Lecture 7

Features 1:
Corner detection

Mon. Sept. 28, 2020

Record today's lecture (reminder)

At the end of lecture, if you want to stay and discuss the assignment in small groups, then I can put you into breakout rooms.

So far … edges and gradients.

Today we start … more complex local intensity structures called "features".

This will lead us to two problems:

- detecting & localizing a feature
- describing the feature

COMP 558 Overview

Part 1 : 2D Vision

RGB

Image filtering

Edge detection

Least Squares Estimation

Robust Estimation: Hough transform & RANSAC

Features 1: corners

Image Registration: the Lucas-Kanade method

Scale spaces (Gaussian and Laplacian)

Histogram-based Tracking:

Features 2: SIFT, HOG

Features 3: CNN's

Object classification and detection

Segmentation (TBD)

Part 2 : 3D Vision

Linear perspective, camera translation

Vanishing points, camera rotation

Homogeneous coordinates, camera intrinsics

Least Squares methods (eigenspaces, SVD)

Camera Calibration

Homographies & rectification

Stereo and Epipolar Geometry

Stereo correspondence

Cameras and Photography

RGBD Cameras

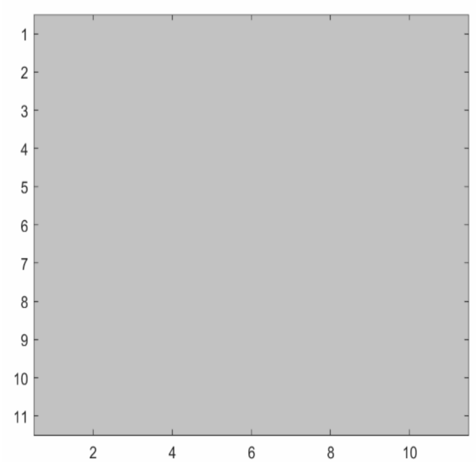
image patch

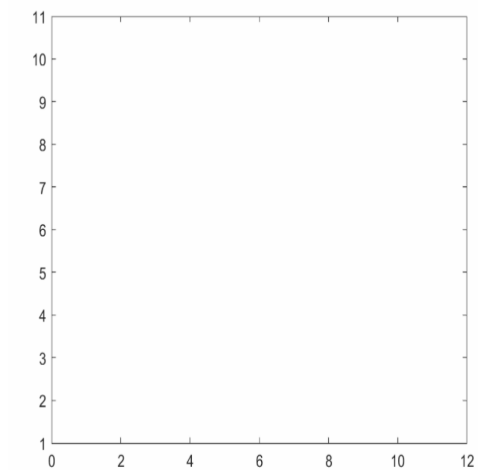
$$I(x, y)$$

gradients

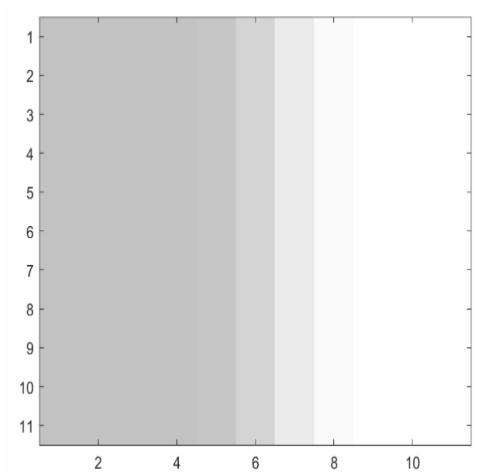
$$\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right)$$

uniform





vertical
step edge



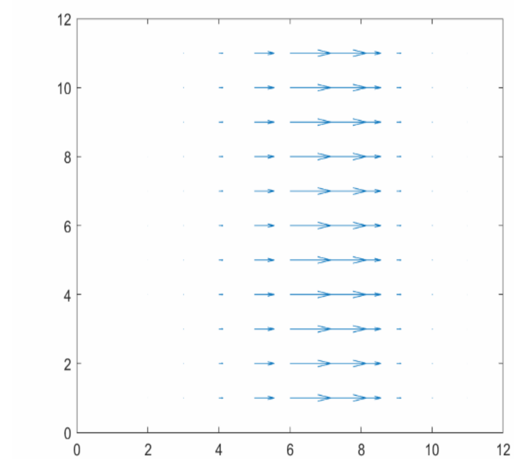


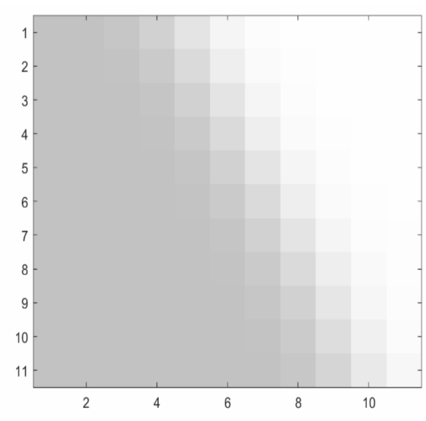
image patch

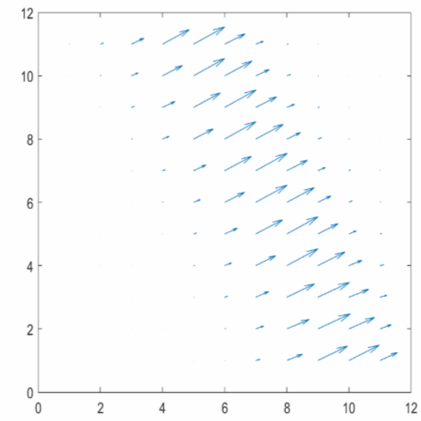
$$I(x, y)$$

gradients

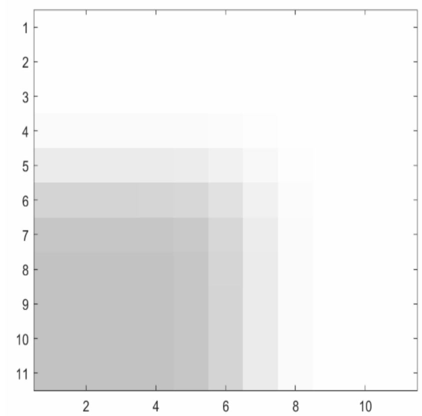
$$\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right)$$

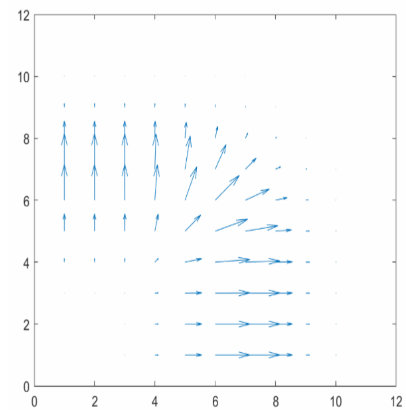
diagonal step edge





corner





Assume we have smoothed the image with a small Gaussian and then taken the local differences to obtain the gradient.

We leave out the Gaussian to avoid cluttering the notation.

$$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

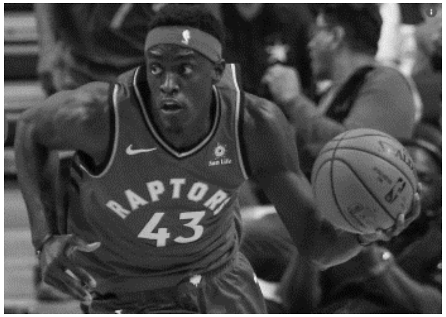


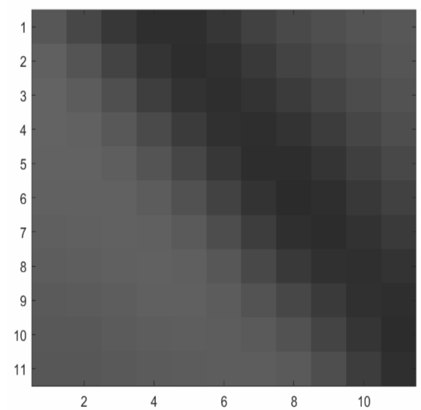
image patch

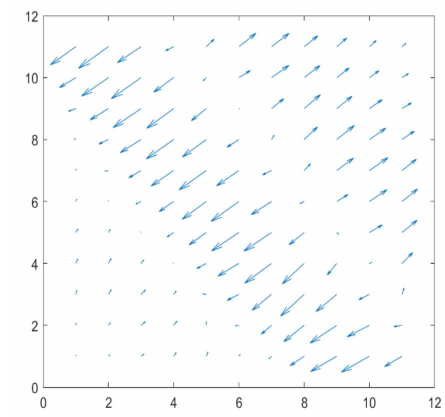
gradients

$$I(x, y)$$

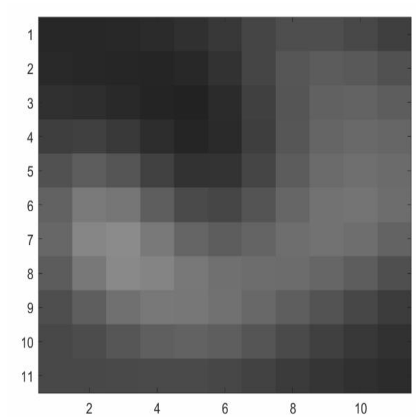
$$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

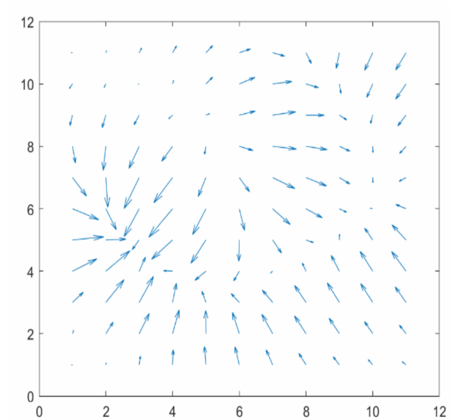
stripe on ball
(edge-like)





eye
(corner-like)





“Corner” detection

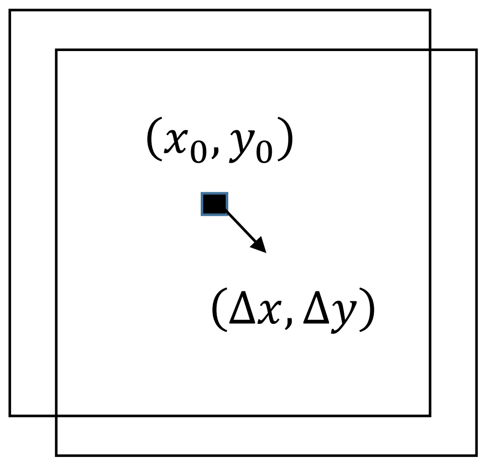
“corner”
= “interest point”
= “keypoint”
= “locally distinctive point”

A crucial difference between a “corner” and an “edge” is that the corner can be localized.

TODAY: How to quantify this idea?

How do the image intensities change when we slightly shift the neighborhood of (x_0, y_0) by $(\Delta x, \Delta y)$?
We examine sum of squared differences.

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y) \in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$

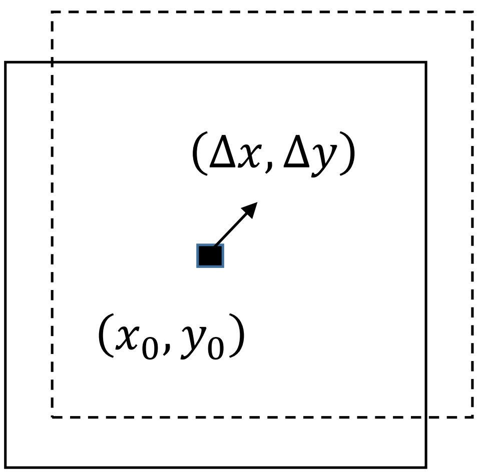


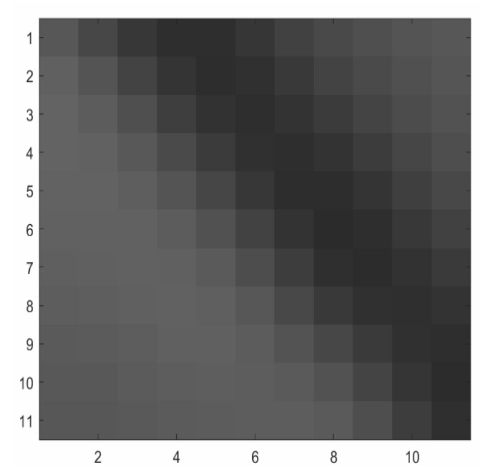
Here we look at discrete pixel steps $(\Delta x, \Delta y)$.
e.g. we could check all 8 immediate neighbors.

How do the image intensities change when we slightly shift the neighborhood of (x_0, y_0) by $(\Delta x, \Delta y)$?

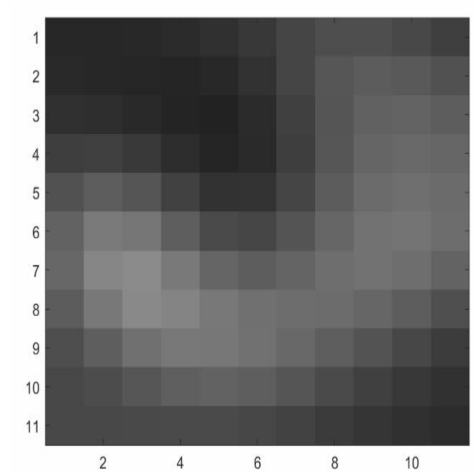
We examine sum of squared differences.

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$





big change
(why?)

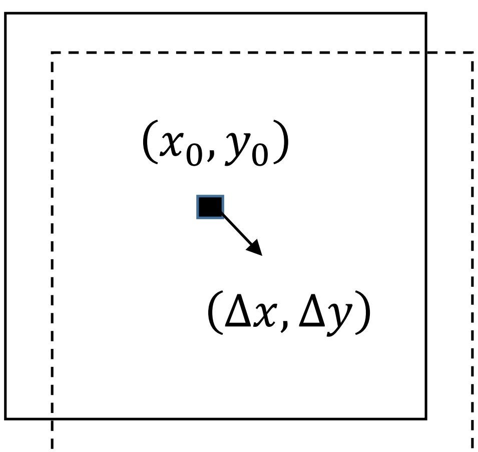


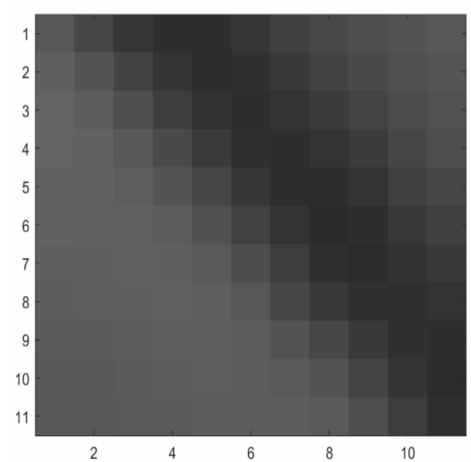
big change
(why?)

How do the image intensities change when we slightly shift the neighborhood of (x_0, y_0) by $(\Delta x, \Delta y)$?

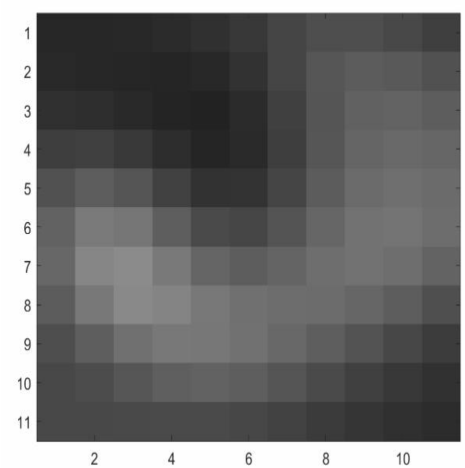
We examine sum of squared differences.

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$



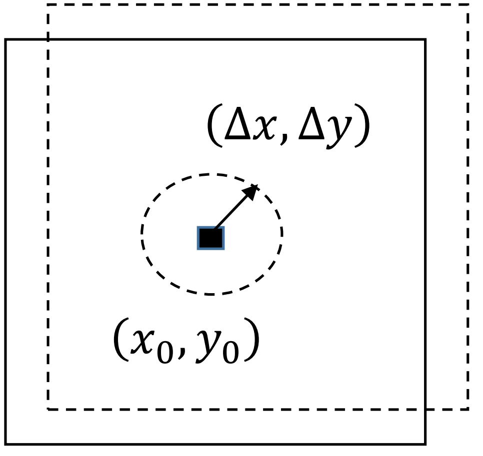


little change
(why?)



big change
(why?)

An image point is *locally distinctive* if a small shift of the neighborhood of the point *in any direction* gives a significantly different image.



Rather than examining discrete $(\Delta x, \Delta y)$ in a 3x3 or 5x5 local neighborhood, we will take a different approach.

First, we approximate $I(x, y)$ locally as follows:

$$I(x + \Delta x, y + \Delta y) \approx I(x, y) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y$$

Recall that we are assuming we have smoothed the image with a small Gaussian and then we obtain the gradient by taking the local differences. We leave out the Gaussian to avoid cluttering the notation.

$$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

$$I(x+\Delta x, y+\Delta y) \approx I(x,y) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y$$

Then, we substitute:

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$

$$\approx \sum_{(x,y)\in Ngd(x_0,y_0)} (\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y)^2$$

squared directional
derivative of image

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$

previous slide

$$\approx \sum_{(x,y)\in Ngd(x_0,y_0)} (\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y)^2$$

$$= \sum_{(x,y)\in Ngd(x_0,y_0)} ((\nabla I)\cdot(\Delta x, \Delta y))^2$$

$$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

1x2

$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$

previous slide

$$\approx \sum_{(x,y)\in Ngd(x_0,y_0)} (\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y)^2$$

$$= \sum_{(x,y)\in Ngd(x_0,y_0)} ((\nabla I)\cdot(\Delta x, \Delta y))^2$$

$$= \sum_{(x,y)\in Ngd(x_0,y_0)} (\Delta x, \Delta y)(\nabla I)^T(\nabla I)(\Delta x, \Delta y)^T$$

$$= (\Delta x, \Delta y)\{\underbrace{\sum_{(x,y)\in Ngd(x_0,y_0)} (\nabla I)^T(\nabla I)}_{2\text{x}2}\}(\Delta x, \Delta y)^T.$$

$$(\nabla I) \cdot (\nabla I)^T = \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & (\frac{\partial I}{\partial y})^2 \end{bmatrix}$$

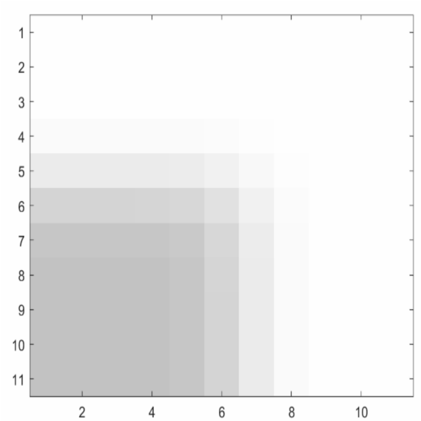
ASIDE: This 2x2 matrix is defined at each pixel, and is called the "structure tensor". (Sometimes in the literature, this term refers to the sum over windows as in previous slide.)

$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} (\nabla I)^T(\nabla I) = \begin{bmatrix} \sum(\frac{\partial I}{\partial x})^2 & \sum(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ \sum(\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & \sum(\frac{\partial I}{\partial y})^2 \end{bmatrix}$$

This matrix $\mathbf{M}$ is called the "second moment matrix".

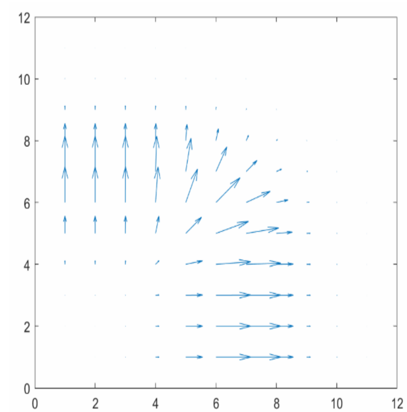
image patch

$$I(x,y)$$



gradient

$$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$$

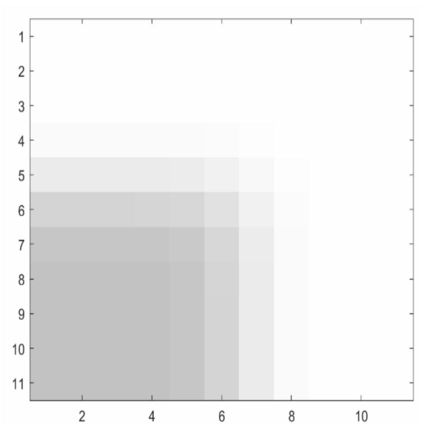


Each matrix element is a sum over the local neighborhood.

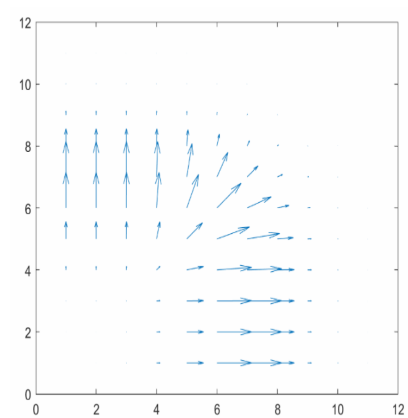
$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} W(x - x_0, y - y_0)\ (\nabla I)(\nabla I)^T$$

Typically we weight the contributions by a 2D Gaussian so that the terms in the center of the neighborhood have the most weight.

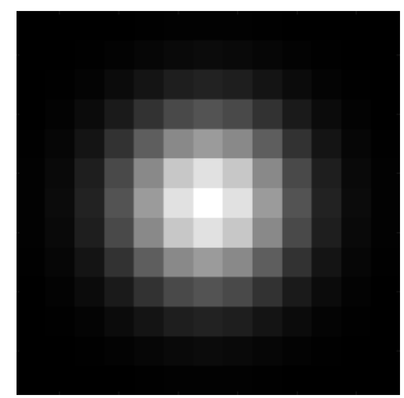
$I(x, y)$



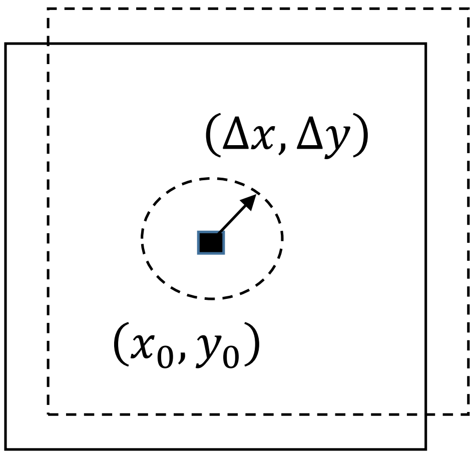
$\nabla I = (\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y})$



$W(\,)$



$$\mathcal{E}(\Delta x, \Delta y) \equiv \sum_{(x,y)\in Ngd(x_0,y_0)} (I(x,y) - I(x+\Delta x, y+\Delta y))^2$$



$$\approx \quad \boldsymbol{u}^T \ \mathbf{M}\boldsymbol{u}$$

where $\boldsymbol{u} = (\Delta x, \Delta y)$

$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} W(x - x_0, y - y_0) \ (\nabla I)(\nabla I)^T$$

An image point is *locally distinctive* if a unit length step in any direction $\boldsymbol{u}$ gives a (sufficiently) different image.

We would like to decide if a point is *locally distinctive* (x_0, y_0) by examining the matrix $\mathbf{M}$.

Linear Algebra

The second moment matrix $\mathbf{M}$ is symmetric. In particular, $\boldsymbol{u}^T \mathbf{M} \boldsymbol{u} \ge 0$ for any vector $\boldsymbol{u}$. (Why?)

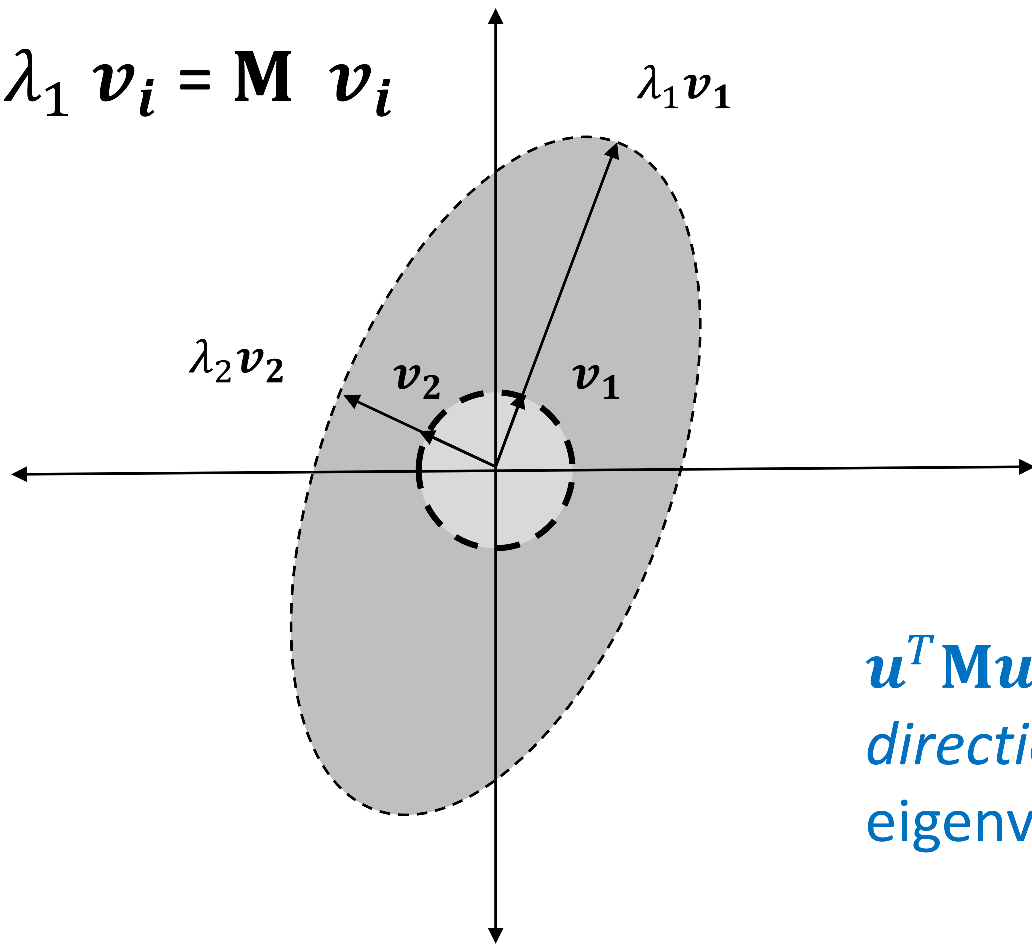
Thus, linear algebra tells us that we can write it as:

$$\mathbf{M} = \boldsymbol{V}\, \Lambda\, \boldsymbol{V}^T$$

where the columns of the 2x2 matrix $\boldsymbol{V}$ are the eigenvectors of $\mathbf{M}$ and Λ is a diagonal matrix, and the diagonal elements of Λ are the eigenvalues of $\mathbf{M}$.

Let $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ be the eigenvectors of $\mathbf{M}$ and let λ_1 and λ_2 be the eigenvalues. We can assume that $\lambda_1 \geq \lambda_2 \geq 0$.

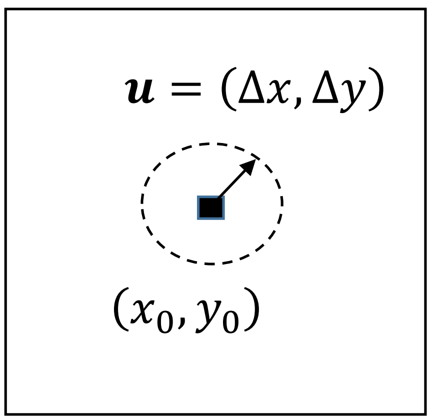
$$\lambda_1 \, \boldsymbol{v_i} = \mathbf{M} \, \boldsymbol{v_i}$$

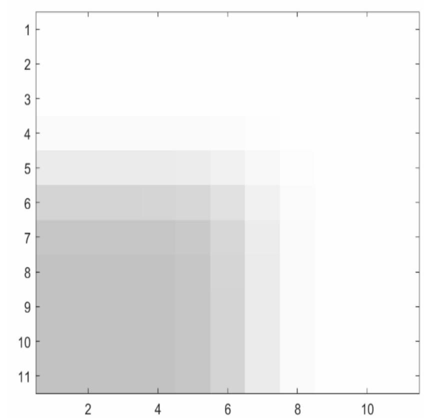


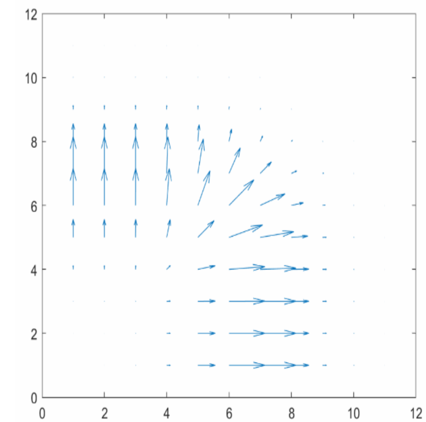
The 2D space is $\boldsymbol{u} = (\Delta x, \Delta y)$ and the ellipsoid shows the magnitude $\boldsymbol{u}^T \mathbf{M} \boldsymbol{u}$ when $\boldsymbol{u}$ is a unit vector.

$\boldsymbol{u}^T \mathbf{M} \boldsymbol{u}$ will be large *in any unit direction* $\boldsymbol{u}$ if and only if both eigenvalues of $\mathbf{M}$ are large.

$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} W(x - x_0, y - y_0)\ (\nabla I)(\nabla I)^T$$

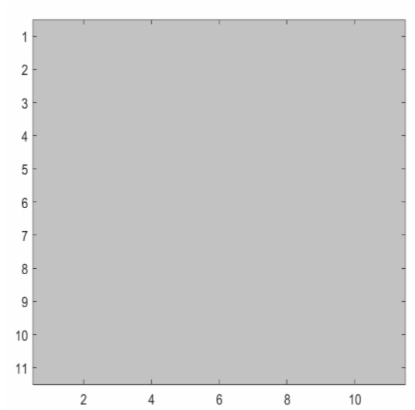
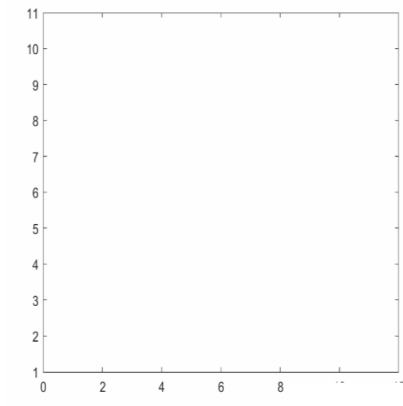
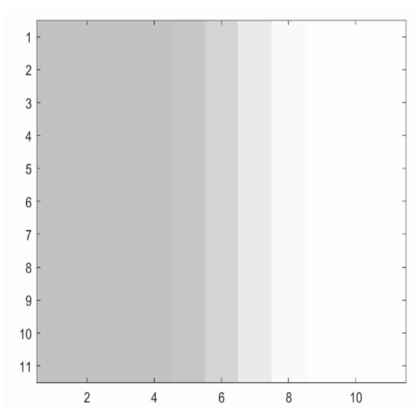
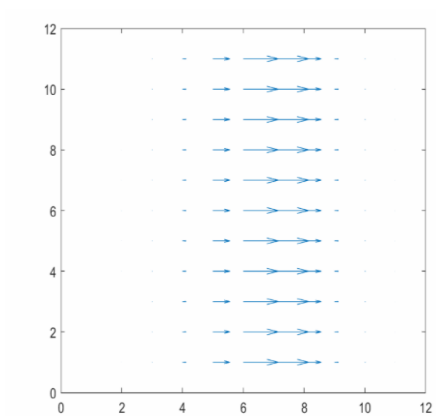






Exercise:

What is the geometric interpretation of $\boldsymbol{u}^T\mathbf{M}\boldsymbol{u}$?

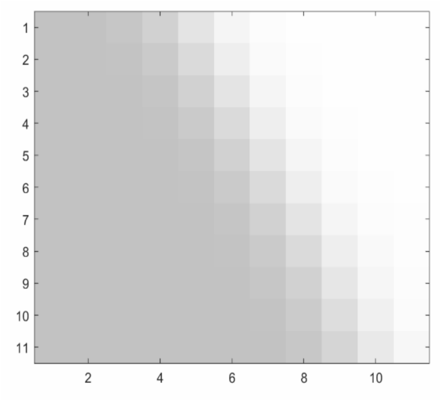
	image	gradient field	2nd moment matrix $\boldsymbol{M}$	eigenvalues λ_1, λ_2
uniform			$\boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}}$	0 0
step edge			$\boxed{\begin{matrix} 620 & 0 \\ 0 & 0 \end{matrix}}$	620, 0

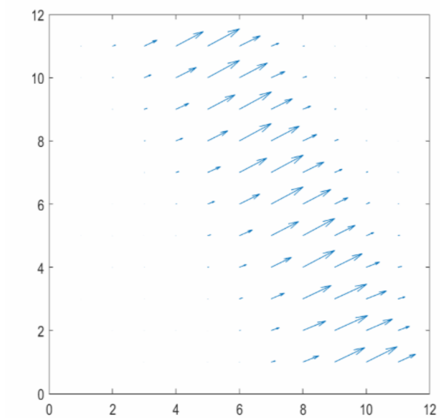
$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} W(x-x_0, y-y_0)\ (\nabla I)(\nabla I)^T$$

Neither are locally distinctive.

image | gradient field | 2nd moment matrix $\boldsymbol{M}$ | eigenvalues λ_1, λ_2

Step edge





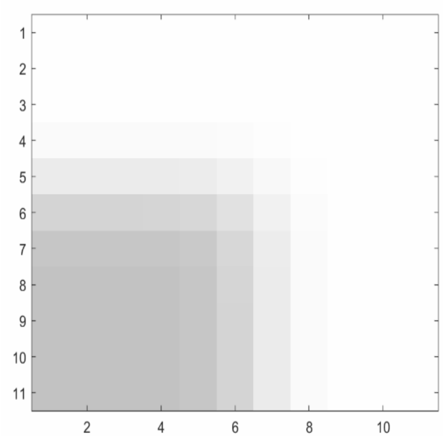
$$\boxed{\begin{matrix} 450 & -230 \\ -230 & 120 \end{matrix}}$$

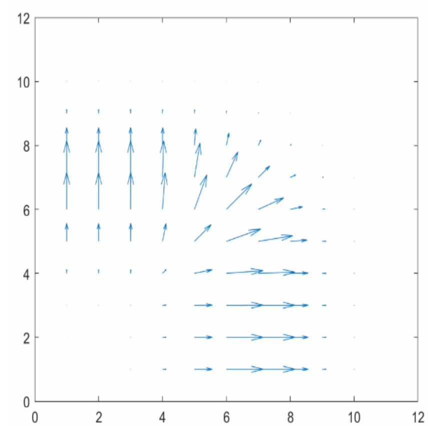
570, 0

✗

Not locally distinctive

Corner





$$\boxed{\begin{matrix} 300 & -120 \\ -120 & 300 \end{matrix}}$$

420, 180

✓

Locally distinctive

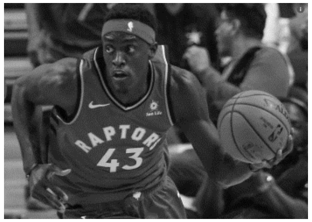
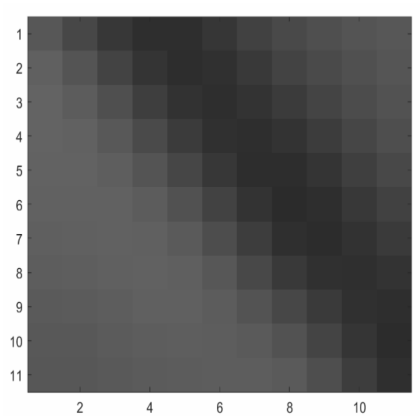
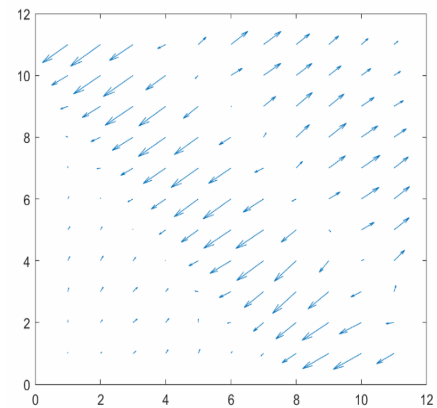
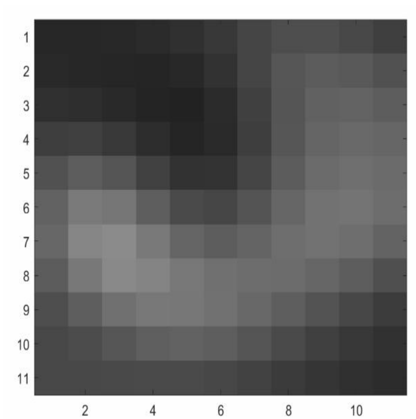
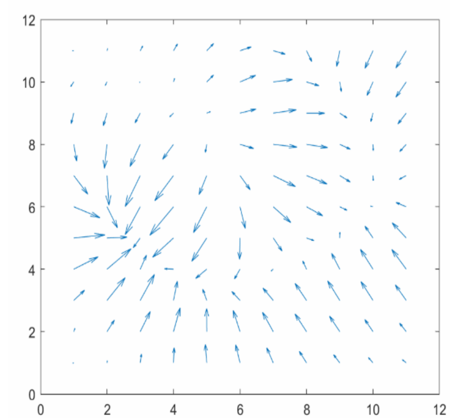


	image	gradient field	2nd moment matrix **M**	eigenvalues λ_1, λ_2
stripe on ball			$\boxed{\begin{matrix} 380 & -280 \\ -280 & 220 \end{matrix}}$	590, 10
eye			$\boxed{\begin{matrix} 580 & -120 \\ -120 & 930 \end{matrix}}$	970, 540

Computing eigenvalues of $\mathbf{M}$ requires solving a quadratic equation, namely

$$\det(\mathbf{M} - \lambda \mathbf{I}) = 0$$

which requires taking a square root.

Historically (30 years ago), this was considered to be unnecessarily expensive.

Recall from linear algebra

tr($\mathbf{M}$) = $M_{11} + M_{22}$

= $\lambda_1 + \lambda_2$

det($\mathbf{M}$) = $M_{11}M_{22} - M_{12}M_{21}$

= $\lambda_1 \lambda_2$

Harris Corner detector (1988)

Harris and Stevens observed that we really just want to separate out corner versus edge versus uniform.

The showed that corners could be detected by combining the determinant and trace of $\mathbf{M}$:

$$Harris(x,y) = \det(\mathbf{M}(x,y)) - k\,\mathrm{tr}(\mathbf{M}(x,y))^{\mathbf{2}}$$

where k ~ 0.1.

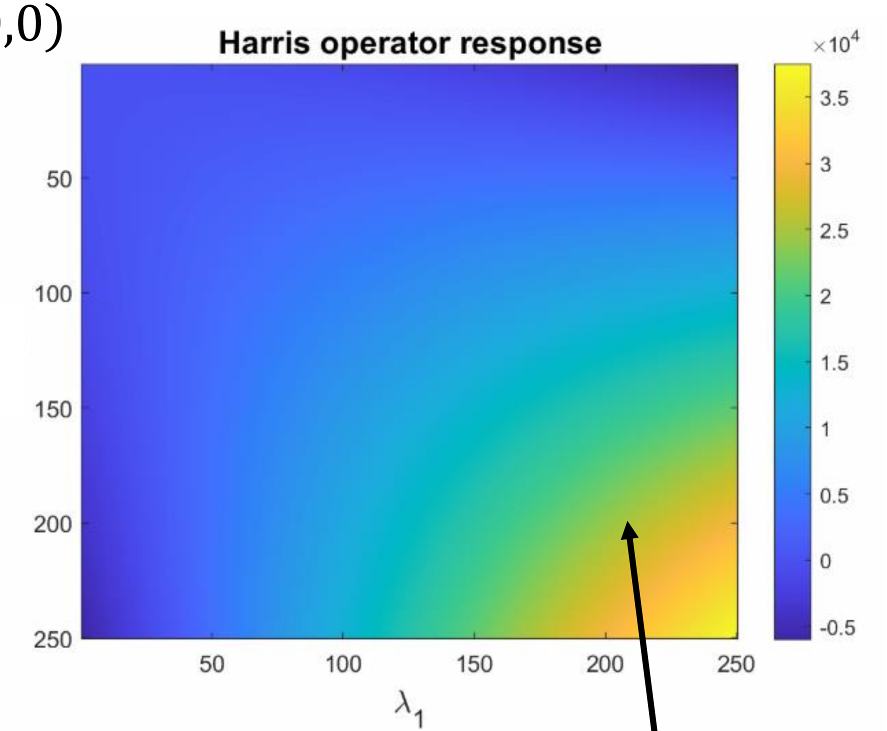
The original slide contained a typo. I forgot to square it.

Harris Corner detector

λ_1

(0,0)

λ_2



$$Harris(x, y)$$

$$= \det(\mathbf{M}(x,y)) - k\,\mathrm{tr}(\mathbf{M}(x,y))^2$$

$$= \lambda_1\,\lambda_2 \ - \ 0.1\,(\lambda_1 + \lambda_2)^2$$

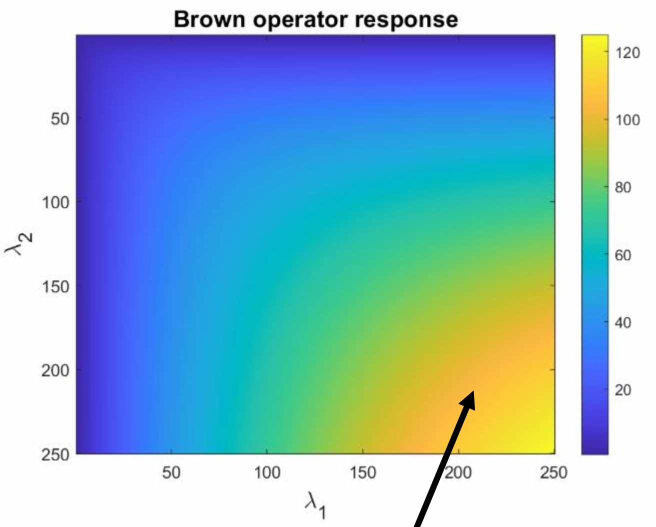
The original slide contained a typo. I forgot to square it. But the plot was correct.

Harris response is large if and only if both eigenvalues of **M** are large

Alternative corner detector(s)...

$$\frac{\det(\mathbf{M}(x,y))}{\operatorname{tr}(\mathbf{M}(x,y)) + \varepsilon} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2 + \varepsilon}$$

Brown, Szeliski, Winder (2004)



Again, operator response is large if and only if both eigenvalues of M are large.

Where are the corners/keypoints/ locally distinctive points ?

Where are the corners/keypoints/ locally distinctive points ?

Top 1% of Harris responses, after non-maximum suppression.

Top 5% of Harris responses, after non-maximum suppression.

One final thought…

Recall that we blurred the image with a Gaussian before estimating the intensity gradient. The sigma of this blur is called the *inner scale*. Typically $\sigma = 1$.

We also used a weighted average of the structure tensor (outer product of gradients), where the weights were a Gaussian. The sigma of this weighting function is called the *outer scale*. Typically $\sigma = 2$ or more.

Other notions of scale will come up in lecture 9

(Scale Space).

- Quiz 1 today

- Assignment 1 -- breakout room discussion
  (for those who wish)