



## Vision & Localization "Seeing Where You Are"

METR 4202: Advanced Control & Robotics

Drs Surya Singh, Paul Pounds, and Hanna Kurniawati

Lecture # 8

September 10, 2012

(with Vision Material from [Richard Szeliski](#), *Computer Vision: Algorithms and Applications*)

[metr4202@itee.uq.edu.au](mailto:metr4202@itee.uq.edu.au)

<http://itee.uq.edu.au/~metr4202/>



© 2012 School of Information Technology and Electrical Engineering at the University of Queensland



### NEW, NEW Schedule

| Week | Date   | Lecture (M: 12-1:30, 43-102)   |
|------|--------|--|
| 1    | 23-Jul | Introduction   |
| 2    | 30-Jul | Representing Position & Orientation & State (Frames, Transformation Matrices & Affine Transformations) |
| 3    | 6-Aug  | Robot Kinematics and Dynamics  |
| 4    | 13-Aug | Robot Dynamics & Control   |
| 5    | 20-Aug | Obstacle Avoidance & Motion Planning   |
| 6    | 27-Aug | Sensors, Measurement and Perception  |
| 7    | 3-Sep  | Perception (+ Prof. S. LaValle)  |
| 8    | 10-Sep | Computer Vision & Localization (SFM/SLAM)  |
| 9    | 17-Sep | State-Space Modelling (+ Prof. M. Srinivasan)  |
|      | 24-Sep | Study break  |
| 10   | 1-Oct  | Public Holiday   |
| 11   | 8-Oct  | Localization and Navigation  |
| 12   | 15-Oct | Control/Planning Under Uncertainty   |
| 13   | 22-Oct | Vision-based control (Prof. P. Corke) & Course Review  |



## Updates: Lab 2 → Free Extension to Oct 2

For Lab 2:

- Project description posted
- Uses a Kinect mostly as a “Camera”
- Submit code via **UQ eLearning Report** via Submission System
- Remember “extension” to **Tuesday, October 2 at 10am**
- Final Exam: 12/11/12 at 11:15 am



## Quick Outline

### 1. Perception II: Computer Vision → Localization

1. Feature extraction
2. Stereopsis and depth

### 2. Camera Calibration

### 3. Lab 2: Q & A

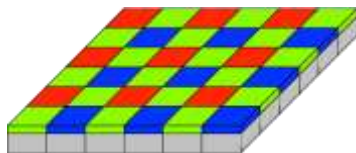


## Features

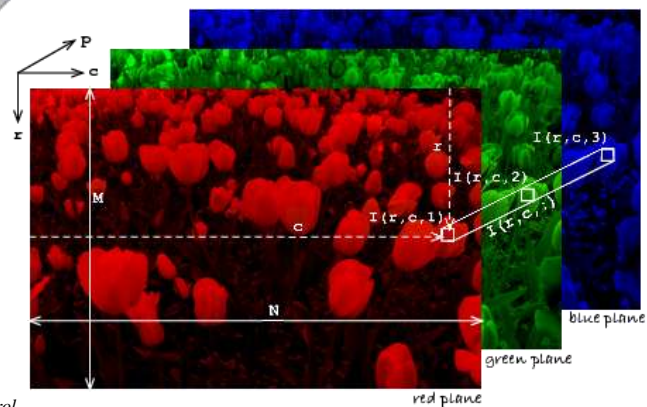
- Dense:
  - Colour
  - Edges
  - Disparity
- Sparse:
  - Corners
  - Lines
  - Statistics on Edges: SIFT, SURF



## Features -- Colour Features



Bayer Patterns



## Edge Detection

- Canny edge detector:
  - Pepsi Sequence:

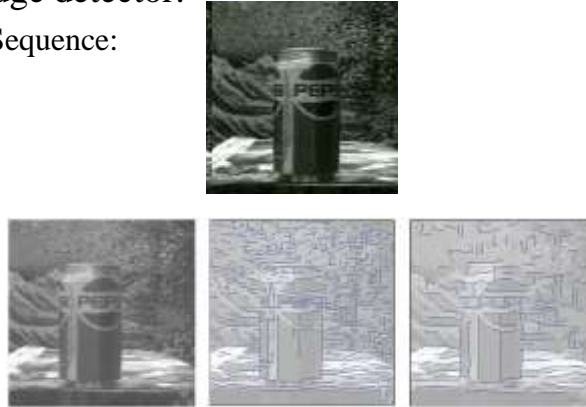


Image Data: <http://www.cs.brown.edu/~black/mixtureOF.html> and Szeliski, CS223B-L9

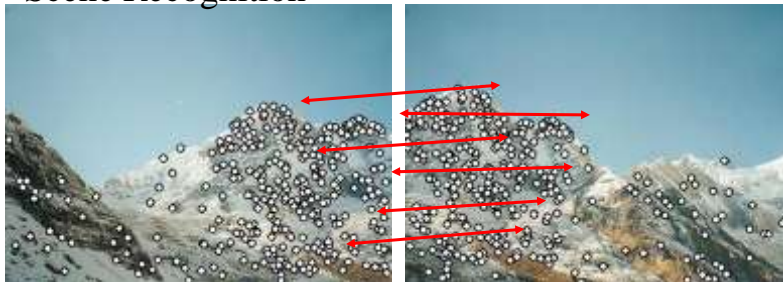
**See also:** Use of Temporal information to aid segmentation:

[http://www.cs.toronto.edu/~babalex/SpatiotemporalClosure/supplementary\\_material.html](http://www.cs.toronto.edu/~babalex/SpatiotemporalClosure/supplementary_material.html)



## Why extract features?

- **Object detection**
- Robot Navigation
- Scene Recognition



- Steps:
  - Extract Features
  - Match Features

Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



## Why extract features? [2]

- Panorama stitching...
  - Step 3: Align images  
(see: Hartley & Zisserman,  
*Multiple View Geometry*)



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



## Characteristics of good features

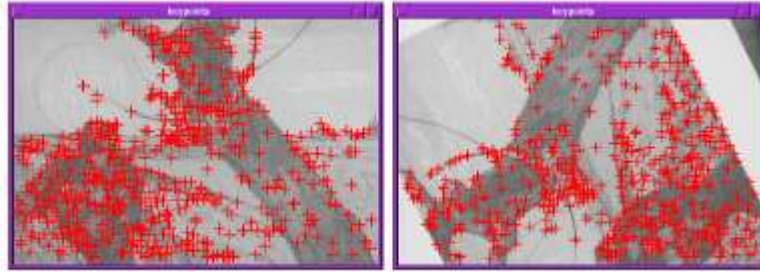
- Repeatability
  - The same feature can be found in several images despite geometric and photometric transformations
- Saliency
  - Each feature is distinctive
- Compactness and efficiency
  - Many fewer features than image pixels
- Locality
  - A feature occupies a relatively small area of the image; robust to clutter and occlusion



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



## Finding Corners



- Key property: in the region around a corner, image gradient has two or more dominant directions
- Corners are repeatable and distinctive

C.Harris and M.Stephens. "[A Combined Corner and Edge Detector.](#)" *Proceedings of the 4th Alvey Vision Conference*: pages 147—151, 1988.

Adopted from S. Lazebnik, Gang Hua ([CS 558](#))

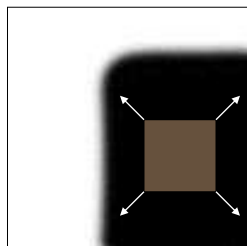


METR 4202: Robotics

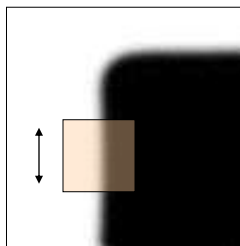
10 September 2012 - 11

## Corner Detection: Basic Idea

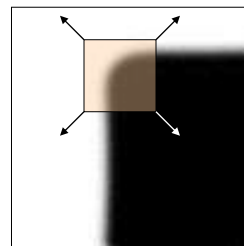
- Look through a window
- Shifting a window in any direction should give a large change in intensity



“flat” region:  
no change in  
all directions



“edge”:  
no change along  
the edge direction



“corner”:  
significant change  
in all directions

Source: A. Efros



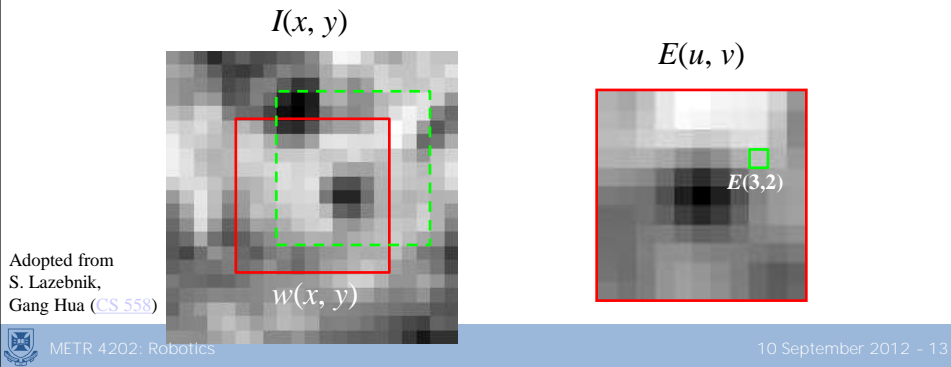
METR 4202: Robotics

10 September 2012 - 12

## Corner Detection: Mathematics

Change in appearance of window  $w(x,y)$   
for the shift  $[u,v]$ :

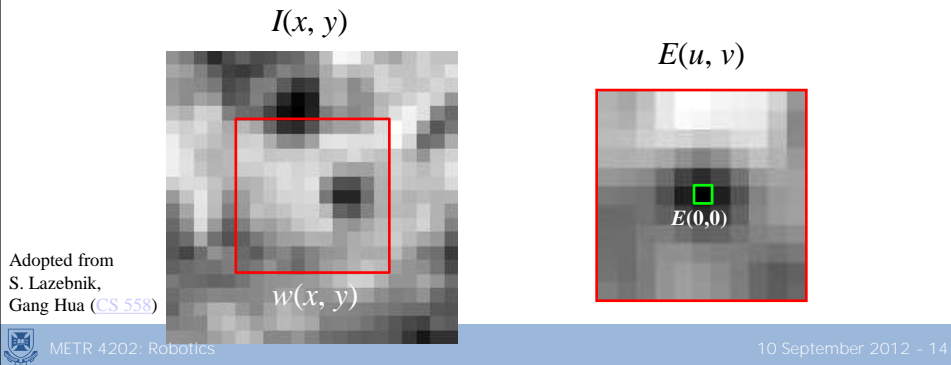
$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$



## Corner Detection: Mathematics

Change in appearance of window  $w(x,y)$   
for the shift  $[u,v]$ :

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$



## Corner Detection: Mathematics

Change in appearance of window  $w(x,y)$   
for the shift  $[u,v]$ :

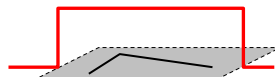
$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$

Window  
function

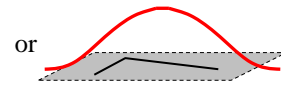
Shifted  
intensity

Intensity

Window function  $w(x,y) =$



1 in window, 0 outside



Gaussian

Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 16

## Corner Detection: Mathematics

Change in appearance of window  $w(x,y)$   
for the shift  $[u,v]$ :

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$

We want to find out how this function behaves for small shifts

$E(u, v)$



Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 16



## Corner Detection: Mathematics

Change in appearance of window  $w(x,y)$   
for the shift  $[u,v]$ :

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$

We want to find out how this function behaves for small shifts

$$E(u,v) \approx E(0,0) + [u \ v] \begin{bmatrix} E_u(0,0) \\ E_v(0,0) \end{bmatrix} + \frac{1}{2} [u \ v] \begin{bmatrix} E_{uu}(0,0) & E_{uv}(0,0) \\ E_{vu}(0,0) & E_{vv}(0,0) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Local quadratic approximation of  $E(u,v)$  in the neighborhood of  $(0,0)$  is given by the *second-order Taylor expansion*:

Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 17

## Corner Detection: Mathematics

$$E(u,v) = \sum_{x,y} w(x,y) [I(x+u, y+v) - I(x,y)]^2$$

Second-order Taylor expansion of  $E(u,v)$  about  $(0,0)$ :

$$E(u,v) \approx E(0,0) + [u \ v] \begin{bmatrix} E_u(0,0) \\ E_v(0,0) \end{bmatrix} + \frac{1}{2} [u \ v] \begin{bmatrix} E_{uu}(0,0) & E_{uv}(0,0) \\ E_{vu}(0,0) & E_{vv}(0,0) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$E_u(u,v) = \sum_{x,y} 2w(x,y) [I(x+u, y+v) - I(x,y)] I_x(x+u, y+v)$$

$$E_{uu}(u,v) = \sum_{x,y} 2w(x,y) I_x(x+u, y+v) I_x(x+u, y+v) + \sum_{x,y} 2w(x,y) [I(x+u, y+v) - I(x,y)] I_{xx}(x+u, y+v)$$

$$E_{uv}(u,v) = \sum_{x,y} 2w(x,y) I_y(x+u, y+v) I_x(x+u, y+v)$$

Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))

$$+ \sum_{x,y} 2w(x,y) [I(x+u, y+v) - I(x,y)] I_{xy}(x+u, y+v)$$



METR 4202: Robotics

10 September 2012 - 18

## Corner Detection: Mathematics

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

Second-order Taylor expansion of  $E(u, v)$  about  $(0, 0)$ :

$$E(u, v) \approx [u \ v] \begin{bmatrix} \sum_{x, y} w(x, y) I_x^2(x, y) & \sum_{x, y} w(x, y) I_x(x, y) I_y(x, y) \\ \sum_{x, y} w(x, y) I_x(x, y) I_y(x, y) & \sum_{x, y} w(x, y) I_y^2(x, y) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$E(0, 0) = 0$$

$$E_u(0, 0) = 0$$

$$E_v(0, 0) = 0$$

$$E_{uu}(0, 0) = \sum_{x, y} 2w(x, y) I_x(x, y) I_x(x, y)$$

$$E_{vv}(0, 0) = \sum_{x, y} 2w(x, y) I_y(x, y) I_y(x, y)$$

$$E_{uv}(0, 0) = \sum_{x, y} 2w(x, y) I_x(x, y) I_y(x, y)$$

Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))



## Harris detector: Steps

- Compute Gaussian derivatives at each pixel
- Compute second moment matrix  $M$  in a Gaussian window around each pixel
- Compute corner response function  $R$
- Threshold  $R$
- Find local maxima of response function (nonmaximum suppression)

C.Harris and M.Stephens. "[A Combined Corner and Edge Detector.](#)"  
*Proceedings of the 4th Alvey Vision Conference*: pages 147—151, 1988.

Adopted from  
S. Lazebnik,  
Gang Hua ([CS 558](#))



## Harris Detector: Steps



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))

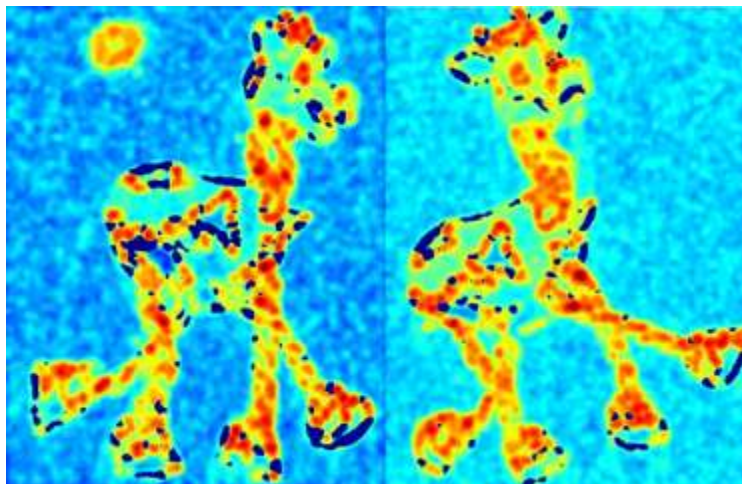


METR 4202: Robotics

10 September 2012 - 21

## Harris Detector: Steps

Compute corner response  $R$



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))

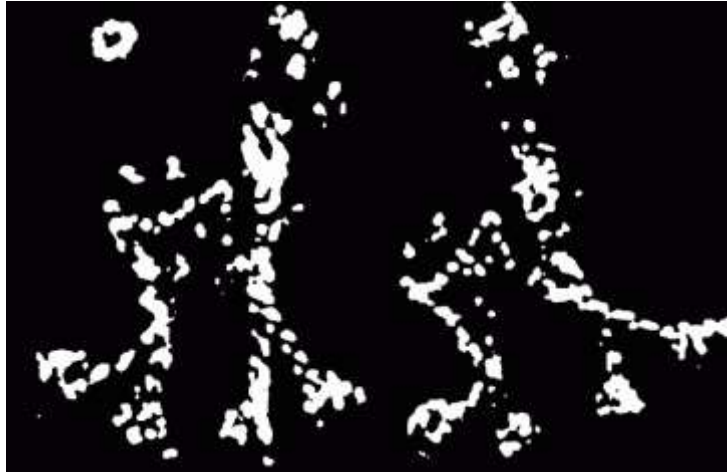


METR 4202: Robotics

10 September 2012 - 22

## Harris Detector: Steps

Find points with large corner response:  $R > \text{threshold}$



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 23

## Harris Detector: Steps

Take only the points of local maxima of  $R$



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 24

## Harris Detector: Steps



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



METR 4202: Robotics

10 September 2012 - 25

## Invariance and covariance

- We want corner locations to be invariant to photometric transformations and covariant to geometric transformations
  - Invariance: image is transformed and corner locations do not change
  - Covariance: if we have two transformed versions of the same image, features should be detected in corresponding locations



Adopted from S. Lazebnik, Gang Hua ([CS 558](#))



METR 4202: Robotics

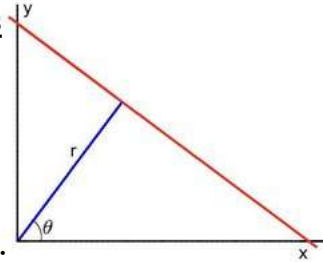
10 September 2012 - 26

## Line Detection – Hough Lines [1]

- A line in an image can be expressed as two variables:
  - Cartesian coordinate system:  $m, b$
  - Polar coordinate system:  $r, \theta$ 
    - avoids problems with vert. lines

$$y = mx + b \rightarrow$$

$$y = \left( -\frac{\cos \theta}{\sin \theta} \right) x + \left( \frac{r}{\sin \theta} \right)$$



- For each point  $(x_1, y_1)$  we can write:

$$r = x_1 \cos \theta + y_1 \sin \theta$$

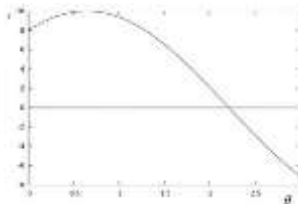
- Each pair  $(r, \theta)$  represents a line that passes through  $(x_1, y_1)$

See also OpenCV documentation ([cv::HoughLines](#))

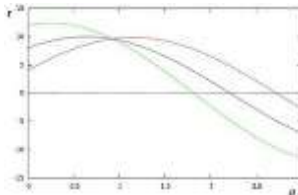


## Line Detection – Hough Lines [2]

- Thus a given point gives a sinusoid



- Repeating for all points on the image

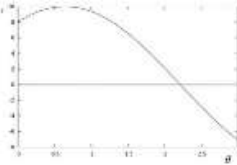


See also OpenCV documentation ([cv::HoughLines](#))

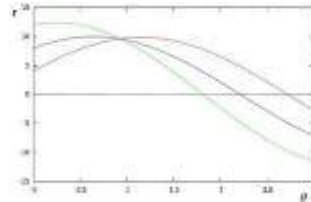


## Line Detection – Hough Lines [3]

- Thus a given point gives a sinusoid



- Repeating for all points on the image



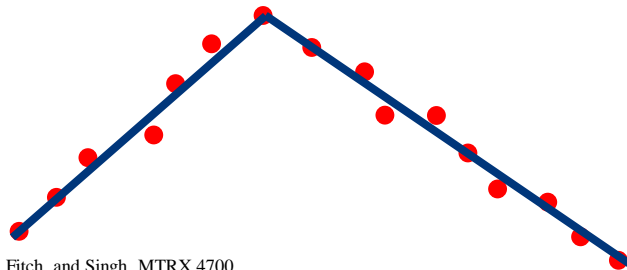
- NOTE that an intersection of sinusoids represents **(a point)** represents **a line** in which pixel points lay.

→ Thus, a line can be *detected* by finding the number of Intersections between curves

See also OpenCV documentation ([cv::HoughLines](#))



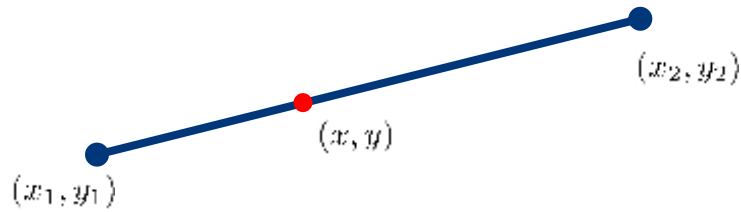
## Line Extraction and Segmentation



Adopted from Williams, Fitch, and Singh, MTRX 4700



## Line Formula



$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$y = mx + b$$

Adopted from Williams, Fitch, and Singh, MTRX 4700



METR 4202: Robotics

10 September 2012 - 31

## Line Estimation



Least squares minimization of the line:

- Line Equation:  $y - mx - b = 0$

- Error in Fit:  $\sum_i (y_i - mx_i - b)^2$

- Solution:  $\begin{pmatrix} \bar{x}\bar{y} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \bar{x}^2 & \bar{x} \\ \bar{x} & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix}$

Adopted from Williams, Fitch, and Singh, MTRX 4700

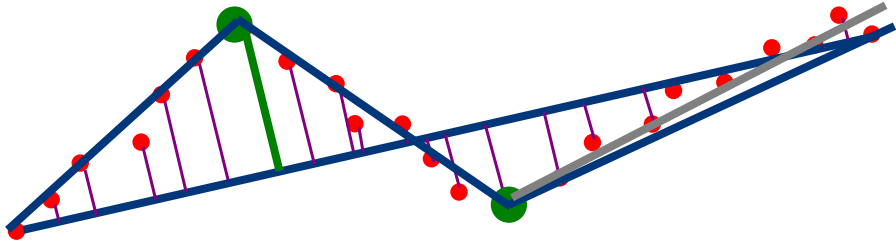


METR 4202: Robotics

10 September 2012 - 32



## Line Splitting / Segmentation



- What about corners?
- Split into multiple lines (via expectation maximization)
1. Expect (assume) a number of lines  $N$  (say 3)
  2. Find “breakpoints” by finding nearest neighbours upto a threshold or simply at random (RANSAC)
  3. How to know  $N$ ? (Also RANSAC)

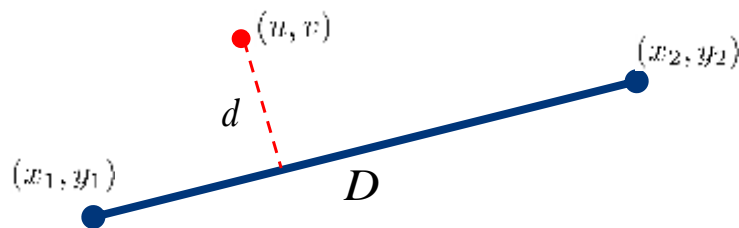
Adopted from Williams, Fitch, and Singh, MTRX 4700



METR 4202: Robotics

10 September 2012 - 33

## ⊥ of a Point from a Line Segment



$$r = u(y_1 - y_2) + v(x_2 - x_1) + y_2x_1 - y_1x_2$$

$$d = \frac{r}{D}$$

Adopted from Williams, Fitch, and Singh, MTRX 4700



METR 4202: Robotics

10 September 2012 - 34

## RANdom SAMple Consensus

1. Repeatedly select a small (minimal) subset of correspondences
  2. Estimate a solution (in this case a the line)
  3. Count the number of “inliers”,  $|e| < \Theta$   
(for LMS, estimate  $\text{med}(|e|)$ )
  4. Pick the *best* subset of inliers
  5. Find a complete least-squares solution
- Related to least median squares
  - See also:  
MAPSAC (Maximum *A Posteriori* SAMple Consensus)

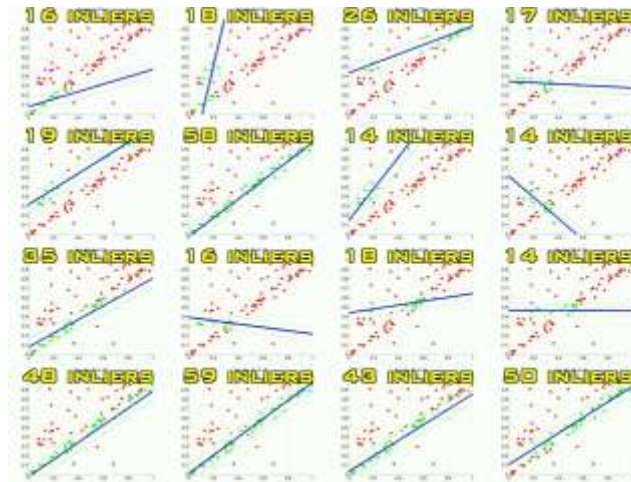
From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 36

## Cool Robotics Share (this week)



D. Wedge, *The RANSAC Song*



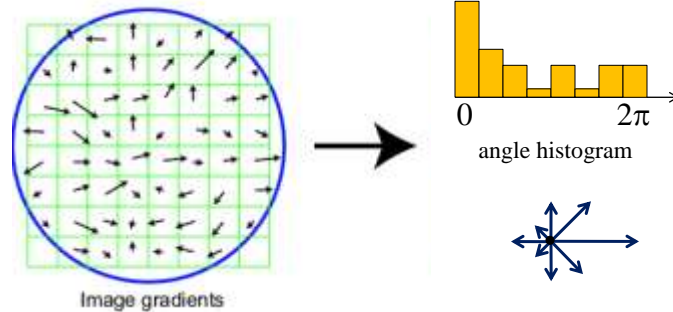
METR 4202: Robotics

10 September 2012 - 37

## Scale Invariant Feature Transform

Basic idea:

- Take 16x16 square window around detected feature
- Compute edge orientation (angle of the gradient -  $90^\circ$ ) for each pixel
- Throw out weak edges (threshold gradient magnitude)
- Create histogram of surviving edge orientations



Adapted from slide by David Lowe



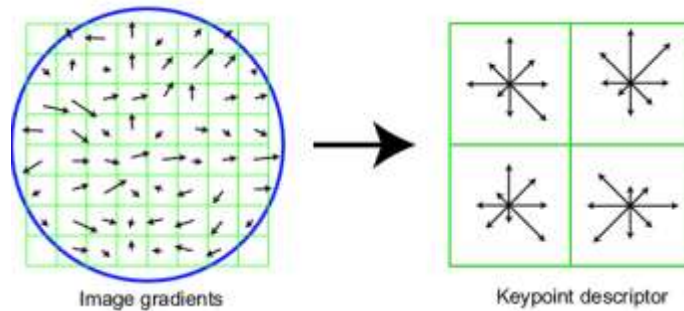
METR 4202: Robotics

10 September 2012 - 38

## SIFT descriptor

Full version

- Divide the 16x16 window into a 4x4 grid of cells (2x2 case shown below)
- Compute an orientation histogram for each cell
- 16 cells \* 8 orientations = 128 dimensional descriptor



Adapted from slide by David Lowe



METR 4202: Robotics

10 September 2012 - 39

## Properties of SIFT

- 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](http://people.csail.mit.edu/albert/ladypack/wiki/index.php/Known_Implementations_of_SIFT)



From David Lowe and Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 40

## Feature matching

- Given a feature in  $I_1$ , how to find the best match in  $I_2$ ?
  1. Define distance function that compares two descriptors
  2. Test all the features in  $I_2$ , find the one with min distance

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 41

## Feature distance

- How to define the difference between two features  $f_1, f_2$ ?
  - Simple approach is  $SSD(f_1, f_2)$ 
    - sum of square differences between entries of the two descriptors
    - can give good scores to very ambiguous (bad) matches



$I_1$

$I_2$

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 42

## Feature distance

- How to define the difference between two features  $f_1, f_2$ ?
  - Better approach: ratio distance =  $SSD(f_1, f_2) / SSD(f_1, f_2')$ 
    - $f_2$  is best SSD match to  $f_1$  in  $I_2$
    - $f_2'$  is 2<sup>nd</sup> best SSD match to  $f_1$  in  $I_2$
    - gives small values for ambiguous matches



$I_1$

$I_2$

From Szeliski, *Computer Vision: Algorithms and Applications*

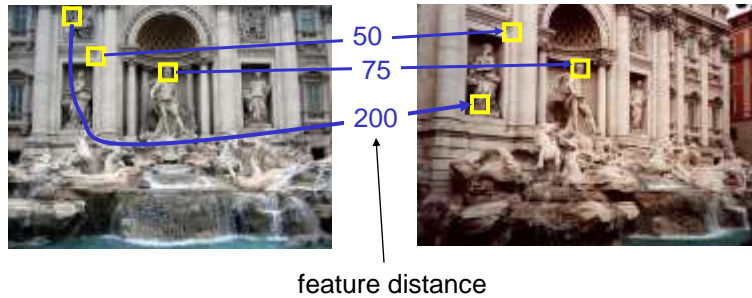


METR 4202: Robotics

10 September 2012 - 43

## Evaluating the results

- How can we measure the performance of a feature matcher?



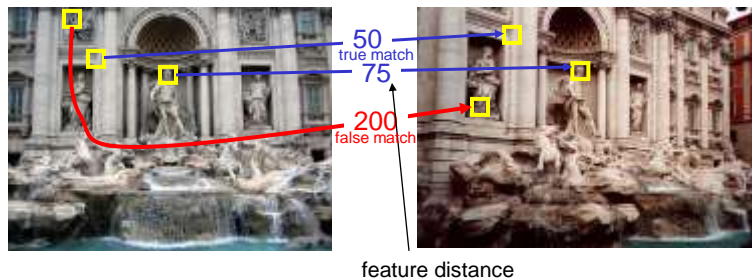
From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 44

## True/false positives



- The distance threshold affects performance
  - True positives = # of detected matches that are correct
    - Suppose we want to maximize these—how to choose threshold?
  - False positives = # of detected matches that are incorrect
    - Suppose we want to minimize these—how to choose threshold?

From Szeliski, *Computer Vision: Algorithms and Applications*

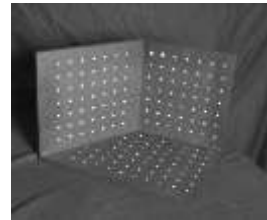


METR 4202: Robotics

10 September 2012 - 45

## Camera calibration

- Determine camera parameters from known 3D points or calibration object(s)
- internal or intrinsic parameters such as focal length, optical center, aspect ratio:  
what kind of camera?
- external or extrinsic (pose) parameters:  
where is the camera?
- How can we do this?



From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 46

## Camera calibration – approaches

- Possible approaches:
  - linear regression (least squares)
  - non-linear optimization
  - vanishing points
  - multiple planar patterns
  - panoramas (rotational motion)

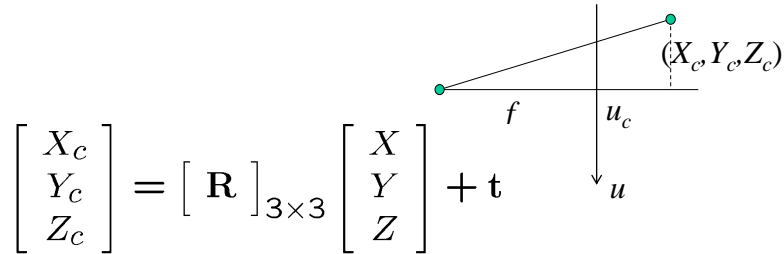
From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 47

## Image formation equations



$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = [\mathbf{R}]_{3 \times 3} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} f & 0 & u_c \\ 0 & f & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 48

## Calibration matrix

- Is this form of  $\mathbf{K}$  good enough?
- non-square pixels (digital video)
- skew
- radial distortion

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} f & 0 & u_c \\ 0 & f & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{K} \mathbf{X}_c$$

$$\begin{bmatrix} fa & s & u_c \\ 0 & f & v_c \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{K}$$

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 49



## Levenberg-Marquardt

- Iterative non-linear least squares [Press'92]
  - Linearize measurement equations

$$\hat{u}_i = f(\mathbf{m}, \mathbf{x}_i) + \frac{\partial f}{\partial \mathbf{m}} \Delta \mathbf{m}$$

$$\hat{v}_i = g(\mathbf{m}, \mathbf{x}_i) + \frac{\partial g}{\partial \mathbf{m}} \Delta \mathbf{m}$$

- Substitute into log-likelihood equation:  
quadratic cost function in  $\Delta \mathbf{m}$

$$\sum_i \sigma_i^{-2} (\hat{u}_i - u_i + \frac{\partial f}{\partial \mathbf{m}} \Delta \mathbf{m})^2 + \dots$$

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 50

## Levenberg-Marquardt

- What if it doesn't converge?
  - Multiply diagonal by  $(1 + l)$ , increase  $l$  until it does
  - Halve the step size  $\Delta \mathbf{m}$  (my favorite)
  - Use line search
  - Other ideas?
- Uncertainty analysis: covariance  $\mathbf{S} = \mathbf{A}^{-1}$
- Is maximum likelihood the best idea?
- How to start in vicinity of global minimum?

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 51

## Camera matrix calibration

- Advantages:
  - very simple to formulate and solve
  - can recover  $K [R | t]$  from  $M$  using QR decomposition [Golub & VanLoan 96]
- Disadvantages:
  - doesn't compute internal parameters
  - more unknowns than true degrees of freedom
  - need a separate camera matrix for each new view

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 52

## Multi-plane calibration

- Use several images of planar target held at unknown orientations [Zhang 99]
  - Compute plane homographies
$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \sim K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \sim HX$$
  - Solve for  $K$ -TK-1 from  $H_k$ 's
    - 1plane if only  $f$  unknown
    - 2 planes if  $(f,uc,vc)$  unknown
    - 3+ planes for full  $K$
  - Code available from Zhang and OpenCV



From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 53

## Rotational motion

- Use pure rotation (large scene) to estimate  $f$ 
  - estimate  $f$  from pairwise homographies
  - re-estimate  $f$  from 360° “gap”
  - optimize over all  $\{K, R_j\}$  parameters  
[Stein 95; Hartley '97; Shum & Szeliski '00; Kang & Weiss '99]



- Most accurate way to get  $f$ , short of surveying distant points

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 54

## SFM: Structure from Motion

(& Cool Robotics Share (this week))



METR 4202: Robotics

10 September 2012 - 55

## Structure [from] Motion

- Given a set of feature tracks, estimate the 3D structure and 3D (camera) motion.
- Assumption: orthographic projection
- Tracks:  $(u_{fp}, v_{fp})$ , f: frame, p: point
- Subtract out **mean** 2D position...

$\mathbf{i}_f$ : rotation,  $\mathbf{s}_p$ : position

$$u_{fp} = \mathbf{i}_f^T \mathbf{s}_p, v_{fp} = \mathbf{j}_f^T \mathbf{s}_p$$

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 56

## Structure from motion

- How many points do we need to match?
- 2 frames:
  - $(\mathbf{R}, \mathbf{t})$ : 5 dof + 3n point locations  $\leq$
  - 4n point measurements  $\Rightarrow$
  - $n \geq 5$
- k frames:
  - $6(k-1) + 3n \leq 2kn$
- always want to use many more

From Szeliski, *Computer Vision: Algorithms and Applications*



METR 4202: Robotics

10 September 2012 - 57

## Measurement equations

- Measurement equations

$$u_{fp} = \mathbf{i}_f^T \mathbf{s}_p \quad \mathbf{i}_f: \text{rotation, } \mathbf{s}_p: \text{position}$$

$$v_{fp} = \mathbf{j}_f^T \mathbf{s}_p$$

- Stack them up...

$$\mathbf{W} = \mathbf{R} \mathbf{S}$$

$$\mathbf{R} = (\mathbf{i}_1, \dots, \mathbf{i}_F, \mathbf{j}_1, \dots, \mathbf{j}_F)^T$$

$$\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_p)$$

From Szeliski, [Computer Vision: Algorithms and Applications](#)



METR 4202: Robotics

10 September 2012 - 58

## Factorization

$$\mathbf{W} = \mathbf{R}_{2F \times 3} \mathbf{S}_{3 \times P}$$

SVD

$$\mathbf{W} = \mathbf{U} \mathbf{A} \mathbf{V} \quad \mathbf{A} \text{ must be rank } 3$$

$$\mathbf{W}' = (\mathbf{U} \mathbf{A}^{1/2})(\mathbf{A}^{1/2} \mathbf{V}) = \mathbf{U}' \mathbf{V}'$$

Make  $\mathbf{R}$  orthogonal

$$\mathbf{R} = \mathbf{Q} \mathbf{U}', \quad \mathbf{S} = \mathbf{Q}^{-1} \mathbf{V}'$$

$$\mathbf{i}_f^T \mathbf{Q}^T \mathbf{Q} \mathbf{i}_f = 1 \dots$$

From Szeliski, [Computer Vision: Algorithms and Applications](#)



METR 4202: Robotics

10 September 2012 - 59

## Results

- Look at paper figures...

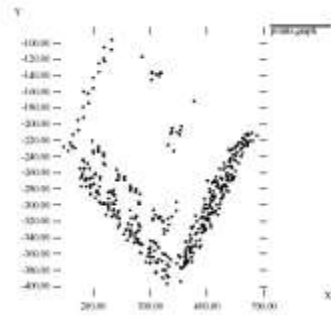


Figure 4.5: A view of the computed shape from approximately above the building (compare with figure 4.6).

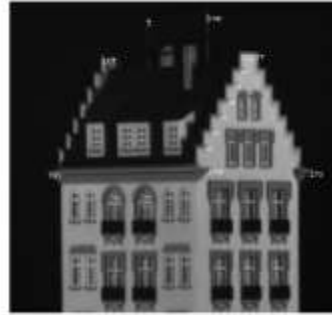


Figure 4.7: For a quantitative evaluation, distances between the features shown in the picture were measured on the actual model, and compared with the computed results. The comparison is shown in figure 4.8.

From Szeliski, *Computer Vision: Algorithms and Applications*



## Cool Robotics Share (replay from last, last week)

### Fast and Accurate Knife-Edge Maneuvers for Autonomous Aircraft

Andrew Barry  
Anirudha Majumdar  
Tim Jenks  
Russ Tedrake

Robot Locomotion Group  
MIT/CSAIL

Huai-Ti Lin  
Ivo Ros  
Andrew Biewener

Concord Field Station  
Harvard University

