Linear Correlation Filters

CS 510

Lecture #19

March 4th, 2013

Colorado State University

Assignment #2

- Rough draft of report due today
  - Mail it to me
- Any questions?



Colorado State University

Where are we?

- We have looked at image-level matching
  - Whole to whole
  - Part to whole
  - True it is only the tip of the iceberg…
- Next step : feature extraction/matching
- Intermediate step: correlation filters
  - Trained to match specific features.

Correlation Filters

- A *correlation filter* is just a template that you correlate with images
- In Assignment #1, the example eye, ear, etc. were correlation filters
  - But they weren't optimal
  - They were just examples
- How do you create an optimal filter?

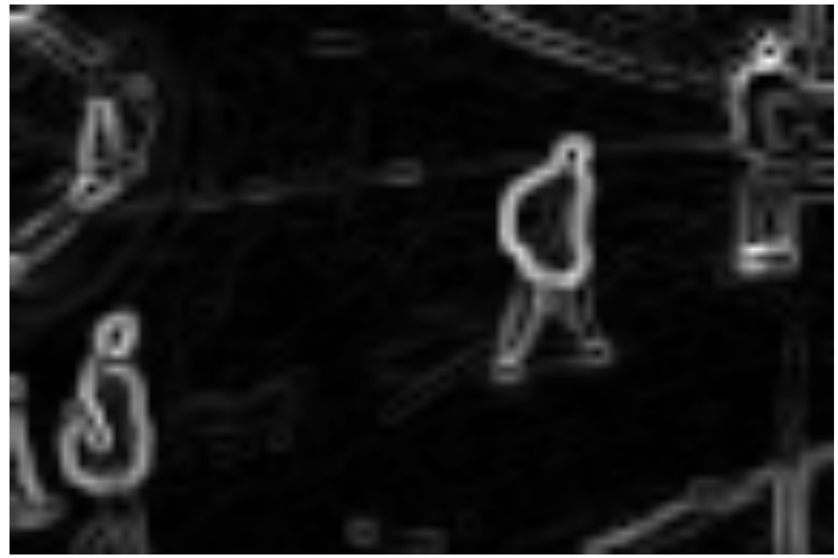
Motivating Example

Colorado State University

Side Issue : Edge Detection

- Clothing appears in all colors/intensities
- To focus on structure, extract edge magnitudes
  - Convolve with Sobel edge masks
  - Compute Dx & Dy for every pixel
  - Edge magnitude is $\sqrt{(Dx^2 + Dy^2)}$
- Remember: linear filter

Edge Detection in Practice



Main Issues

- Intra-class variation
- Changes in illumination
- Changes in pose
- Differing backgrounds

The challenge: find the pixels the separate targets from background

Simple Example

- Linear correlation is just a dot product
- One goal is for positive samples to produce a positive score (1)
- The other is for negative samples to produce zero scores (0)
- So to find eyes…

 $\cdot W = 1$

 $\cdot W = 0$

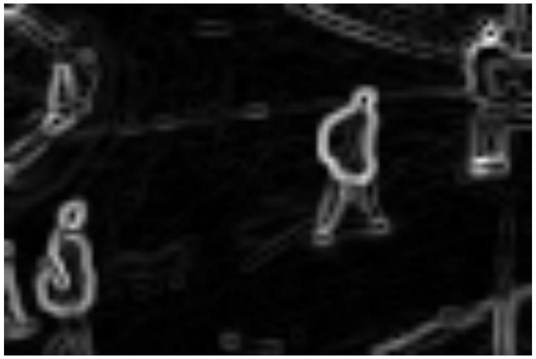
Example (cont.)

- How many solutions are there to the previous example?
- How could you make the solution unique?
- Would it necessarily generalize?
- There is a family of techniques that take
  - Positive examples
  - Negative examples
  - Additional constraint

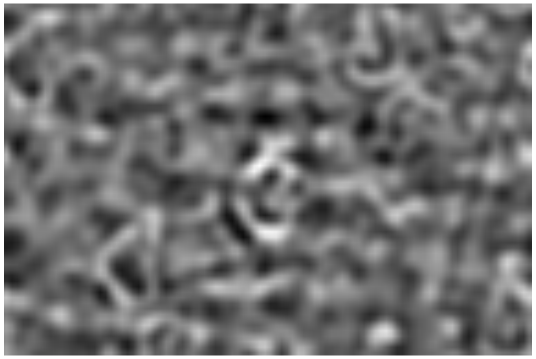
CSU Approach (Bolme)

- Don’t pick a handful of examples
- Use the (whole) desired response image as the training image
- Put a Gaussian at every target location
  - Sigma 2 is typical
  - Allows response to degrade smoothly with distance

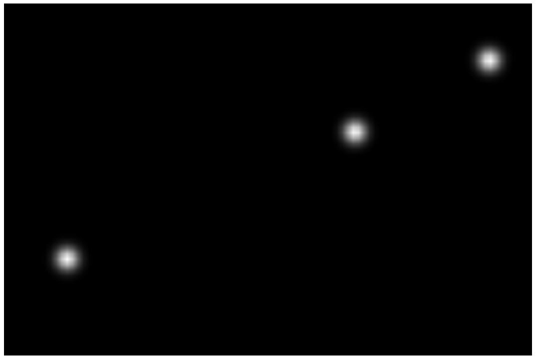
Exact Filter



Edge Image (f)



Exact Filter (h)



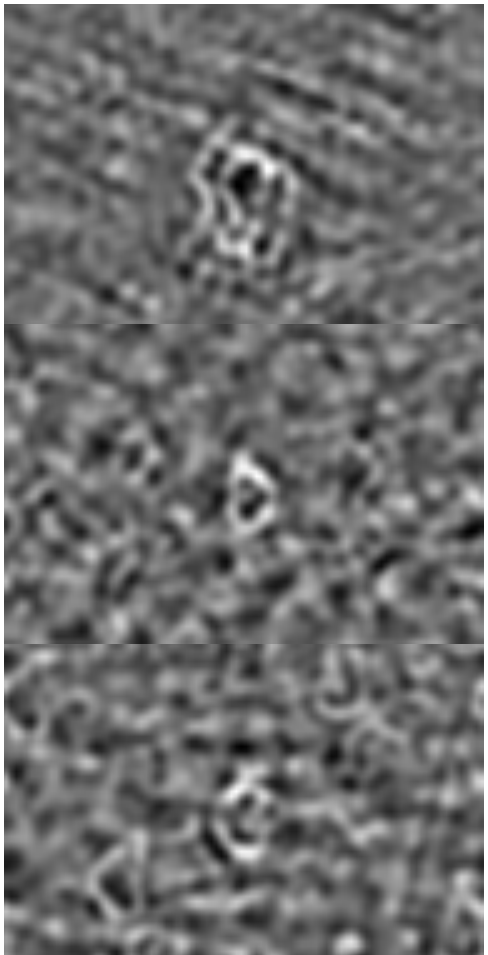
Output (g)

$$F \cdot H^{*} = G$$

$$H^{*} = \frac{F}{G}$$

Average of Exact Synthetic Filters

Exact filters in spatial domain



$$H^* = \frac{1}{N}\sum_i \frac{F_i}{G_i}$$

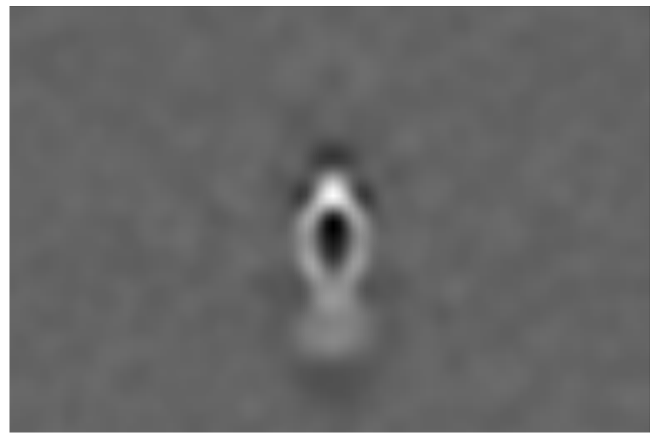


Image is of h, not H

Training statistics

- 25 annotated target frames
- Training time: 12 seconds
- Result: template
  - In frequency space

Detection

- Correlate trained template to every video frame
  - Use frequency domain to speed computation
- Find peaks in correlation images
  - Keep peeks that exceed a threshold

Show Video

MOSSE filter

- Minimize Output Sum of Squared Errors

$$H^* = \min_{H^*} \sum_i \left| F_i \cdot H^* - G_i \right|^2$$

- This form is more stable for small numbers of training samples

$$H = \frac{\sum_i G_i \cdot F^*{}_i}{\sum_i F_i \cdot F^*{}_i + \varepsilon}$$