CS664 Lecture #16: Image registration, robust statistics, motion

Some material taken from:

- **Alyosha Efros, CMU**
  http://www.cs.cmu.edu/~efros

- **Xenios Papademetris**
  http://noodle.med.yale.edu/~papad/various/papademetris_Image_Registration.ppt

- **Rick Szeliski, Microsoft Research**
  http://research.microsoft.com/~szeliski/

- **Steve Seitz, University of Washington**
  http://www.cs.washington.edu/homes/seitz/

Announcements

- Next quiz: 10/25 (Tuesday)
  - Coverage through last lecture
- 1-page paper report due on 11/15
- PS2 out later today, due on Tuesday 11/8
  - Experiment with energy minimization

Next topic: Image registration

- Origins in aerial surveillance
  - Now available with any digital camera

+

+ … +

=

Another example

Registration algorithms

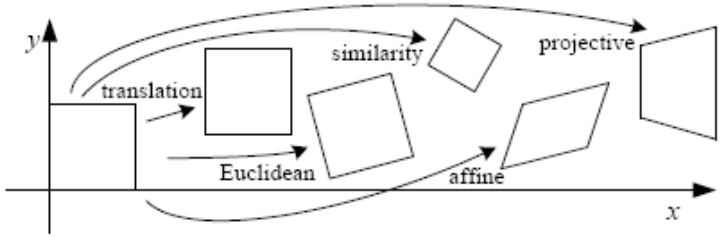
- Consider many different translations τ of image I_2
  - Compare $\tau(I_2)$ with I_1

$$\hat{\tau} = \arg \min_{\tau \in \text{Translations}} \text{dist}(I_1, \tau(I_2))$$

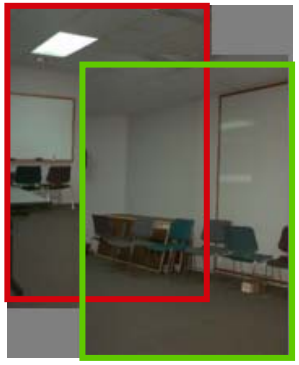
$$\text{dist}(I_1, I_2) = \sum_p \left(I_1(p) - I_2(p) \right)^2$$

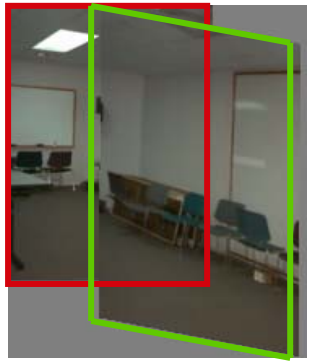
- Issues:
  - Are translations enough?
  - Right image comparison function (dist)?

Beyond translations



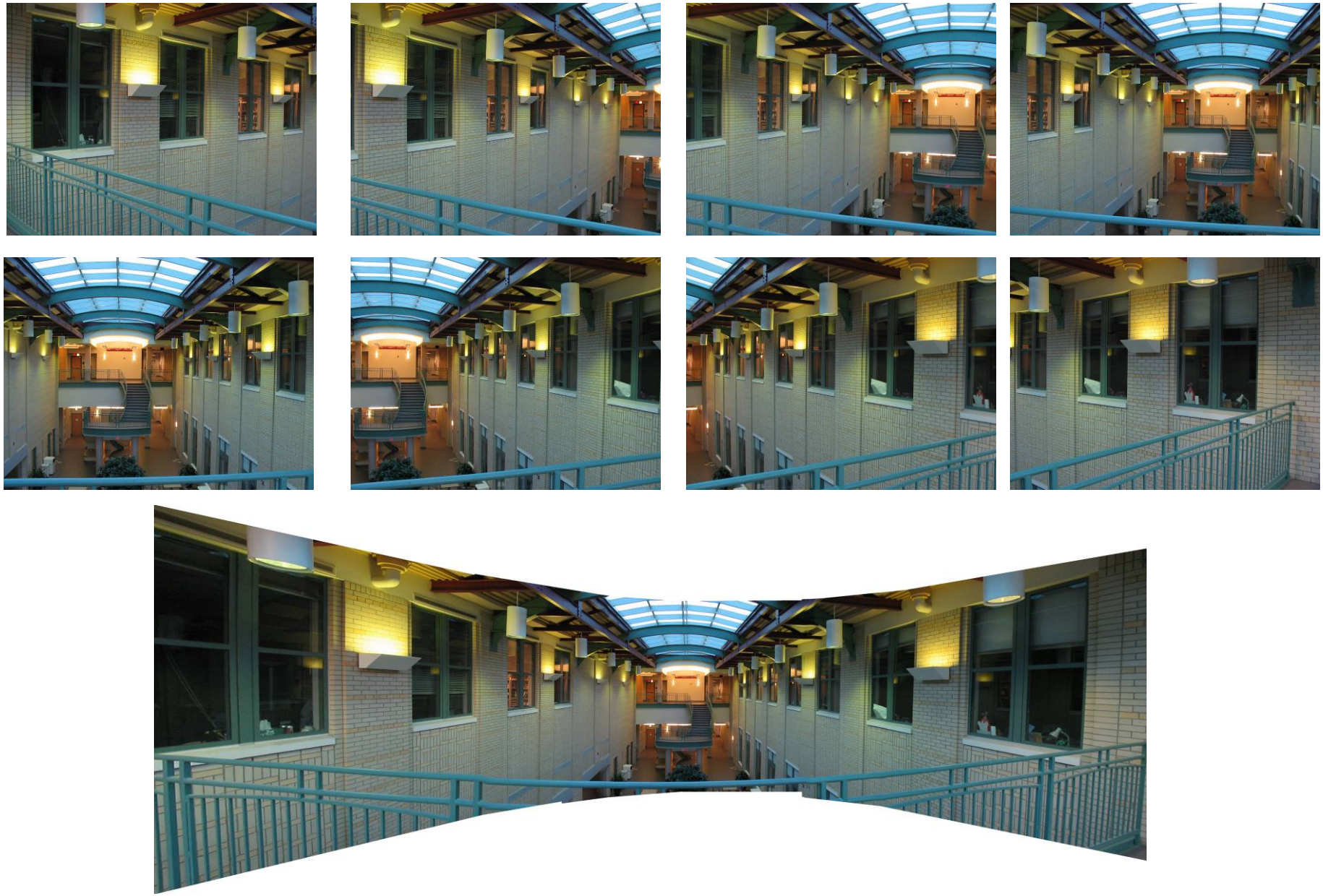
Translation **Affine** **Perspective**





2 unknowns **6 unknowns** **8 unknowns**

Affine registration in action



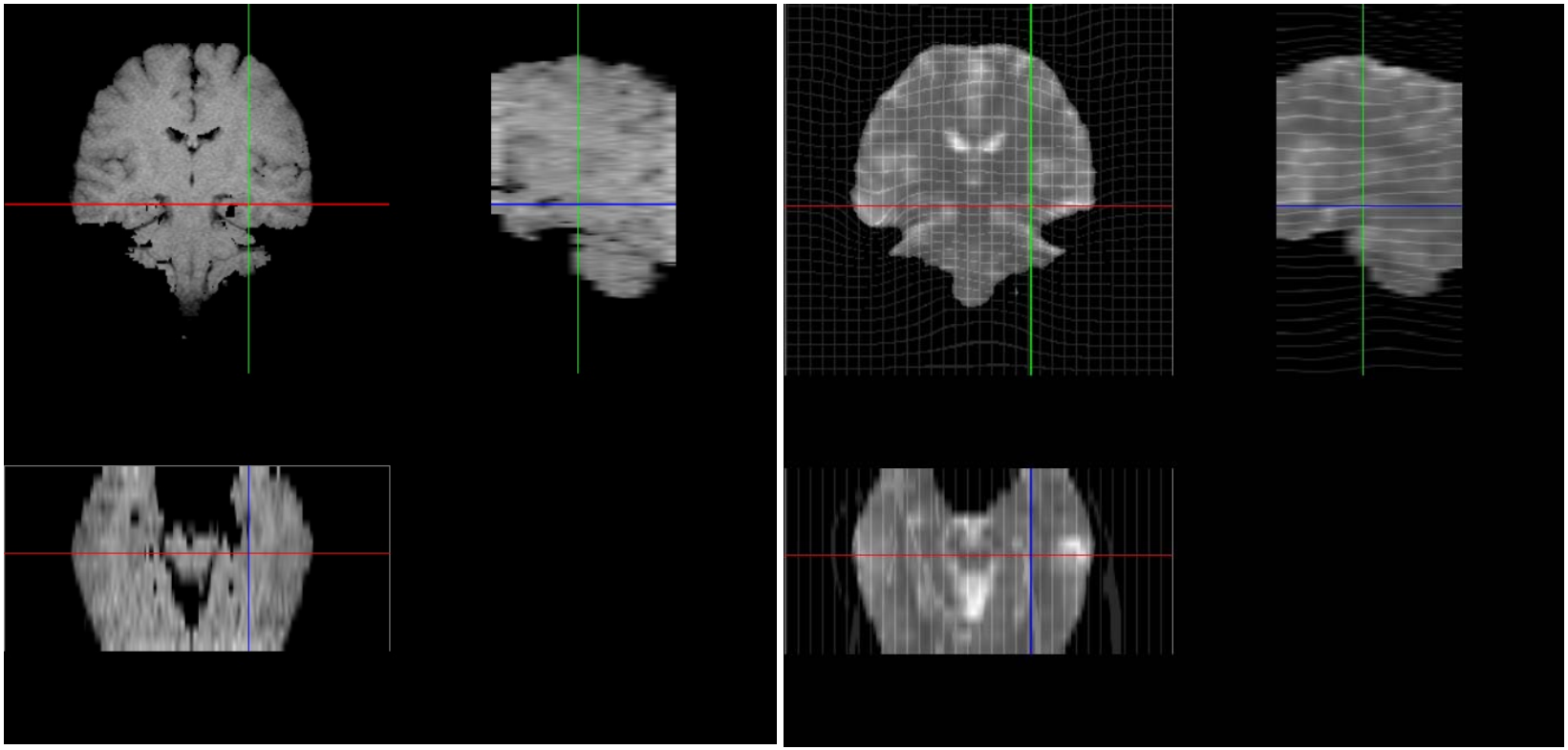
Algorithms to compare images

- We need a way to compare two images
  - A common problem in vision (and beyond)
  - Depends on what changes you expect to see between “similar” images
- Alternative: color histograms
  - Compute vector for each image, use L_2 distance on this vector
  - No spatial information
    - There are papers on how to fix this
  - Widely used in image databases

What if the cameras differ?

- Suppose one camera is twice as bright as the other, or some such
  - We know how to do this already
- Mutual information is widely used in medical imaging for registration
  - Not quite the right thing, for various reasons
    - Allows arbitrary intensity mapping
    - Maximized by gradient descent
  - Usually combined with a complicated non-rigid registration scheme (think: splines)

Non-rigid registration example



Local matching

- Image restration algorithms can also be used to compute dense stereo or motion
  - Locally at each pixel, take a square window and register it w.r.t. the other image
  - If we assume only translations, this is fast
    - Especially in stereo, where we can often assume purely horizontal translations
  - Natural generalization to handle geometric distortions (rotations, foreshortening, etc.)
    - Usually we consider affine transformations
- Used for many real-time systems

What else can go wrong?

- Some things that are present in one image will not be present in the other!
  - Obviously happens where the registered images do not overlap
  - Easier to think about inside the image
    - Imagine a person who is present in one image but not in the other, or who just moves a bit
  - What does this do to the L_2 distance? How can we fix it?

General model fitting problem

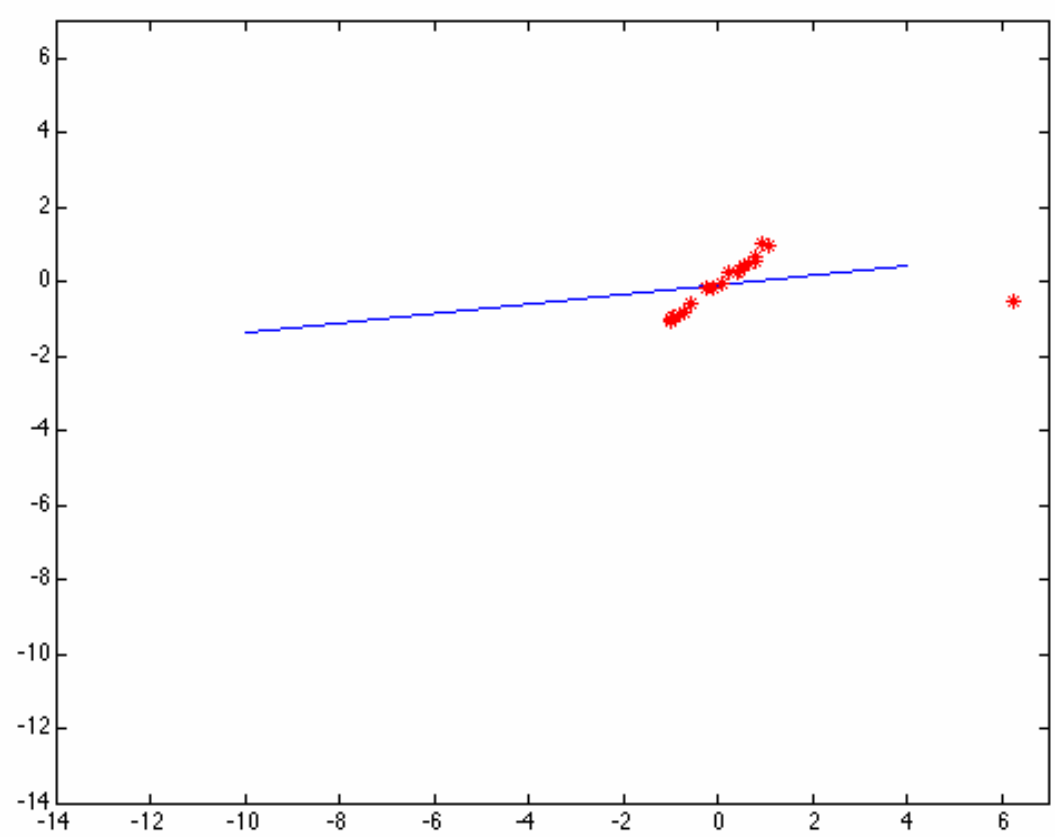
- We have some data points (x_i, y_i) and some possible models, each of which has some parameters θ
  - Example: line fitting, y = mx + b
  - A model predicts $M(x;\theta)$
- What set of parameters gives the best fit to the data?
- For a particular θ, the residuals are

$$r_i = y_i - M(x_i;\theta)$$

Least squares fit

- The least squares fit says that the best fit minimizes $\sum_i r_i^2$
  - Sum of the squared residuals
- At the correct alignment of the two images, what are the residuals?
  - If the only change between images is camera noise, they are generally small
    - Think about the joint histogram
  - What if a person moves?
- Same issue arises in line fitting!

1 bad point can ruin your whole line



Example c/o Kim Boyer, OSU

Problem is subtle

- You can't simply do an LS fit and then declare the worst-fitting point to be "bad"
  - There are examples where the bad data is fit better than the good data
- Robust statistics addresses this problem
  - Data is "outliers" and "inliers"
  - A robust fitting method tolerates outliers
    - Obviously, LS is not robust
  - Note that in vision, the term "robust" sometimes simply means "good"

Robust model fitting

- There are two problems with the LS fit
  - We square the residuals
  - We sum up these squares
- The main approaches in robust statistics address each of these problems
  - The problem with squaring the residuals is that the squared values get too large
    - I.e., we will never allow intensity 0 to turn into intensity 255; instead, we'll prefer a translation where every pixel is off by by 25 (50x50 image)
  - We saw this problem in lecture 7 or so

M-estimation

- Suppose that our measure of goodness of fit is $\sum_i \rho(r_i)$, where
$$\rho(r_i) = \min(r_i^2, s^2)$$
  - Here, s is a scale parameter
  - All residuals worse than s count like s
- The scale parameter essentially controls the boundary between inliers and outliers
  - We expect outliers to have residuals larger than s, but not inliers
  - How do we pick s?