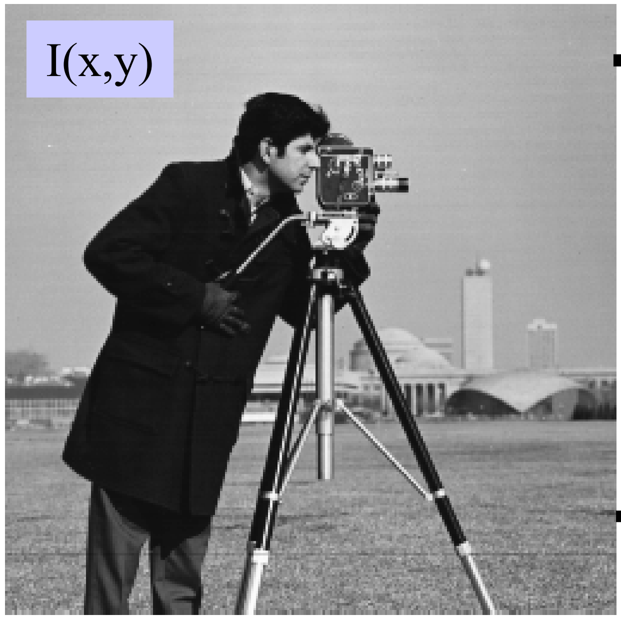
Robert Collins
CSE486, Penn State

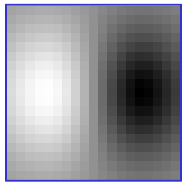
Lecture 7: Correspondence Matching

Reading: T&V Section 7.2

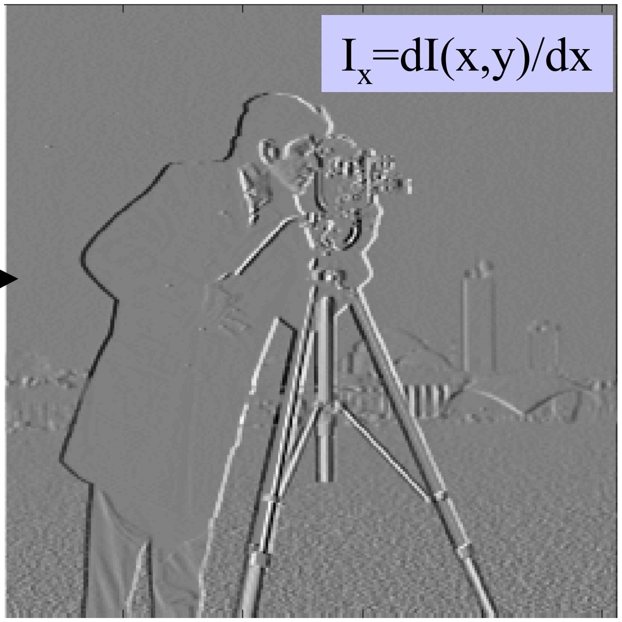
Recall: Derivative of Gaussian Filter



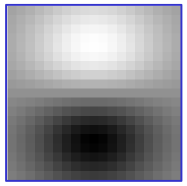
G_x



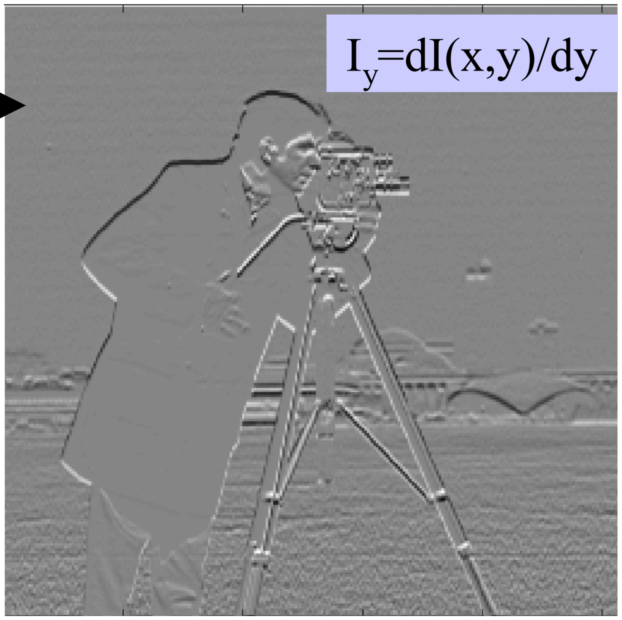
convolve



G_y

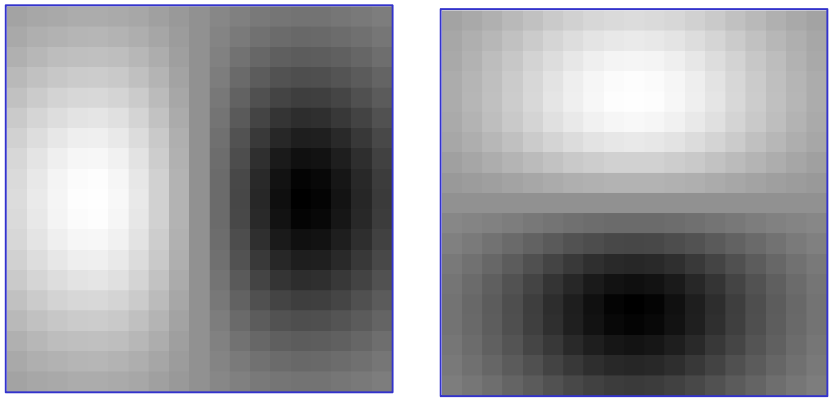


convolve



Observe and Generalize

Derivative of Gaussian



Looks like vertical and horizontal step edges

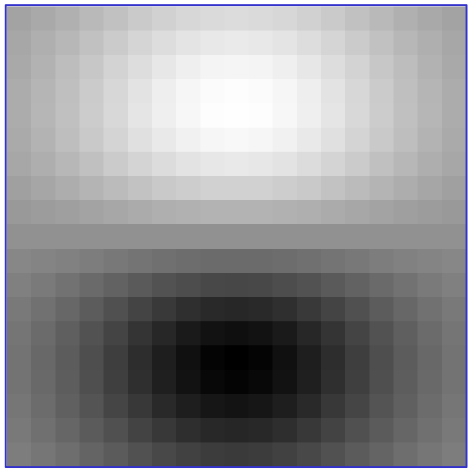
Key idea: Convolution (and cross correlation) with a filter can be viewed as comparing a little "picture" of what you want to find against all local regions in the image.

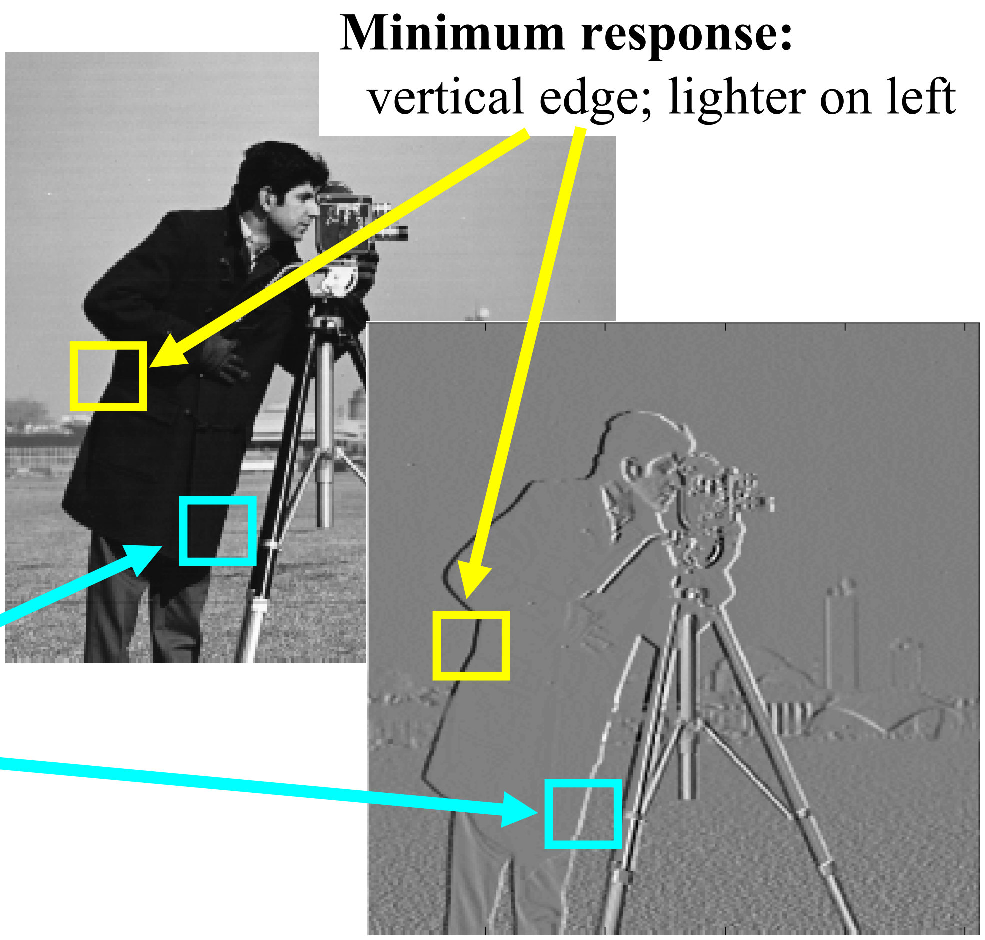
Robert Collins
CSE486, Penn State

Observe and Generalize

Key idea: <u>Cross correlation</u> with a filter can be viewed as comparing a little "picture" of what you want to find against all local regions in the image.

looks like vertical edge; lighter on right





Minimum response:
vertical edge; lighter on left

Maximum response:
vertical edge; lighter on right

Observe and Generalize

Key idea: Cross correlation with a filter can be viewed as comparing a little "picture" of what you want to find against all local regions in the image.

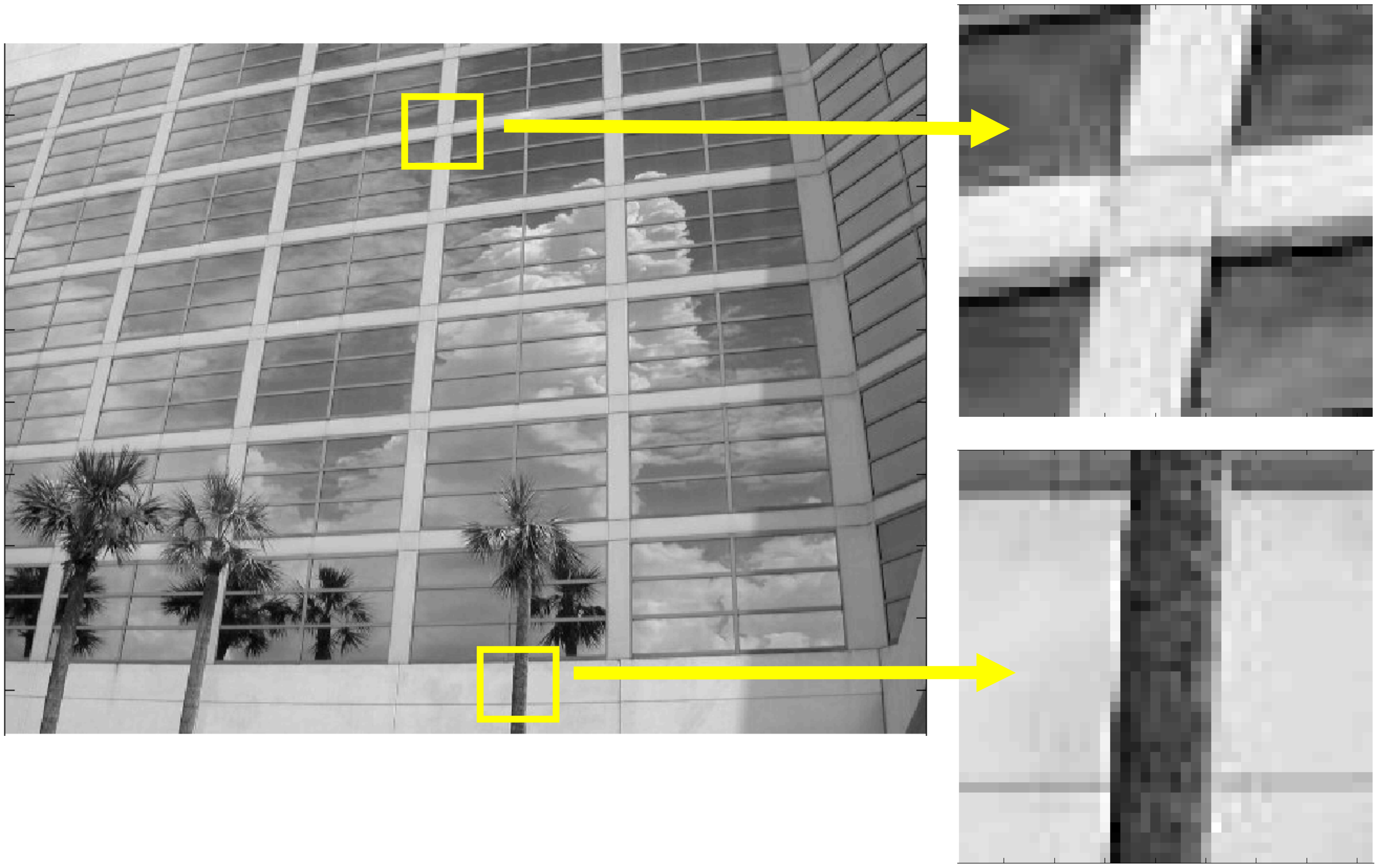
For this reason, it is sometimes called
"matched filtering"

In fact, you can prove that the best linear operator for finding an image patch is essentially the patch itself

(using variational calculus, outside scope of our course).

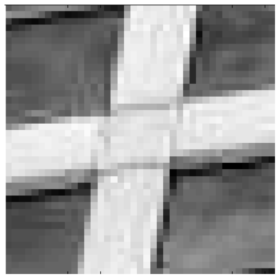
Template Matching

What if we cut little pictures out from an image, then tried convolve them with the same or other images?

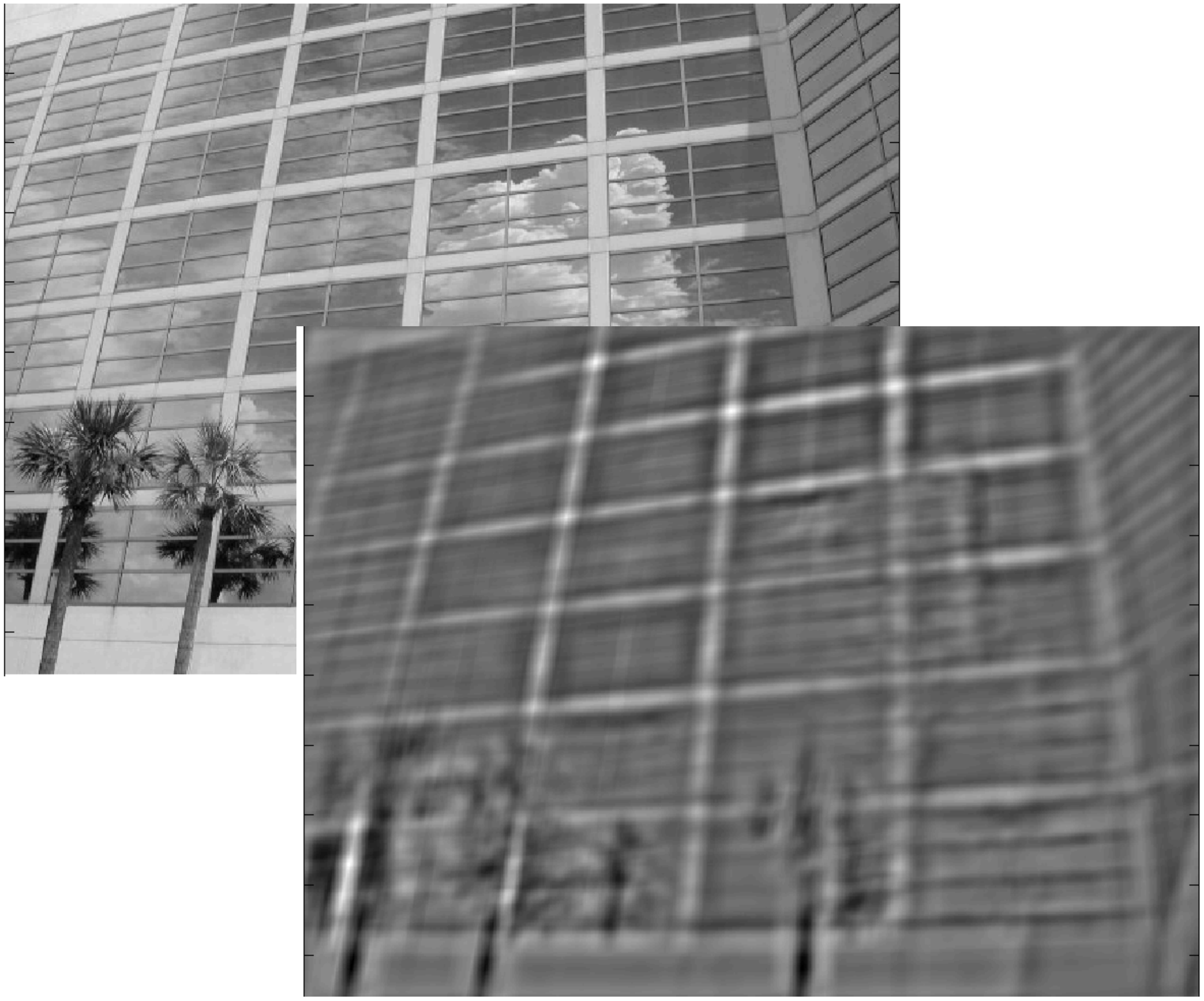


Robert Collins
CSE486, Penn State

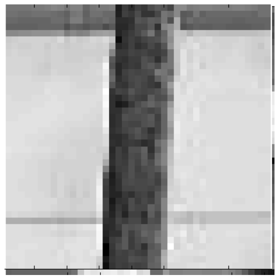
Template Matching



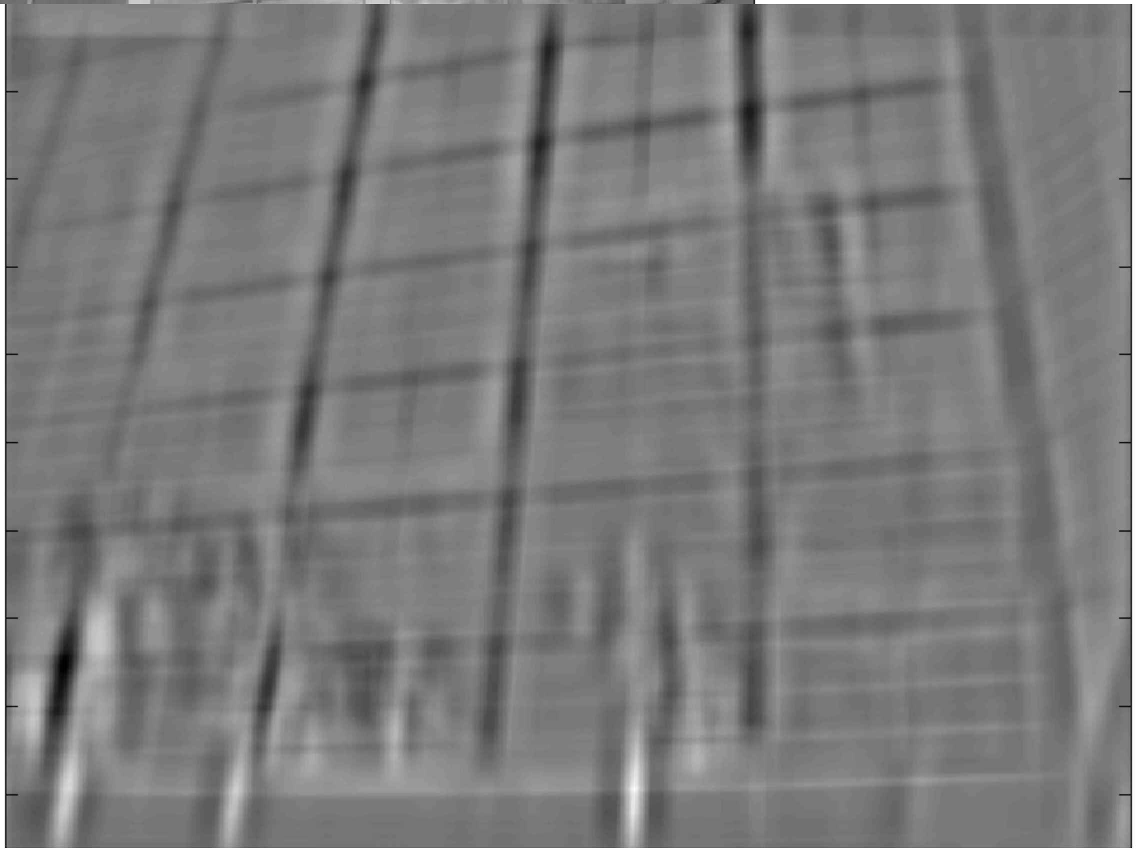
*



Template Matching



*



Confession

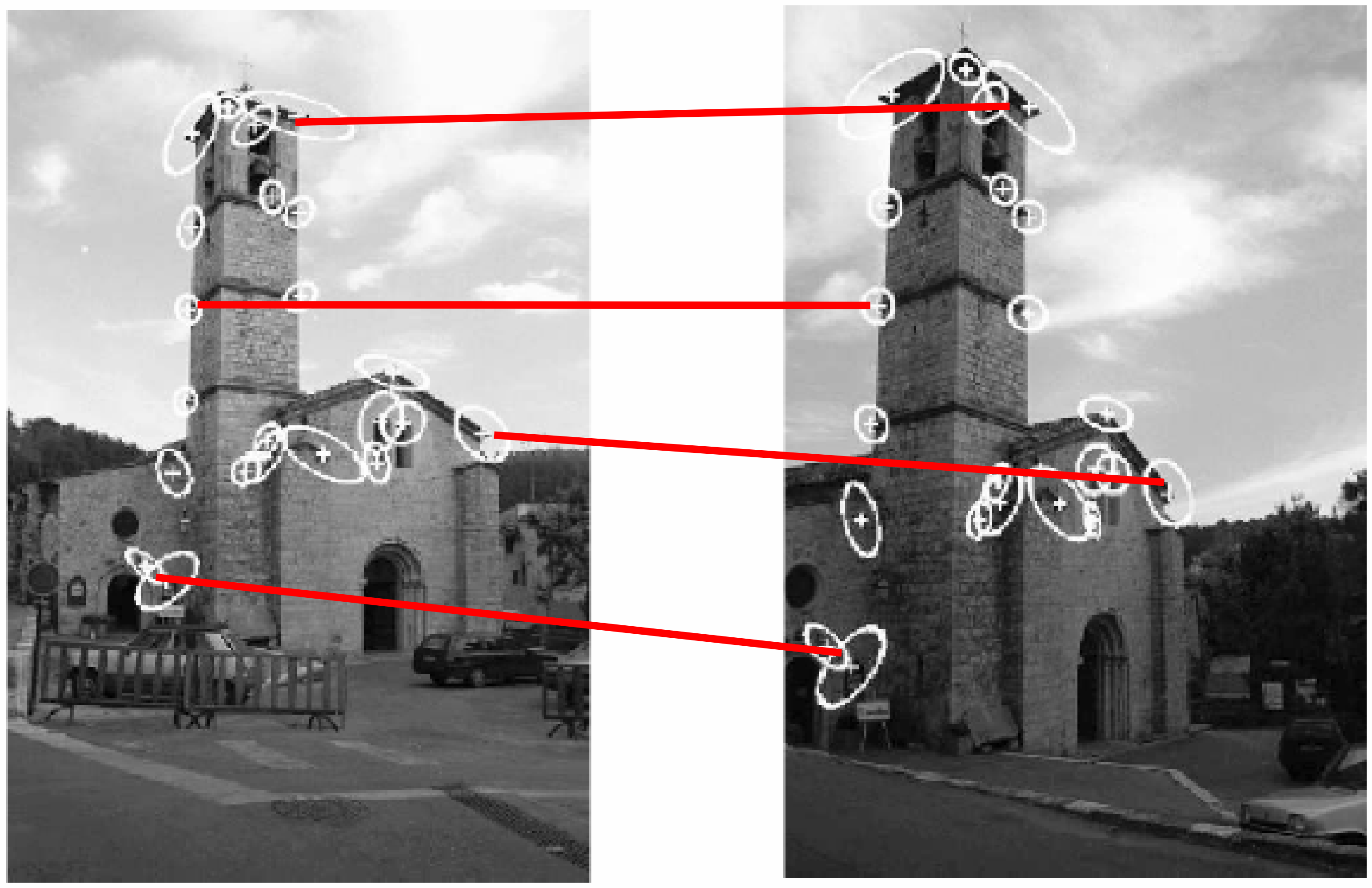
I've cheated a little bit.

I subtracted the mean greyvalue from both the image and the template before doing cross correlation.

Why? (we will discuss later, but think about it first)

Correspondence Problem

Vision tasks such as stereo and motion estimation require finding corresponding features across two or more views.



The Correspondence Problem

- Basic assumptions:
  - Most scene points are visible in both images
  - Corresponding image regions are similar
- These assumptions hold if:
  - The distance of points from the cameras is much larger than the distance between cameras

The Correspondence Problem

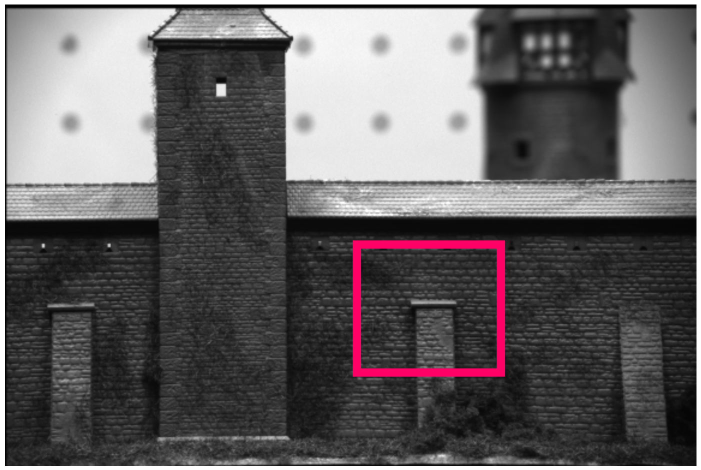
- Is a "search" problem:
  - Given an element in the left image, search for the corresponding element in the right image.
  - We will typically need geometric constraints to reduce the size of the search space
- We must choose:
  - Elements to match
  - A similarity measure to compare elements

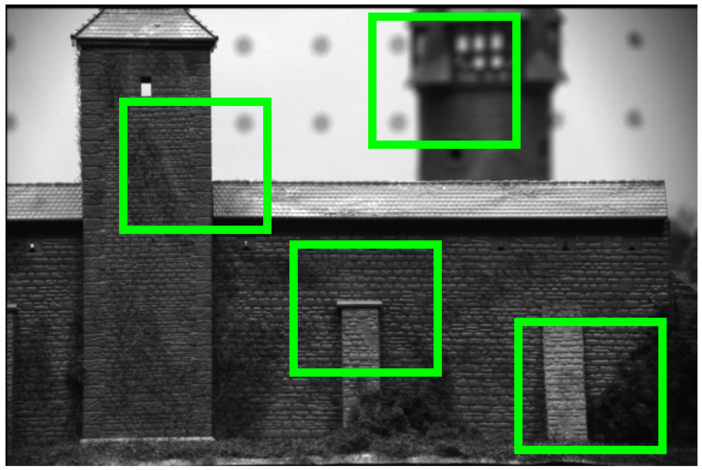
Correspondence Problem

- Two classes of algorithms:
  - Correlation-based algorithms
    - Produce a DENSE set of correspondences
  - Feature-based algorithms
    - Produce a SPARSE set of correspondences

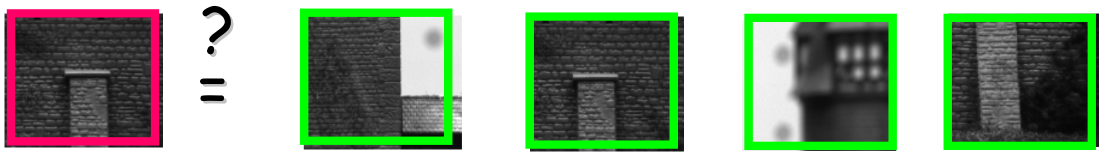
Correlation-based Algorithms

Elements to be matched are image patches of fixed size





Task: what is the corresponding patch in a second image?



Robert Collins
CSE486, Penn State

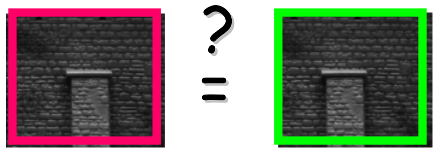
Correlation-based Algorithms

Task: what is the corresponding patch in a second image?

1) Need an appearance similarity function.

2) Need a search strategy to find location with highest similarity. Simplest (but least efficient) approach is exhaustive search.

Comparing Windows:



Some possible measures:

$$\max_{[i,j]\in R} |f(i,j) - g(i,j)|$$

$$\sum_{[i,j]\in R} |f(i,j) - g(i,j)|$$

$$\left.\begin{aligned} SSD &= \sum_{[i,j]\in R} (f(i,j) - g(i,j))^2 \\ C_{fg} &= \sum_{[i,j]\in R} f(i,j)g(i,j) \end{aligned}\right\} \text{Most popular}$$

Correlation C_{fg}

$$C_{fg} = \sum_{[i,j] \in R} f(i,j) g(i,j)$$

If we are doing exhaustive search over all image patches in the second image, this becomes cross-correlation of a template with an image. We have seen this before – imfilter(im,template,'corr').

Example

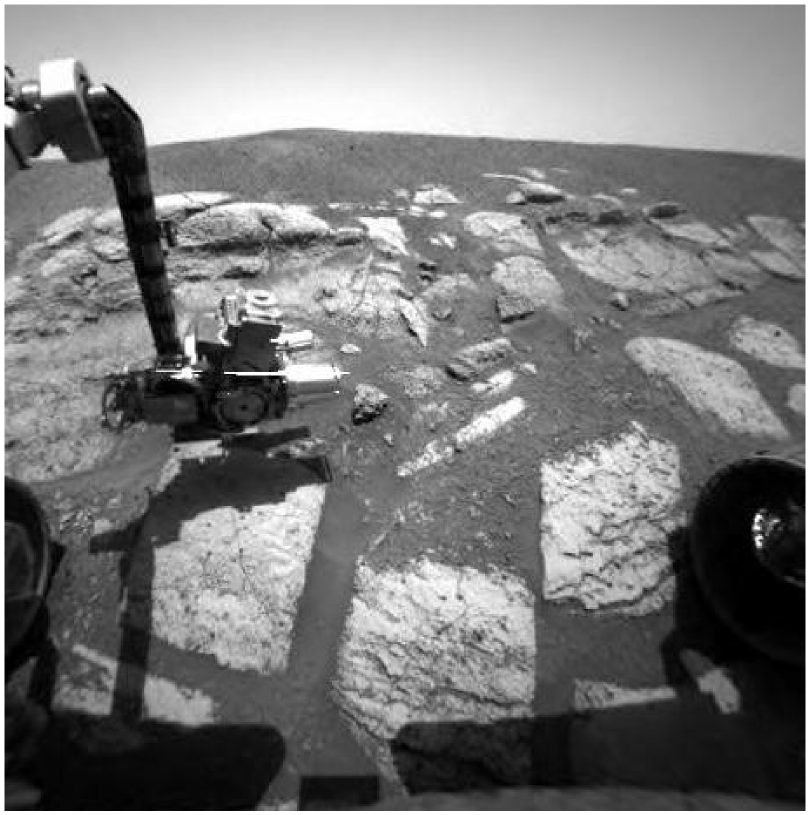


Image 1

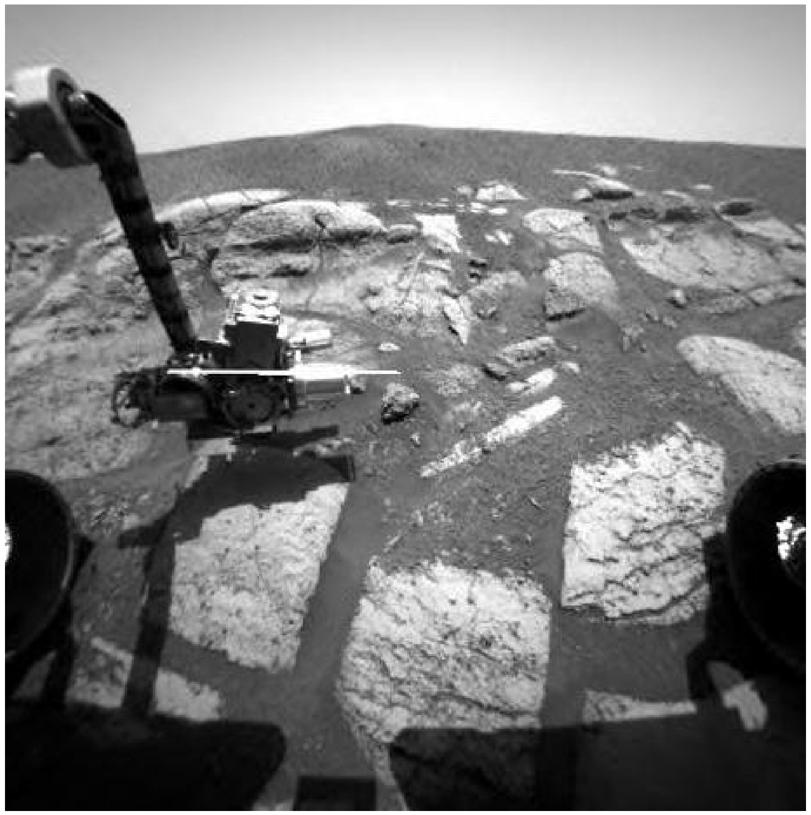


Image 2

Note: this is a stereo pair from the NASA mars rover.
The rover is exploring the "El Capitan" formation.

Example

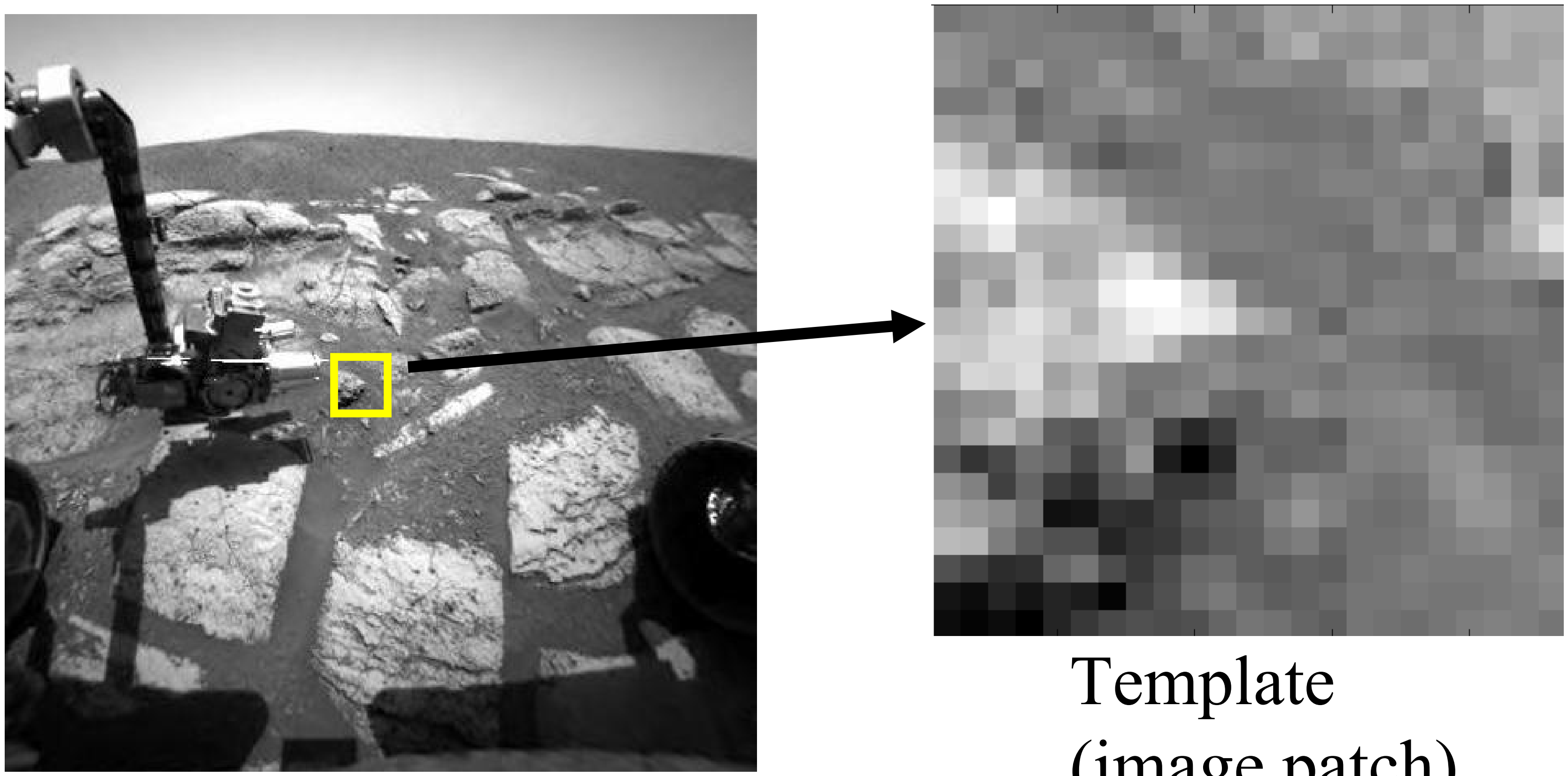
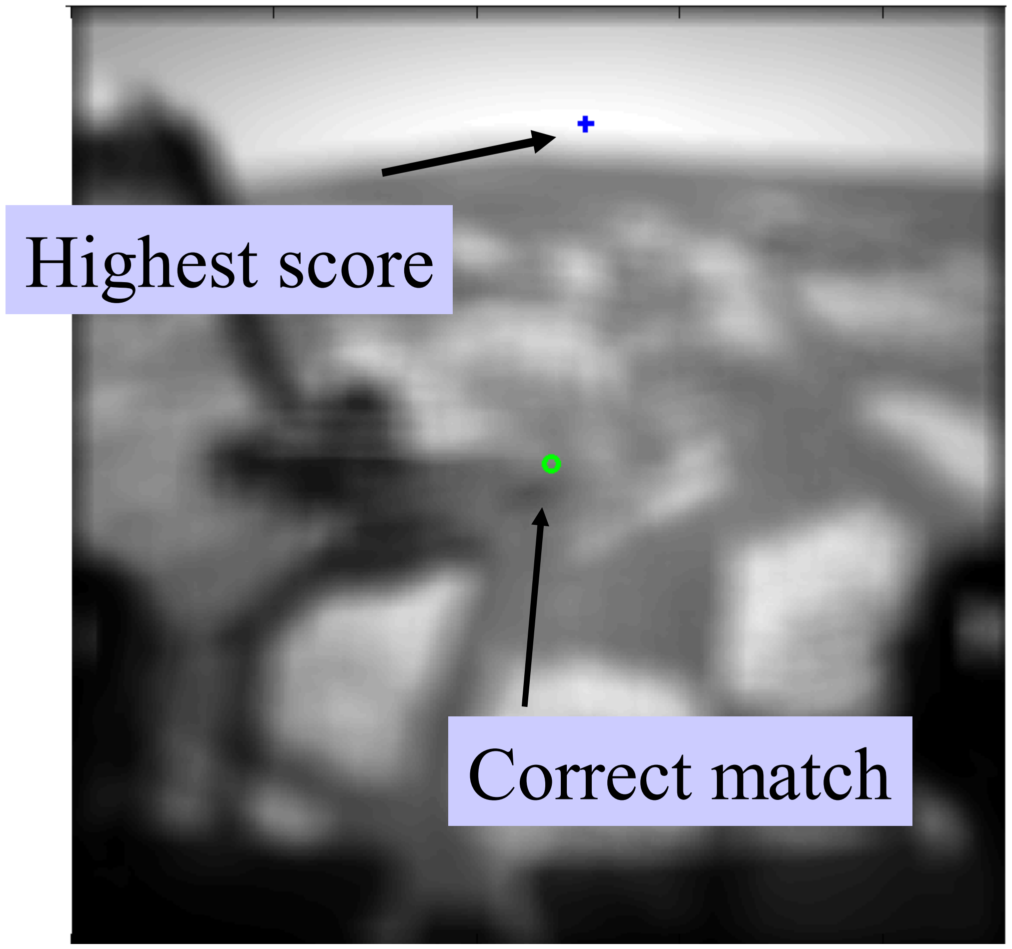


Image 1

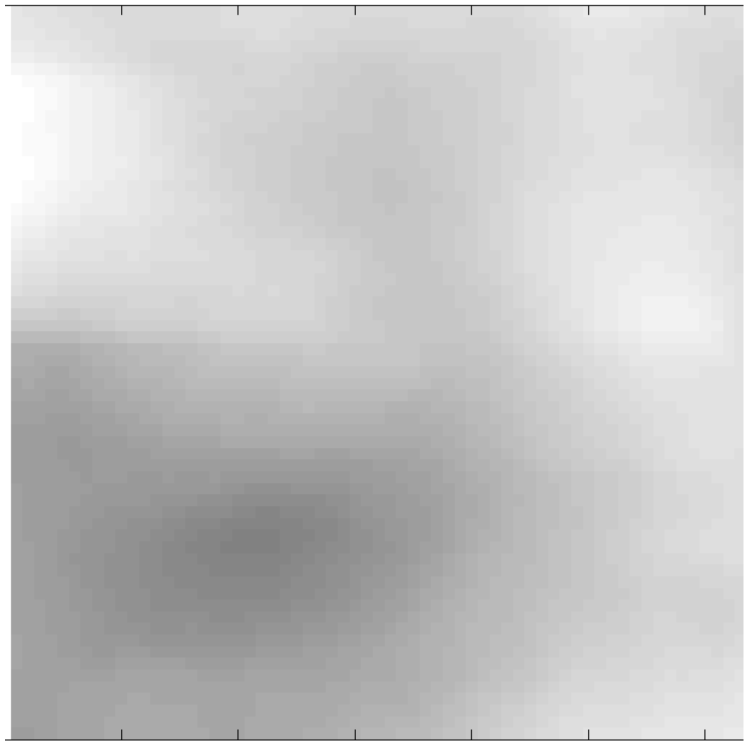
Template
(image patch)

Example: Raw Cross-correlation

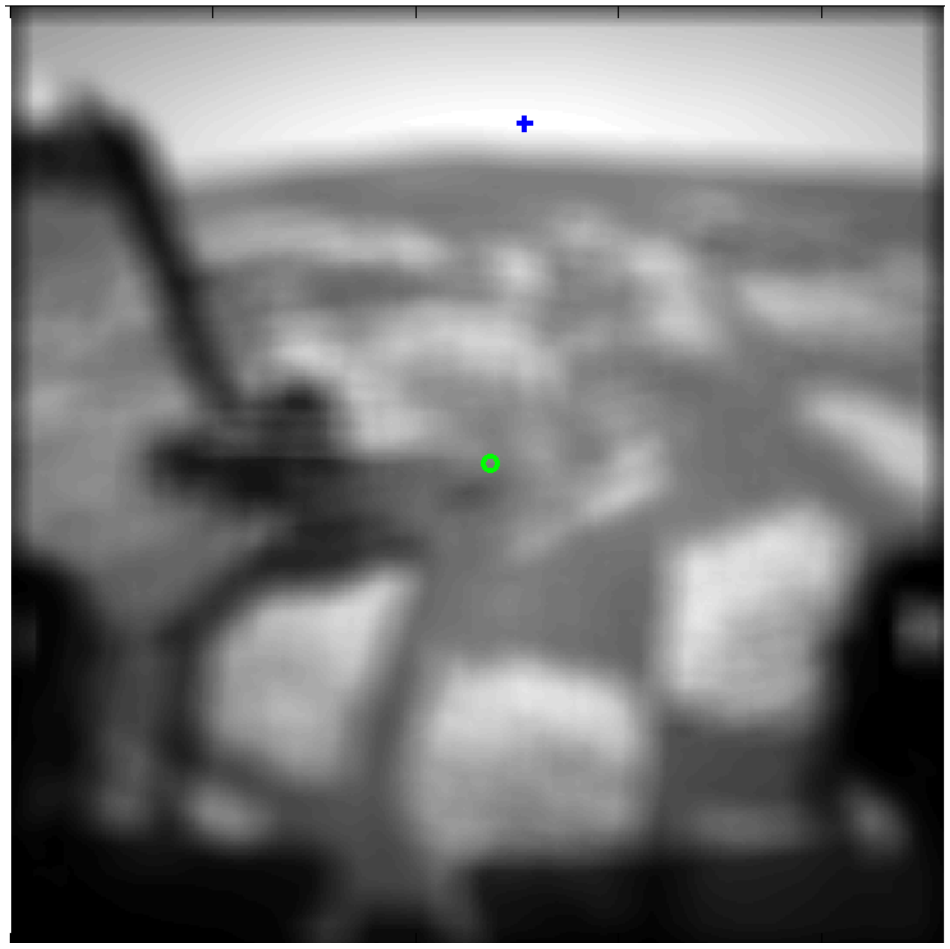
score = imfilter(image2,tmpl)



Score around correct match



Example: Cross-correlation



Note that score image looks a lot like a blurry version of image 2.

This clues us in to the problem with straight correlation with an image template.

Problem with Correlation of Raw Image Templates

Consider correlation of template with an image of constant grey value:

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline d & e & f \\ \hline g & h & i \\ \hline \end{array} \otimes \begin{array}{|c|c|c|} \hline v & v & v \\ \hline v & v & v \\ \hline v & v & v \\ \hline \end{array}$$

Result: $v*(a+b+c+d+e+f+g+h+i)$

Problem with Correlation of Raw Image Templates

Now consider correlation with a constant image that is twice as bright.

$$\begin{array}{|c|c|c|}\hline a & b & c \\ \hline d & e & f \\ \hline g & h & i \\ \hline \end{array} \otimes \begin{array}{|c|c|c|}\hline 2v & 2v & 2v \\ \hline 2v & 2v & 2v \\ \hline 2v & 2v & 2v \\ \hline \end{array}$$

Result: $2*v*(a+b+c+d+e+f+g+h+i)$

$> v*(a+b+c+d+e+f+g+h+i)$

Larger score, regardless of what the template is!

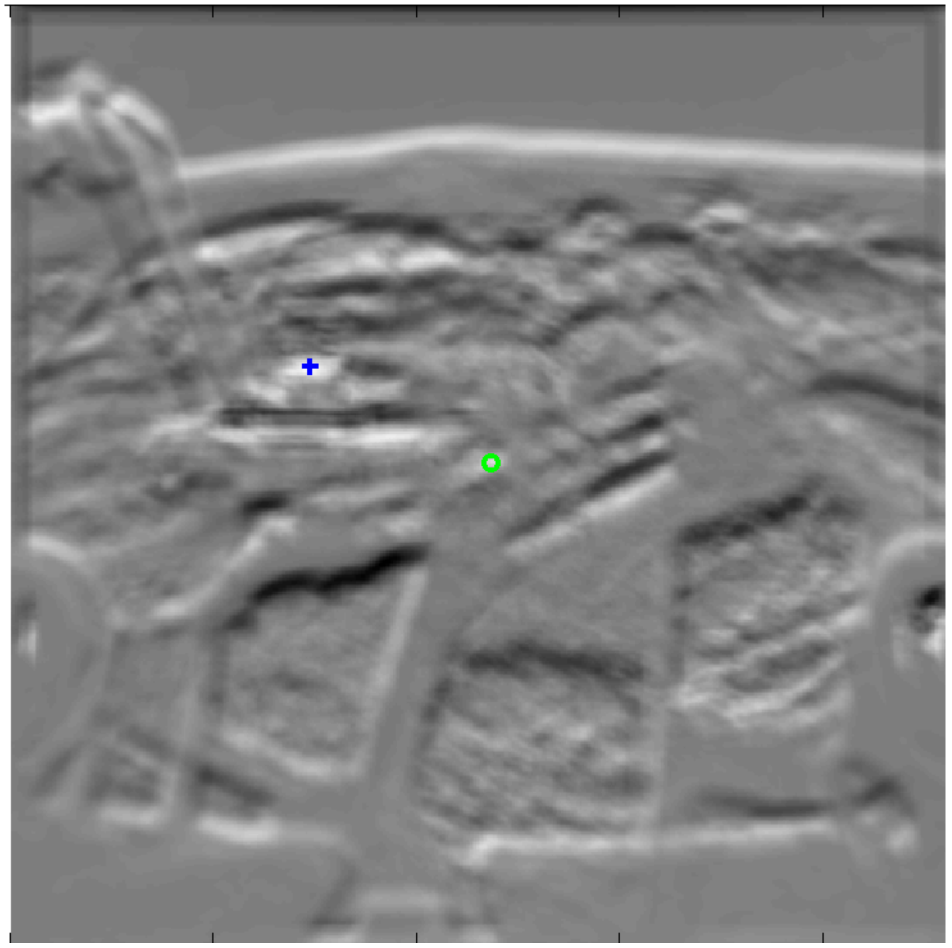
Solution

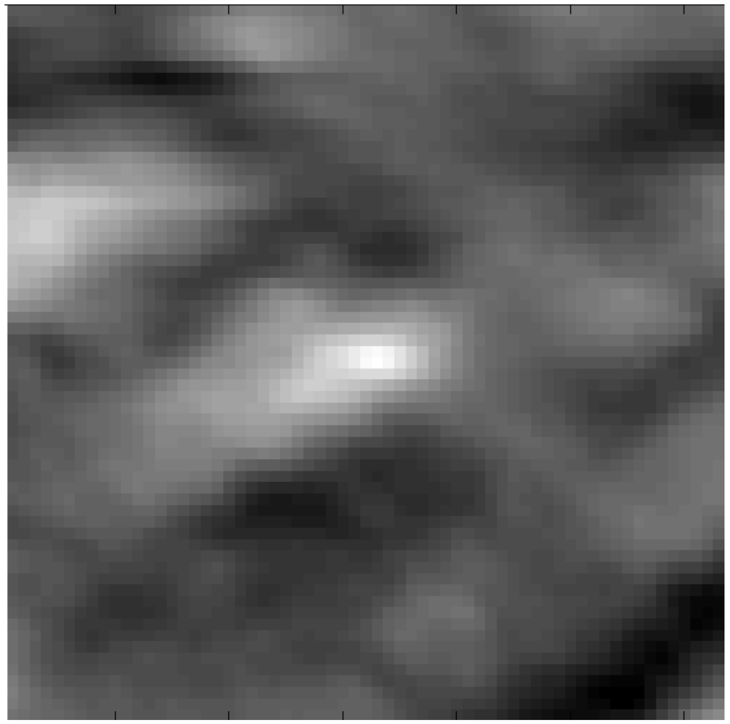
Subtract off the mean value of the template.

In this way, the correlation score is higher only when darker parts of the template overlap darker parts of the image, and brighter parts of the template overlap brighter parts of the image.

Robert Collins
CSE486, Penn State

Correlation, zero-mean template





Better! But highest score is still not the correct match.

Note: highest score IS best within local neighborhood of correct match.

“SSD” or “block matching” (Sum of Squared Differences)

$$\sum_{[i,j]\in R} (f(i,j) - g(i,j))^2$$

1) The most popular matching score.
2) We used it when deriving Harris corners
3) T&V claim it works better than cross-correlation

Relation between SSD and Correlation

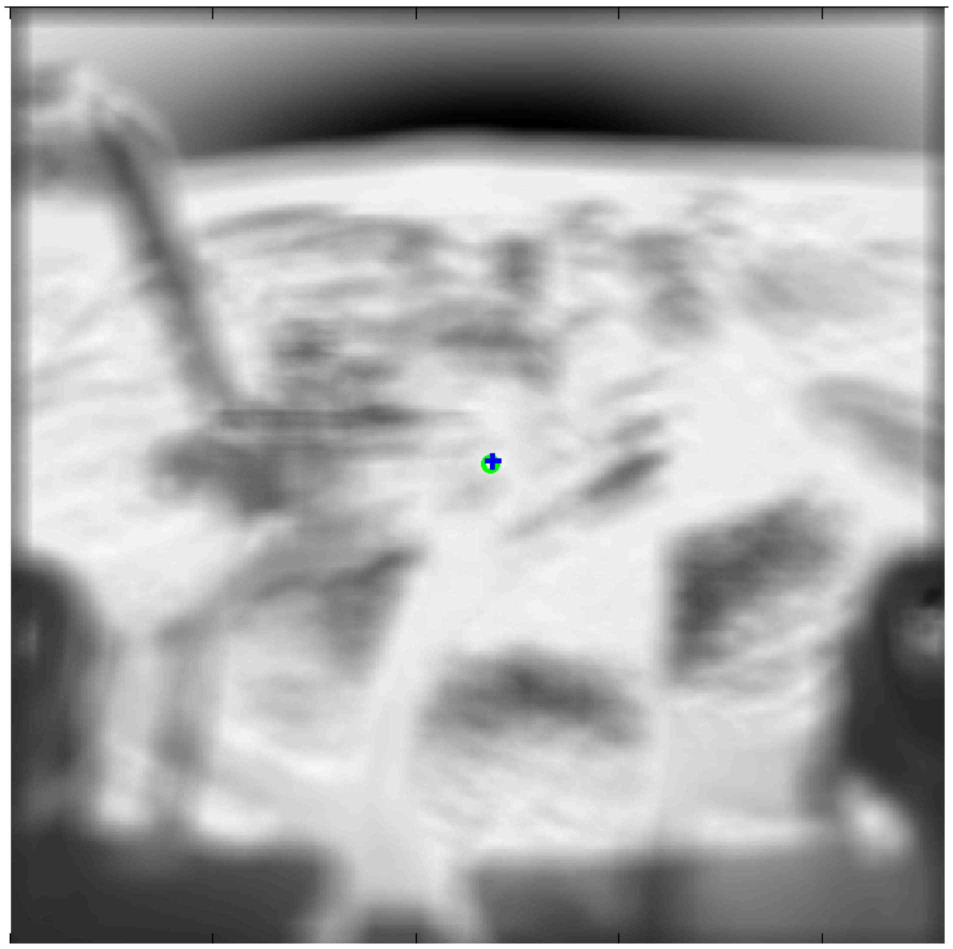
$$SSD = \sum_{[i,j]\in R} (f - g)^2$$

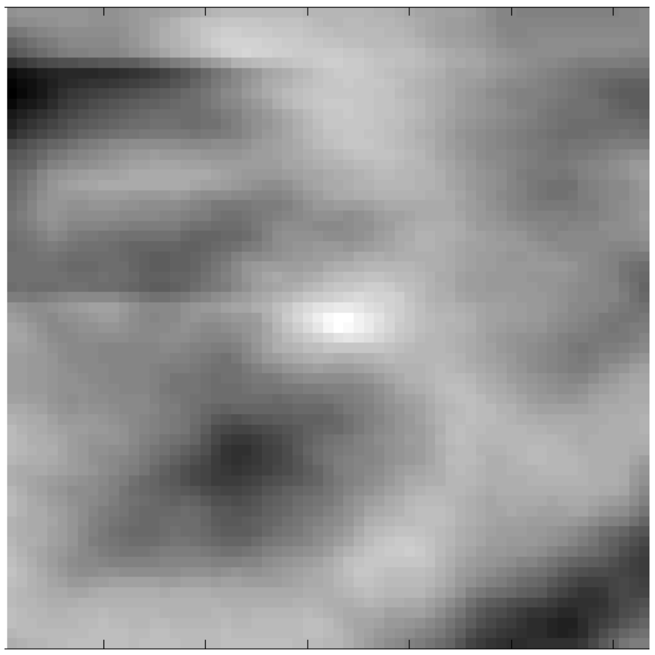
$$= \sum_{[i,j]\in R} f^2 + \sum_{[i,j]\in R} g^2 - 2 \boxed{\sum_{[i,j]\in R} fg}$$

$$C_{fg} = \sum_{[i,j]\in R} f(i,j)g(i,j)$$

Correlation!

SSD





Best match (highest score) in image coincides with correct match in this case!

Handling Intensity Changes

Intensity Changes:

•the camera taking the second image might have different intensity response characteristics than the camera taking the first image

•Illumination in the scene could change

•The camera might have auto-gain control set, so that it's response changes as it moves through the scene.

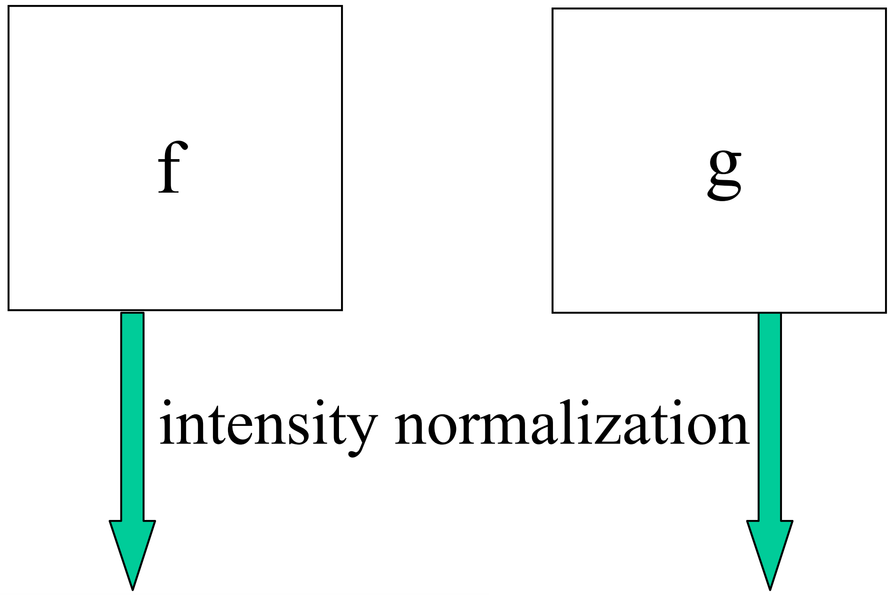
Robert Collins
CSE486, Penn State

Intensity Normalization

- When a scene is imaged by different sensors, or under different illumination intensities, both the SSD and the C_{fg} can be large for windows representing the same area in the scene!
- A solution is to **NORMALIZE** the pixels in the windows before comparing them by subtracting the mean of the patch intensities and dividing by the std.dev.

$$\hat{f} = \frac{f - \bar{f}}{\sqrt{\sum (f - \bar{f})^2}} \qquad \hat{g} = \frac{g - \bar{g}}{\sqrt{\sum (g - \bar{g})^2}}$$

Normalized Cross Correlation

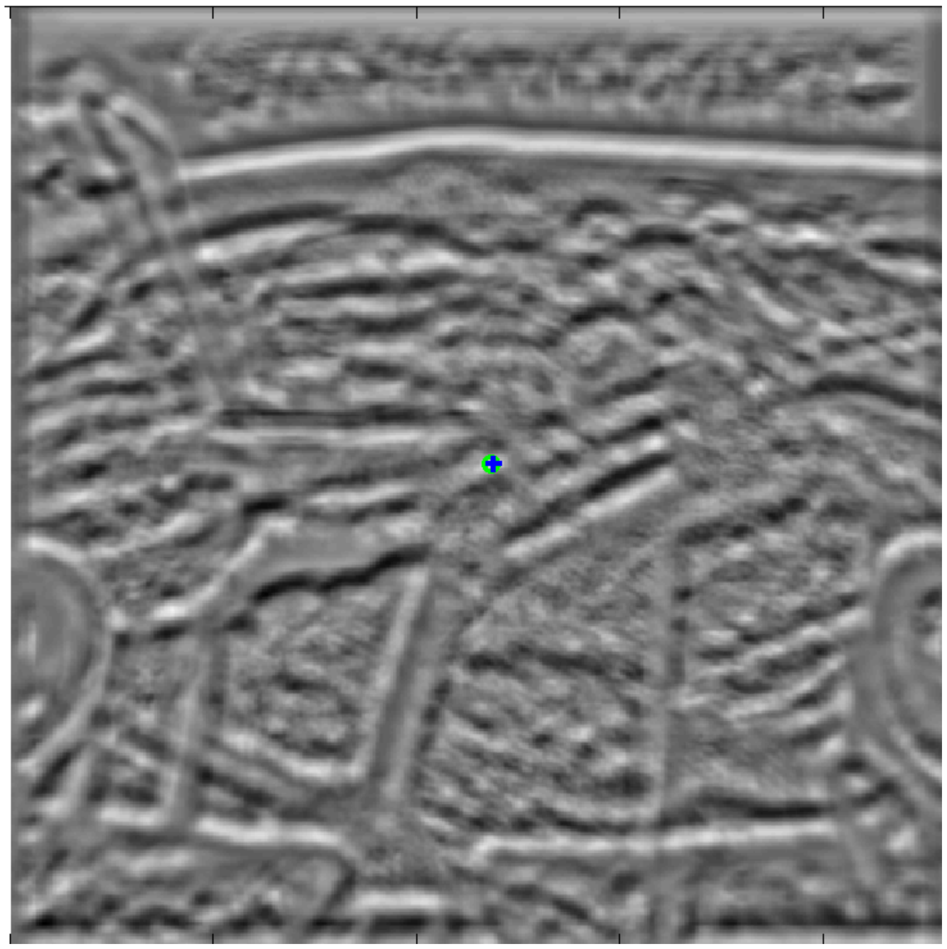


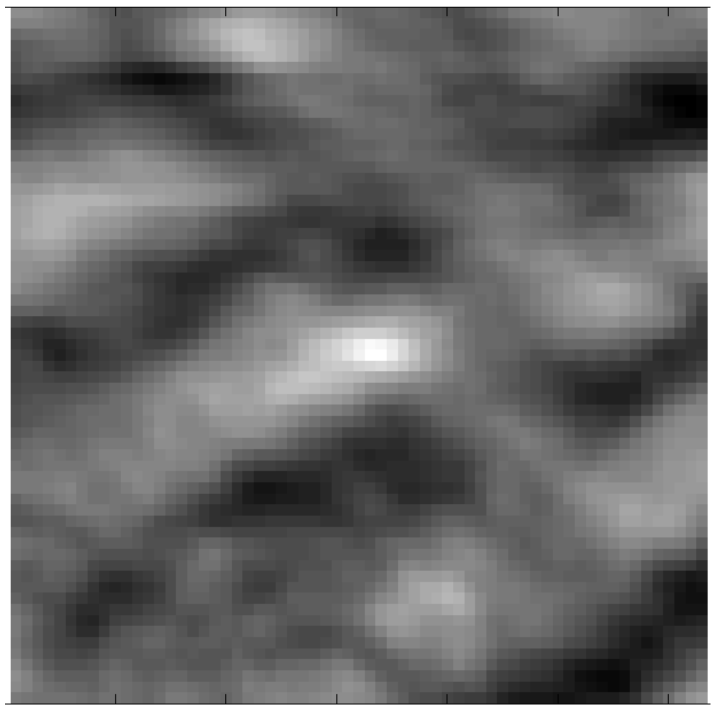
$$\hat{f} = \frac{f - \bar{f}}{\sqrt{\sum (f - \bar{f})^2}} \qquad \hat{g} = \frac{g - \bar{g}}{\sqrt{\sum (g - \bar{g})^2}}$$

$$\boxed{\mathrm{NCC(f,g)} = C_{fg}(\hat{f}, \hat{g}) = \sum_{[i,j] \in R} \hat{f}(i,j)\hat{g}(i,j)}$$

Robert Collins
CSE486, Penn State

Normalized Cross Correlation





Highest score also coincides with correct match.
Also, looks like less chances of getting a wrong match.

Normalized Cross Correlation

Important point about NCC:
Score values range from 1 (perfect match)
to -1 (completely anti-correlated)

Intuition: treating the normalized patches as vectors, we see they are unit vectors. Therefore, correlation becomes dot product of unit vectors, and thus must range between -1 and 1.