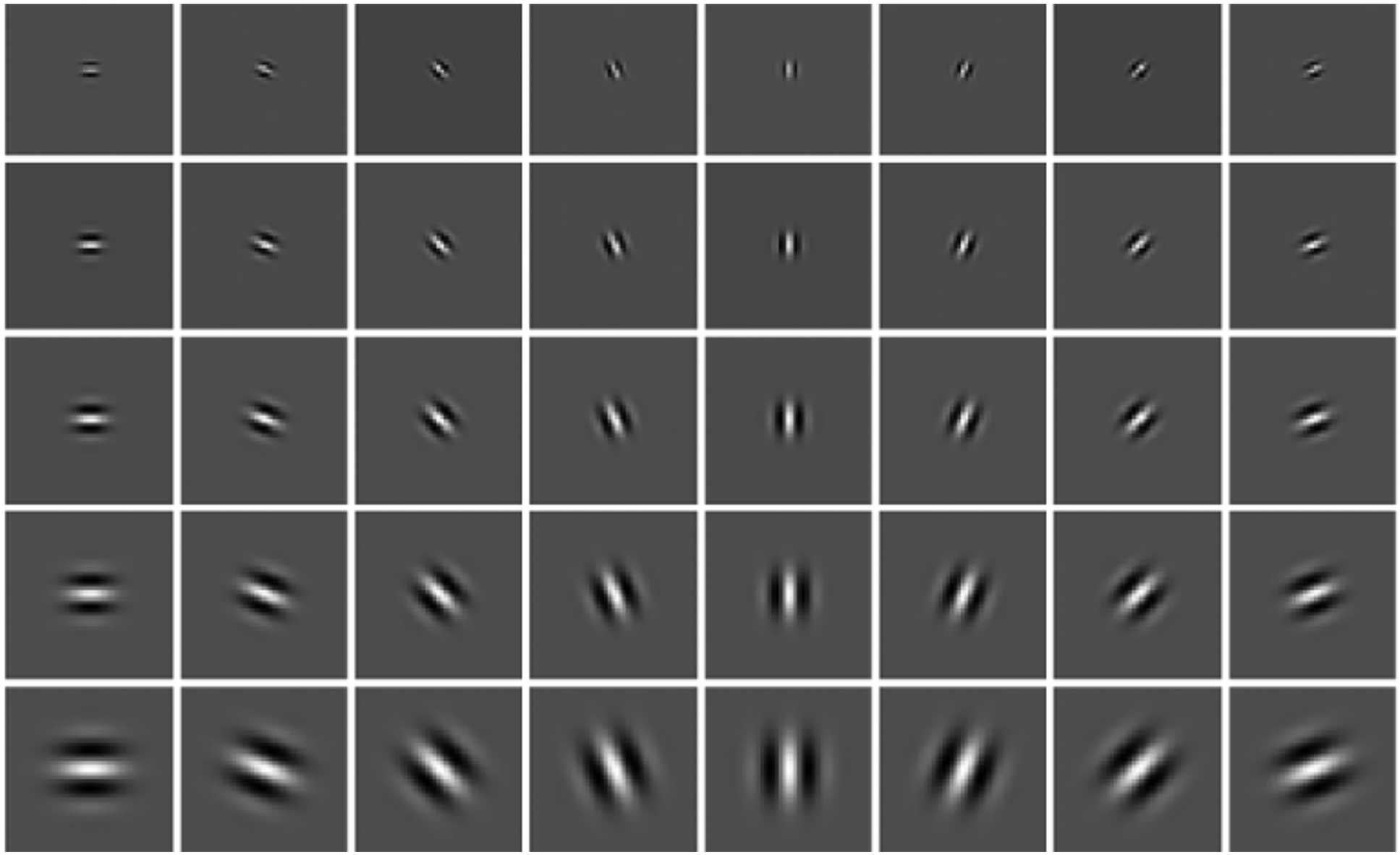


GIST

Computer Vision

Carnegie Mellon University (Kris Kitani)

GIST

1. Compute filter responses (filter bank of Gabor filters)

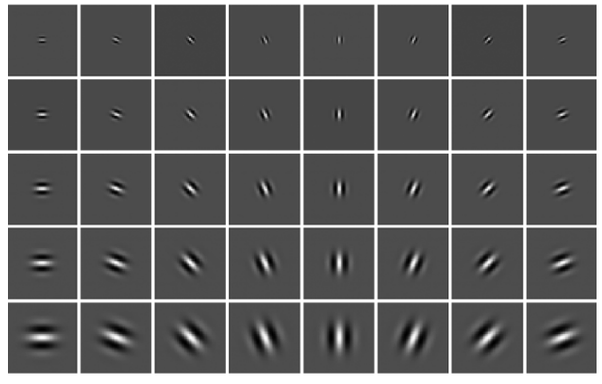
2. Divide image patch into 4 x 4 cells

3. Compute filter response averages for each cell

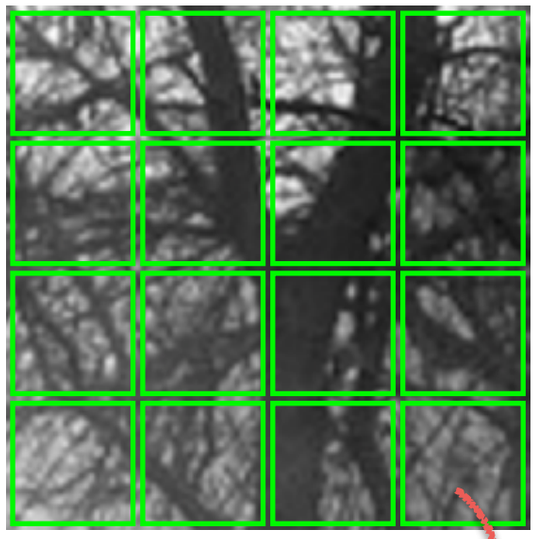
4. Size of descriptor is 4 x 4 x N, where N is the size of the filter bank

Originally designed to describe entire images but ideas also apply to feature descriptors

Filter bank



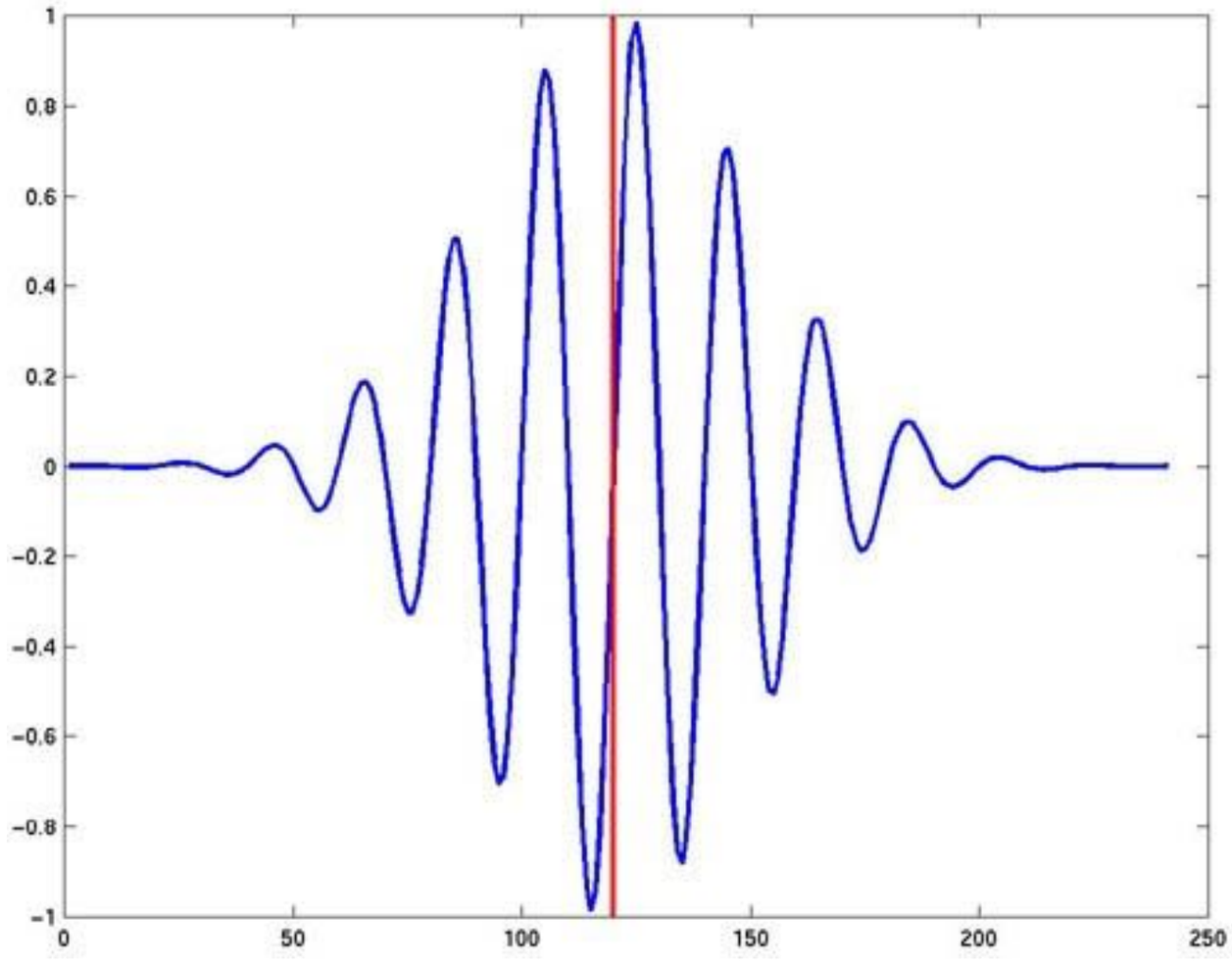
4 x 4 cell



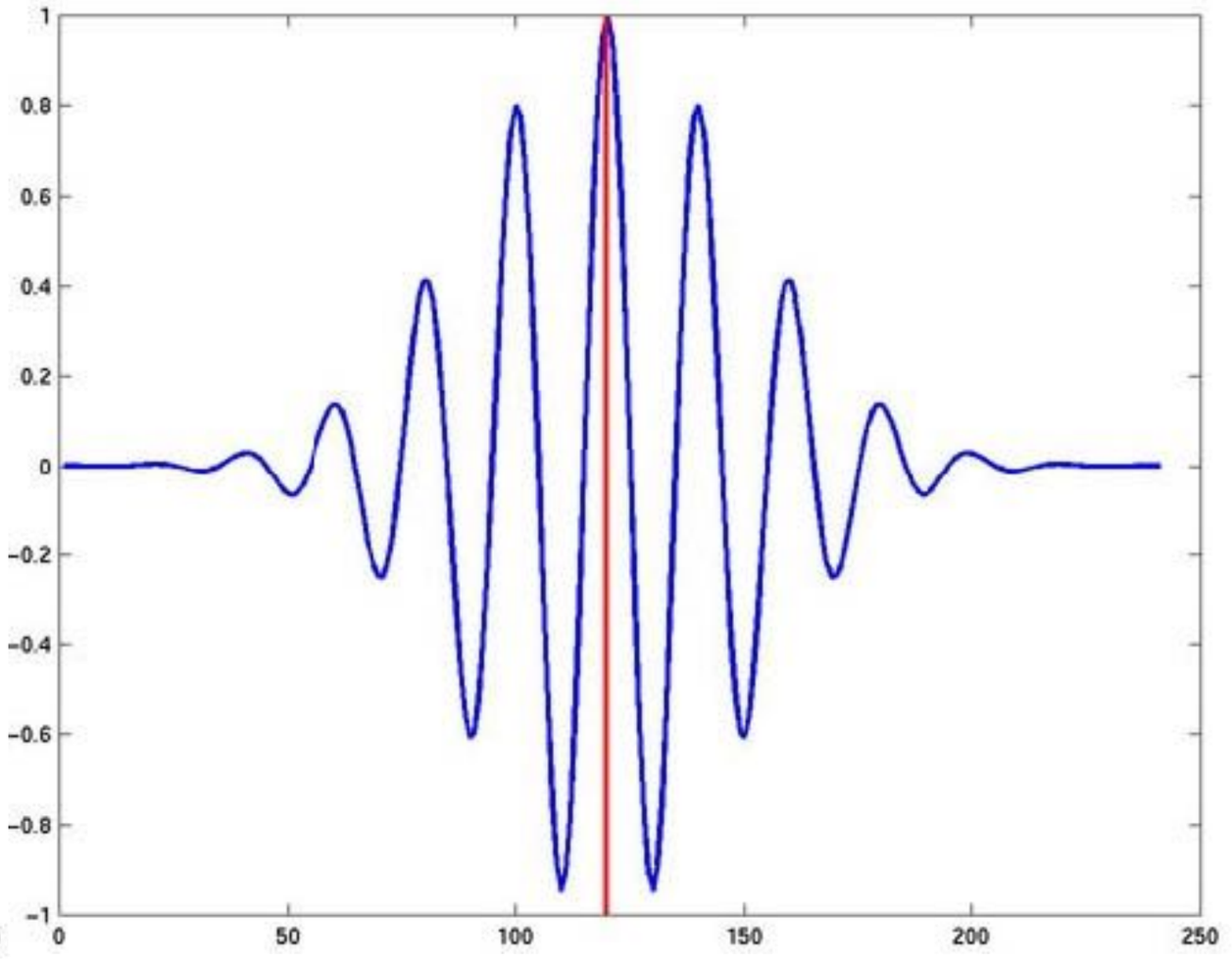
averaged filter responses

Gabor Filters

(1D examples)



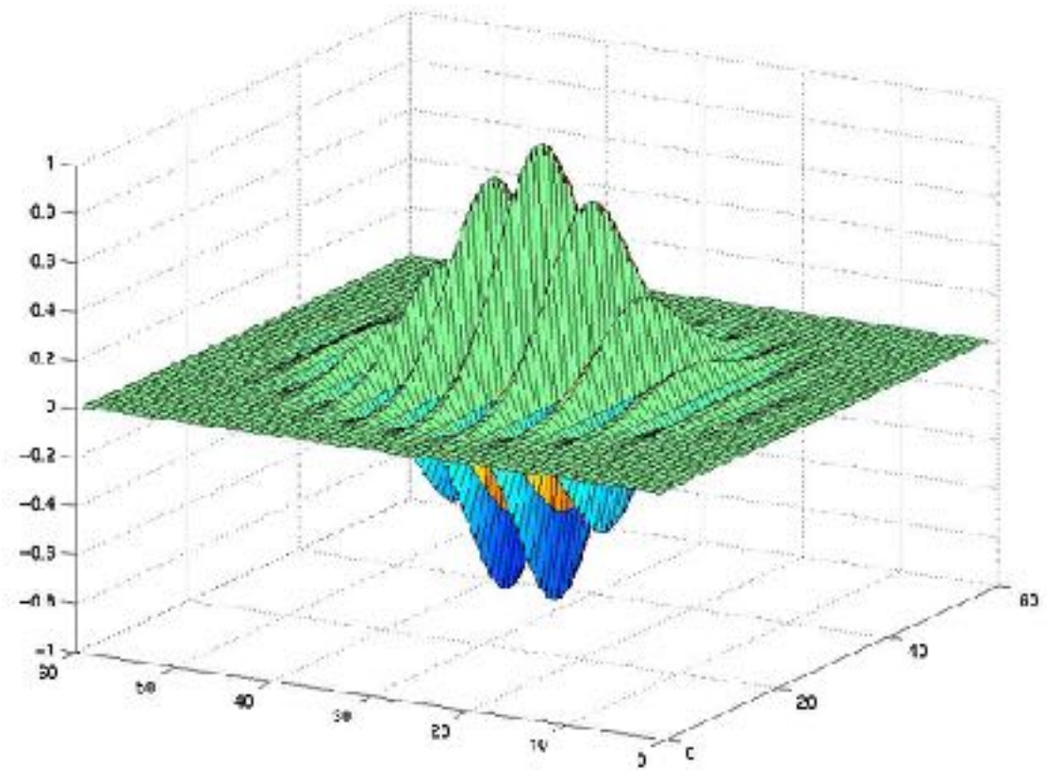
$$e^{-\frac{x^2}{2\sigma^2}} \sin(2\pi\omega x)$$

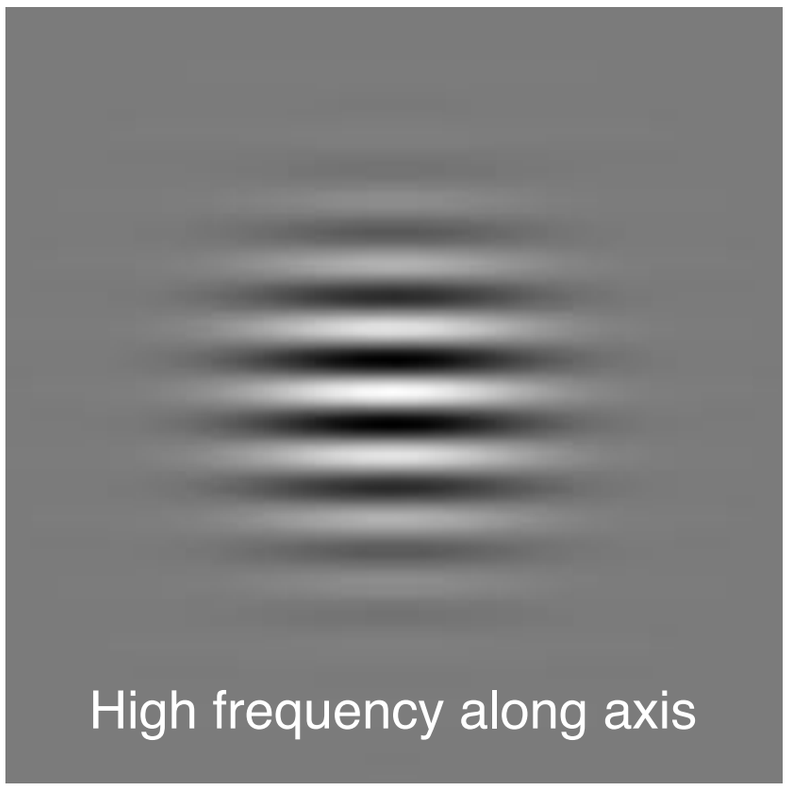


$$e^{-\frac{x^2}{2\sigma^2}} \cos(2\pi\omega x)$$

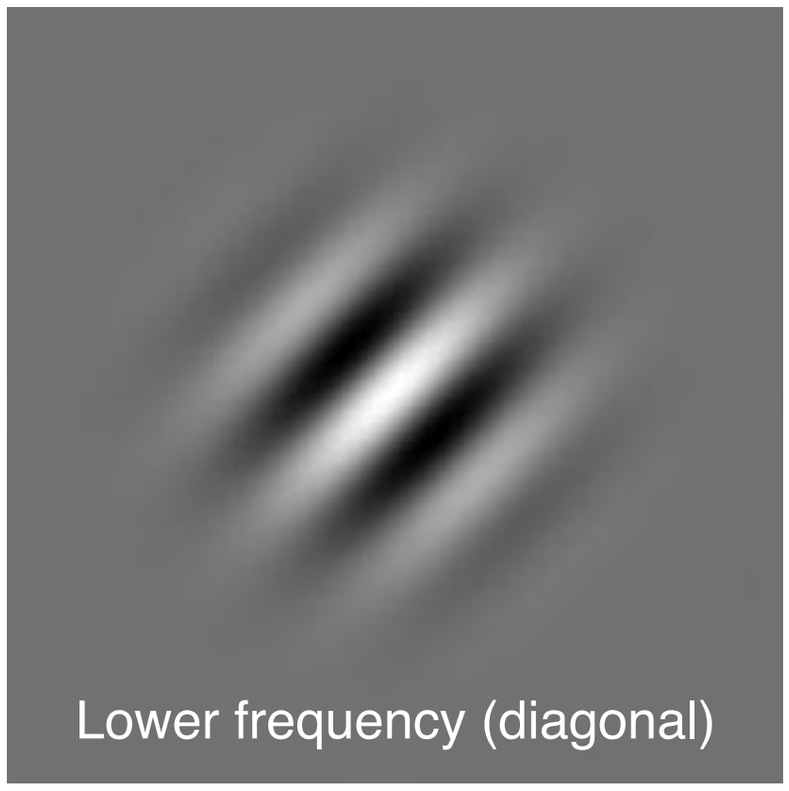
2D Gabor Filters

$$e^{-\frac{x^2+y^2}{2\sigma^2}} \cos\left(2\pi\left(k_x x + k_y y\right)\right)$$

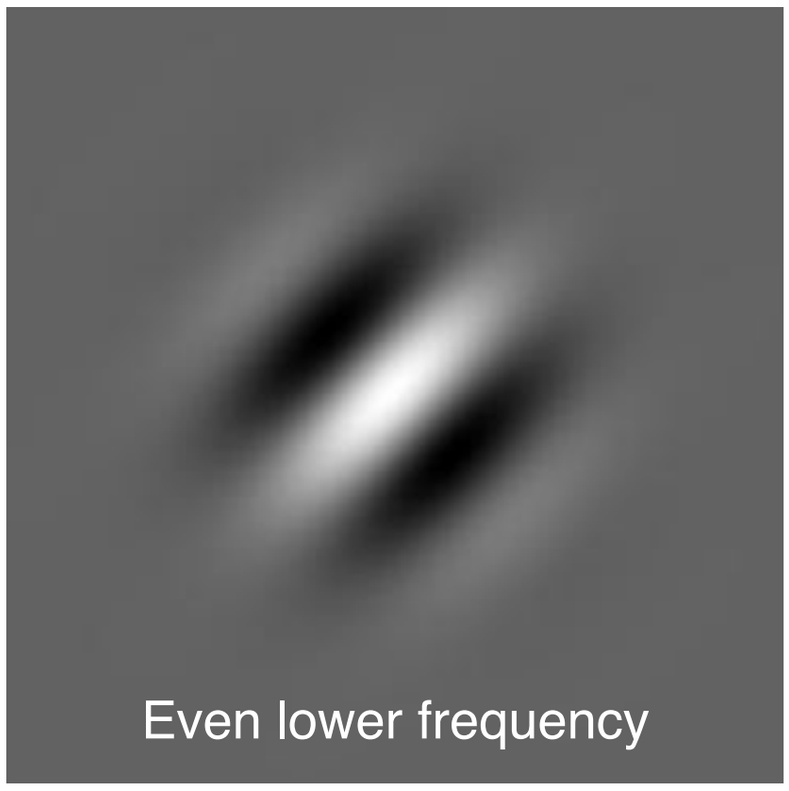




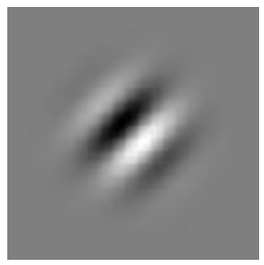
High frequency along axis

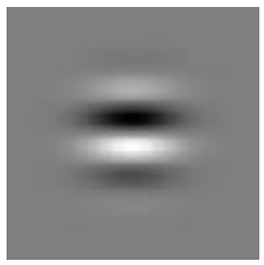


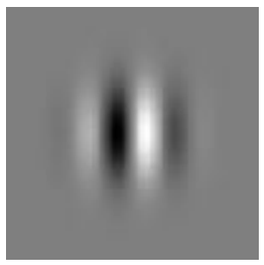
Lower frequency (diagonal)

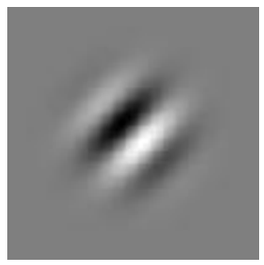


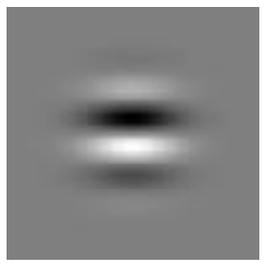
Even lower frequency

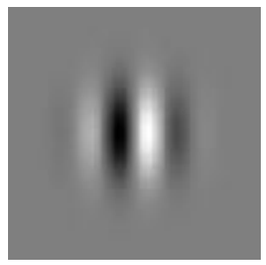


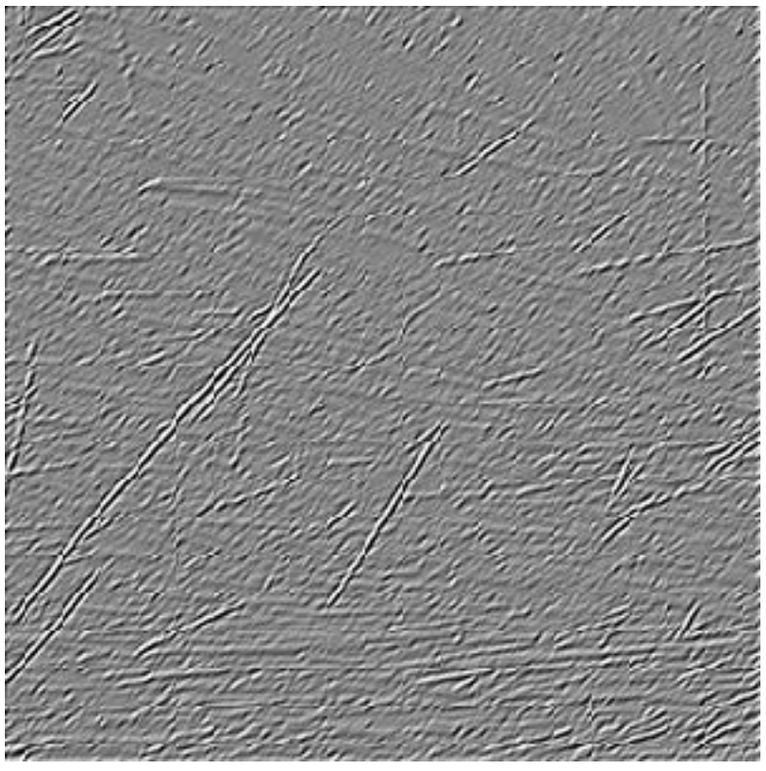


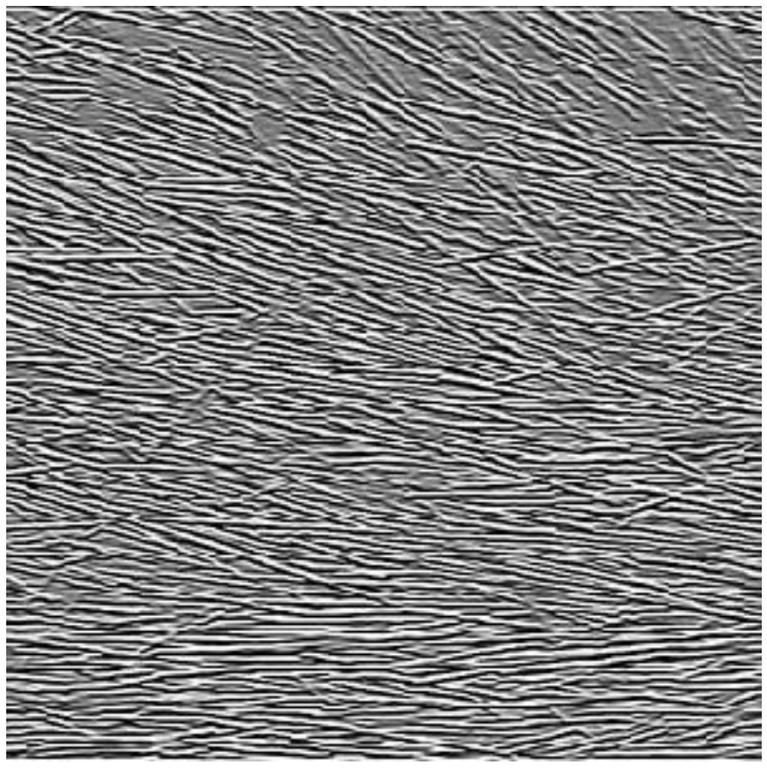


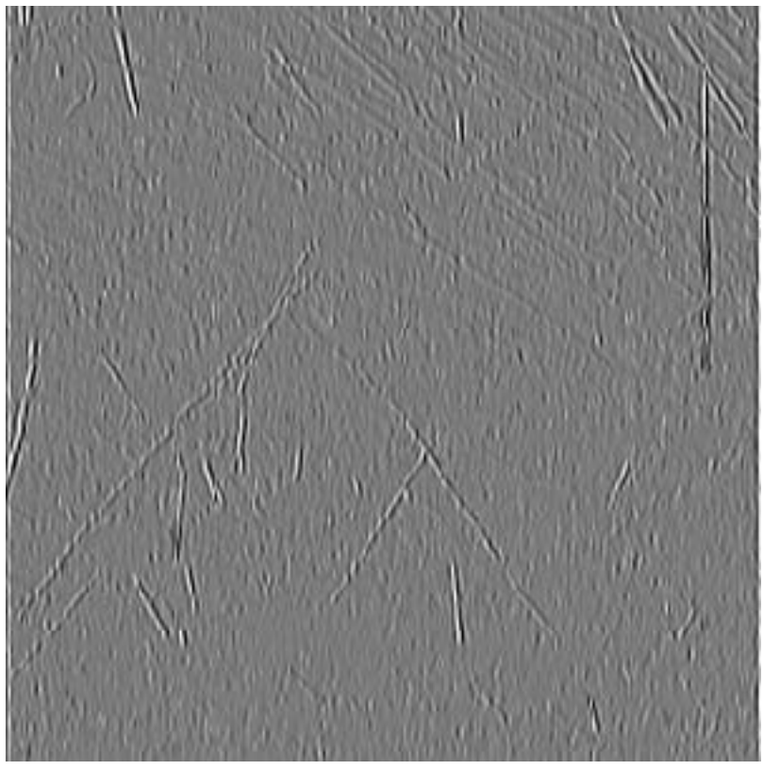


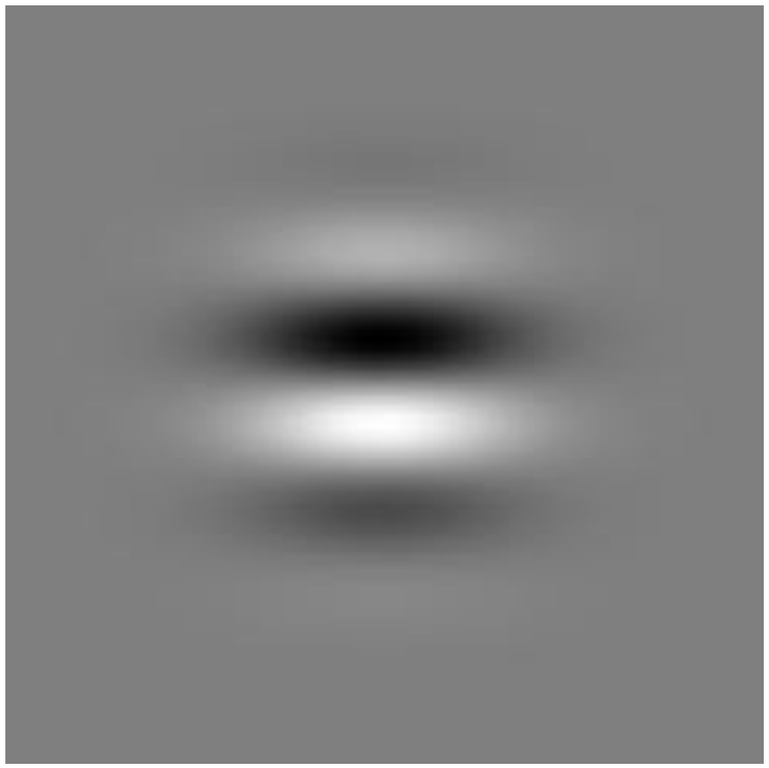




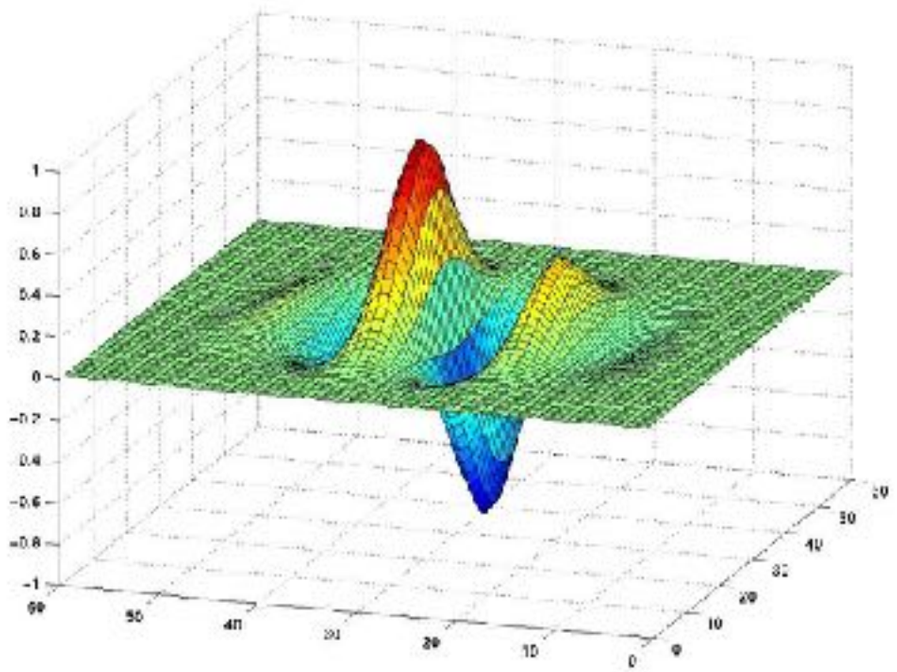




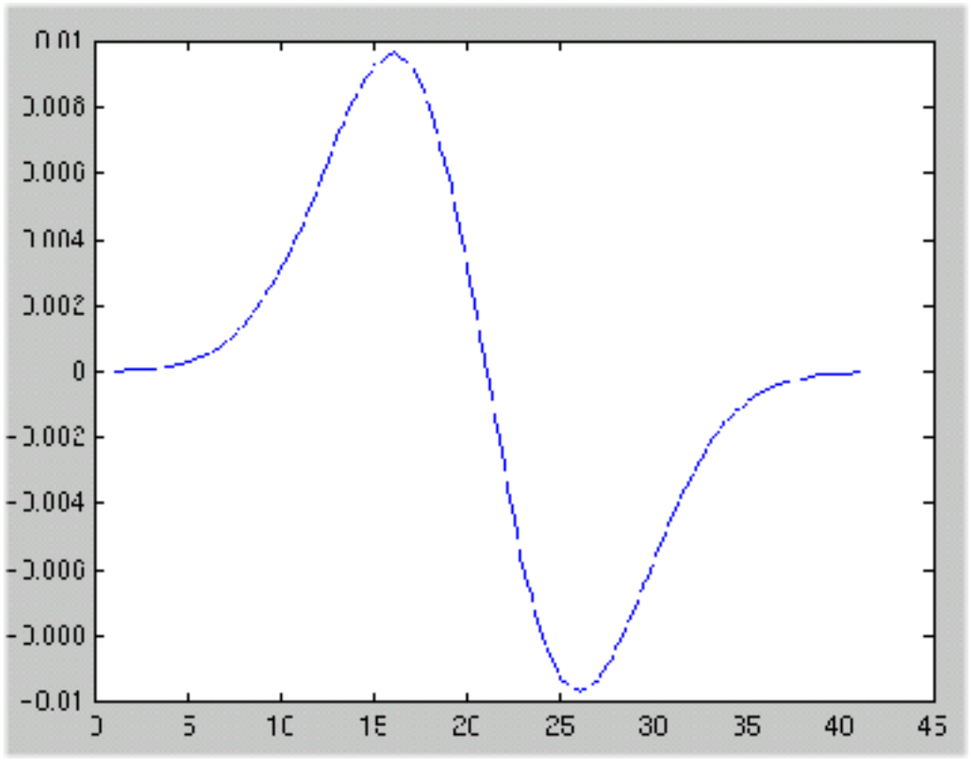




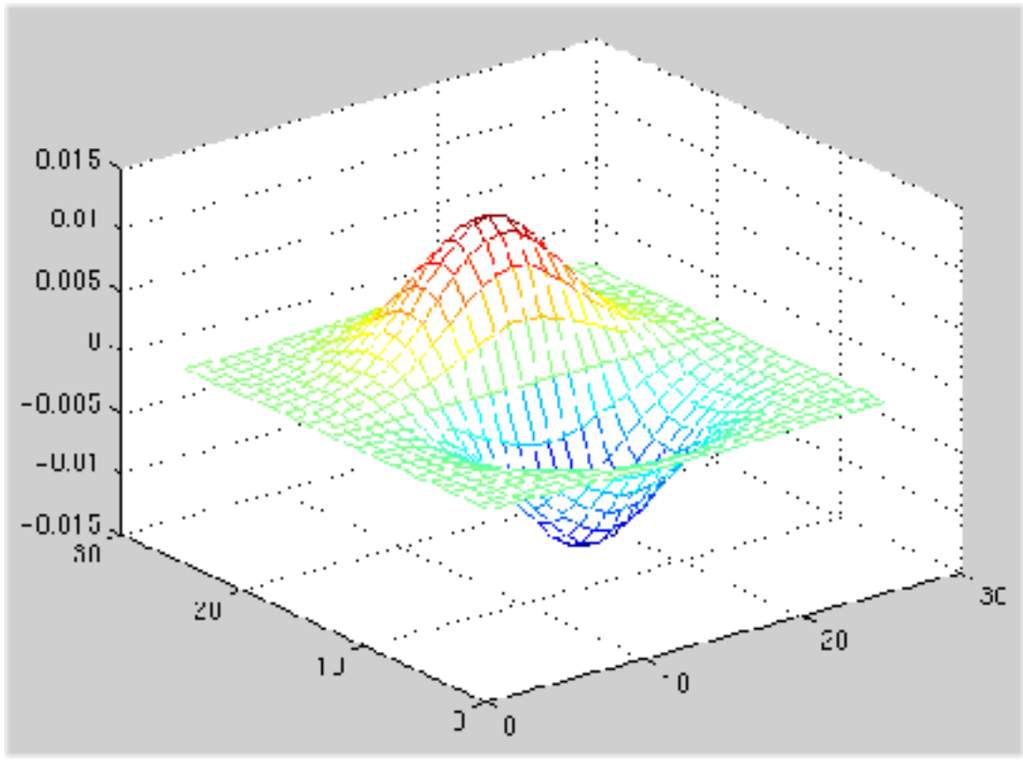
Odd Gabor filter

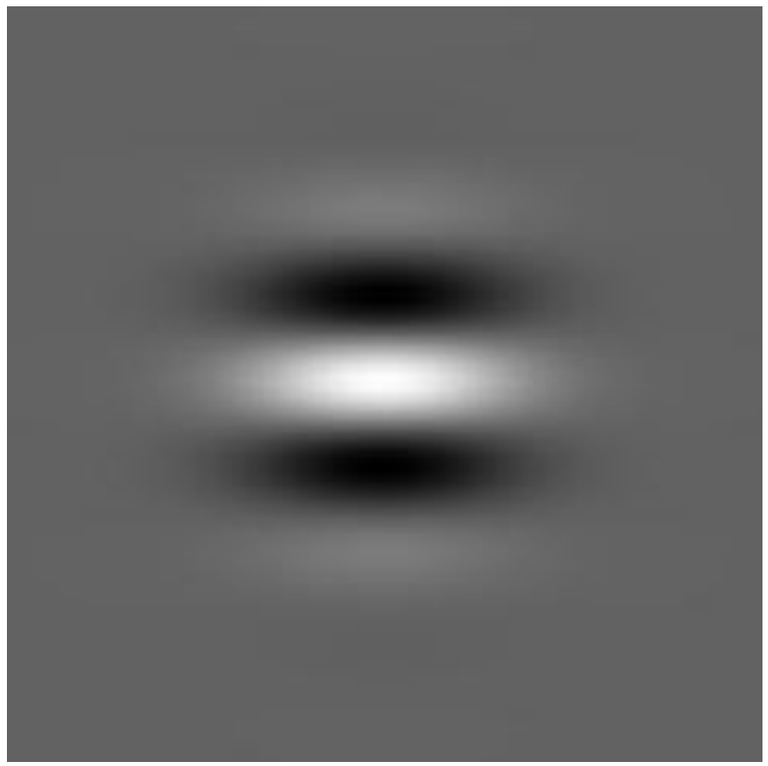


… looks a lot like…

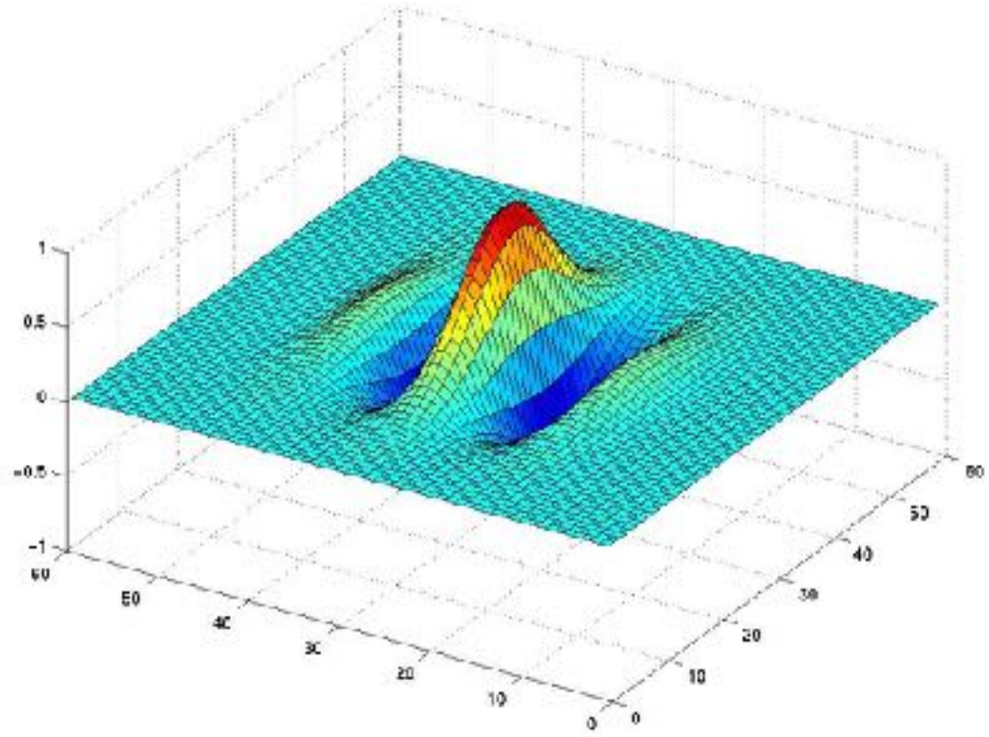


Gaussian Derivative

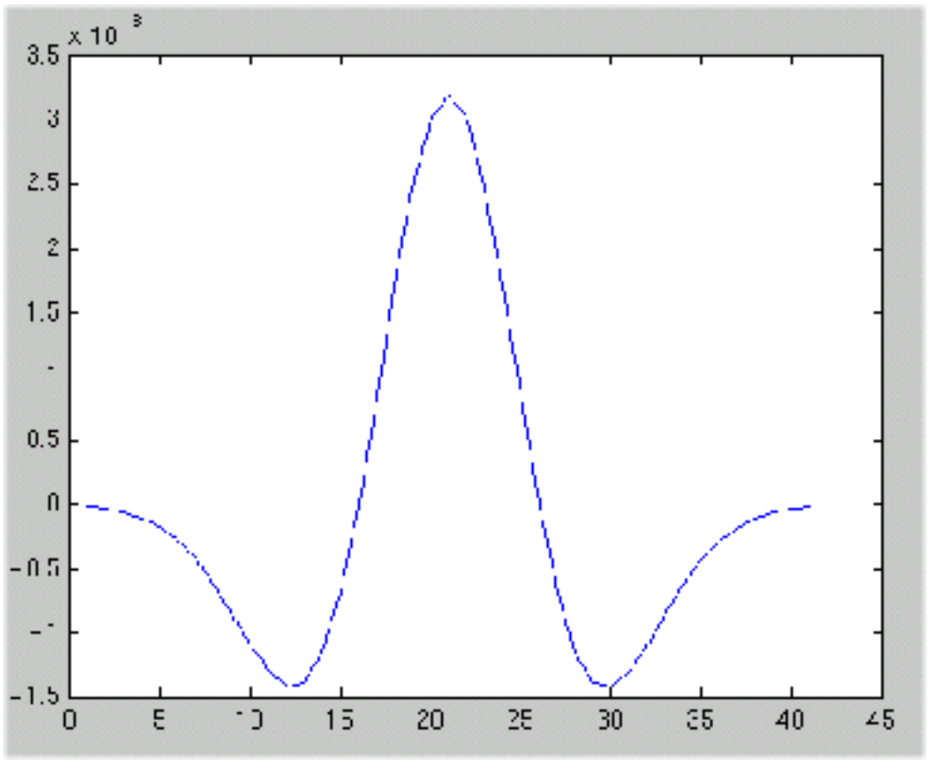




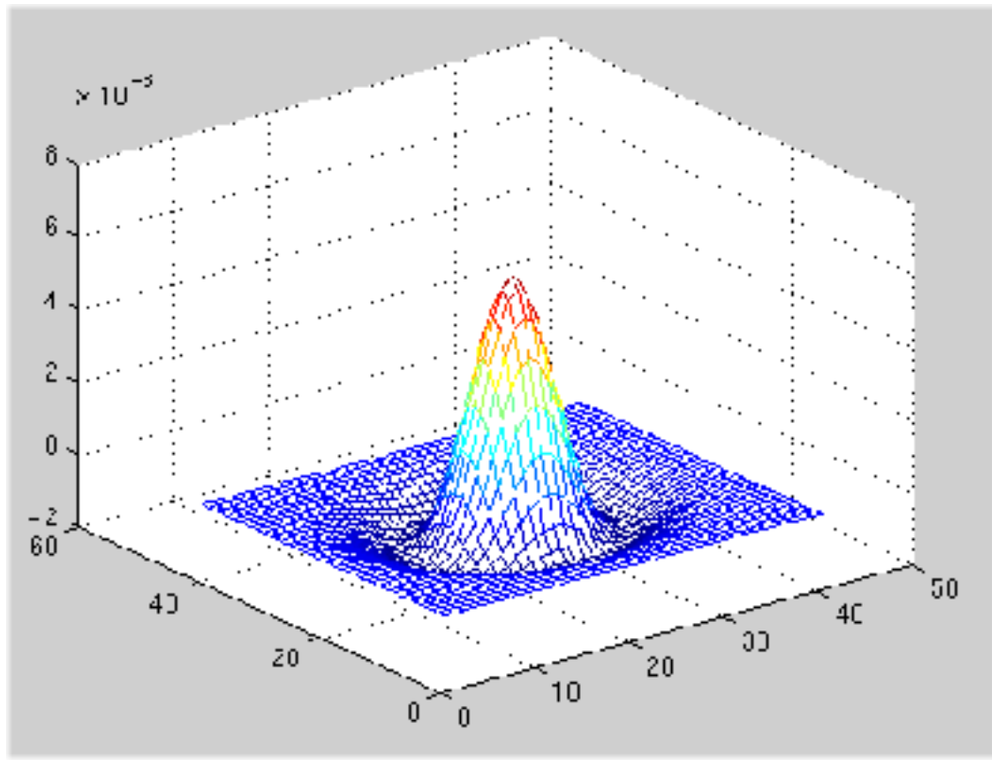
Even Gabor filter



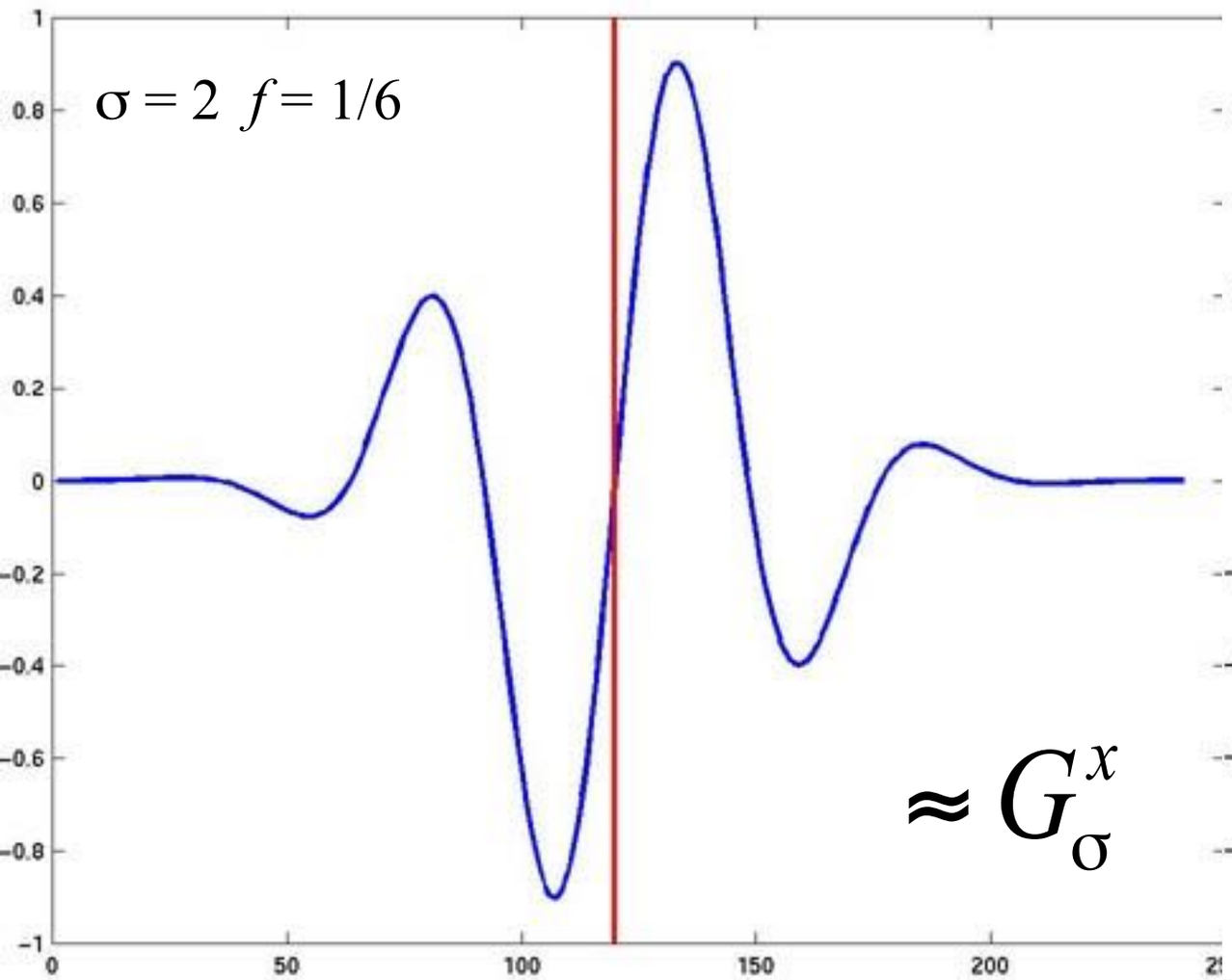
… looks a lot like…

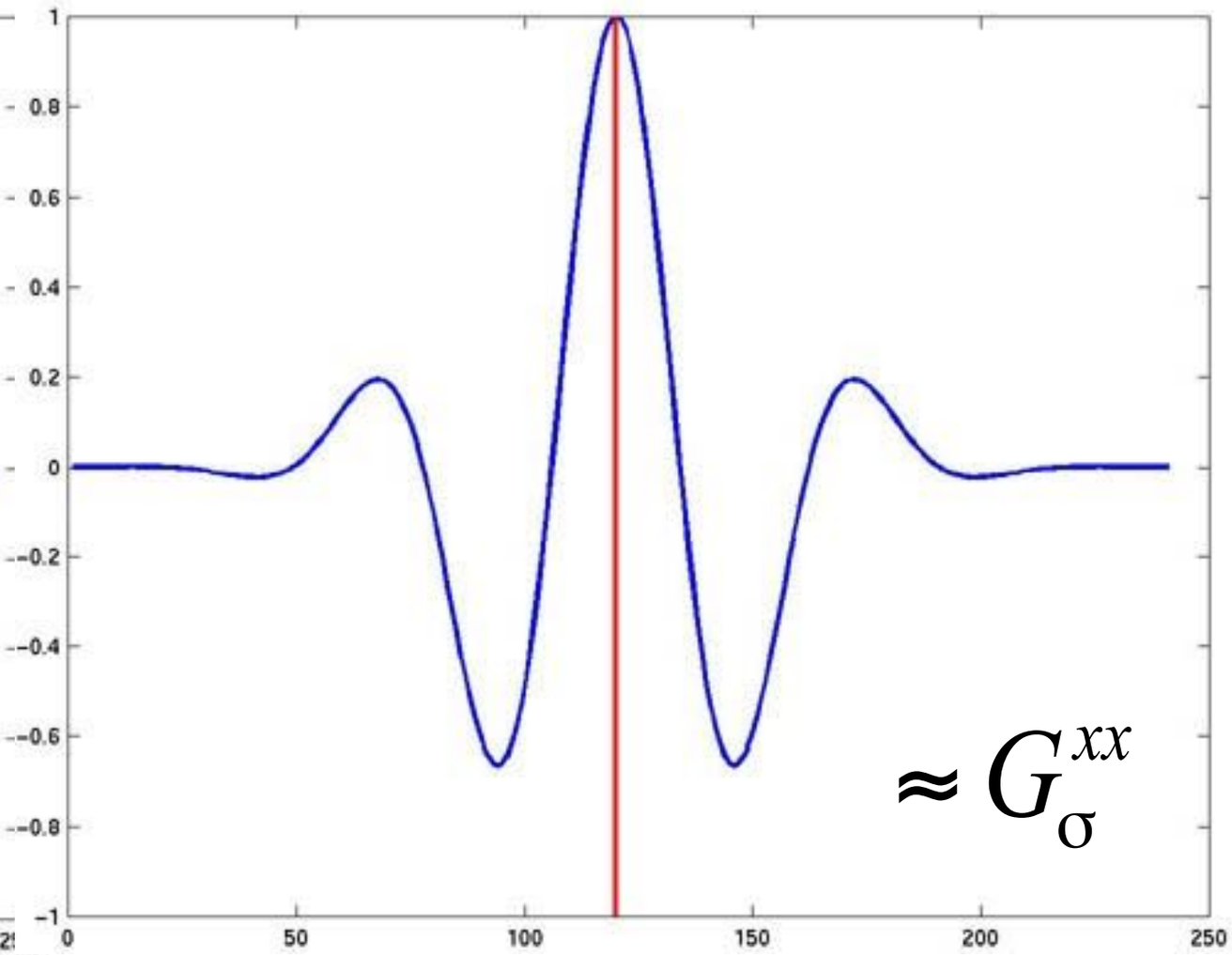


Laplacian

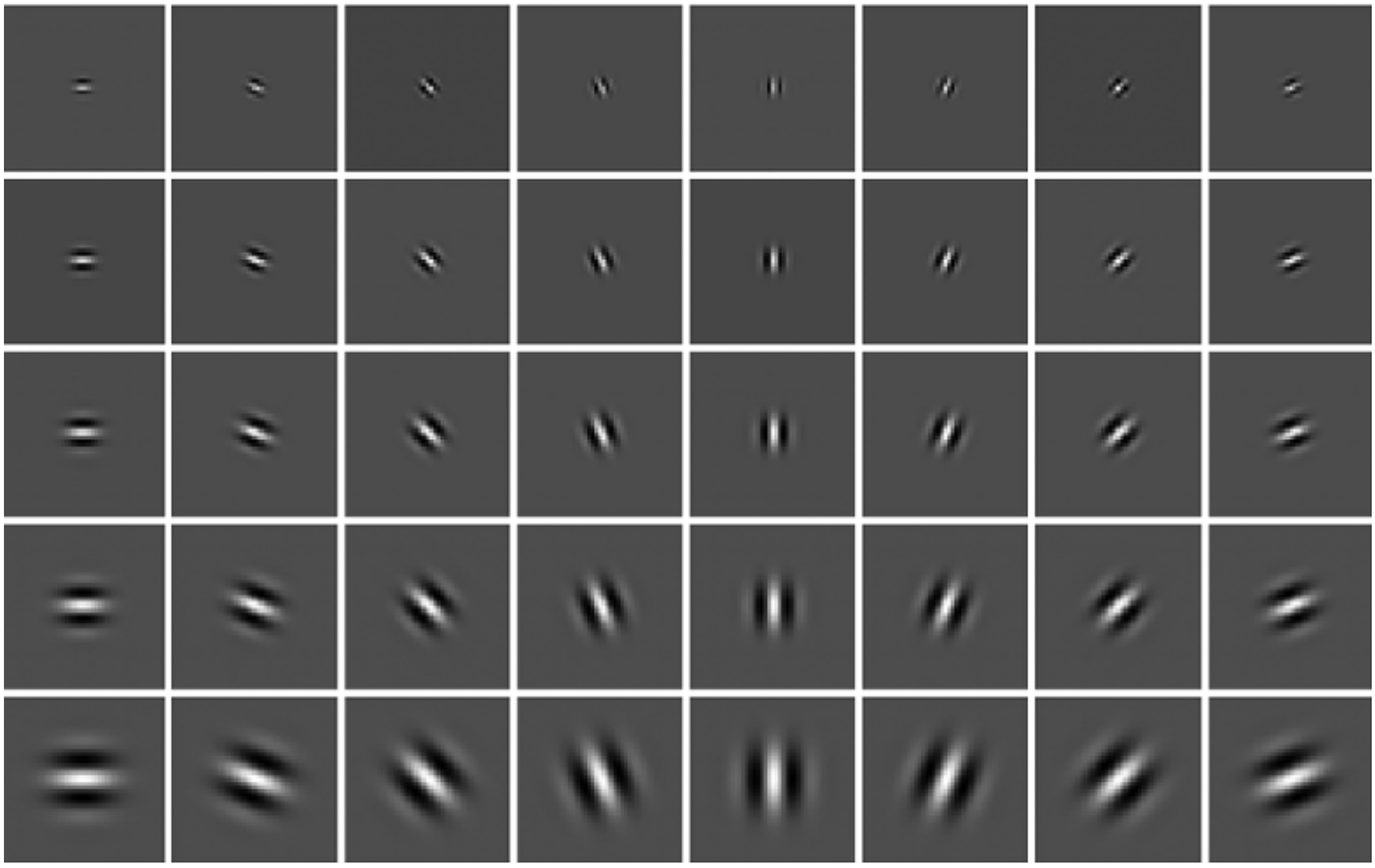


If scale small compared to inverse frequency,
the Gabor filters become derivative operators





Directional edge detectors

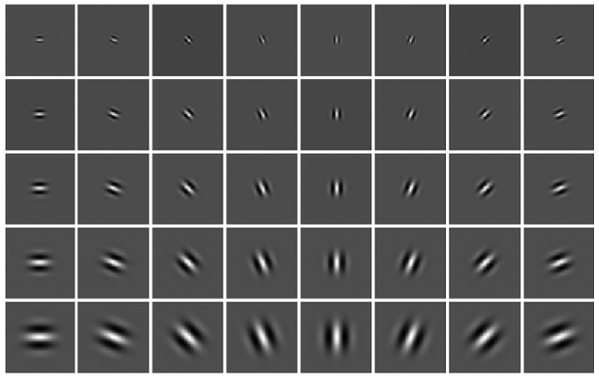


GIST

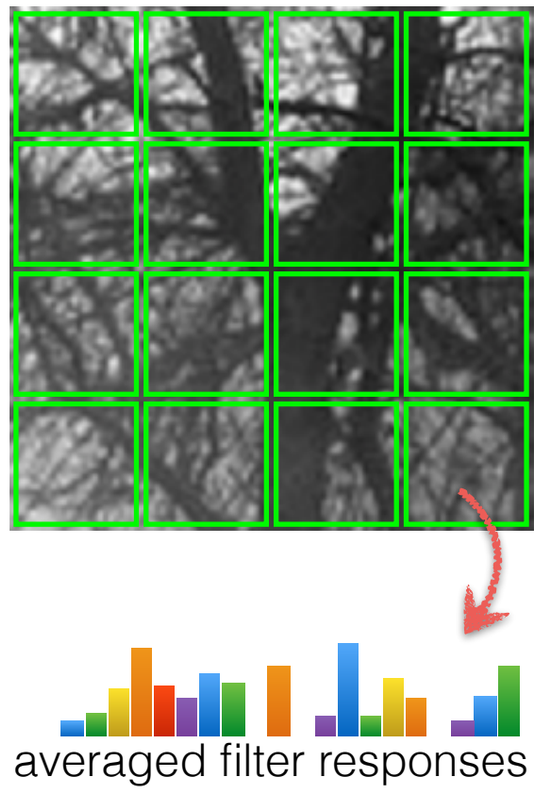
1. Compute filter responses (filter bank of Gabor filters)
2. Divide image patch into 4 x 4 cells
3. Compute filter response averages for each cell
4. Size of descriptor is 4 x 4 x N, where N is the size of the filter bank

What is the GIST descriptor encoding?

Filter bank



4 x 4 cell

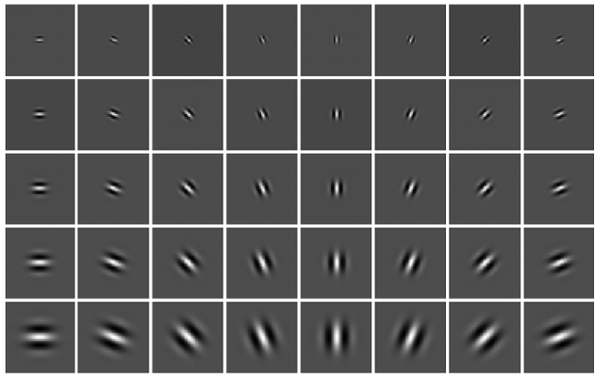


averaged filter responses

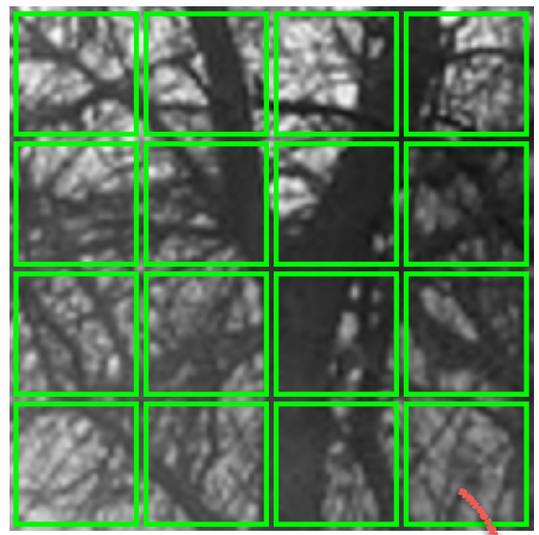
GIST

1. Compute filter responses (filter bank of Gabor filters)
2. Divide image patch into 4 x 4 cells
3. Compute filter response averages for each cell
4. Size of descriptor is 4 x 4 x N, where N is the size of the filter bank

Filter bank



4 x 4 cell



averaged filter responses

What is the GIST descriptor encoding?

Rough spatial distribution of image gradients

When will this feature descriptor fail?