MOSSE

CS 510
Lecture #13
February 22, 2017

Colorado State University

Simple Tracking

- To track an object from one frame to the next:
  - Find the moving object in frame t
  - Cross-correlation the attention window against frame t+1
  - The best-matching position is where the object went
    - Drop track if below minimum threshold.
  - Image window at t+1 position becomes new target; iterate for frame t+1

Colorado State University

Can we do better?

- Is there something better we can compare to than the raw attention window?
  - *Sometimes edges are better to track than pixels*
  - *Is there something better still?*
- Do we have to search all of frame t+1?
  - Can we limit the search?
  - Predict where the object is headed?
  - Describe the object's motion?
  - Exploit foreground information in frame t+1?
- What if we don't find the target in frame t+1?

Colorado State University

Optimized Correlation Output Filters

Ph.D. thesis by David Bolme
Colorado State University,
Dec., 2010.

US Patent 8,520,956. David S. Bolme, J. Ross Beveridge and Bruce A. Draper, Optimized Correla- tion Filters for Signal Processing, August 27, 2013.

Colorado State University

Intra-class variation

- Challenges for matching & tracking
  - Changes in shape/pose/viewpoint
  - Changes in apparent color
- Goal: learn a general template
  - Capable of matching many samples
  - Within constraints of a linear filter

Colorado State University

Step #1: Edge Detection

- 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

Colorado State University

Edge Detection in Practice

Colorado State University

Simple Templates

- Cutting a template from an example doesn't work…

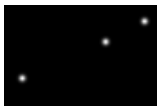
Edge Image | Edge Template | Correlation

Colorado State University

Training

Edge Image (f)

Unknown Template (h)



Desired Output (g)

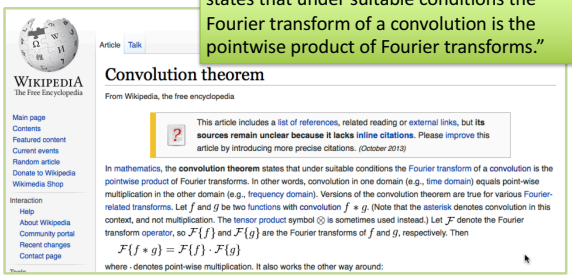
$$f \otimes h = g$$

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

Colorado State University

Convolution Theorem (review)

"In mathematics, the convolution theorem states that under suitable conditions the Fourier transform of a convolution is the pointwise product of Fourier transforms."



WIKIPEDIA
The Free Encyclopedia

Article Talk

Convolution theorem

From Wikipedia, the free encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikimedia Shop

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

This article includes a list of references, related reading or external links, but its sources remain unclear because it lacks inline citations. Please improve this article by introducing more precise citations. (October 2013)

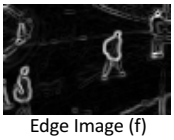
In mathematics, the **convolution theorem** states that under suitable conditions the Fourier transform of a convolution is the pointwise product of Fourier transforms. In other words, convolution in one domain (e.g., time domain) equals point-wise multiplication in the other domain (e.g., frequency domain). Versions of the convolution theorem are true for various Fourier-related transforms. Let f and g be two functions with convolution $f * g$. (Note that the asterisk denotes convolution in this context, and not multiplication. The tensor product symbol $\otimes$ is sometimes used instead.) Let $\mathcal{F}$ denote the Fourier transform operator, so $\mathcal{F}(f)$ and $\mathcal{F}(g)$ are the Fourier transforms of f and g, respectively. Then

$$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$$

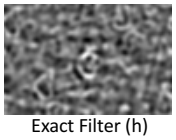
where · denotes point-wise multiplication. It also works the other way around:

Colorado State University

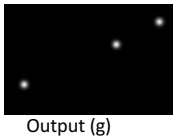
Exact Filter



Edge Image (f)



Exact Filter (h)



Output (g)

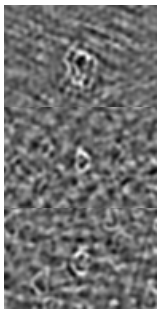
$$F \cdot H^* = G$$

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

Colorado State University

Average of Exact Synthetic Filters

Exact filters in spatial domain



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

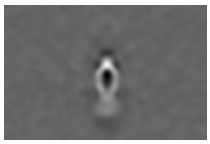


Image is of h, not H

Colorado State University

Detection

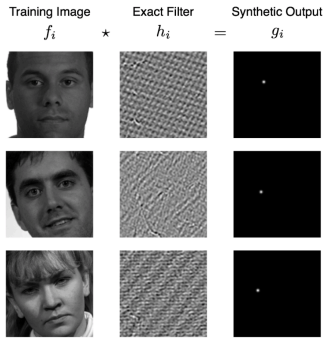
- Correlate trained template to every video frame
  - Use frequency domain to speed computation
- Find peaks in correlation images
  - Keep peeks that exceed a threshold

Colorado State University

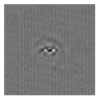
First Example Video - YouTube

Colorado State University

First Application: Eyes



$$\frac{1}{N}\sum_i h_i =$$

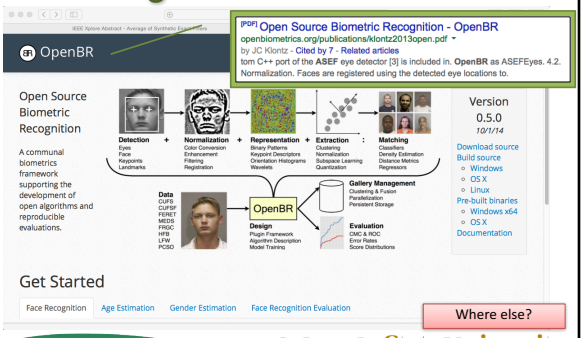


ASEF Filter

Bolme, D.S.; Draper, B.A.; Beveridge, J.R., "Average of Synthetic Exact Filters," *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* , vol., no., pp.2105,2112, 20-25 June 2009

Colorado State University

Is it being used …



Colorado State University

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

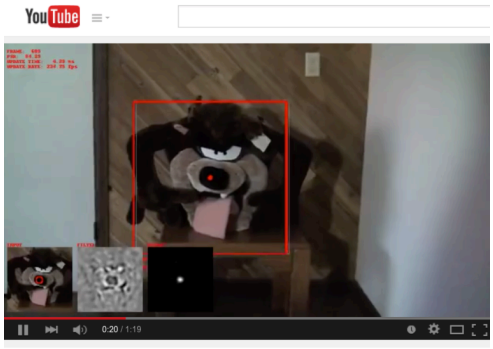
$$H = \frac{\sum_i G_i \cdot F^*_i}{\sum_i F_i \cdot F^*_i + \varepsilon}$$

Colorado State University

MOSSE Filter Tracking

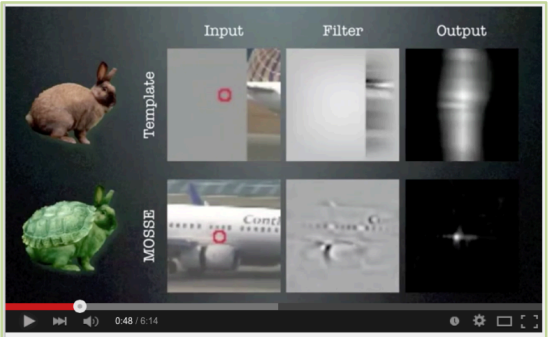
- User selects initial window to track
- Train filter on
  - Initial window
  - Small affine transformations of initial window
- Update filter using
  - Previous filter
  - If tracked, add small transformations of current window

Colorado State University



MOSSE Track : Correlation filter based tracking

Colorado State University



CVPR2010 Tracking Highlights

Colorado State University