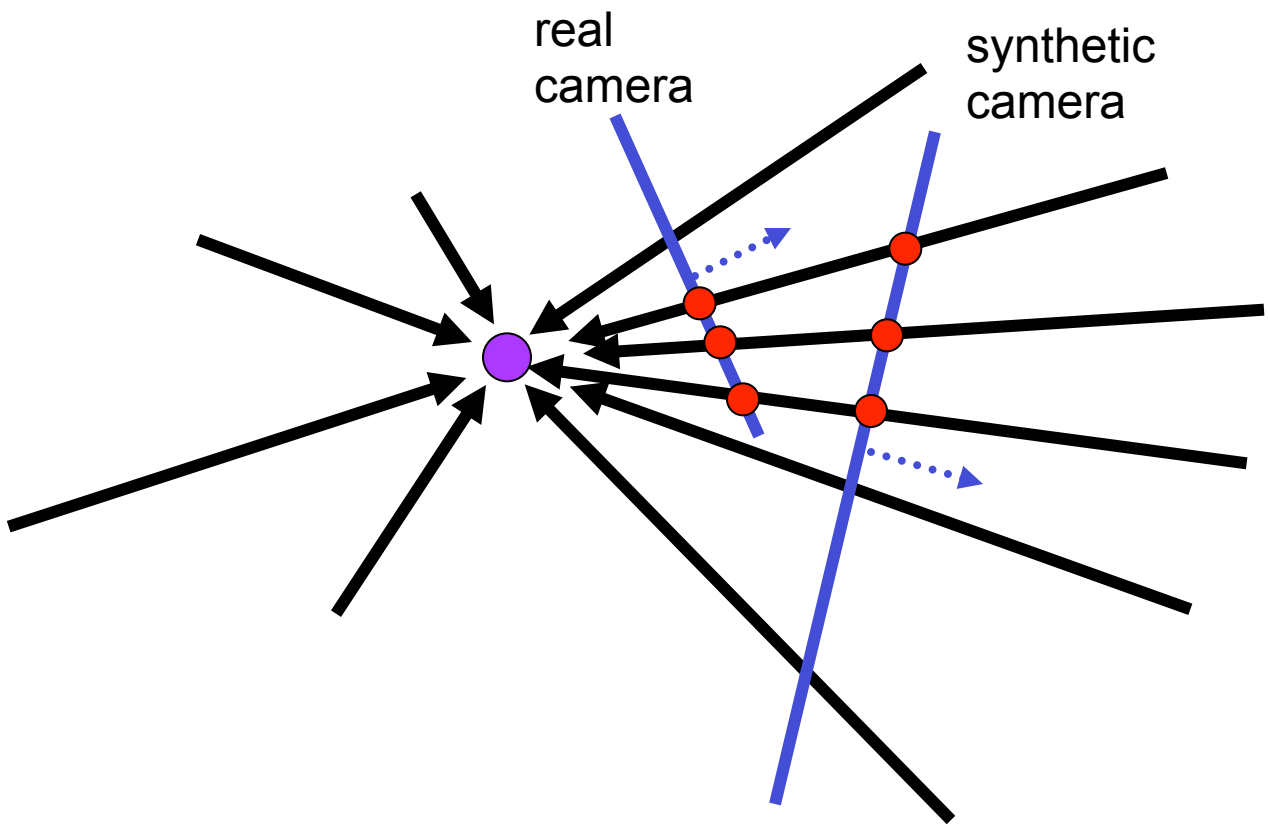
What's your focal length ?

- Focal length is (highly!) depends on picture/camera
  - Can get a rough estimate by measuring FOV:



  - Can use the EXIF data tag (might not give the right thing)
  - Can use several images together and try to find f that would make them match
  - Can use a known 3D object and its projection to solve for f
  - Can use vanishing points Etc.
- There are other camera parameters too:
  - Optical center, non-square pixels, lens distortion, etc.

A pencil of rays contains all views



Can generate any synthetic camera view
as long as it has **the same center of projection**!

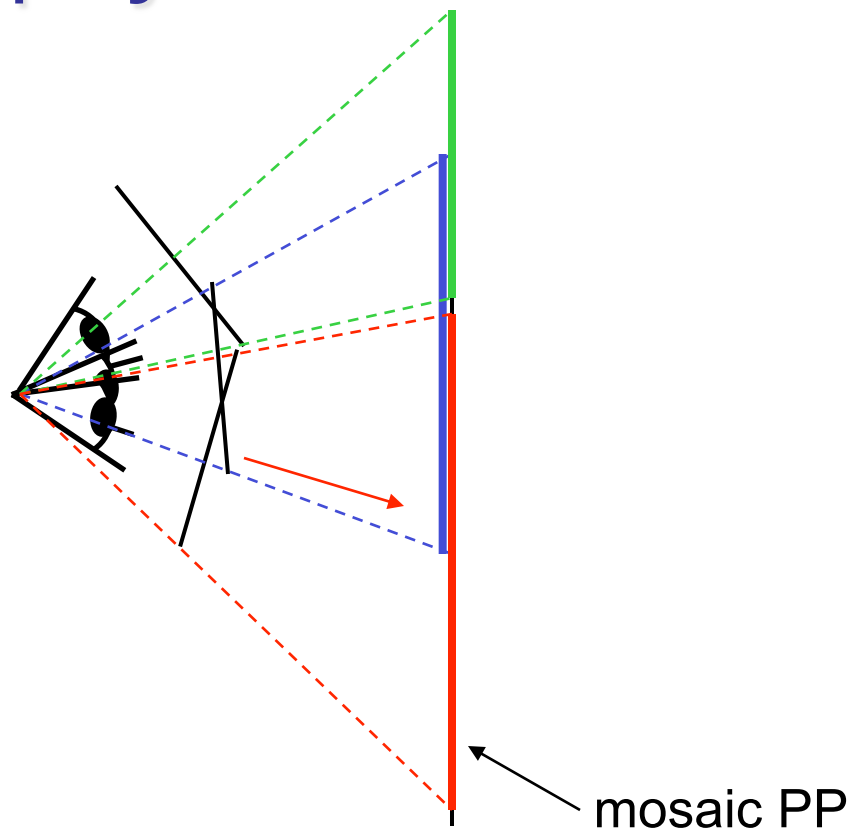
Aligning images

left on top

right on top

Translations are not enough to align the images

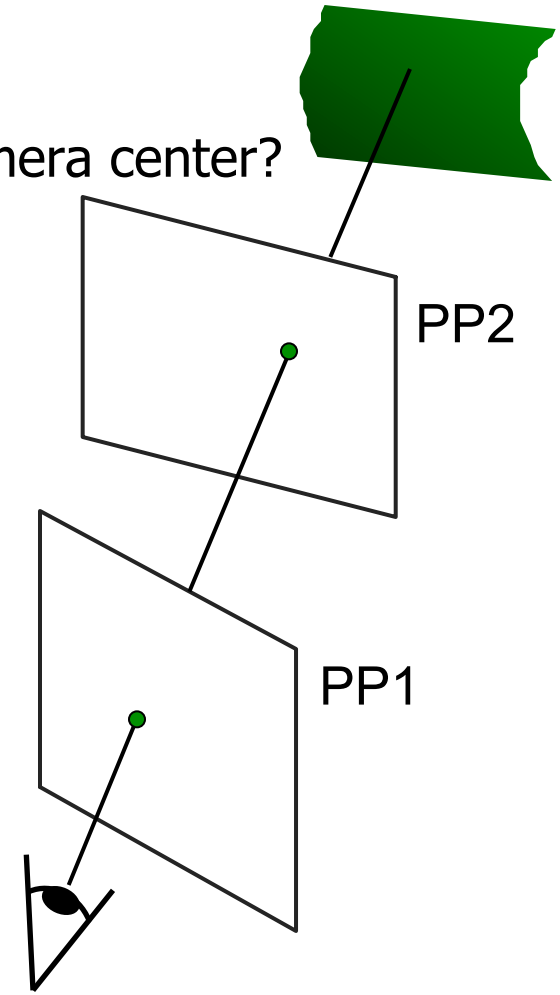
Image reprojection



- The mosaic has a natural interpretation in 3D
  - The images are reprojected onto a common plane
  - The mosaic is formed on this plane
  - Mosaic is a *synthetic wide-angle camera*

Image reprojection

- Basic question
  - How to relate two images from the same camera center? how to map a pixel from PP1 to PP2
  - Image warping
- Answer
  - Cast a ray through each pixel in PP1
  - Draw the pixel where that ray intersects PP2

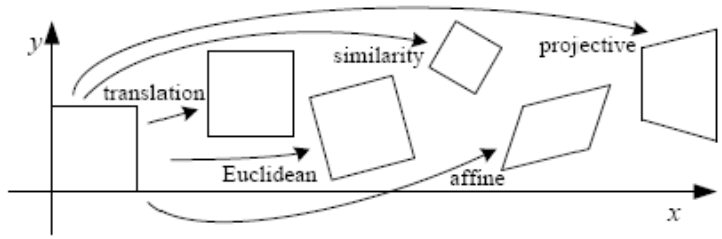


But don't we need to know the geometry of the two planes in respect to the eye?

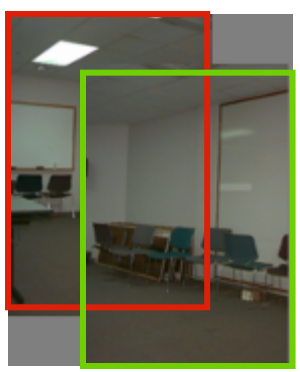
Observation:
Rather than thinking of this as a 3D reprojection, think of it as a 2D **image warp** from one image to another

Back to Image Warping

- Which transform is the right one for warping PP1 into PP2
e.g. translation, Euclidean, affine, projective
- Previously in the context of optical flow – local warping

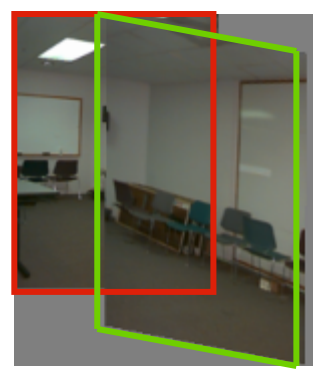


Translation



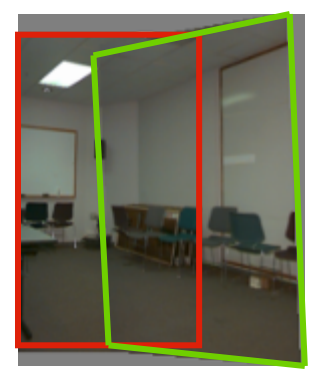
2 unknowns

Affine



6 unknowns

Perspective



8 unknowns

Homography

- A: Projective – mapping between any two PPs with the same center of projection
  - rectangle should map to arbitrary quadrilateral
  - parallel lines aren't
  - but must preserve straight lines
  - same as: project, rotate, reproject
- Homography
- One-to-one image mapping

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$\mathbf{p}'$ $\mathbf{H}$ $\mathbf{p}$

In homogeneous coordinates

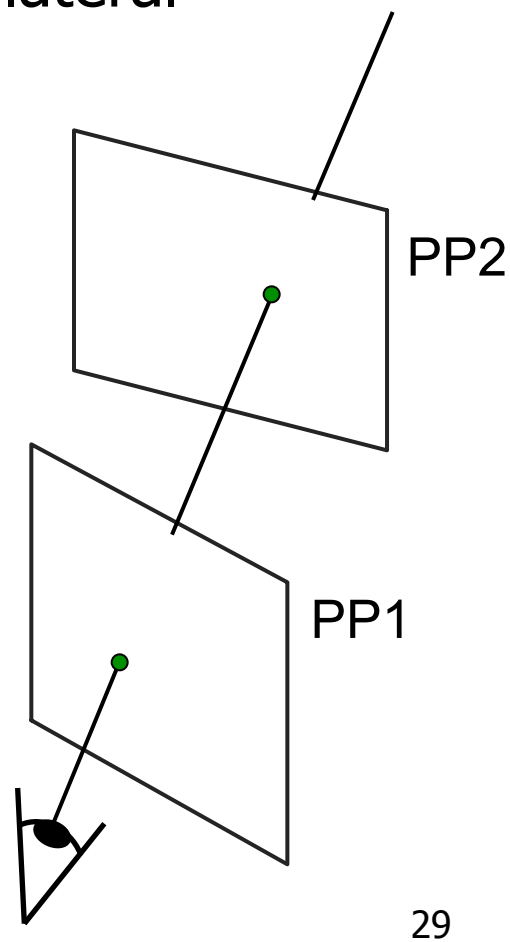
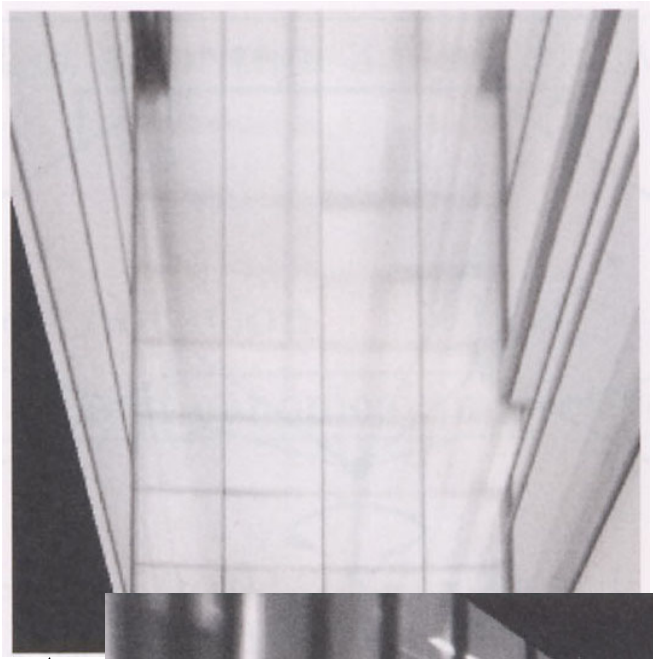


Image warping with homographies



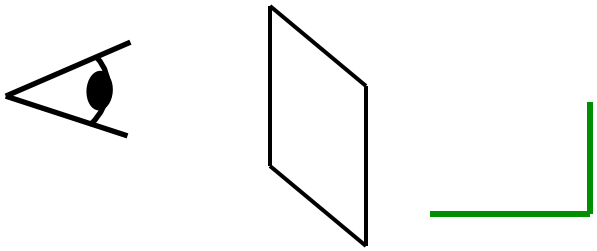
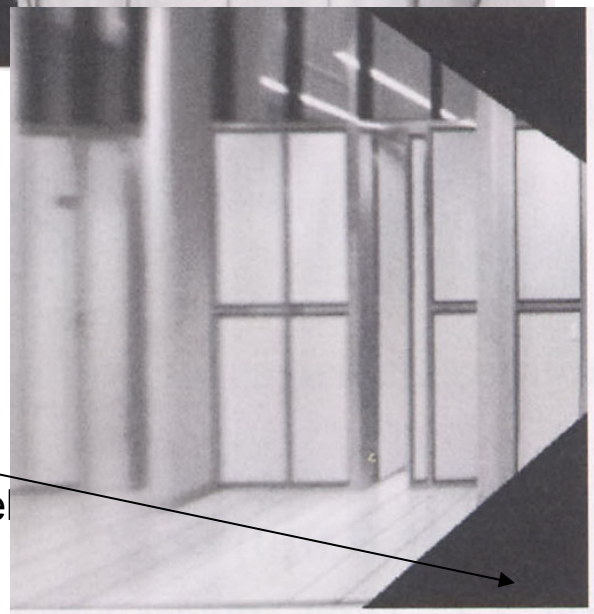


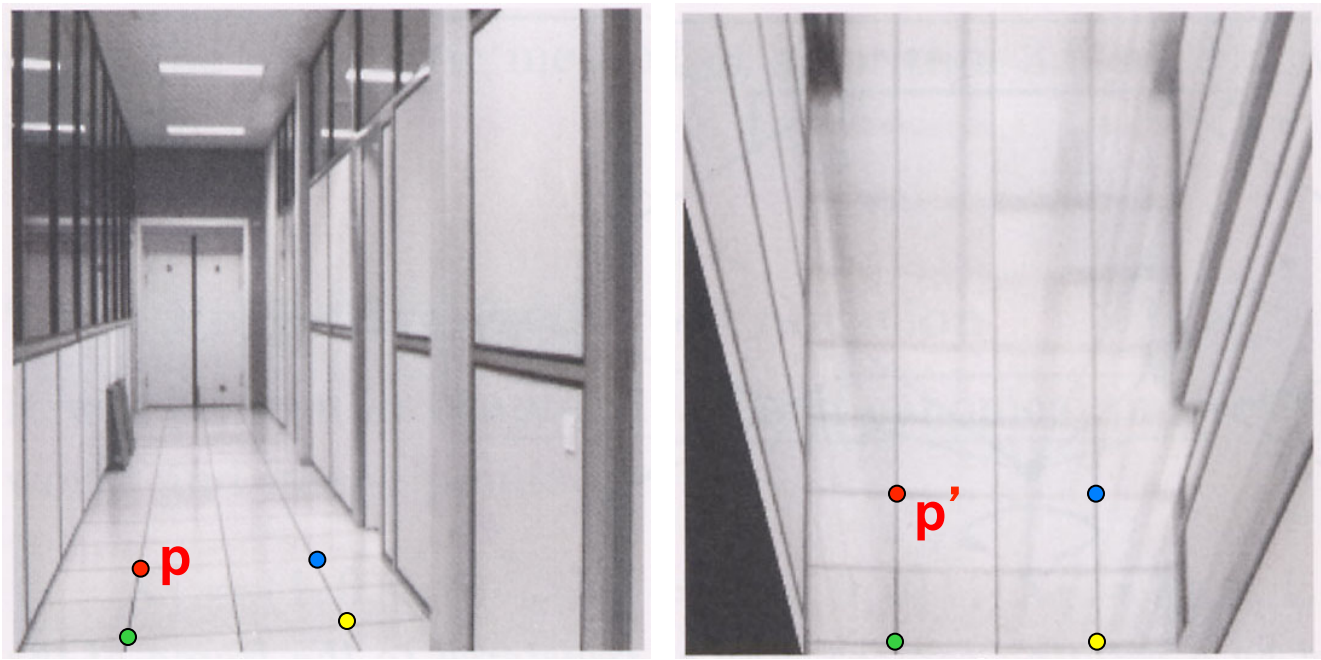
image plane in front



black area where no pixel maps to

- Warp to virtual ground plane
- How to compute the homography ?

Image rectification



To unwarp (rectify) an image

- Find the homography **H** given a set of **p** and **p'** pairs
- How many correspondences are needed?
- Tricky to write H analytically, but we can <u>solve</u> for it!
  - Find such H that "best" transforms points p into p'
  - Use least-squares!

Fun with homographies

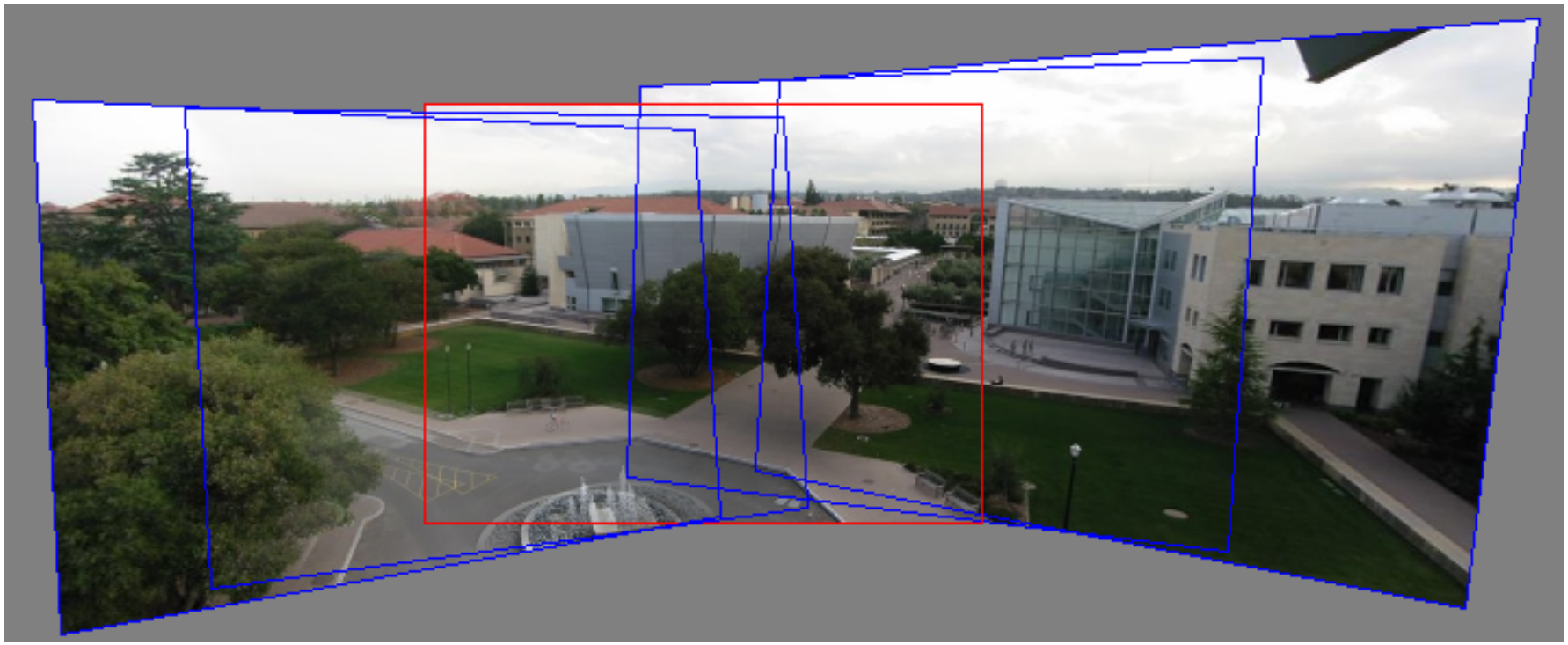
Original image

St.Petersburg
photo by A. Tikhonov

Virtual camera rotations

k

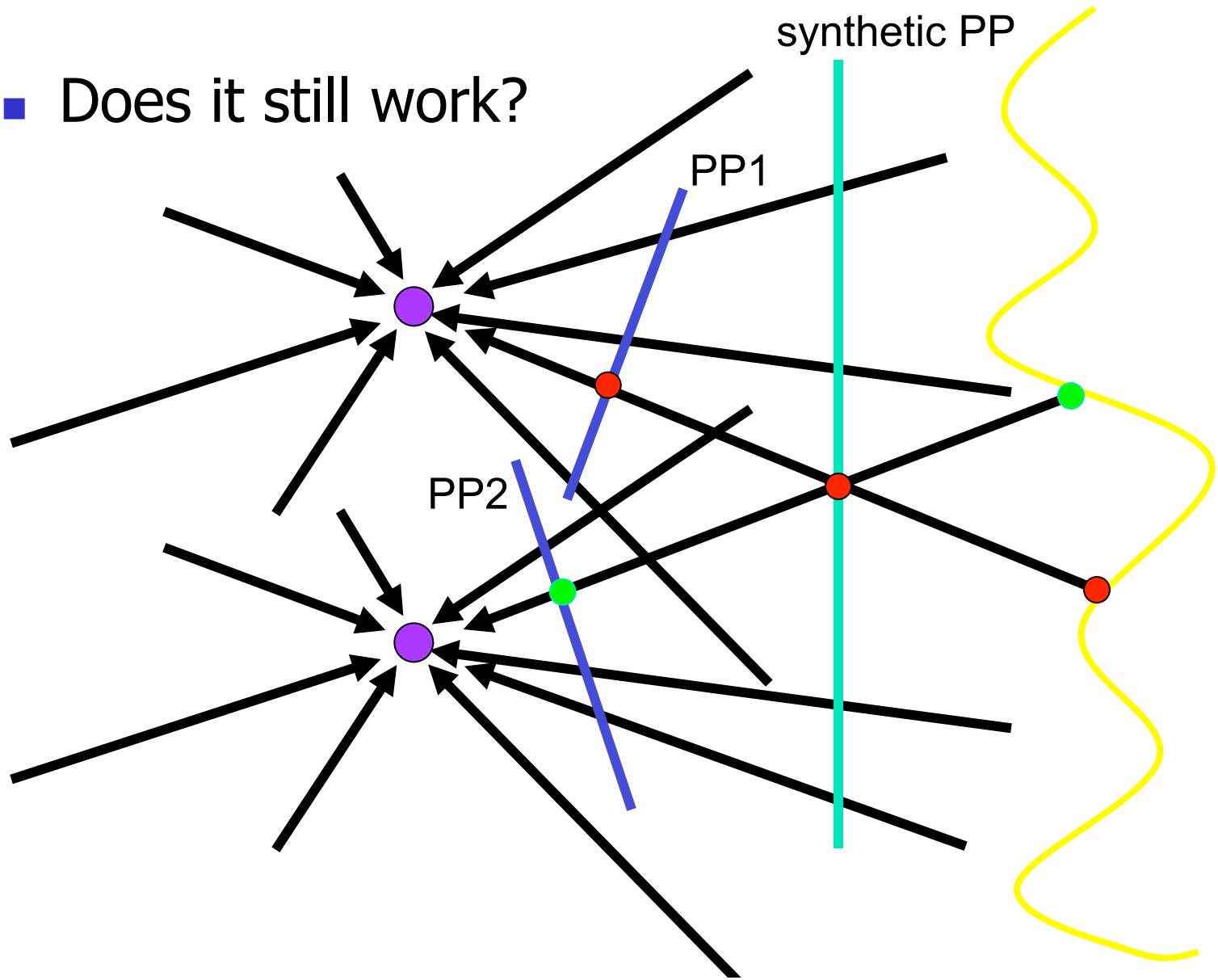
Panoramas



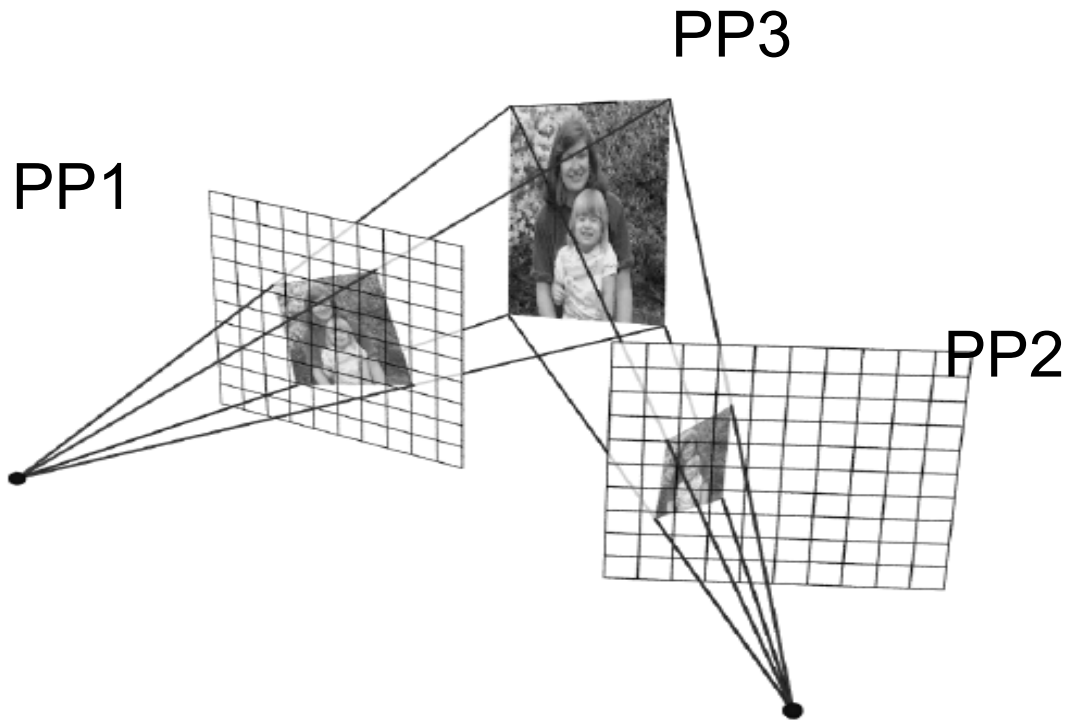
1. Pick one image (red)
2. Warp the other images towards it (usually, one by one)
3. blend

Changing camera center

- Does it still work?

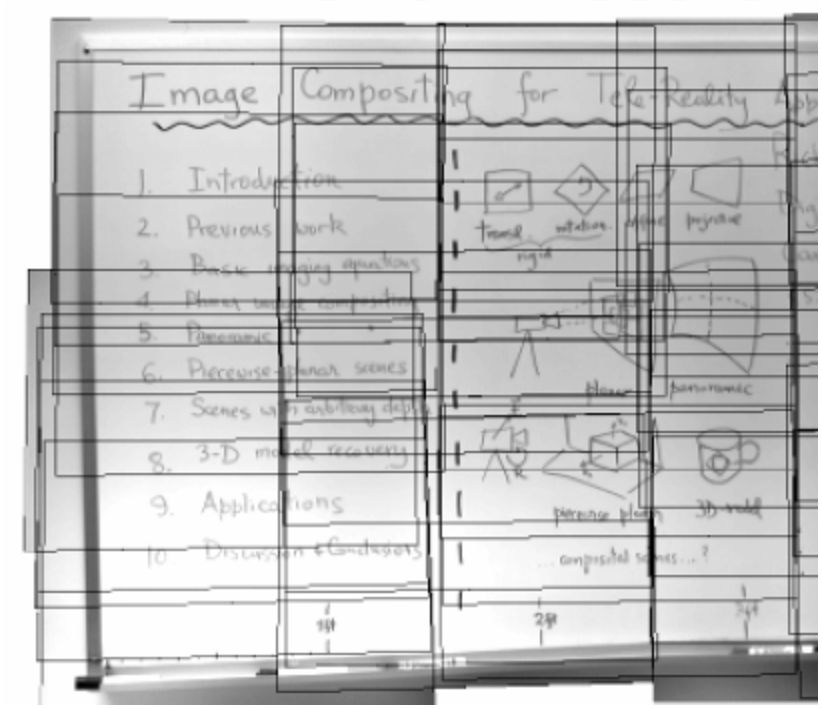


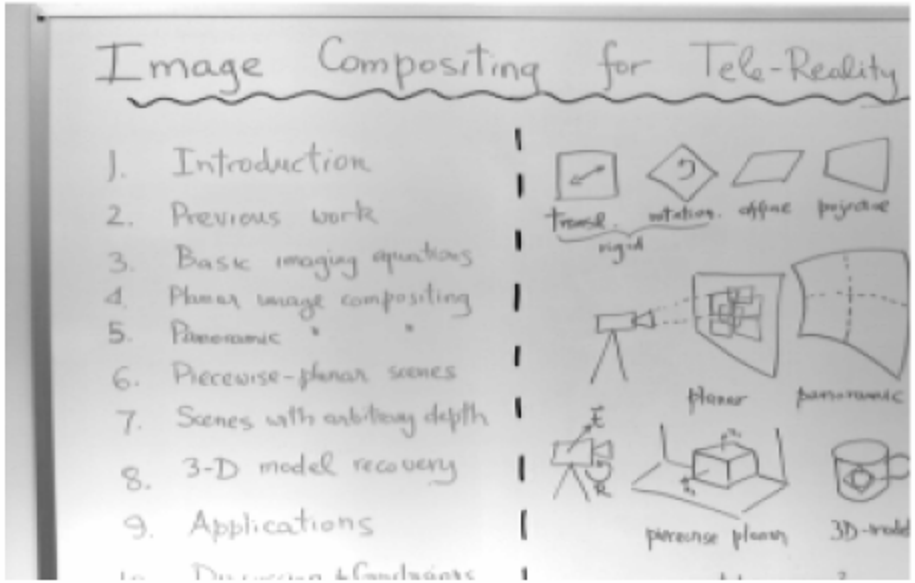
Planar scene (or far away)



- PP3 is a projection plane of both centers of projection, so we are OK!
- This is how big aerial photographs are made

Planar mosaic






Map
Satellite
Hybrid
200 ft
100 m
©2007 Google - Terms of Use

Bells and Whistles

- Blending and Compositing
  - use homographies to combine images or video and images together in an interesting (fun) way. E.g.
    - put fake graffiti on buildings or chalk drawings on the ground
    - replace a road sign with your own poster
    - project a movie onto a building wall

Bells and Whistles

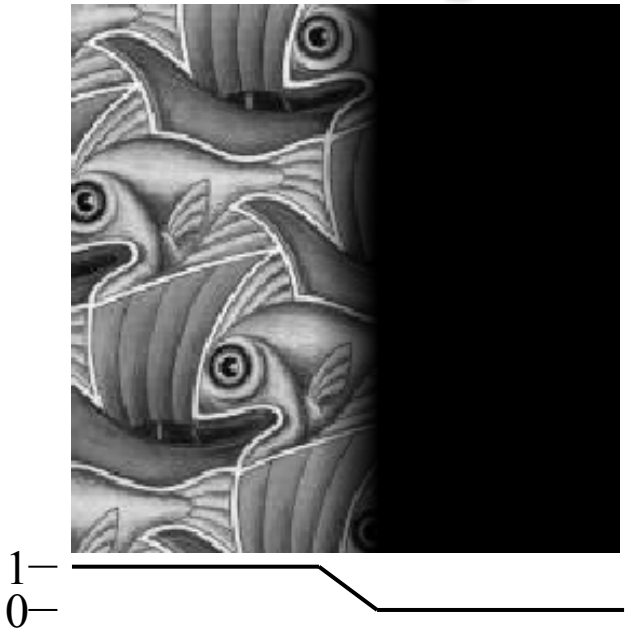
- Capture creative/cool/bizzare panoramas
  - Example from UW (by Brett Allen):

  - Ever wondered what is happening inside your fridge while you are not looking?

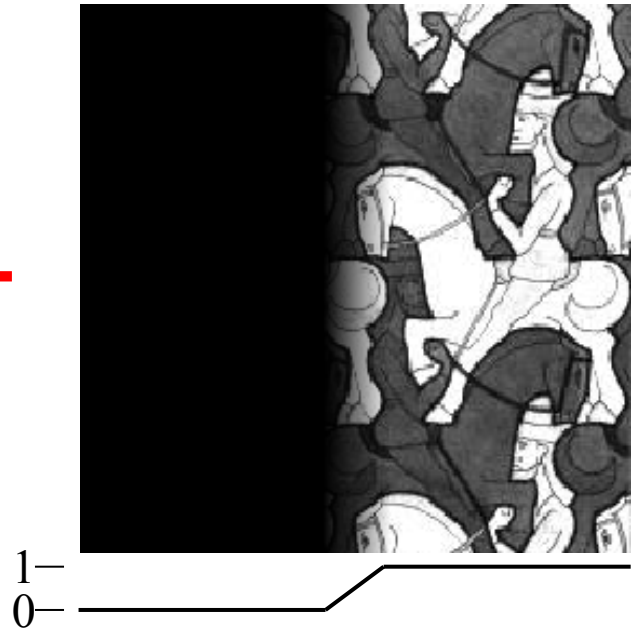
Blending the mosaic

An example of image compositing: the art (and sometime science) of combining images together…

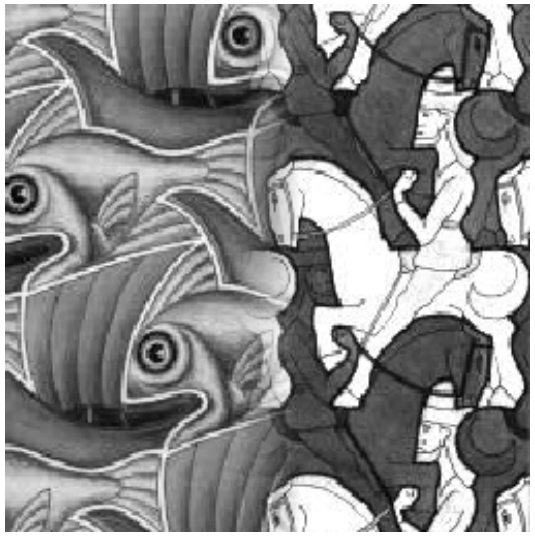
Feathering



+



=



Encoding transparency

$I(x,y) = (\alpha R, \alpha G, \alpha B, \alpha)$

$I_{blend} = I_{left} + I_{right}$

Limitations

- Lens distortion and vignetting
- Off-centered camera motion
- Moving objects
- Single perspective may not be enough!

- Current research

Panoramic Video Textures

Output Video

http://grail.cs.washington.edu/projects/panovidtex/

[Agarwala et al, 2005]

Multi-perspective Panoramas

Input Video

[Roman 2006]