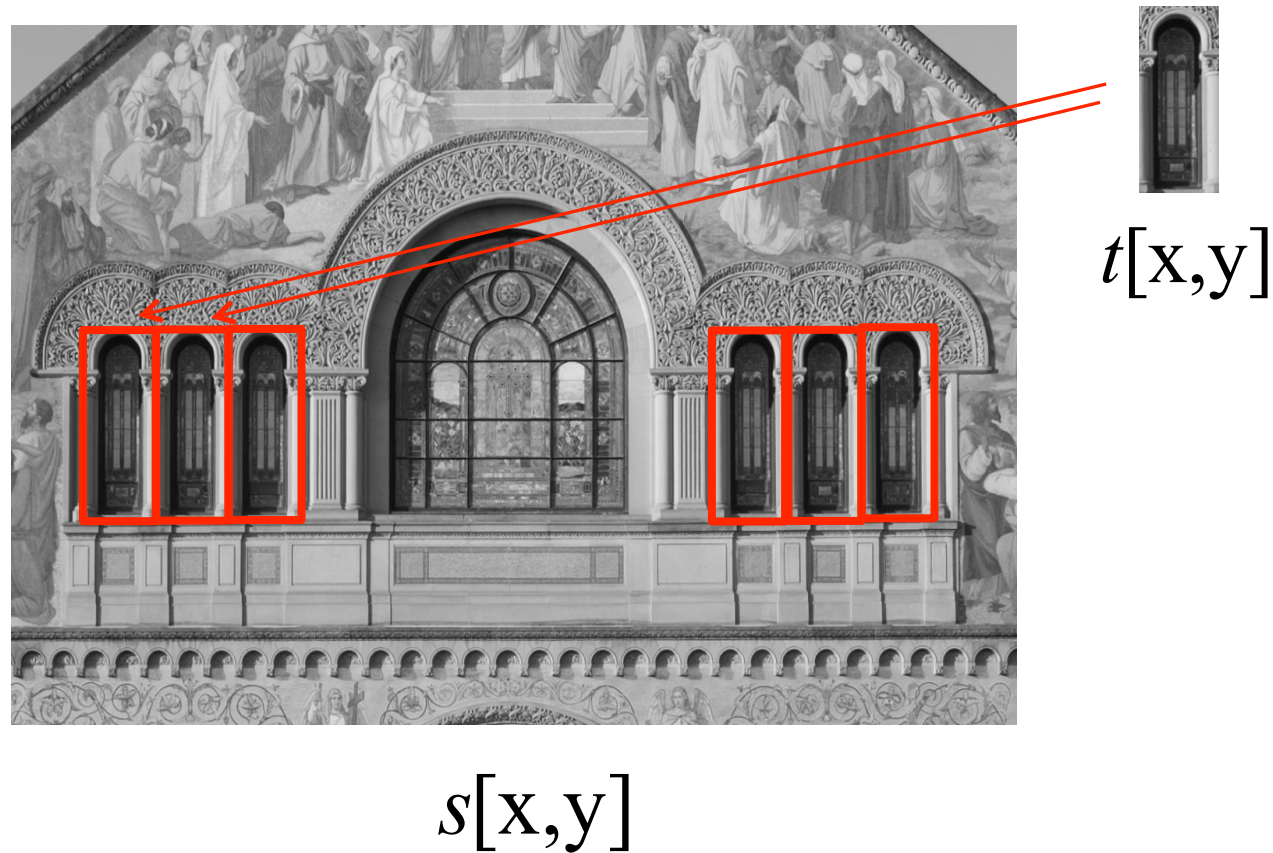


Template matching

- Problem: locate an object, described by a template $t[x,y]$, in the image $s[x,y]$
- Example



Template matching (cont.)

- Search for the best match by minimizing mean-squared error

$$\begin{aligned} E[p, q] &= \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} [s[x, y] - t[x - p, y - q]]^2 \\ &= \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} |s[x, y]|^2 + \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} |t[x, y]|^2 - 2 \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} s[x, y] \cdot t[x - p, y - q] \end{aligned}$$

- Equivalently, maximize *area correlation*

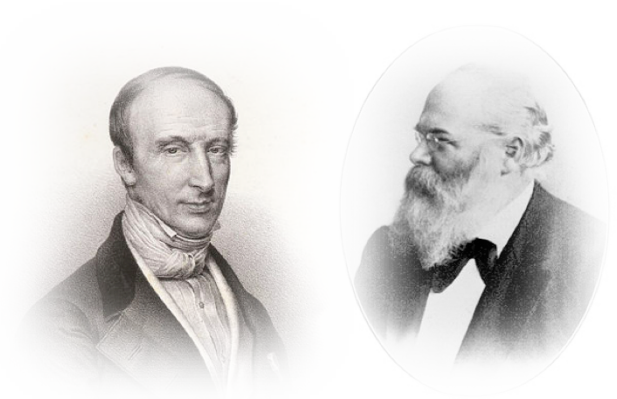
$$r[p, q] = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} s[x, y] \cdot t[x - p, y - q] = s[p, q] * t[-p, -q]$$

- Area correlation is equivalent to convolution of image $s[x, y]$ with impulse response $t[-x, -y]$

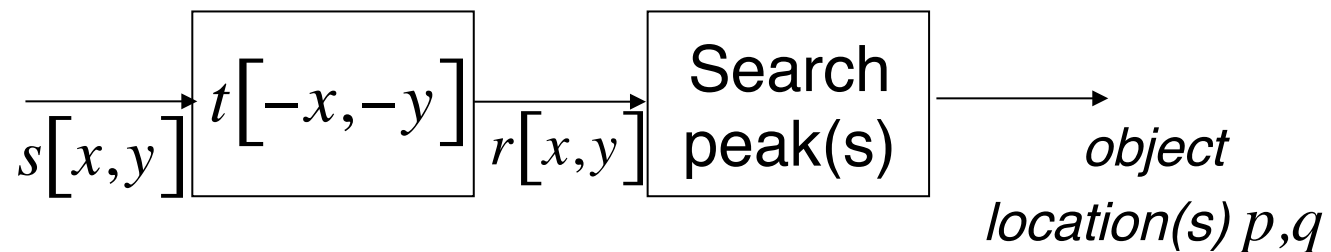
Template matching (cont.)

- From Cauchy-Schwarz inequality

$$r[p, q] = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} s[x, y] \cdot t[x - p, y - q] \leq \sqrt{\left(\sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} |s[x, y]|^2 \right) \left(\sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} |t[x, y]|^2 \right)}$$



- Equality, iff $s[x, y] = \alpha \cdot t[x - p, y - q]$ with $\alpha \geq 0$
- Block diagram of template matcher

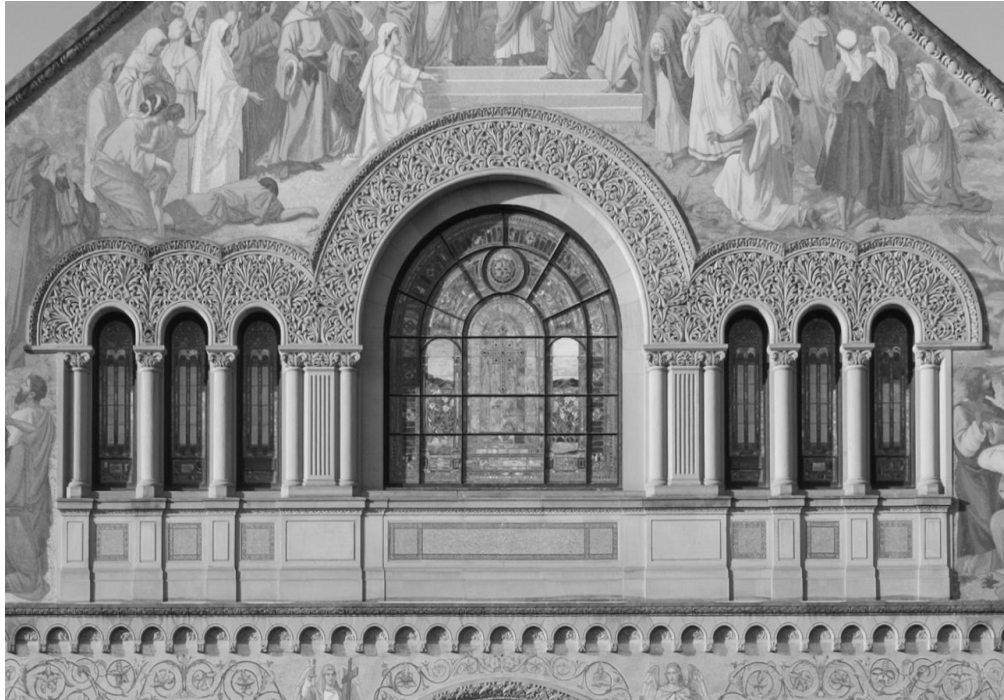


- Remove mean before template matching to avoid bias towards bright image areas

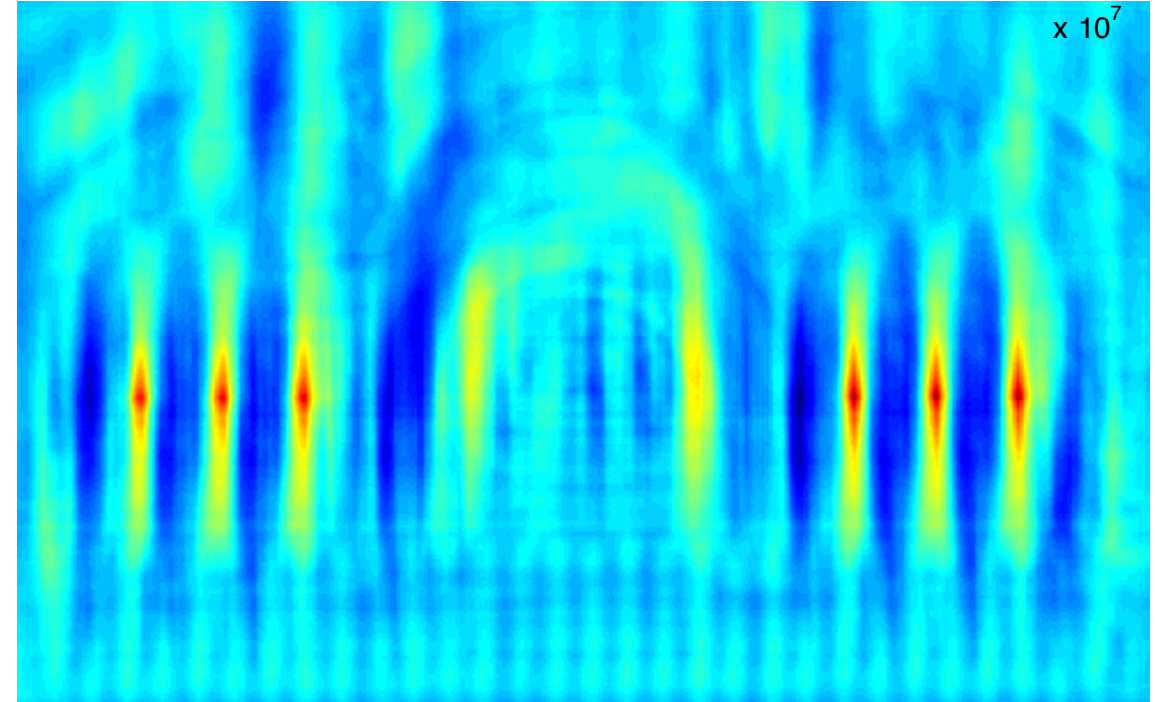
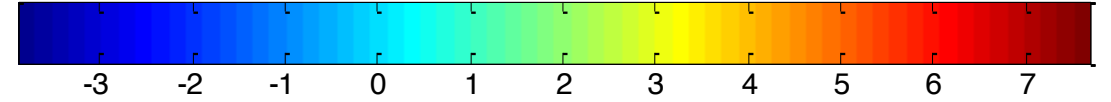
Template matching example



$t[x,y]$



$s[x,y]$

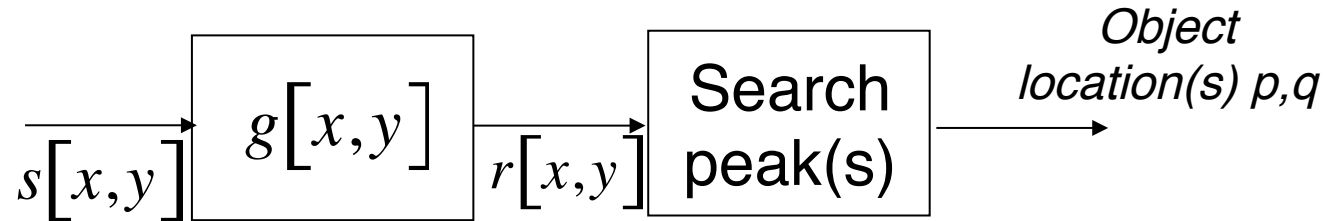


$r[p,q]$



Matched filtering

- Consider signal detection problem



- Signal model

$$s[x,y] = t[x-p, y-q] + n[x,y]$$

Annotations: $t[x-p, y-q]$ is labeled 'shifted template'. $n[x,y]$ is labeled 'Other objects: "noise" or "clutter" psd $\Phi_{nn}(e^{j\omega_x}, e^{j\omega_y})$ '.

- Problem: design filter $g[x,y]$ to maximize

$$SNR = \frac{|r[p,q]|^2}{E\left\{ |n[x,y] * g[x,y]|^2 \right\}}$$

Annotations: $|r[p,q]|^2$ is labeled 'correct peak'. $E\left\{ |n[x,y] * g[x,y]|^2 \right\}$ is labeled 'false readings'.

Vector-matrix formulation

$$r[p,q] = \vec{g}^H \vec{s}$$

$$R_{nn} = E\left\{ \vec{n}\vec{n}^H \right\}$$

Annotation: 'covariance'

$$\vec{s} = \vec{t} + \vec{n}$$

$$SNR = \frac{|r[p,q]|^2}{E\left\{ |\vec{g}^H \vec{n}|^2 \right\}}$$

Matched filtering (cont.)

- Optimum filter has frequency response

$$G(e^{j\omega_x}, e^{j\omega_y}) = \frac{T^*(e^{j\omega_x}, e^{j\omega_y})}{\Phi_{nn}(e^{j\omega_x}, e^{j\omega_y})}$$

- Proof:

$$SNR = \frac{|r[p, q]|^2}{E\{|n[x, y] * g[x, y]|^2\}} \approx \frac{\left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} G(e^{j\omega_x}, e^{j\omega_y}) T(e^{j\omega_x}, e^{j\omega_y}) d\omega_x d\omega_y \right|^2}{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |G(e^{j\omega_x}, e^{j\omega_y})|^2 \Phi_{nn}(e^{j\omega_x}, e^{j\omega_y}) d\omega_x d\omega_y}$$

$$= \frac{\left| \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [G \Phi_{nn}^{1/2}] [\Phi_{nn}^{-1/2} T] d\omega_x d\omega_y \right|^2}{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |G|^2 \Phi_{nn} d\omega_x d\omega_y} \leq \frac{\left[\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |G|^2 \Phi_{nn} d\omega_x d\omega_y \right] \left[\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |T|^2 \Phi_{nn}^{-1} d\omega_x d\omega_y \right]}{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |G|^2 \Phi_{nn} d\omega_x d\omega_y}$$

Cauchy-Schwarz inequality,
with equality, iff $G \Phi_{nn}^{1/2} = \alpha \cdot [\Phi_{nn}^{-1/2} T]^*$

$$= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |T|^2 \Phi_{nn}^{-1} d\omega_x d\omega_y$$

max. SNR

Vector-matrix formulation

$$\vec{g} = R_{nn}^{-1} \vec{t}$$

Matched filtering (cont.)

- Optimum filter corresponds to projection on

$$\vec{g} = R_{nn}^{-1} \vec{t}$$

- Proof:

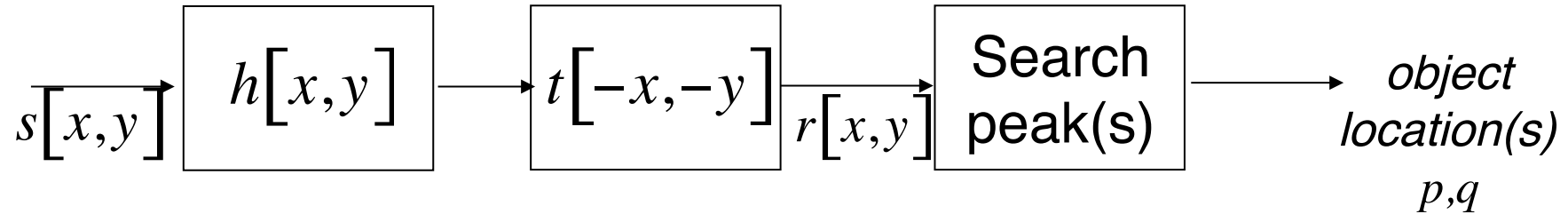
$$\begin{aligned}
 SNR &= \frac{|r[p, q]|^2}{E\left\{|\vec{g}^H \vec{n}|^2\right\}} \approx \frac{|\vec{g}^H \vec{t}|^2}{\vec{g}^H R_{nn} \vec{g}} \\
 &= \frac{\left| \left(R_{nn}^{1/2} \vec{g} \right)^H \left(R_{nn}^{-1/2} \vec{t} \right) \right|^2}{\left(R_{nn}^{1/2} \vec{g} \right)^H \left(R_{nn}^{1/2} \vec{g} \right)} \leq \frac{\left[\left(R_{nn}^{1/2} \vec{g} \right)^H \left(R_{nn}^{1/2} \vec{g} \right) \right] \left[\left(R_{nn}^{-1/2} \vec{t} \right)^H \left(R_{nn}^{-1/2} \vec{t} \right) \right]}{\left(R_{nn}^{1/2} \vec{g} \right)^H \left(R_{nn}^{1/2} \vec{g} \right)} \\
 &= \vec{t}^H R_{nn}^{-1} \vec{t}
 \end{aligned}$$

↑ max. SNR

Cauchy-Schwarz inequality,
 with equality, iff $R_{nn}^{1/2} \vec{g} = \alpha \cdot R_{nn}^{-1/2} \vec{t}$

Matched filtering (cont.)

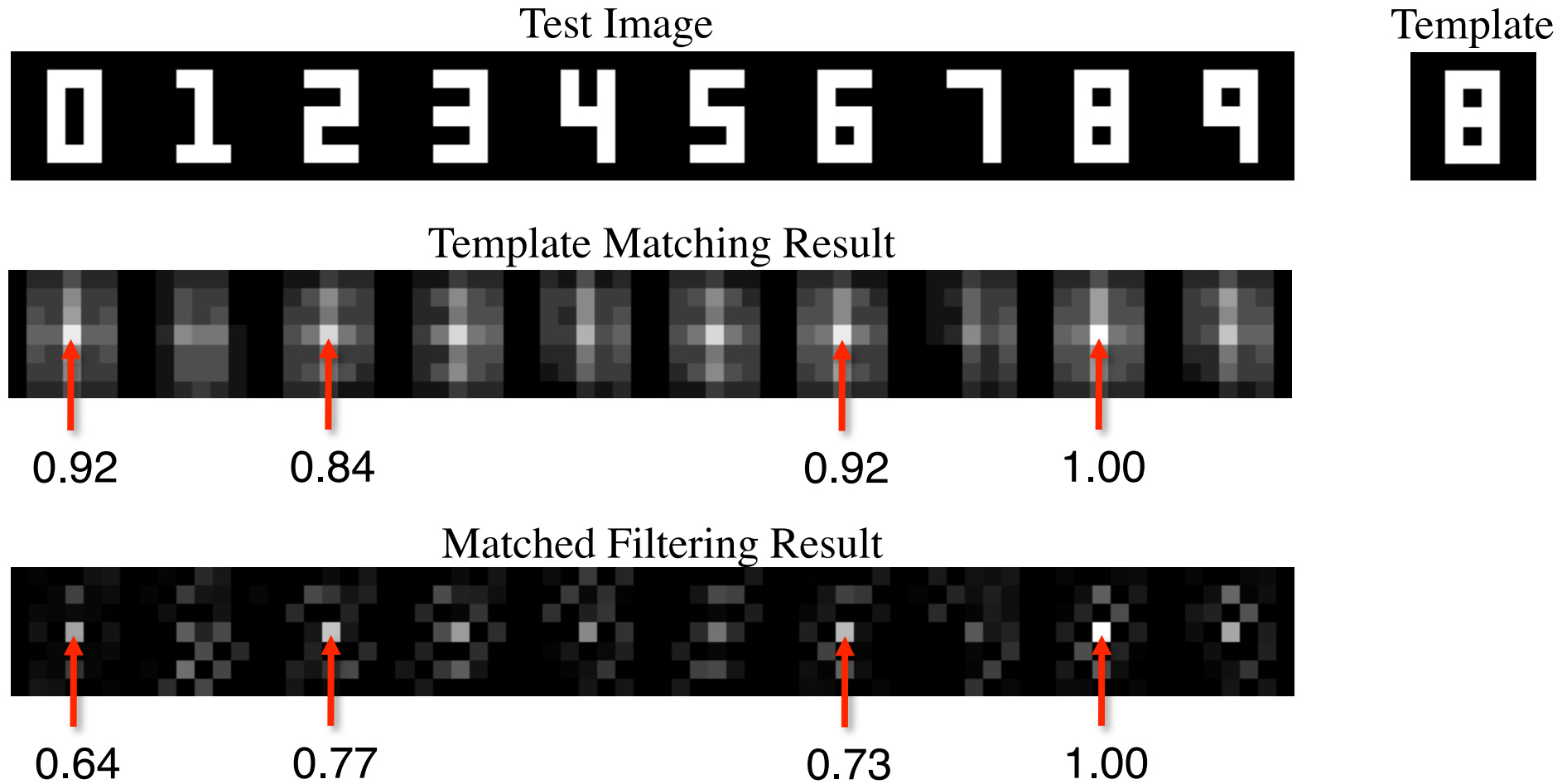
- Optimum detection: prefiltering & template matching



$$h[x, y] = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{\Phi_{nn}(e^{j\omega_x}, e^{j\omega_y})} e^{j\omega_x x + j\omega_y y} d\omega_x d\omega_y$$

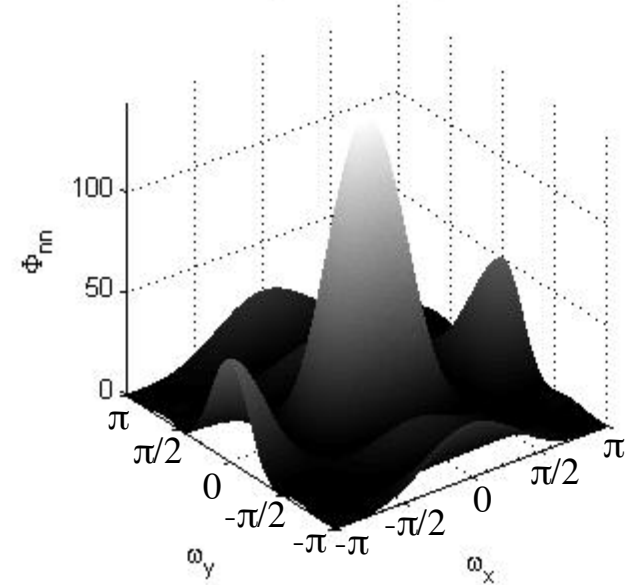
- For white noise $n[x, y]$, no prefiltering $h[x, y]$ required
- Low frequency clutter: highpass prefilter

Matched filtering example

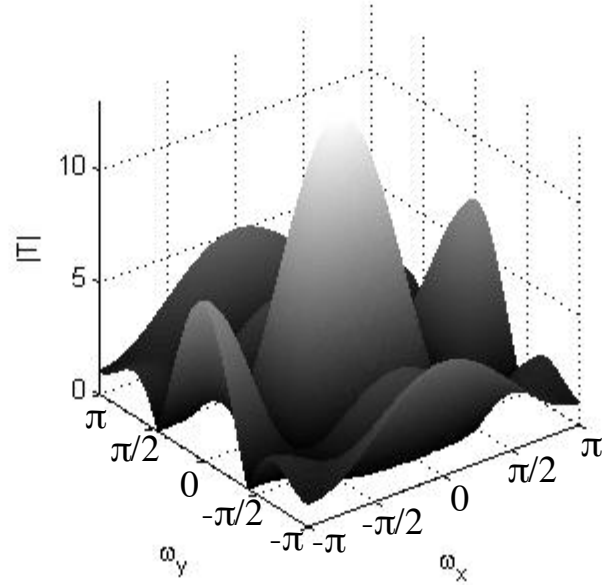


Matched filtering example (cont.)

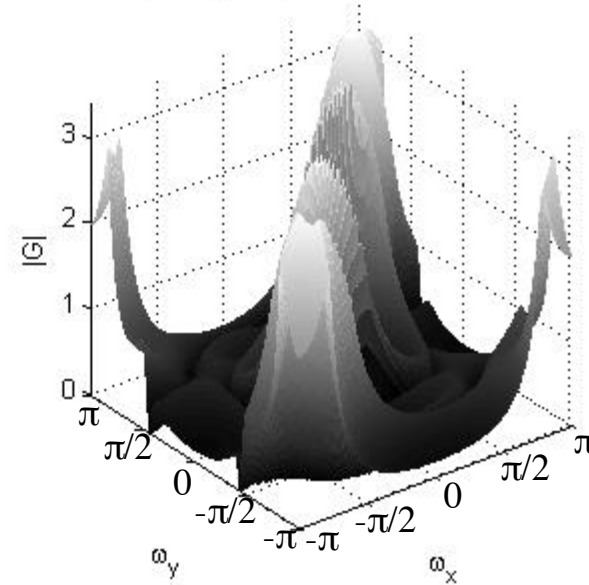
Power Spectral Density of Clutter



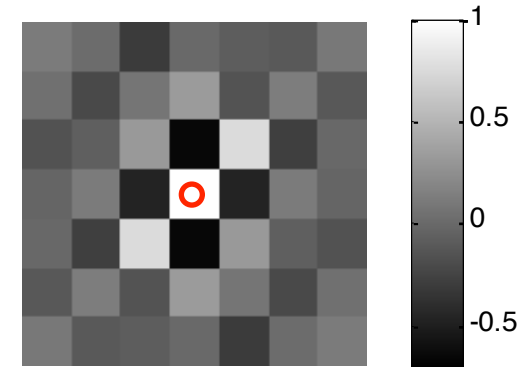
Frequency Response of Template



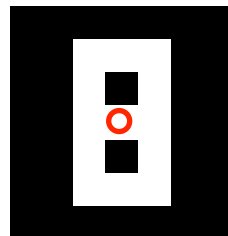
Frequency Response of Matched Filter



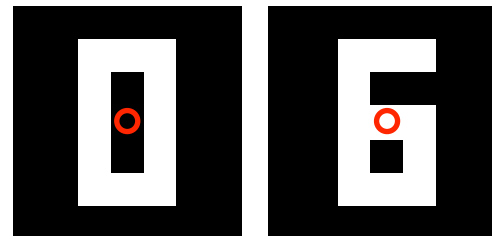
Matched Filter
Impulse Response
(180° rotated)



Template

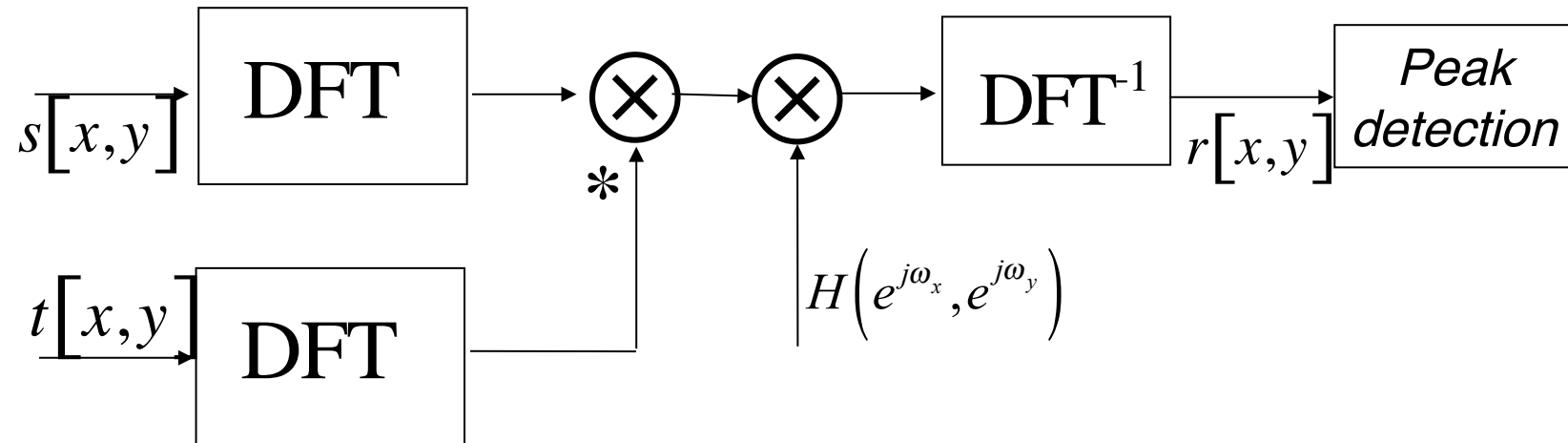


Clutter



Phase correlation

- Efficient implementation employing the Discrete Fourier Transform

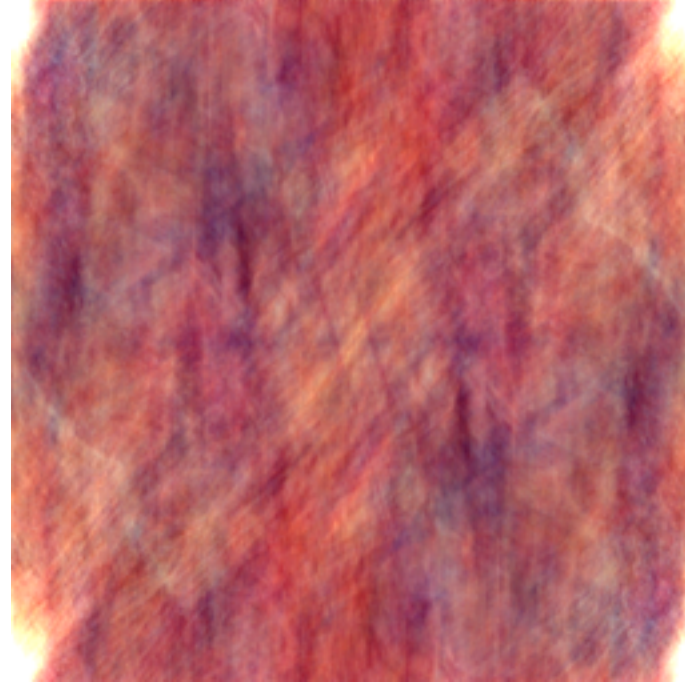


- Phase correlation

$$H(e^{j\omega_x}, e^{j\omega_y}) = \frac{1}{\left| S(e^{j\omega_x}, e^{j\omega_y}) \right| \left| T(e^{j\omega_x}, e^{j\omega_y}) \right|}$$



Original image

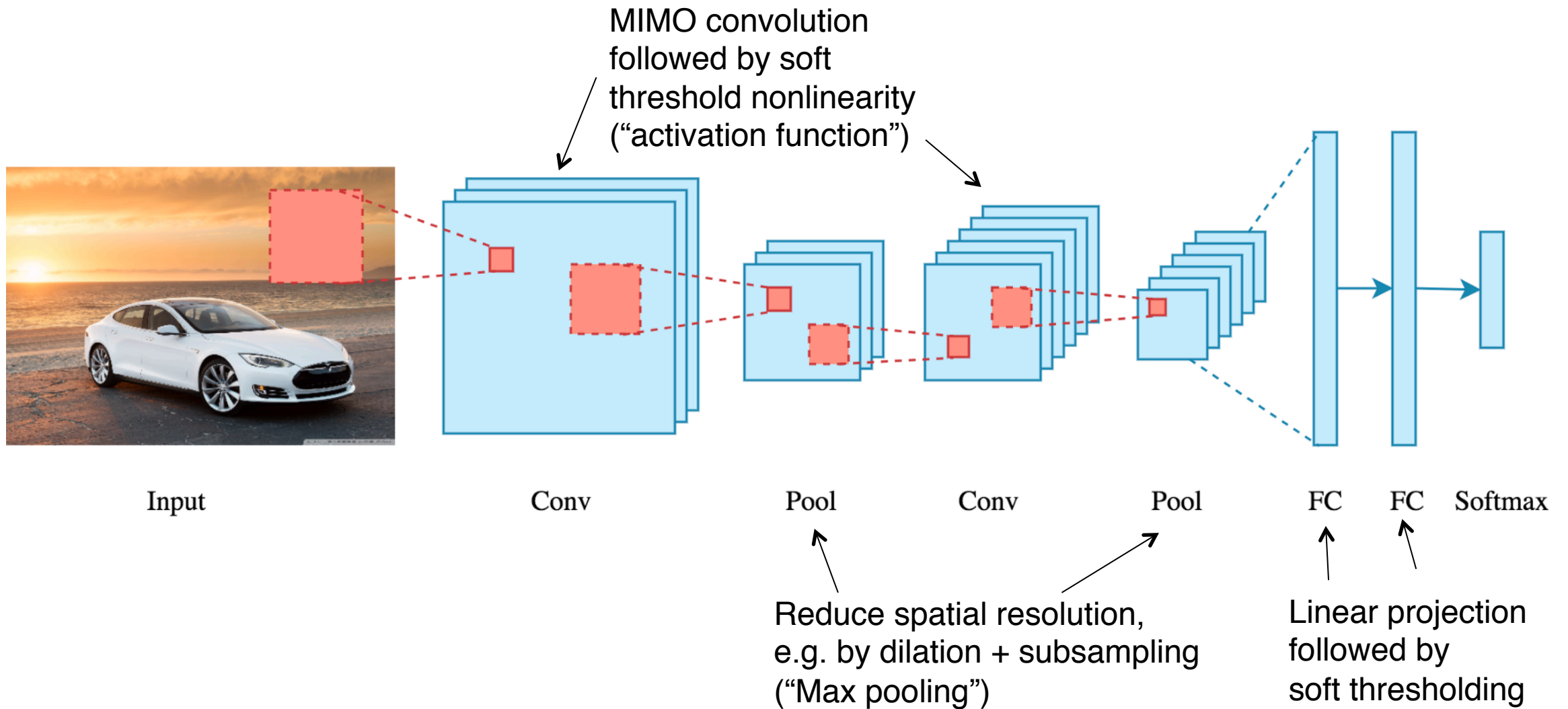


Magnitude only



Phase only

Convolutional neural networks



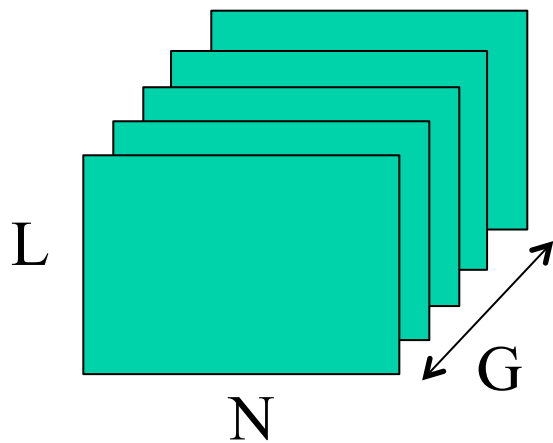
MIMO convolution

- Single-input-single-output: $f[x,y]$ and $g[x,y]$ are arrays of scalar values

$$g[x,y] = \sum_{x'=0}^{N-1} \sum_{y'=0}^{L-1} f[x',y'] \cdot h[x-x',y-y']$$

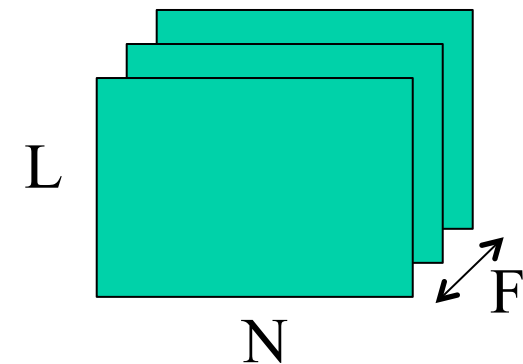
- Multiple-input-multiple-output convolution:

$$\mathbf{g}[x,y] = \sum_{x'=0}^{N-1} \sum_{y'=0}^{L-1} \mathbf{f}[x',y'] \cdot \mathbf{h}[x-x',y-y']$$



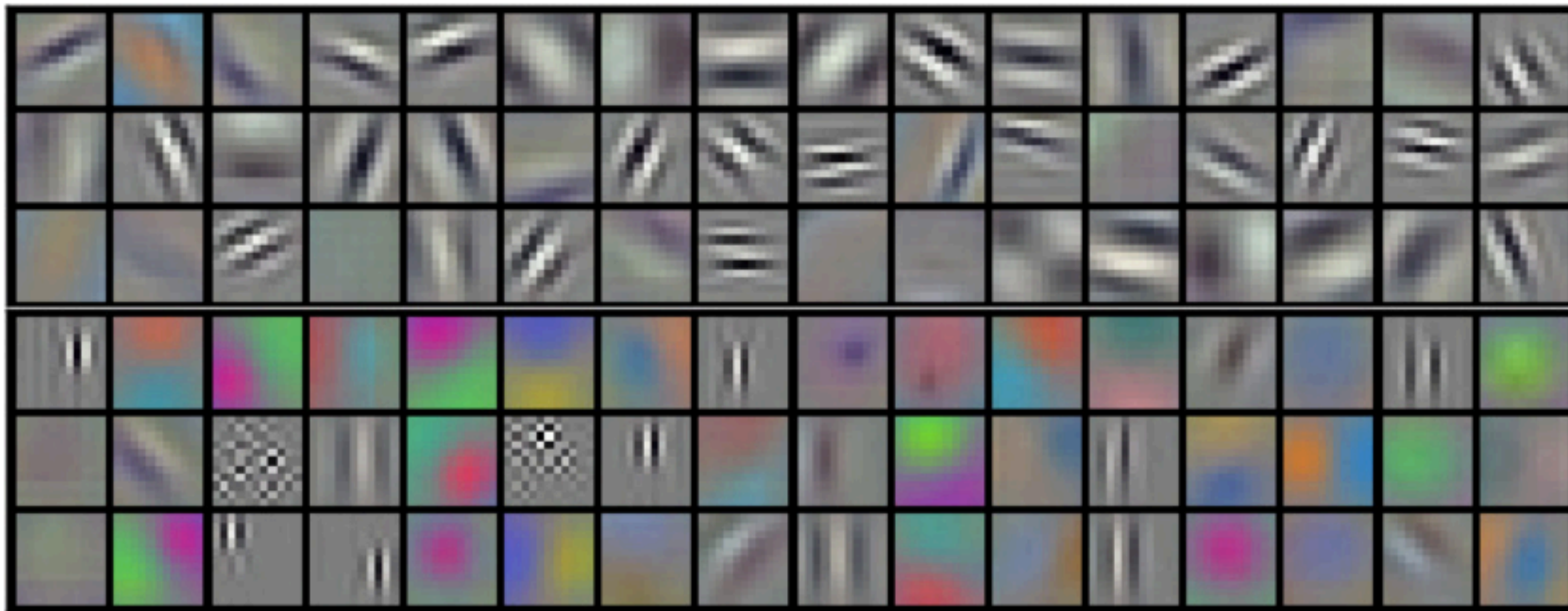
$$\mathbf{g}[x,y] = \begin{pmatrix} g_1[x,y] \\ \vdots \\ g_G[x,y] \end{pmatrix}$$

$$\mathbf{f}[x,y] = \begin{pmatrix} f_1[x,y] \\ \vdots \\ f_F[x,y] \end{pmatrix}$$



Example templates of first convolutional layer

AlexNet, $F=3$, $G=96$

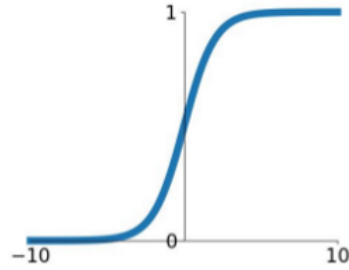


[Krizhevsky et al., 2012]

Activation function

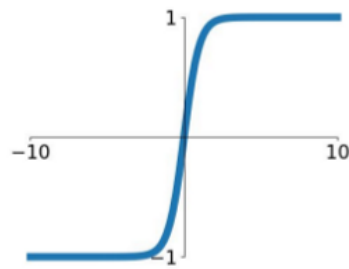
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



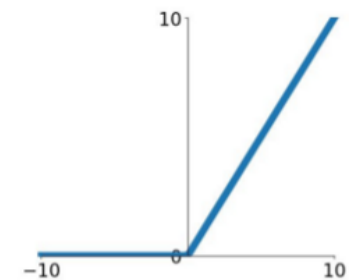
tanh

$$\tanh(x)$$



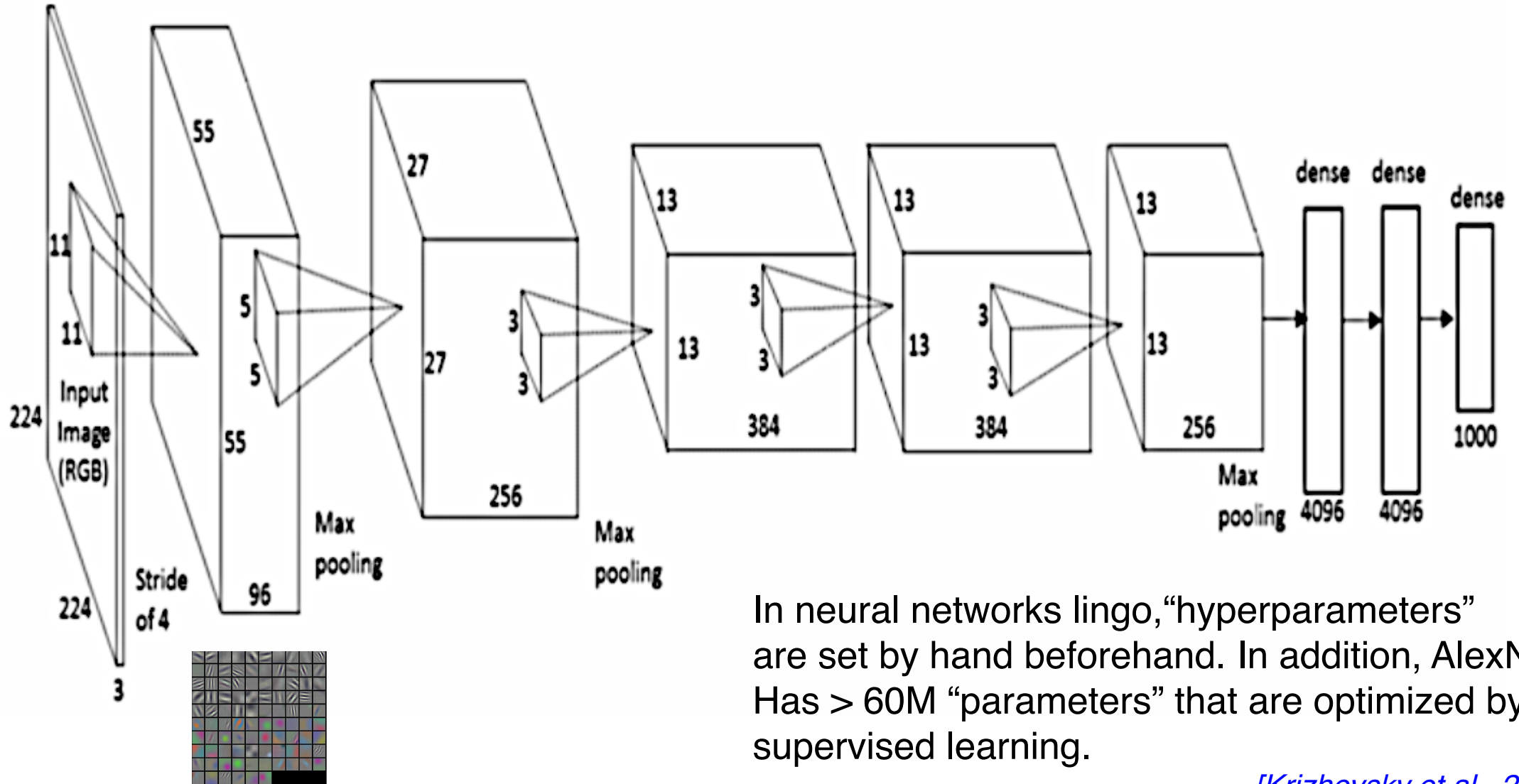
ReLU

$$\max(0, x)$$



- Sigmoid and tanh traditionally used
- ReLU (rectified linear unit) simpler and improves convergence of training
- Trained bias is added before activation function to set the best threshold

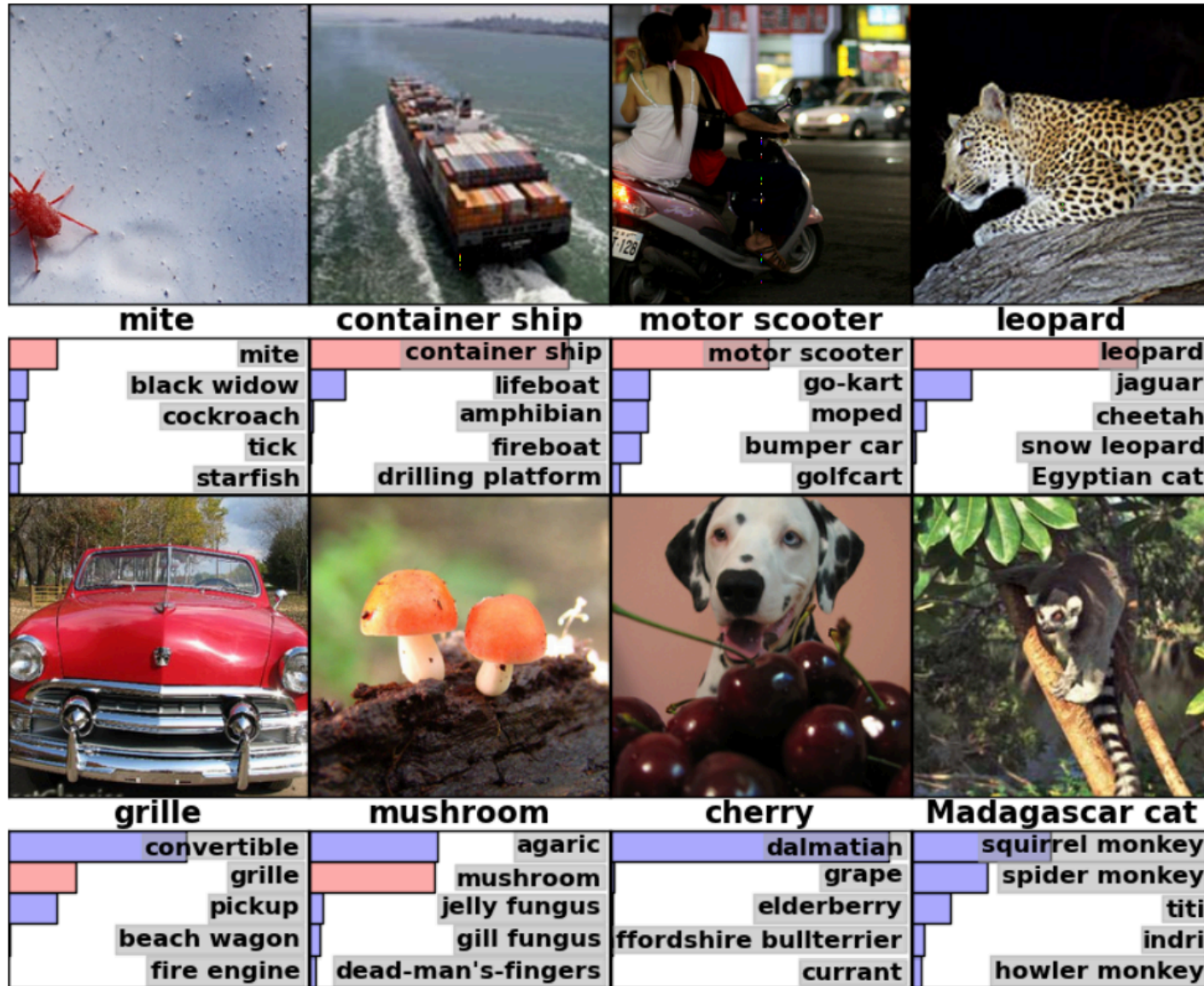
AlexNet hyperparameters



In neural networks lingo, “hyperparameters” are set by hand beforehand. In addition, AlexNet Has > 60M “parameters” that are optimized by supervised learning.

[Krizhevsky et al., 2012]

AlexNet Image Classification Results



[Krizhevsky et al., 2012]

AlexNet Image-based Retrieval Results

Query



Most
similar
images
in database

[Krizhevsky et al., 2012]