Announcements

Project 1 due April 16 (tomorrow midnight)
Project 2 will be assigned on Wed.

Mosaics

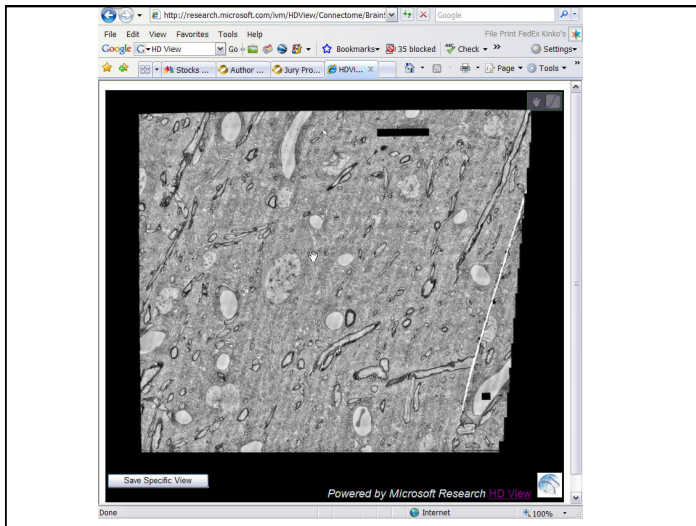
StreetView
Photosynth App

Today's Readings
- Chapters 5.1, 8.1 in Szeliski
- Szeliski and Shum paper (sections 1 and 2, skim the rest)

Image Mosaics

Goal
- Stitch together several images into a seamless composite

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
- Shift the second image to overlap with the first
- Blend the two together to create a mosaic
- If there are more images, repeat

Aligning images
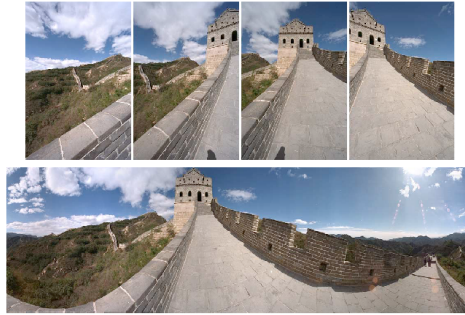
How to account for warping?

- Translations are not enough to align the images

Alignment Demo

Motion models

What happens when we take two images with a camera and try to align them?

- Translation
- Rotation
- Scale
- Perspective

8

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 PP

Image reprojection

Basic question
- How to relate two images from the same camera center?
  - how to map a pixel from PP1 to PP2

Answer
- Cast a ray through each pixel in PP1
- Draw the pixel where that ray intersects PP2

PP2

PP1

Don't need to know what's in the scene!

Image reprojection

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

Image Warping

image filtering: change *range* of image
$$g(x) = h(f(x))$$

f → h → g

x x

image warping: change *domain* of image
$$g(x) = f(h(x))$$

f → h → g

x x

Image Warping

image filtering: change *range* of image

$$g(x) = h(f(x))$$
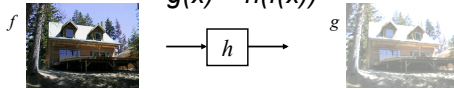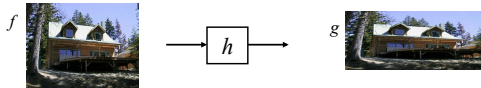

image warping: change *domain* of image

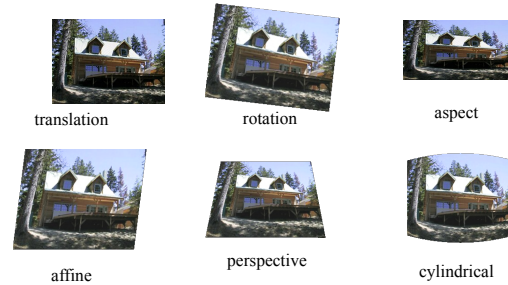$$g(x) = f(h(x))$$

13

Parametric (global) warping

Examples of parametric warps:

translation   rotation   aspect

affine   perspective   cylindrical

14

2D coordinate transformations

translation: $x' = x + t$ $x = (x,y)$

rotation: $x' = R\,x + t$

similarity: $x' = s\,R\,x + t$

affine: $x' = A\,x + t$

perspective: $\underline{x}' \cong H\,\underline{x}$ $\underline{x} = (x,y,1)$
    ($\underline{x}$ is a *homogeneous* coordinate)
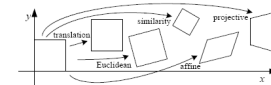
These all form a nested *group* (closed w/ inv.)

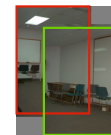Richard Szeliski   Image Stitching   15
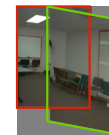
Motion models
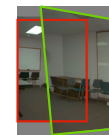

Translation   **Affine**   **Perspective**   **3D rotation**
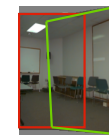
2 unknowns   **6 unknowns**   **8 unknowns**   **3 unknowns**

Richard Szeliski   Image Stitching   16

4

Homographies

Perspective projection of a plane
- Lots of names for this:
  - homography, texture-map, colineation, planar projective map
- Modeled as a 2D warp using homogeneous coordinates

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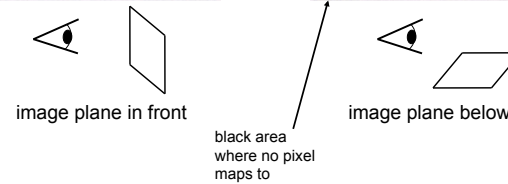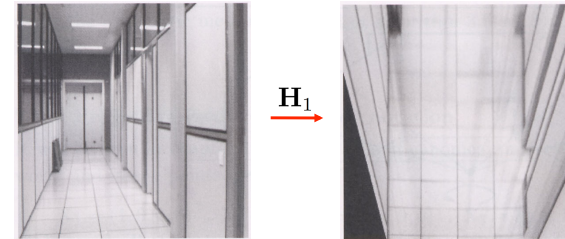
p'   H   p

To apply a homography H
- Compute   p' = Hp   (regular matrix multiply)
- Convert p' from homogeneous to image coordinates
  - divide by w (third) coordinate

Image warping with homographies

$\mathbf{H}_1$

image plane in front

image plane below

black area
where no pixel
maps to

Homographies

Idea: projecting images onto a common plane

each image is warped
with a homography $\mathbf{H}$

We'll see what this homograph means later.

First -- Can't create a 360 panorama this way...

mosaic PP
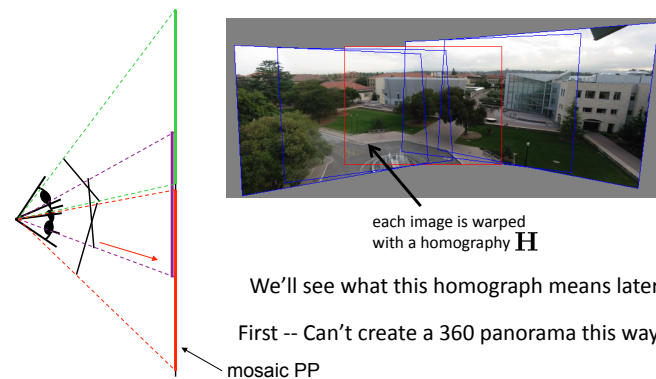
Panoramas

What if you want a 360° field of view?

mosaic Projection Sphere

Panoramas

mosaic Projection Sphere
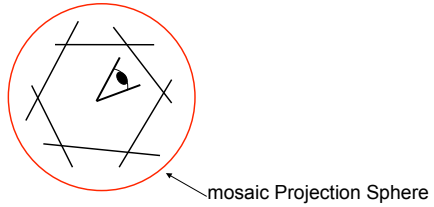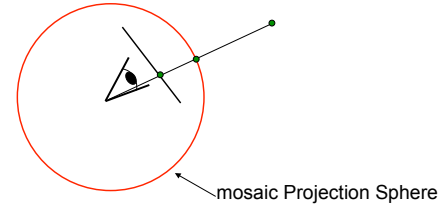
Spherical projection systems

Omnimax

CAVE (UI Chicago)

Spherical projection

(X, Y, Z)

$(\hat{x}, \hat{y}, \hat{z})$

Y

X

Z

unit sphere

ϕ

$(\tilde{x}_c, \tilde{y}_c)$ θ

unwrapped sphere

$\tilde{y}$
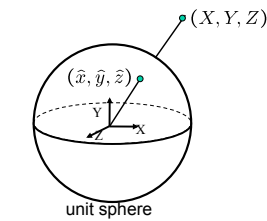
$\tilde{x}$ Spherical image

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

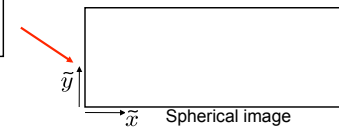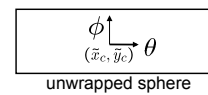$$(\hat{x}, \hat{y}, \hat{z}) = \frac{1}{\sqrt{X^2+Y^2+Z^2}}(X, Y, Z)$$

- 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

$$(\tilde{x}, \tilde{y}) = (s\theta, s\phi) + (\tilde{x}_c, \tilde{y}_c)$$

  – s defines size of the final image
    » often convenient to set s = camera focal length

Spherical reprojection

side view

top-down view

How to map sphere onto a flat image?

- $(\hat{x}, \hat{y}, \hat{z})$ to (x', y')

Spherical reprojection

side view

top-down view

How to map sphere onto a flat image?

- $(\hat{x}, \hat{y}, \hat{z})$ to (x', y')
- Use image projection matrix!
- or use the version of projection that properly accounts for radial distortion, as discussed in projection slides. This is what you'll do for project 2.

Radial distortion

Correct for "bending" in wide field of view lenses

Pro$(\hat{x}, \hat{y}, \hat{z})$ to "normalized" image coordinates

$$x'_n = \hat{x}/\hat{z}$$
$$y'_n = \hat{y}/\hat{z}$$

Apply radial distortion

$$r^2 = {x'_n}^2 + {y'_n}^2$$
$$x'_d = x'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$
$$y'_d = y'_n(1 + \kappa_1 r^2 + \kappa_2 r^4)$$

Apply focal length translate image center

$$x' = f x'_d + x_c$$
$$y' = f y'_d + y_c$$

To model lens distortion

- Use above projection operation instead of standard projection matrix multiplication

Spherical reprojection

| input | f = 200 (pixels) | f = 400 | f = 800 |

Map image to spherical coordinates
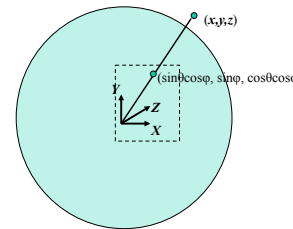
- need to know the focal length

Aligning spherical images

Suppose we rotate the camera by θ about the vertical axis
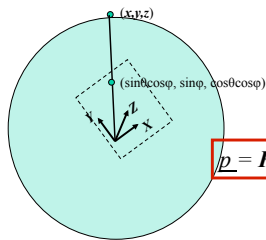- How does this change the spherical image?

Spherical warping

Given focal length *f* and
image center (x_c,y_c)

$$\theta = (x_{cyl} - x_c)/f$$
$$\varphi = (y_{cyl} - y_c)/f$$
$$\widehat{x} = \sin\theta\,\cos\varphi$$
$$\widehat{y} = \sin\varphi$$
$$\widehat{z} = \cos\theta\,\cos\varphi$$
$$x = f\widehat{x}/\widehat{z} + x_c$$
$$y = f\widehat{y}/\widehat{z} + y_c$$

3D rotation

Rotate image before
placing on unrolled sphere

$$\theta = (x_{cyl} - x_c)/f$$
$$\varphi = (y_{cyl} - y_c)/f$$
$$\widehat{x} = \sin\theta\,\cos\varphi$$
$$\widehat{y} = \sin\varphi$$
$$\widehat{z} = \cos\theta\,\cos\varphi$$
$$\underline{p} = R\,p$$
$$x = f\underline{\widehat{x}}/\underline{\widehat{z}} + x_c$$
$$y = f\underline{\widehat{y}}/\underline{\widehat{z}} + y_c$$

Aligning spherical images

Suppose we rotate the camera by θ about the vertical axis
- How does this change the spherical image?
- Translation by θ
- This means that we can align spherical images by translation
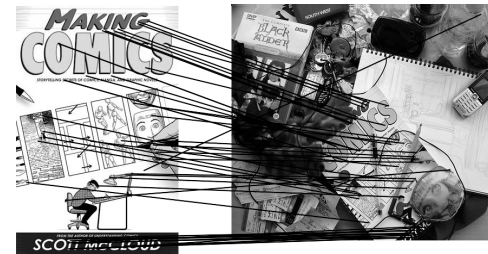
Spherical image stitching

What if you don't know the camera rotation?
- Solve for the camera rotations
  - Note that a pan (rotation) of the camera is a translation of the sphere!
  - **Use feature matching to solve for translations of spherical-warped images**
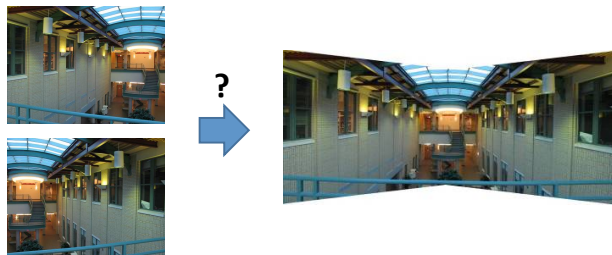
Computing transformations

- Given a set of matches between images A and B
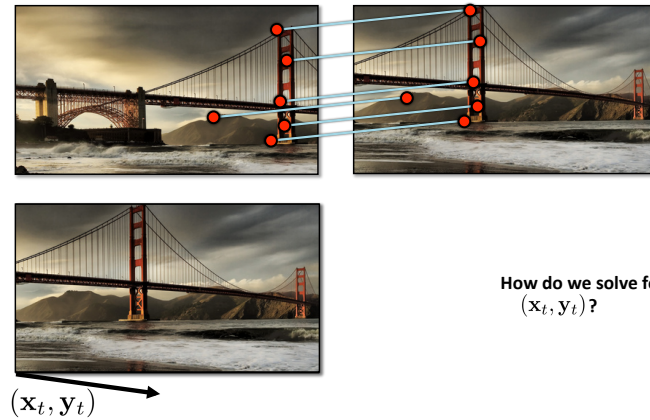 - How can we compute the transform T from A to B?

  - Find transform T that best "agrees" with the matches
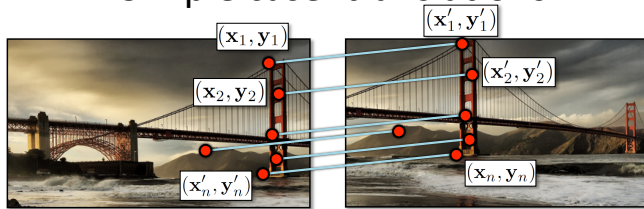
Computing transformations

?
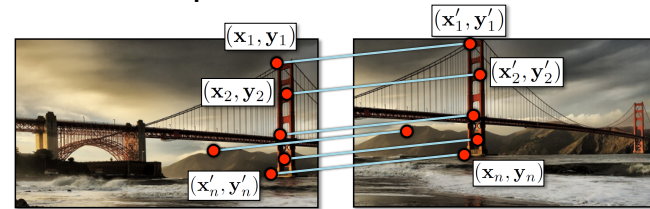
Simple case: translations

How do we solve for $(\mathbf{x}_t, \mathbf{y}_t)$?

$(\mathbf{x}_t, \mathbf{y}_t)$

Simple case: translations

Displacement of match i = $\left(\mathbf{x}'_i - \mathbf{x}_i, \mathbf{y}'_i - \mathbf{y}_i\right)$

$$\left(\mathbf{x}_t, \mathbf{y}_t\right) = \left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}'_i - \mathbf{x}_i, \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}'_i - \mathbf{y}_i\right)$$
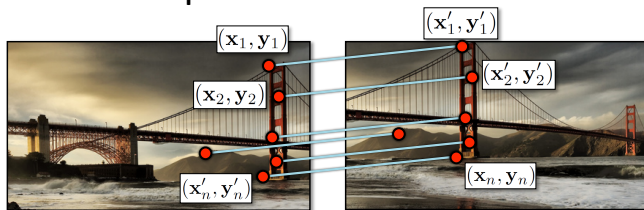
Simple case: translations

$$\mathbf{x}_i + \mathbf{x_t} = \mathbf{x}'_i$$
$$\mathbf{y}_i + \mathbf{y_t} = \mathbf{y}'_i$$

- System of linear equations
  - What are the knowns? Unknowns?
  - How many unknowns? How many equations (per match)?

Simple case: translations

$$\mathbf{x}_i + \mathbf{x_t} = \mathbf{x}'_i$$
$$\mathbf{y}_i + \mathbf{y_t} = \mathbf{y}'_i$$

- Problem: more equations than unknowns
  - "Overdetermined" system of equations
  - We will find the *least squares* solution

Least squares formulation

- For each point $\left(\mathbf{x}_i, \mathbf{y}_i\right)$

$$\mathbf{x}_i + \mathbf{x_t} = \mathbf{x}'_i$$
$$\mathbf{y}_i + \mathbf{y_t} = \mathbf{y}'_i$$

- we define the *residuals* as

$$r_{\mathbf{x}_i}(\mathbf{x}_t) = \left(\mathbf{x}_i + \mathbf{x}_t\right) - \mathbf{x}'_i$$
$$r_{\mathbf{y}_i}(\mathbf{y}_t) = \left(\mathbf{y}_i + \mathbf{y}_t\right) - \mathbf{y}'_i$$

Least squares formulation

- Goal: minimize sum of squared residuals

$$C(\mathbf{x}_t, \mathbf{y}_t) = \sum_{i=1}^{n} \left(r_{\mathbf{x}_i}(\mathbf{x}_t)^2 + r_{\mathbf{y}_i}(\mathbf{y}_t)^2 \right)$$

- "Least squares" solution
- For translations, is equal to mean displacement

Least squares formulation

- Can also write as a matrix equation

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ \vdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} x'_1 - x_1 \\ y'_1 - y_1 \\ x'_2 - x_2 \\ y'_2 - y_2 \\ \vdots \\ x'_n - x_n \\ y'_n - y_n \end{bmatrix}$$

$$\underset{2n \times 2}{\mathbf{A}} \quad \underset{2 \times 1}{\mathbf{t}} = \underset{2n \times 1}{\mathbf{b}}$$

Least squares

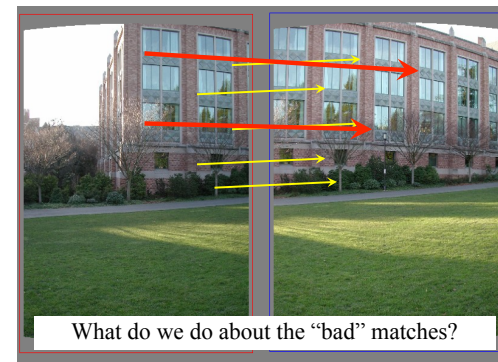$$\mathbf{A}\mathbf{t} = \mathbf{b}$$

- Find **t** that minimizes

$$||\mathbf{A}\mathbf{t} - \mathbf{b}||^2$$
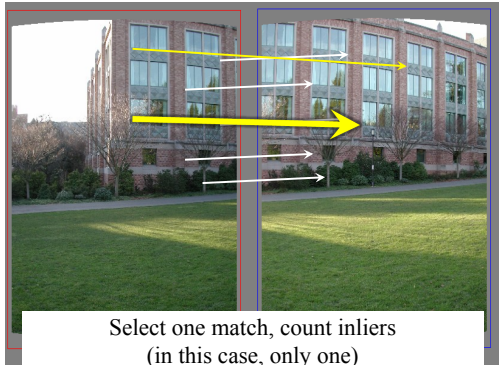
- To solve, form the *normal equations*

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{t} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\mathbf{t} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

But not all matches are good

What do we do about the "bad" matches?

4/14/13

RAndom SAmple Consensus

Select one match, count inliers
(in this case, only one)

RAndom SAmple Consensus

Select one match, count inliers
(4 inliers)

Least squares fit

Find "average" translation vector
for largest set of inliers
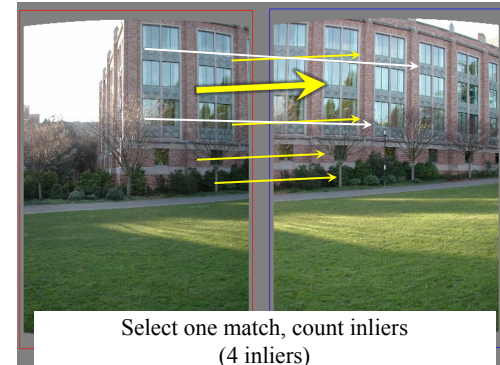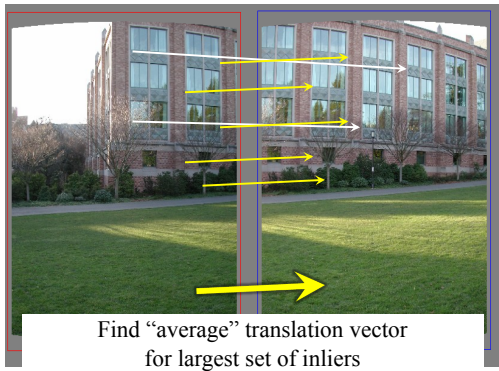
RANSAC

- Idea:
  - All the inliers will agree with each other on the translation vector; the (hopefully small) number of outliers will (hopefully) disagree with each other
    - RANSAC only has guarantees if there are < 50% outliers

12

RANSAC

Same basic approach works for any transformation
- Translation, rotation, homographies, etc.
- Very useful tool

General version
- Randomly choose a set of K correspondences
  - Typically K is the minimum size that lets you fit a model
- Fit a model (e.g., homography) to those correspondences
- Count the number of inliers that "approximately" fit the model
  - Need a threshold on the error
- Repeat as many times as you can
- Choose the model that has the largest set of inliers
- Refine the model by doing a least squares fit using ALL of the inliers

How many rounds?

- If we have to choose s samples each time
  - with an outlier ratio e
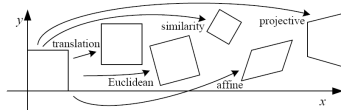 - and we want the right answer with probability p

	proportion of outliers e						
s	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177

$p = 0.99$
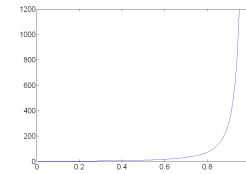
Source: M. Pollefeys

How big is s?

- For alignment, depends on the motion model
 - Here, each sample is a correspondence (pair of matching points)

Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\begin{bmatrix} I & t \end{bmatrix}_{2\times3}$	2	orientation $+\cdots$	□
rigid (Euclidean)	$\begin{bmatrix} R & t \end{bmatrix}_{2\times3}$	3	lengths $+\cdots$	◇
similarity	$\begin{bmatrix} sR & t \end{bmatrix}_{2\times3}$	4	angles $+\cdots$	◇
affine	$\begin{bmatrix} A \end{bmatrix}_{2\times3}$	6	parallelism $+\cdots$	▱
projective	$\begin{bmatrix} \tilde{H} \end{bmatrix}_{3\times3}$	8	straight lines	▱

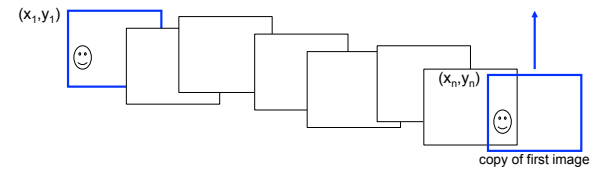## Assembling the panorama

Stitch pairs together, blend, then crop

Problem: Drift
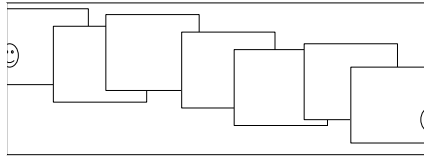
Error accumulation
- small errors accumulate over time

## Problem: Drift

(x_1,y_1)

(x_n,y_n)

copy of first image

Solution
- add another copy of first image at the end
- this gives a constraint: $y_n = y_1$
- there are a bunch of ways to solve this problem
  - add displacement of $(y_1 - y_n)/(n - 1)$ to each image after the first
  - compute a global warp: $y' = y + ax$
  - run a big optimization problem, incorporating this constraint
    - best solution, but more complicated
    - known as "bundle adjustment"

Full-view Panorama


Different projections are possible
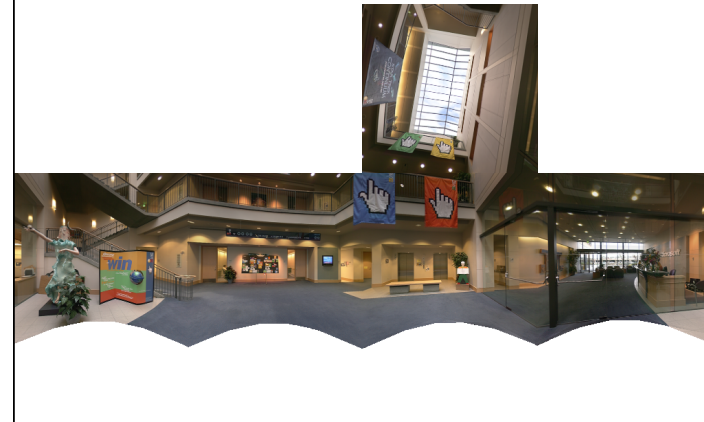
Project 2

Take pictures on a tripod (or handheld)

Warp to spherical coordinates

Extract features

Align neighboring pairs using RANSAC

Write out list of neighboring translations

Correct for drift

Read in warped images and blend them

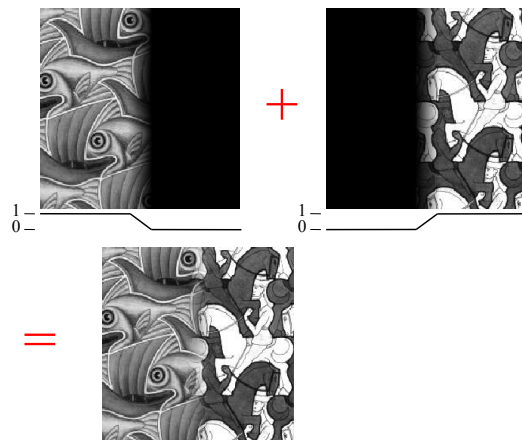Crop the result and import into a viewer

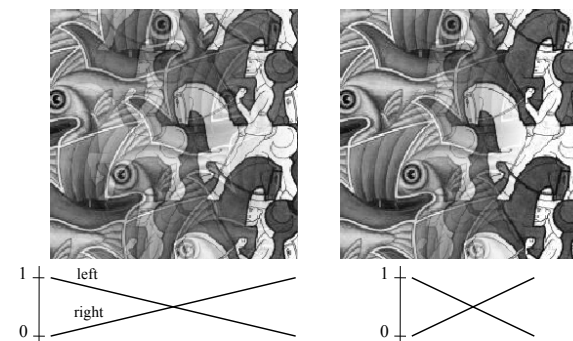Roughly based on Autostitch
- By Matthew Brown and David Lowe
- http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html
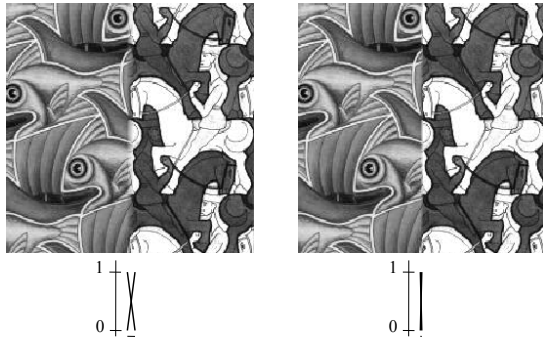
Image Blending
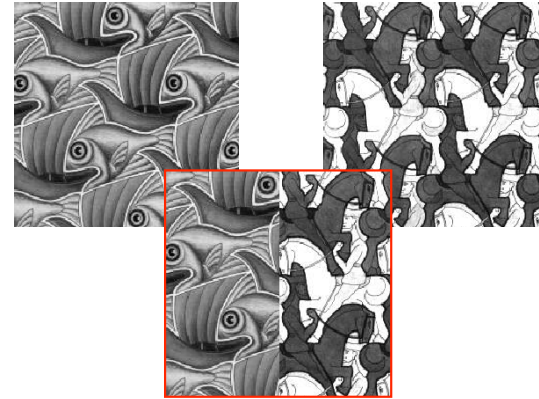
Feathering
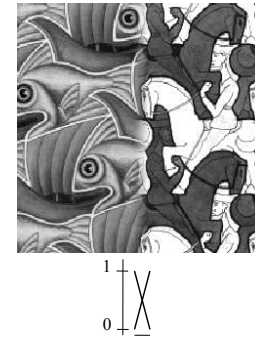

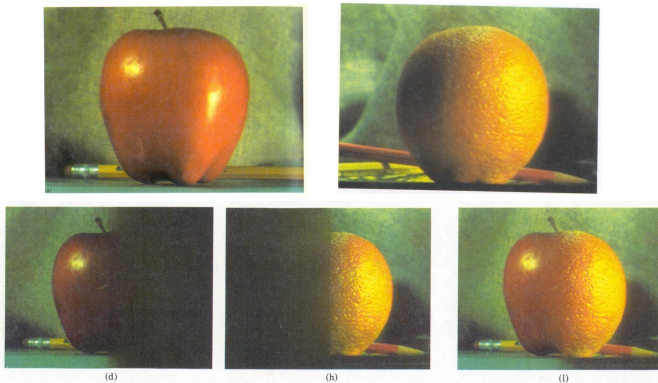## Effect of window (ramp-width) size

Effect of window size

Good window size

"Optimal" window: smooth but not ghosted
• Doesn't always work...

Pyramid blending

Create a Laplacian pyramid, blend each level
• Burt, P. J. and Adelson, E. H., A multiresolution spline with applications to image mosaics, ACM Transactions on Graphics, 42(4), October 1983, 217-236.
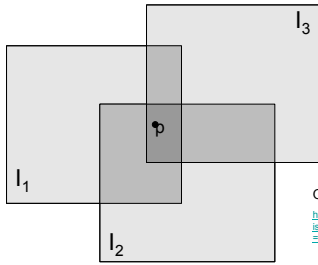
Poisson Image Editing

sources/destinations   cloning   seamless cloning

For more info: Perez et al, SIGGRAPH 2003
• http://research.microsoft.com/vision/cambridge/papers/perez_siggraph03.pdf

Alpha Blending

Optional: see Blinn (CGA, 1994) for details:
http://ieeexplore.ieee.org/iel1/38/7531/00310740.pdf?
isNumber=7531&prod=JNL&arnumber=310740&arSt=83&ared
=87&arAuthor=Blinn%2C+J.F.

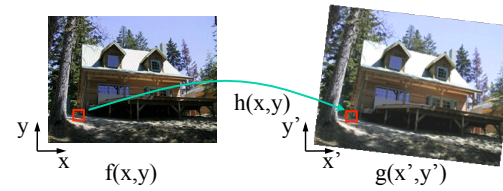Encoding blend weights: $I(x,y) = (\alpha R, \alpha G, \alpha B, \alpha)$

color at p = $\dfrac{(\alpha_1 R_1,\ \alpha_1 G_1,\ \alpha_1 B_1) + (\alpha_2 R_2,\ \alpha_2 G_2,\ \alpha_2 B_2) + (\alpha_3 R_3,\ \alpha_3 G_3,\ \alpha_3 B_3)}{\alpha_1 + \alpha_2 + \alpha_3}$

Implement this in two steps:

1. accumulate: add up the (α premultiplied) RGBα values at each pixel
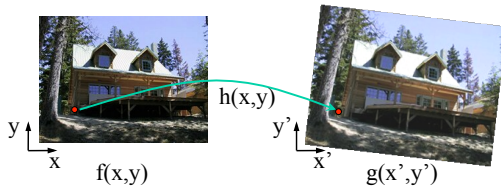2. normalize: divide each pixel's accumulated RGB by its α value

Q: what if $\alpha = 0$?

Image warping

h(x,y)

f(x,y)   g(x',y')

Given a coordinate transform $(x',y') = h(x,y)$ and a source image $f(x,y)$, how do we compute a transformed image $g(x',y') = f(h(x,y))$?

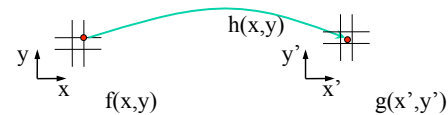## Forward warping

h(x,y)

f(x,y)   g(x',y')

Send each pixel $f(x,y)$ to its corresponding location

$(x',y') = h(x,y)$ in the second image

Q: what if pixel lands "between" two pixels?
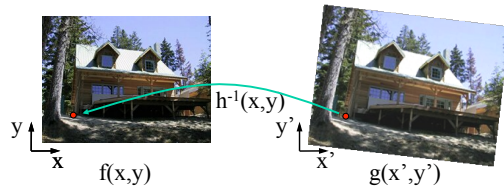
Forward warping

h(x,y)

f(x,y)   g(x',y')

Send each pixel $f(x,y)$ to its corresponding location

$(x',y') = h(x,y)$ in the second image

Q: what if pixel lands "between" two pixels?

A: distribute color among neighboring pixels (x',y')
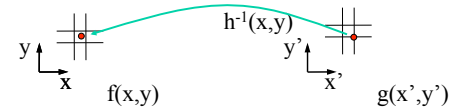
– Known as "splatting"

Inverse warping

h^{-1}(x,y)

y
x
f(x,y)

y'
x'
g(x',y')

Get each pixel g(x',y') from its corresponding location
(x,y) = h^{-1}(x',y') in the first image

Q: what if pixel comes from "between" two pixels?

Inverse warping

h^{-1}(x,y)

y
x
f(x,y)

y'
x'
g(x',y')

Get each pixel g(x',y') from its corresponding location
(x,y) = h^{-1}(x',y') in the first image

Q: what if pixel comes from "between" two pixels?

A: resample color value

– We discussed resampling techniques before

• nearest neighbor, bilinear, Gaussian, bicubic

Forward vs. inverse warping

Q: which is better?

A: usually inverse—eliminates holes
  • however, it requires an invertible warp function—not always possible...

Other types of mosaics

Can mosaic onto any surface if you know the geometry
  • See NASA's Visible Earth project for some stunning earth mosaics
    – http://earthobservatory.nasa.gov/Newsroom/BlueMarble/
    – Click for images…