MOSSE

CS 510

Lecture #12

February 14, 2018

Colorado State University

Simple Tracking

To track an object from one frame to the next:

- Find window with moving object in frame t
- Correlate window to all locations in frame t+1
- Move tracking window to best match
  - Drop track correlation too low.
- Image window at t+1 position becomes new target; iterate for frame t+1

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?

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

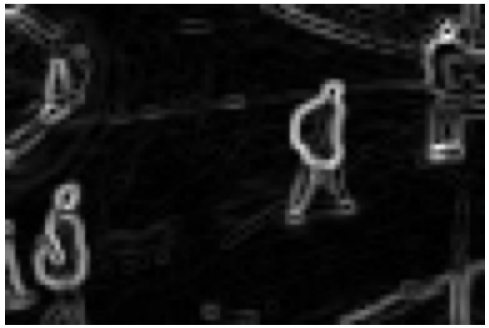
Step #1: Edge Detection

- 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

Edge Detection in Practice

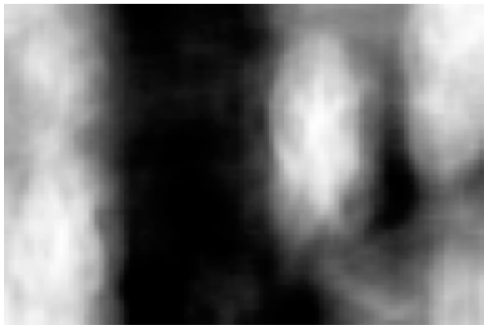
Simple Templates

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



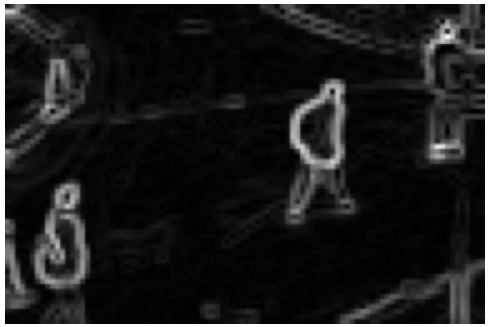
Edge Image

Edge Template

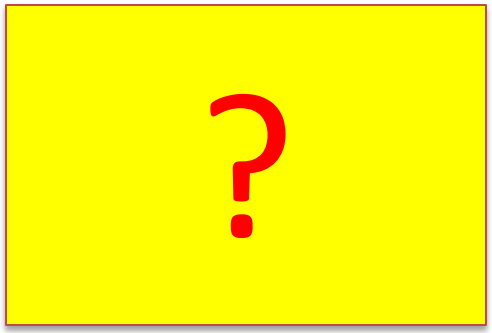


Correlation

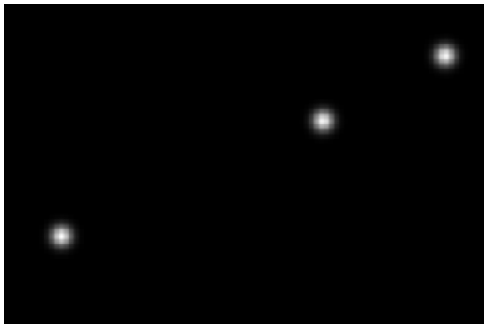
Training



Edge Image (f)



Unknown Template (h)



Desired Output (g)

$$f \otimes h = g$$

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

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

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

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

Article Talk

Convolution theorem

From Wikipedia, the free encyclopedia

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)*

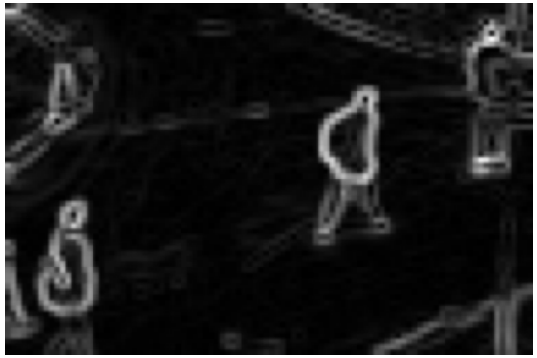
In mathematics, the **convolution theorem** states that under suitable conditions the Fourier transform of a convolution is the pointwise product of Fourier transforms. In other words, convolution in one domain (e.g., time domain) equals point-wise multiplication in the other domain (e.g., frequency 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

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

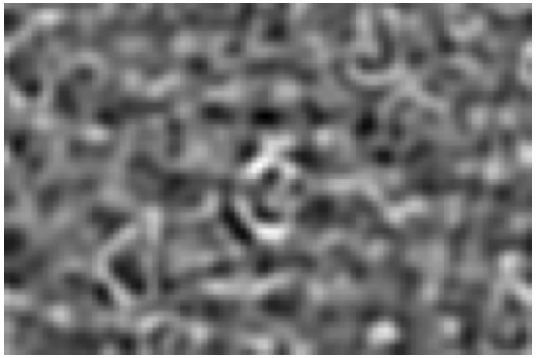
where $\cdot$ denotes point-wise multiplication. It also works the other way around:

Colorado State University

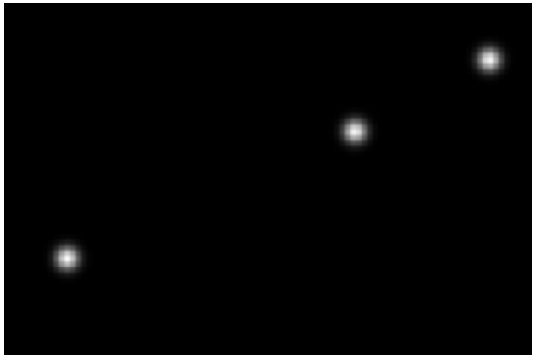
Exact Filter



Edge Image (f)



Exact Filter (h)



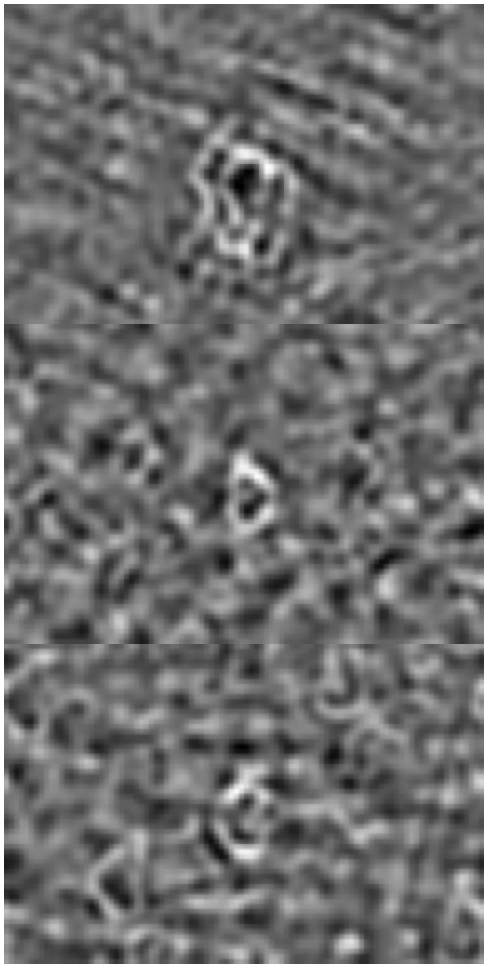
Output (g)

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

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

Average of Exact Synthetic Filters

Exact filters in spatial domain



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

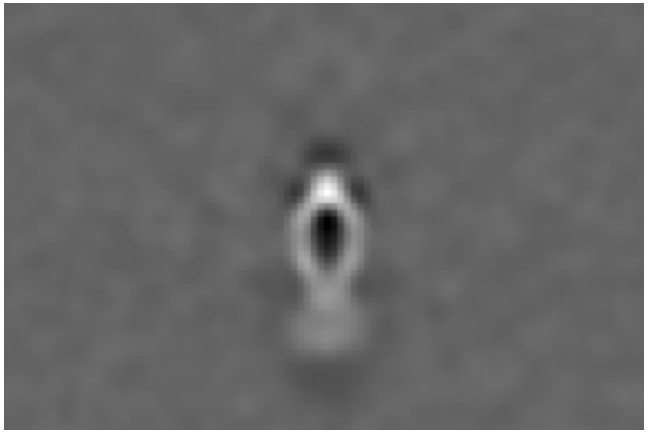


Image is of h, not H

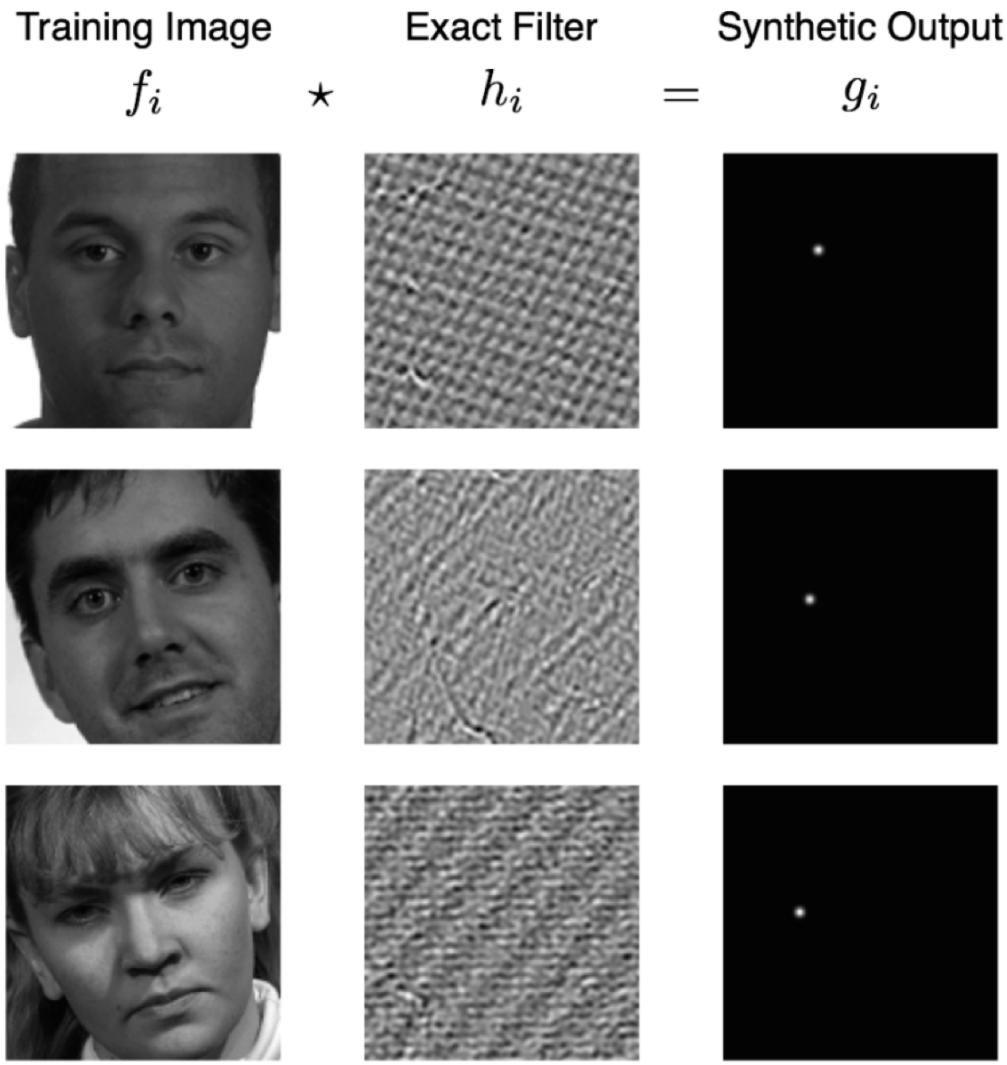
Detection

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

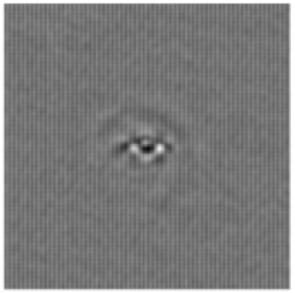
First Example Video - YouTube

Person Detection ASEF S2L1_T1234

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

Is it being used …

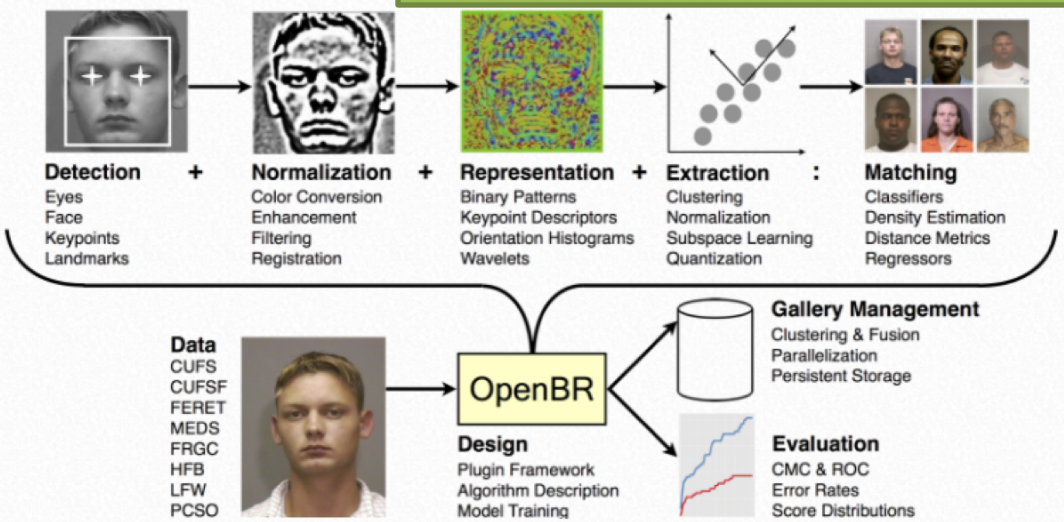
IEEE Xplore Abstract - Average of Synthetic Exact Filters

OpenBR

[PDF] Open Source Biometric Recognition - OpenBR
openbiometrics.org/publications/klontz2013open.pdf
by JC Klontz - Cited by 7 - Related articles
tom C++ port of the **ASEF** eye detector [3] is included in. **OpenBR** as ASEFEyes. 4.2. Normalization. Faces are registered using the detected eye locations to.

Open Source Biometric Recognition

A communal biometrics framework supporting the development of open algorithms and reproducible evaluations.



Version
0.5.0
10/1/14

Download source
Build source
- Windows
- OS X
- Linux

Pre-built binaries
- Windows x64
- OS X

Documentation

Get Started

Face Recognition | Age Estimation | Gender Estimation | Face Recognition Evaluation

Where else?

Colorado State University

MOSSE filter

- Minimize Output Sum of Squared Errors

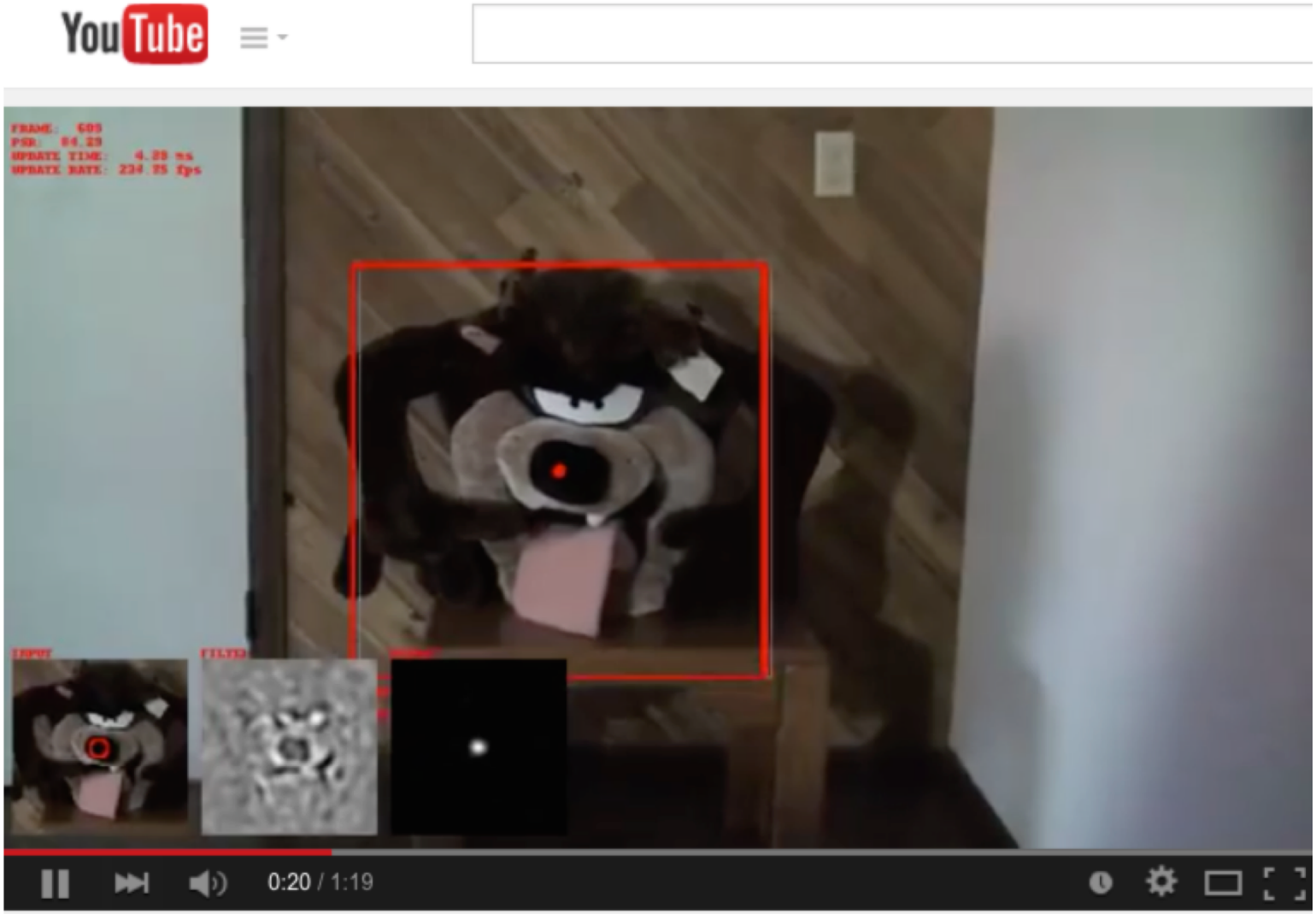
$$H^* = \min_{H^*} \sum_i \left| F_i \cdot H^* - G_i \right|^2$$

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

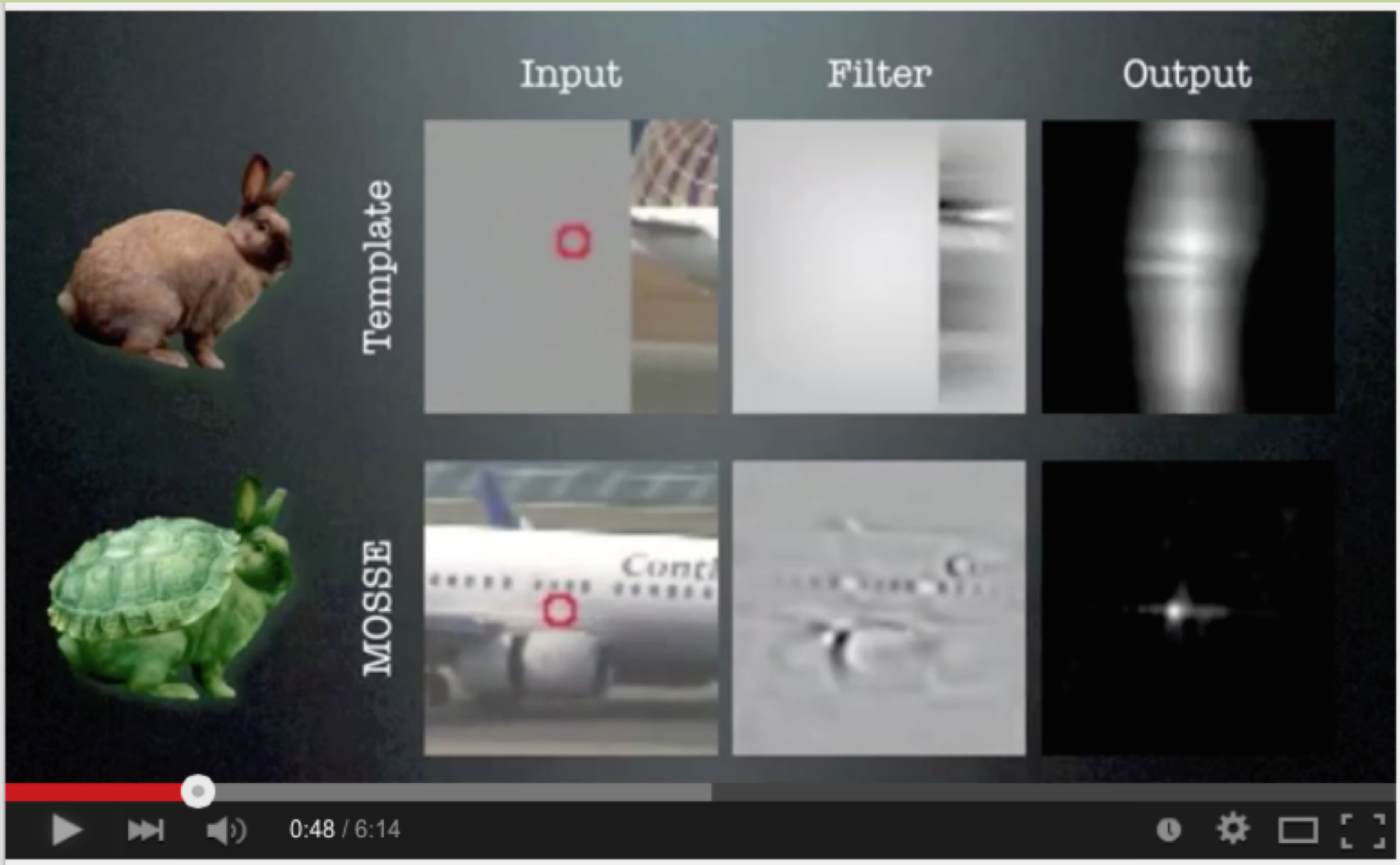
$$H = \frac{\sum_i G_i \cdot F^*_i}{\sum_i F_i \cdot F^*_i + \varepsilon}$$

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



MOSSE Track : Correlation filter based tracking



CVPR2010 Tracking Highlights