More Feature Descriptors

16-385 Computer Vision (Kris Kitani)

Carnegie Mellon University

GIST

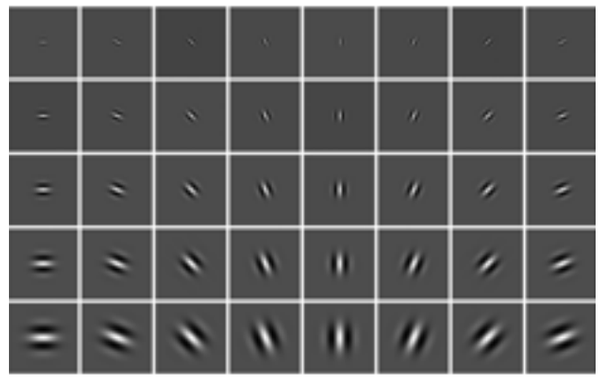
1. Compute filter responses (filter bank of Gabor filters)

2. Divide image patch into 4 x 4 cells

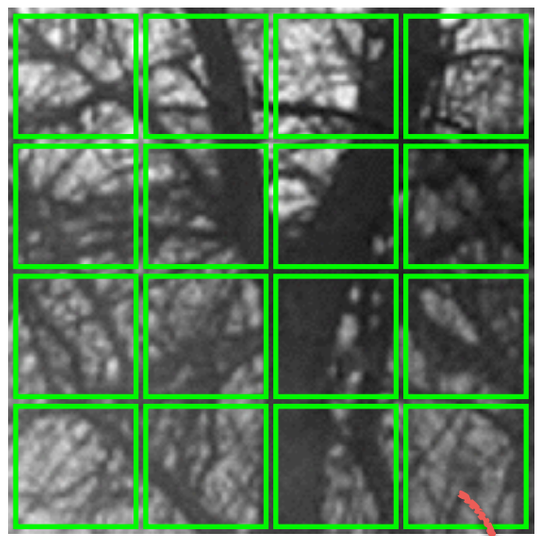
3. Compute filter response averages for each cell

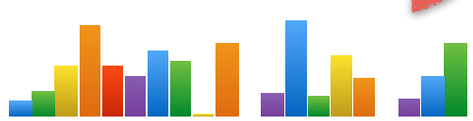
4. Size of descriptor is 4 x 4 x N, where N is the size of the filter bank

Filter bank



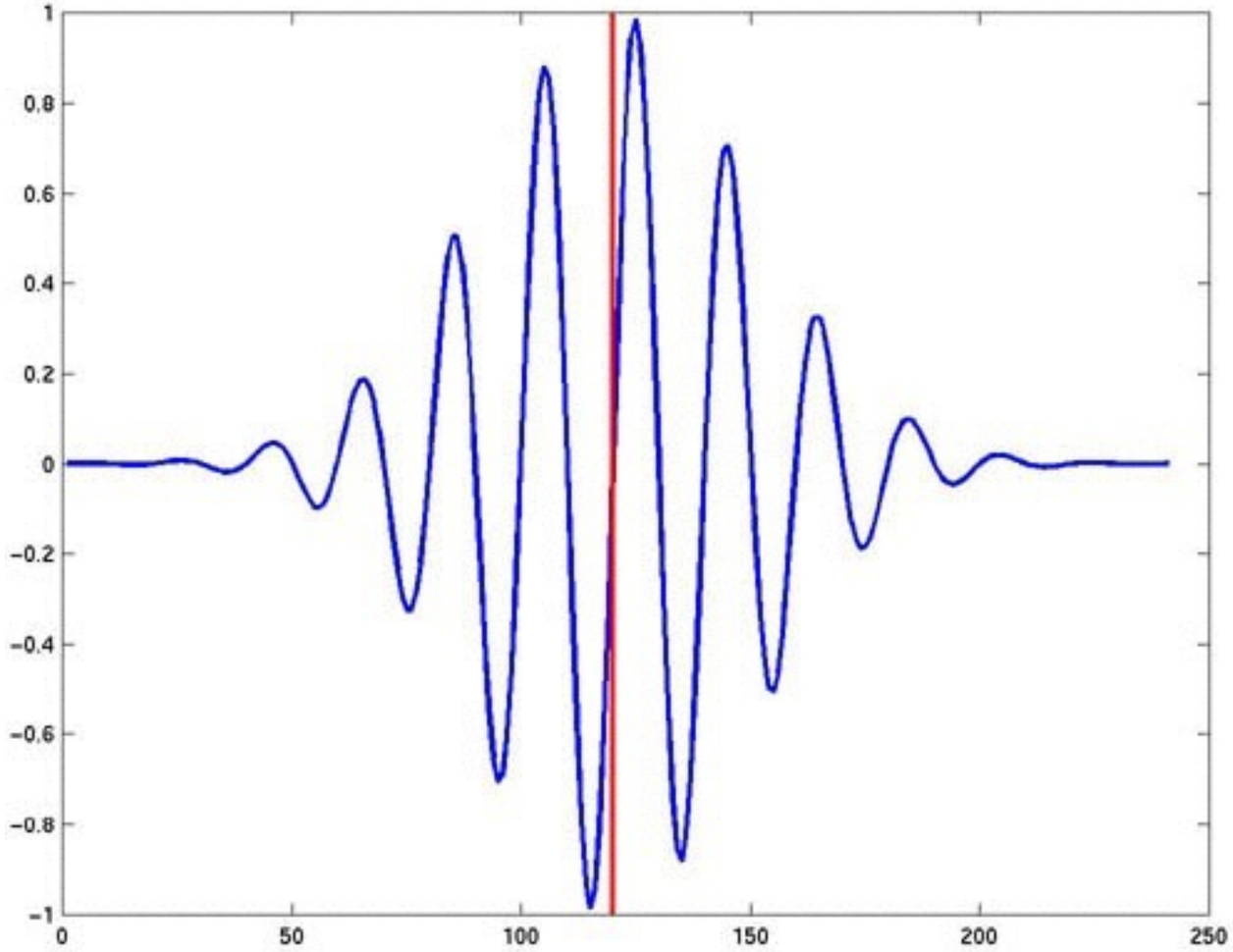
4 x 4 cell



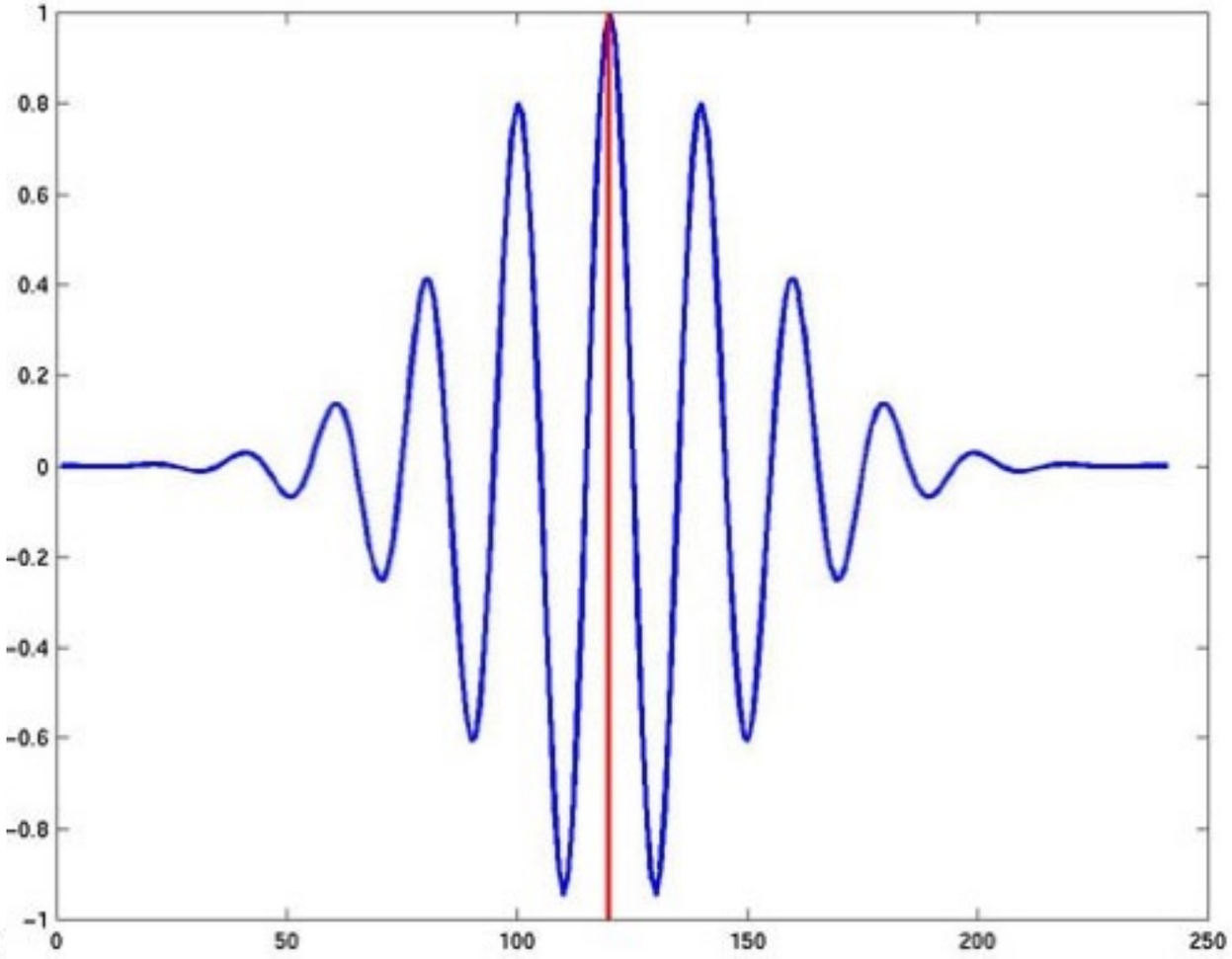


averaged filter responses

Gabor Filters

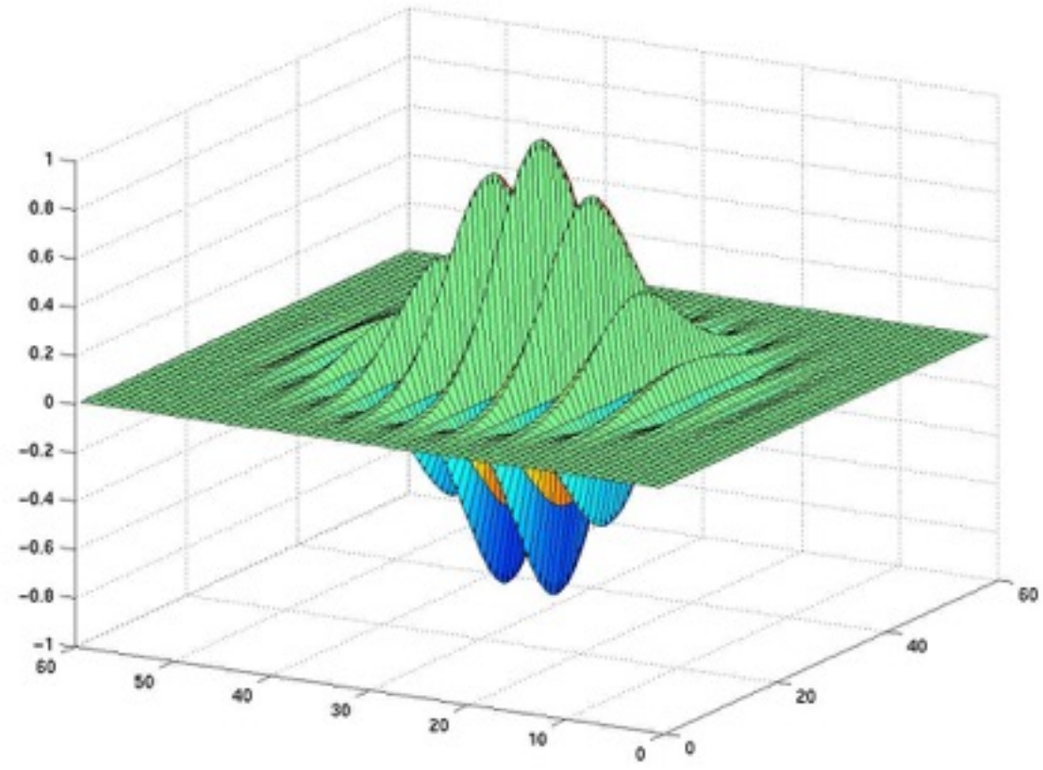


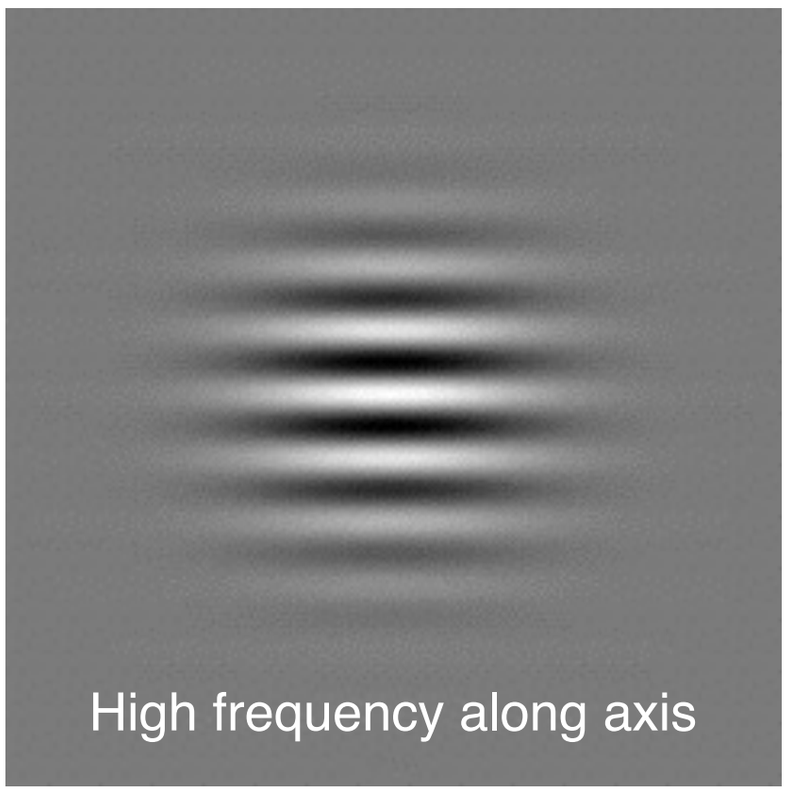
$$e^{-\frac{x^2}{2\sigma^2}} \sin(2\pi\omega x)$$

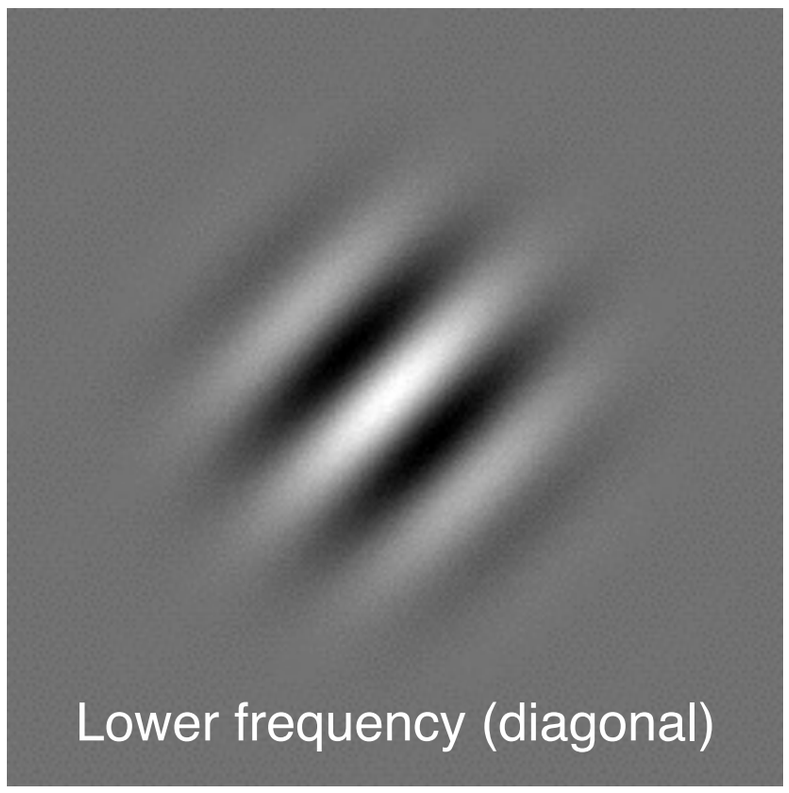


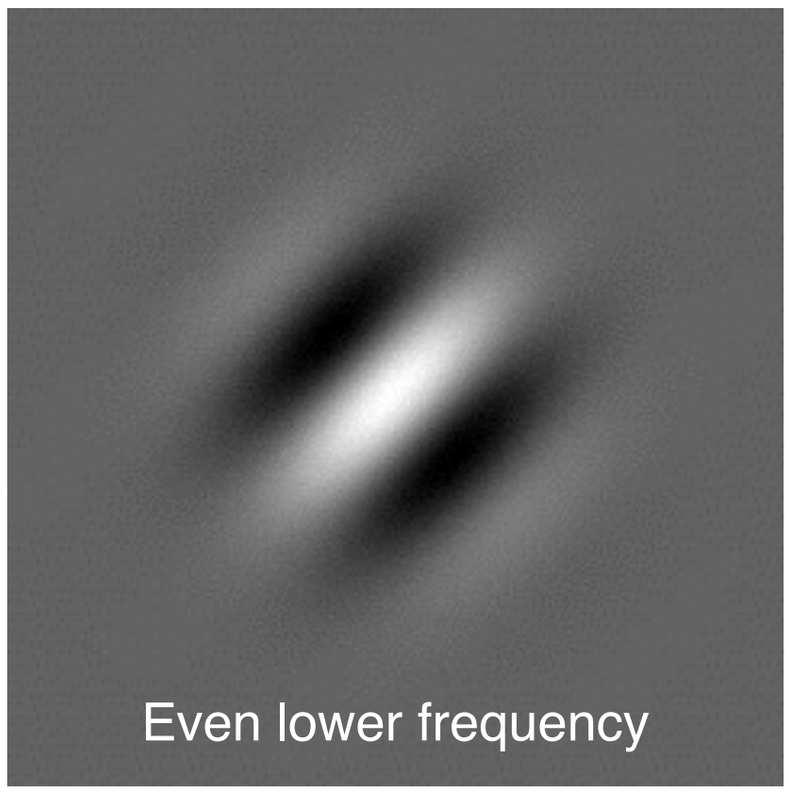
$$e^{-\frac{x^2}{2\sigma^2}} \cos(2\pi\omega x)$$

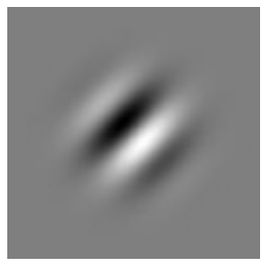
$$e^{-\frac{x^2+y^2}{2\sigma^2}} \cos\left(2\pi\left(k_x x + k_y y\right)\right)$$

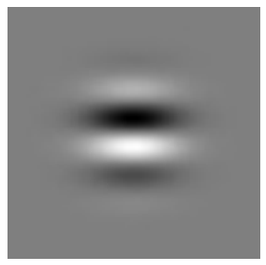


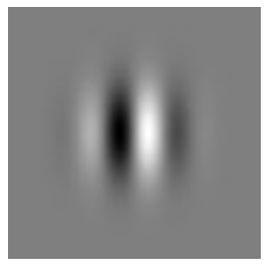


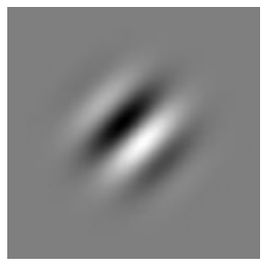


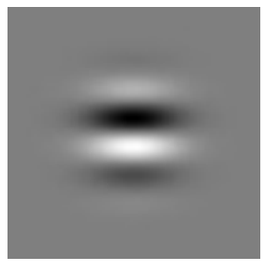


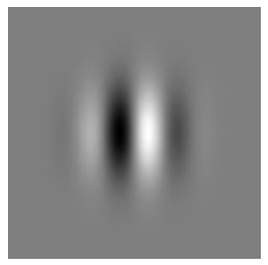


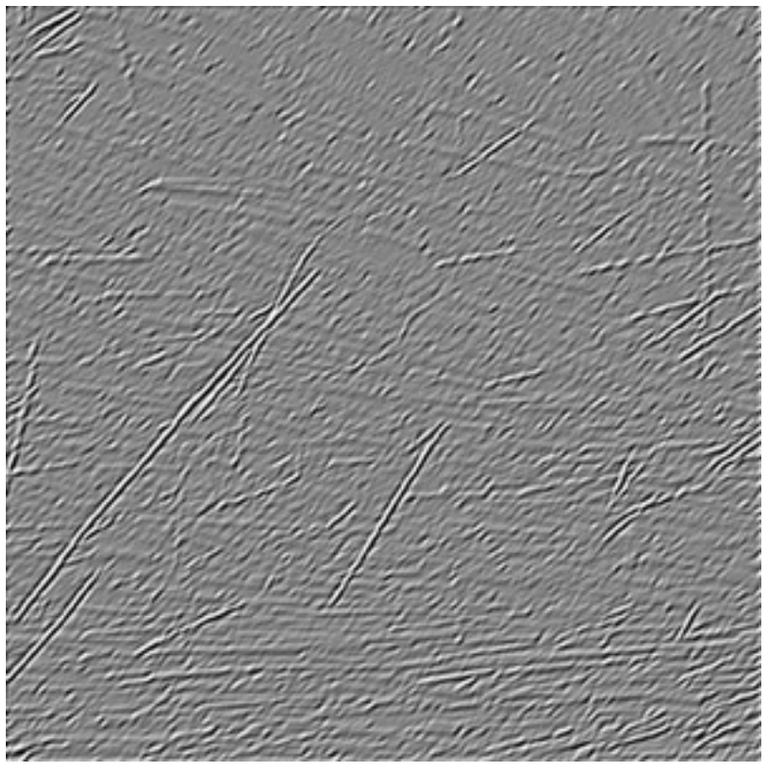


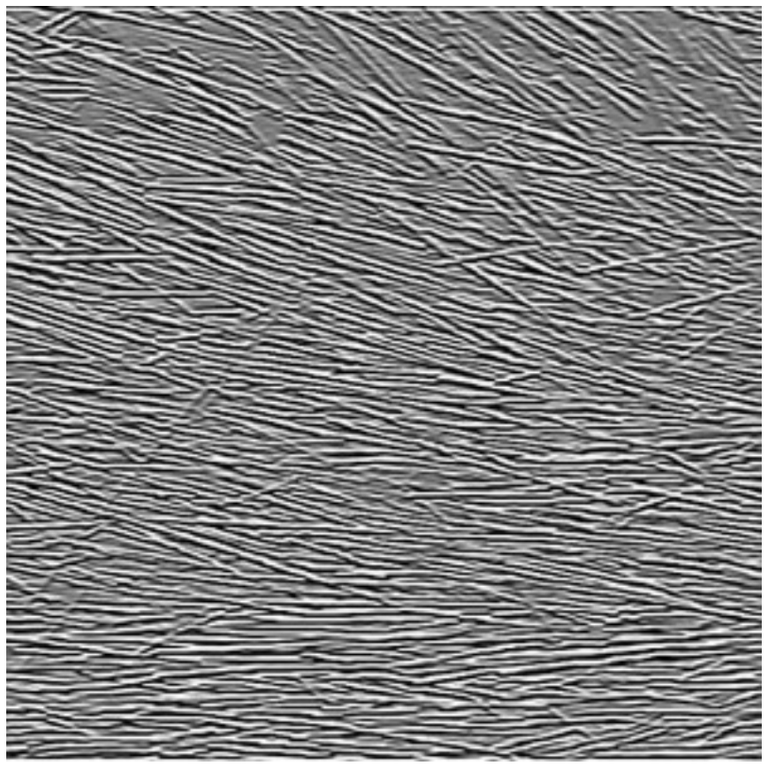


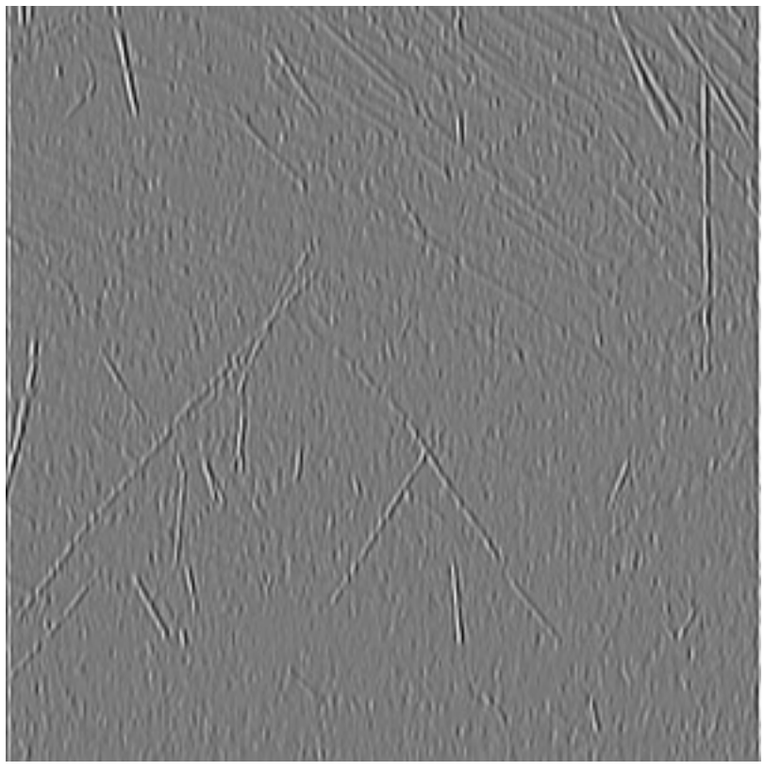


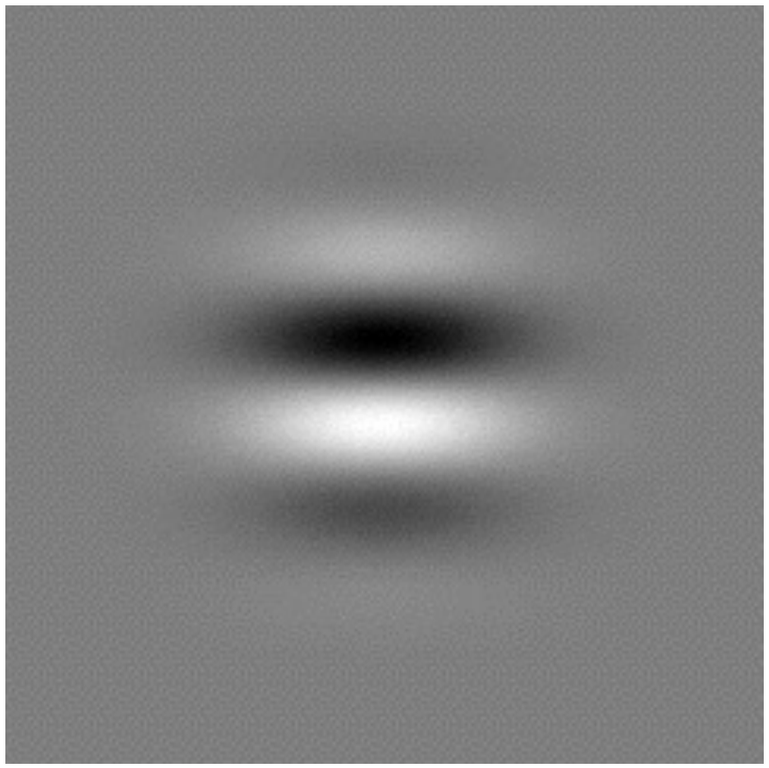




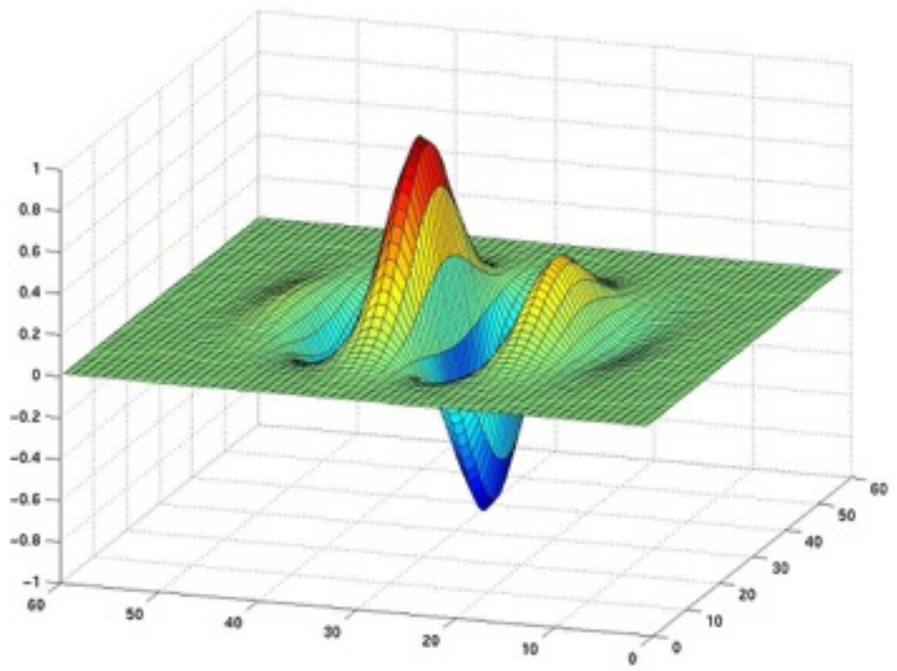




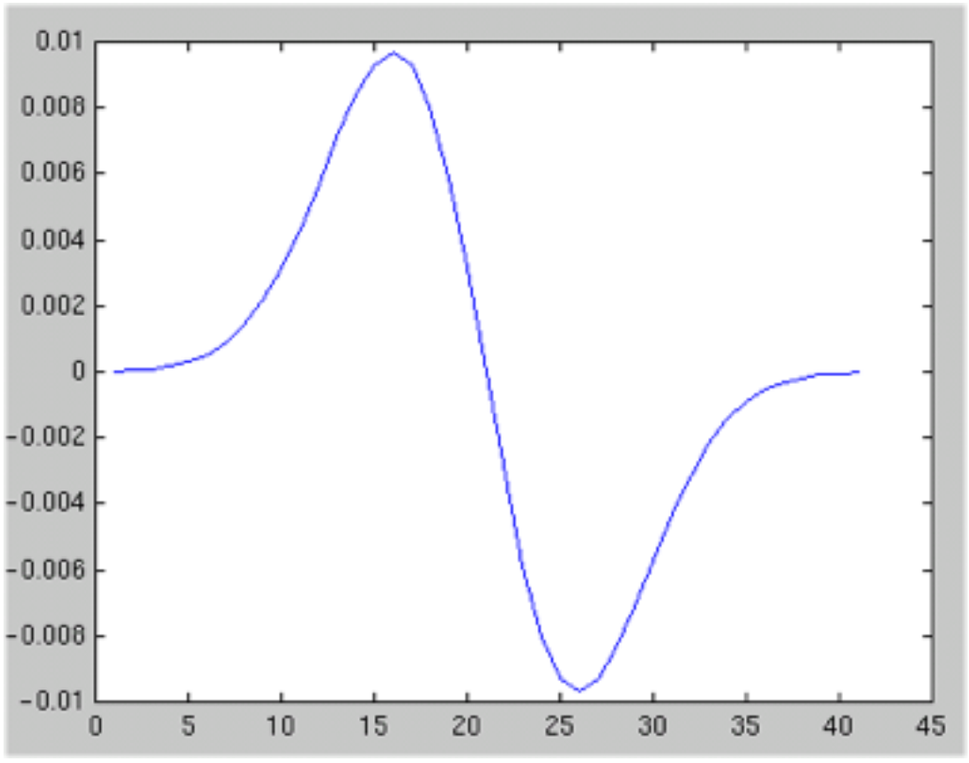




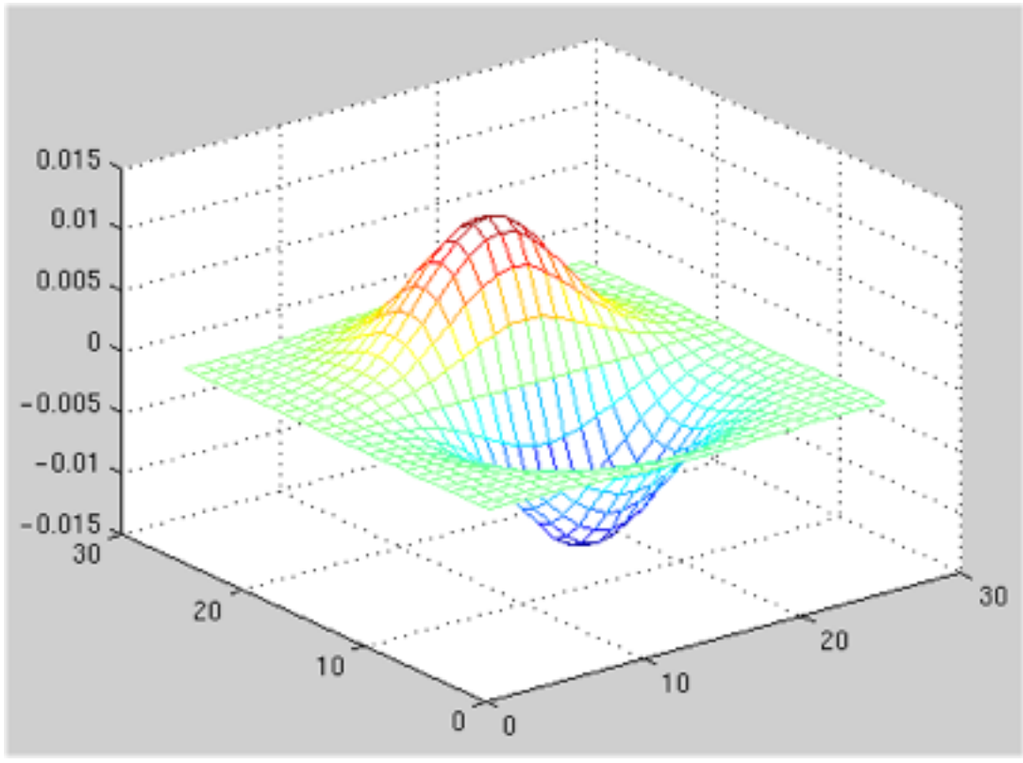
Odd Gabor filter

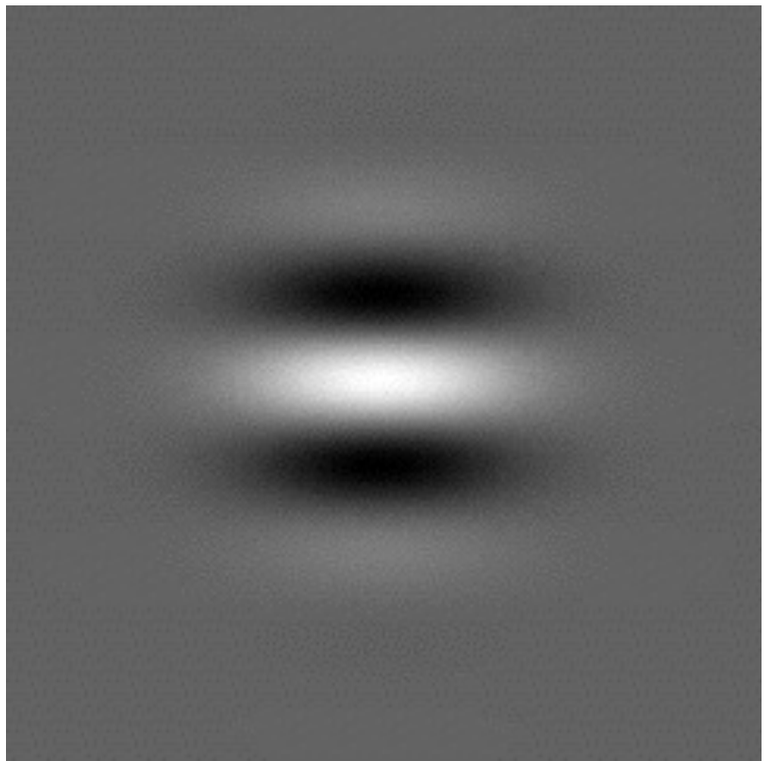


… looks a lot like…

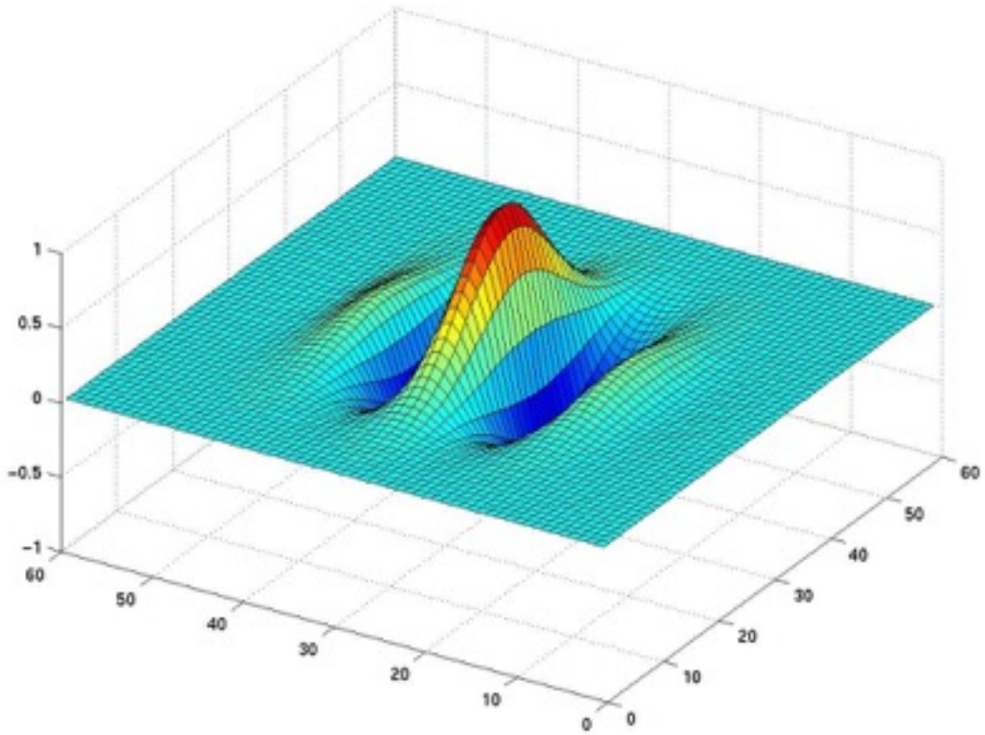


Gaussian Derivative

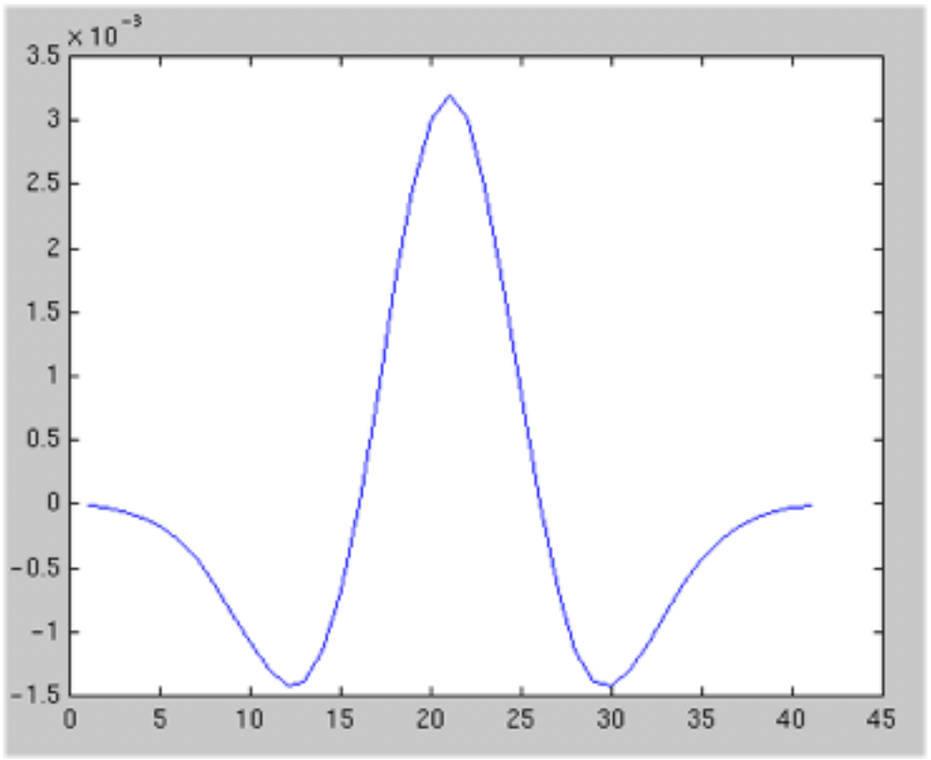




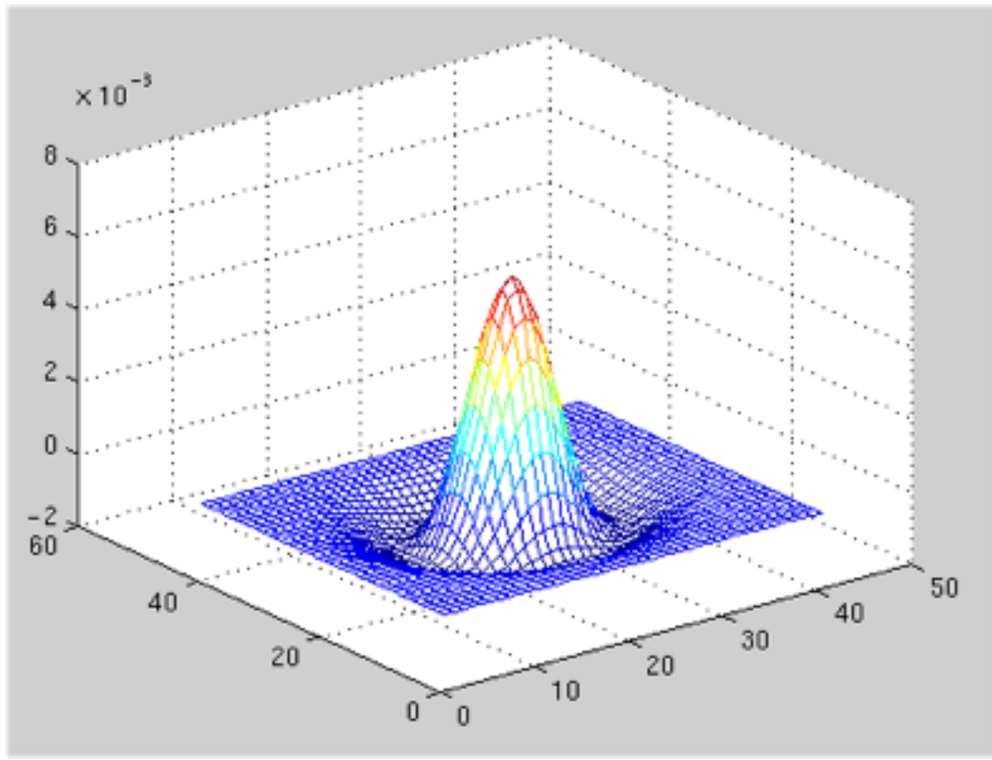
Even Gabor filter



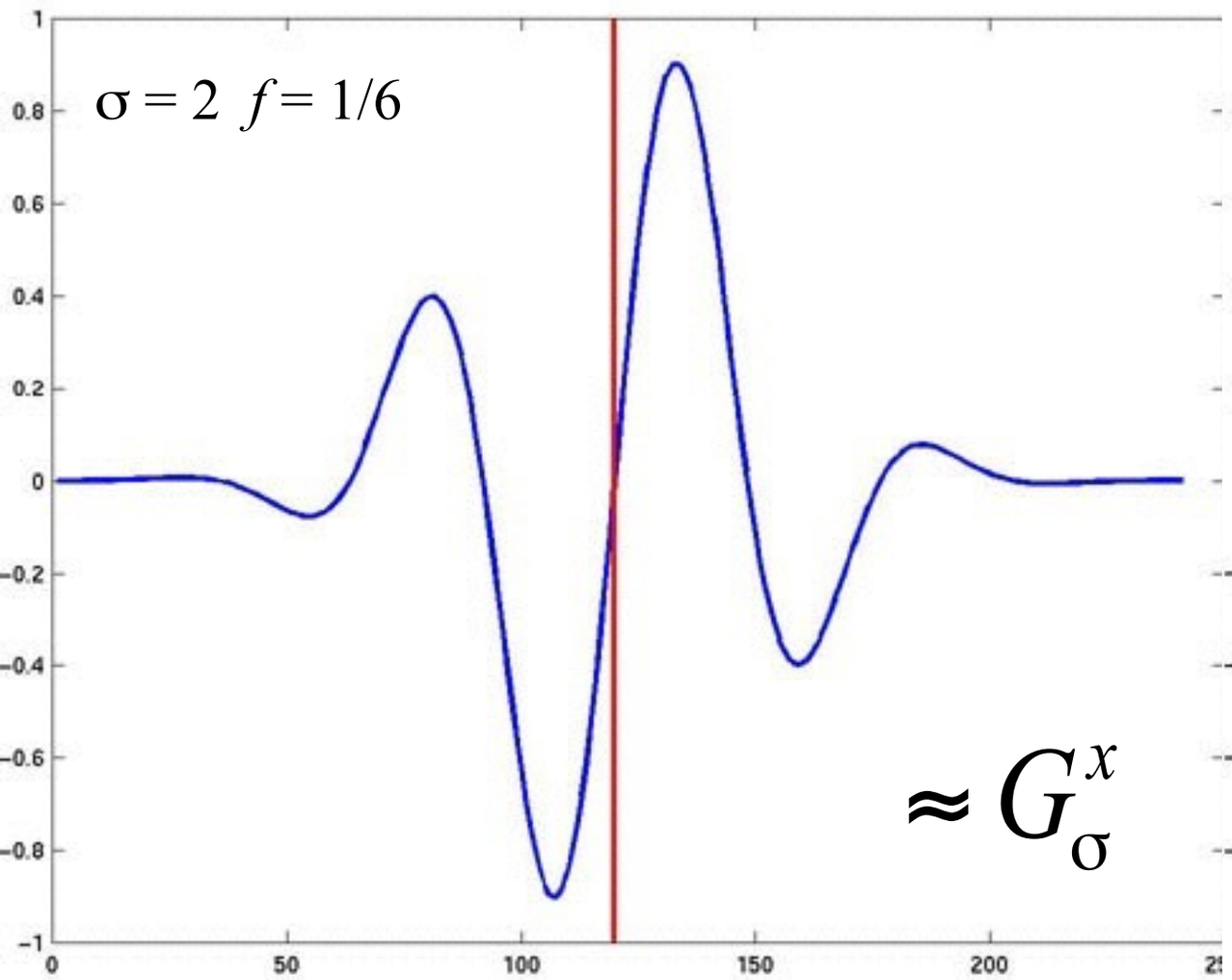
… looks a lot like…

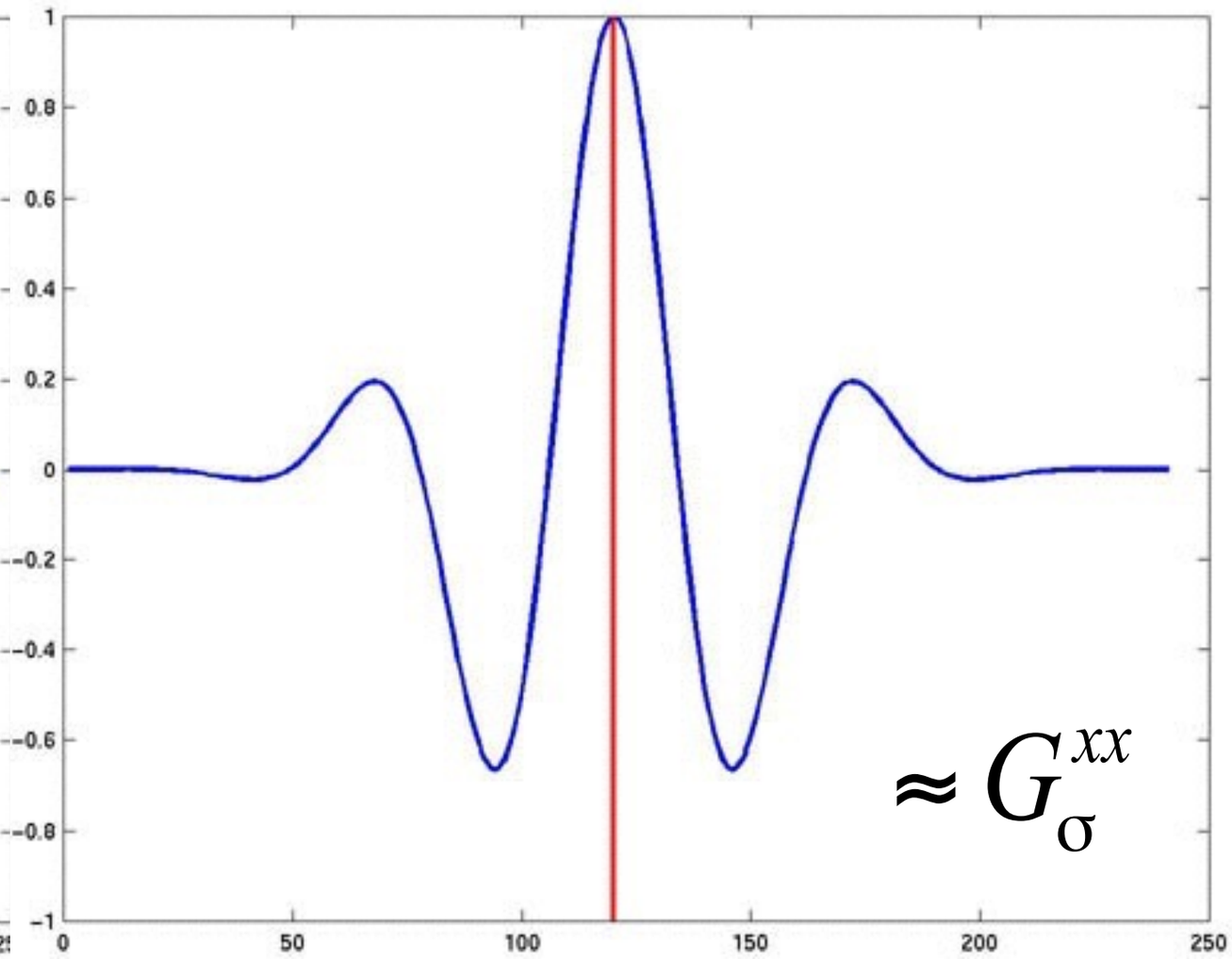


Laplacian

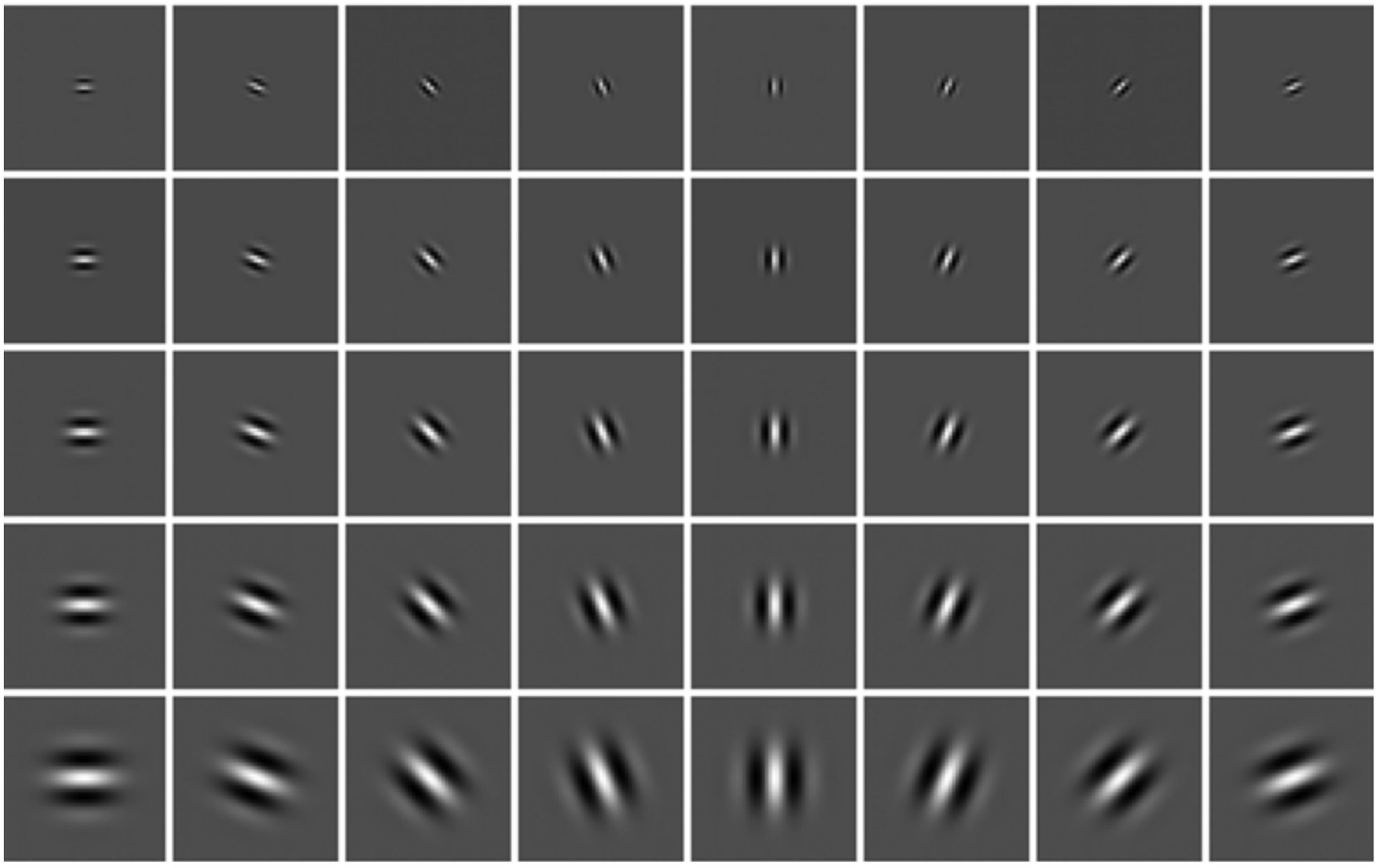


If scale small compared to inverse frequency,
the Gabor filters become derivative operators





Directional edge detectors

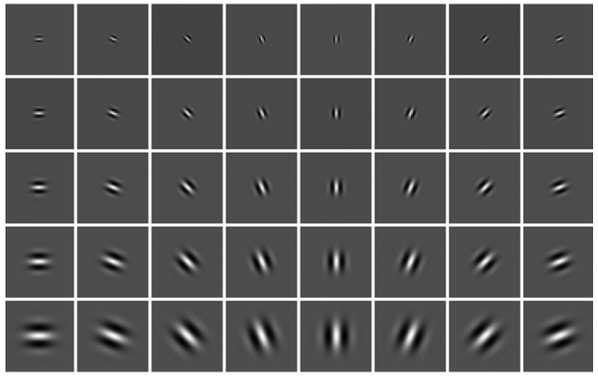


GIST

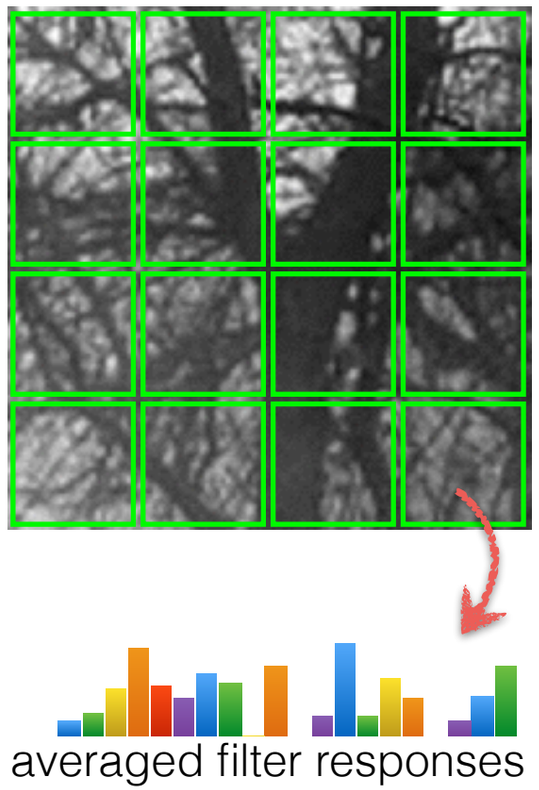
1. Compute filter responses (filter bank of Gabor filters)
2. Divide image patch into 4 x 4 cells
3. Compute filter response averages for each cell
4. Size of descriptor is 4 x 4 x N, where N is the size of the filter bank

What is the GIST descriptor encoding?

Filter bank



4 x 4 cell

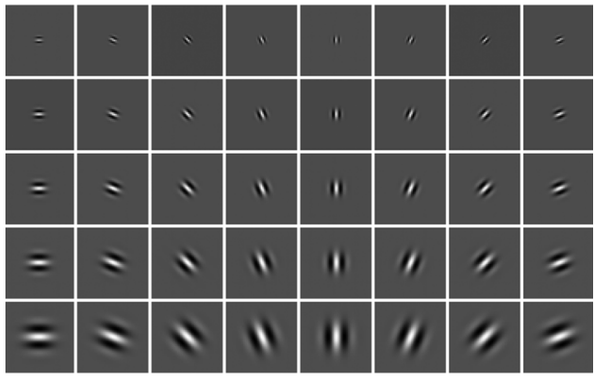


averaged filter responses

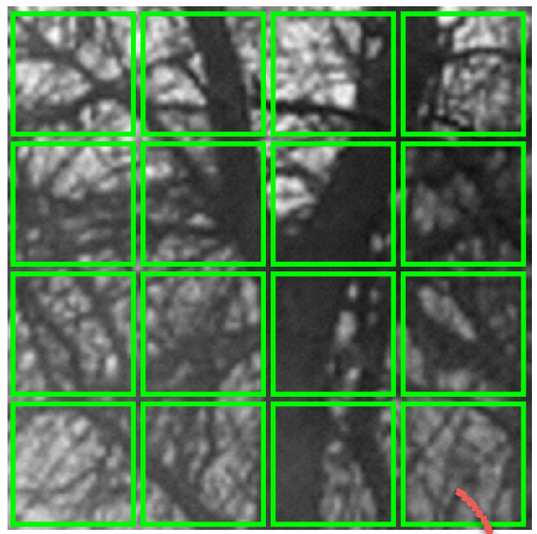
GIST

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?

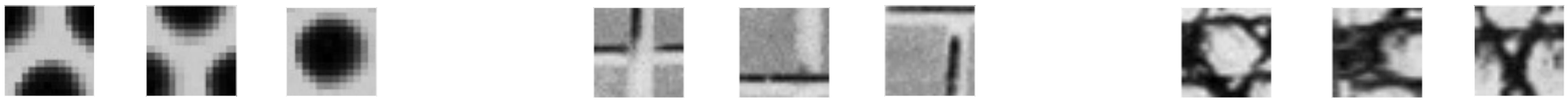
Rough spatial distribution of image gradients

When will this feature descriptor fail?

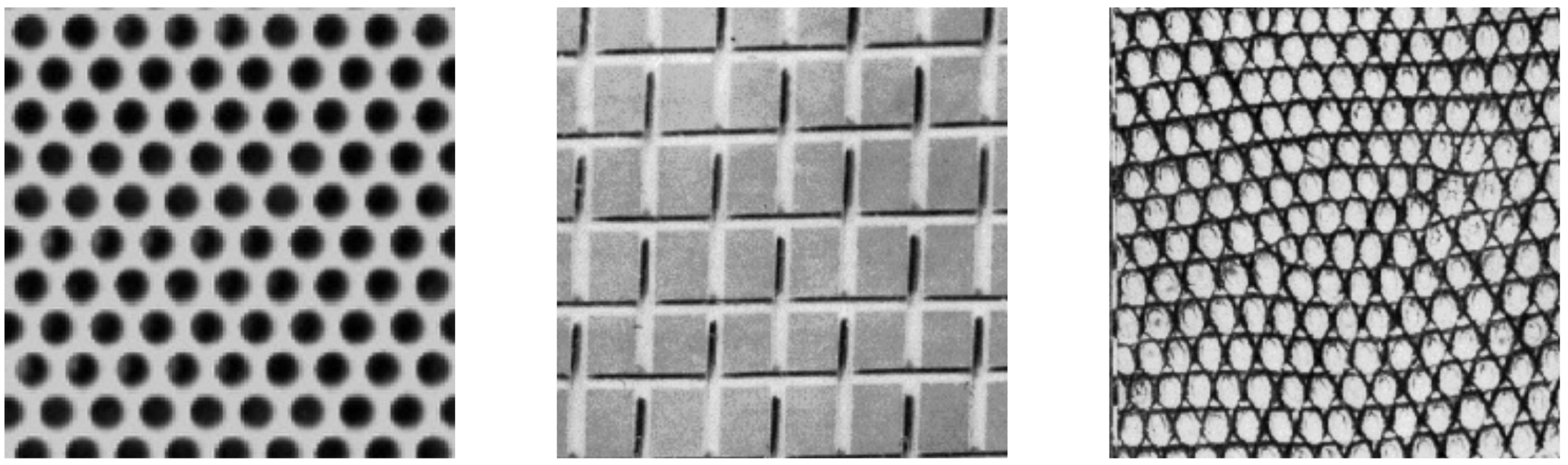
Textons

Julesz. Textons, the elements of texture perception, and their interactions. Nature 1981

Texture is characterized by the repetition of basic elements or ***textons***

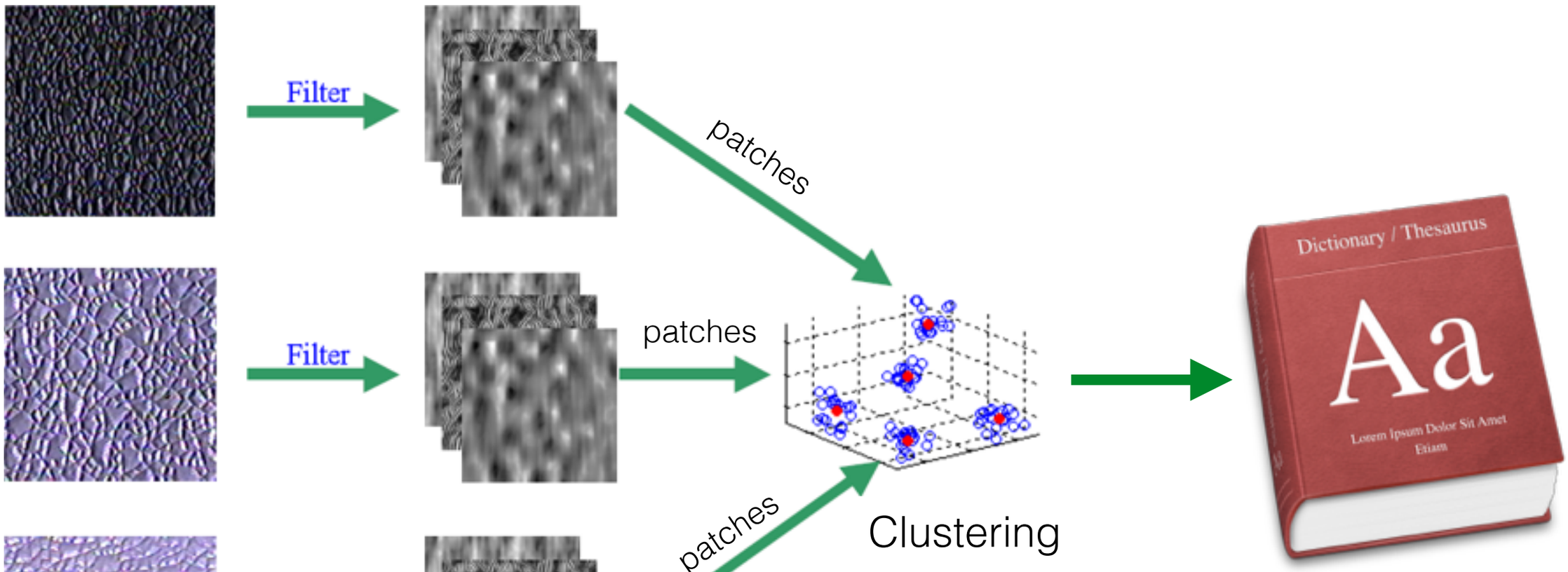


For stochastic textures, it is the identity of the ***textons***, not their spatial arrangement, that matters



Malik, Belongie, Shi, Leung. Textons, Contours and Regions: Cue Integration in Image Segmentation. ICCV 1999.

Learning Textons from data



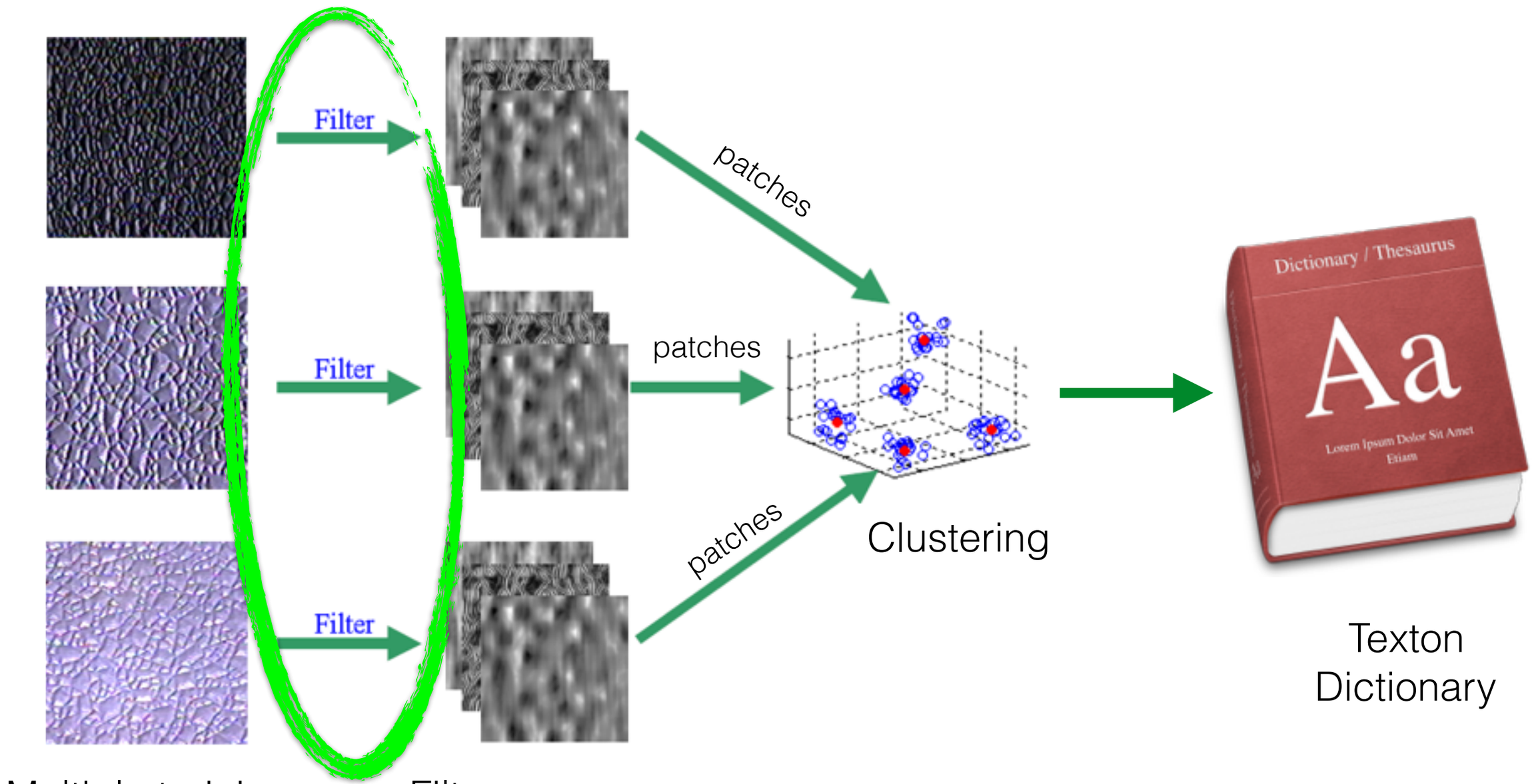
Multiple training images of the **same texture**

Filter response over a bank of filters

Clustering

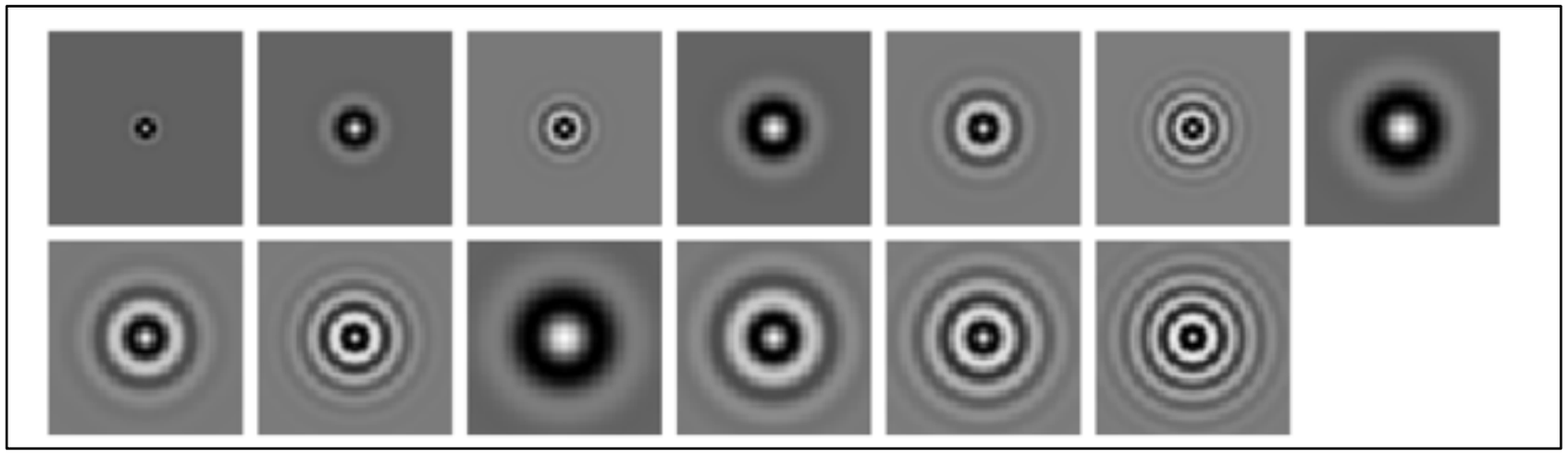
Texton Dictionary

Learning Textons from data



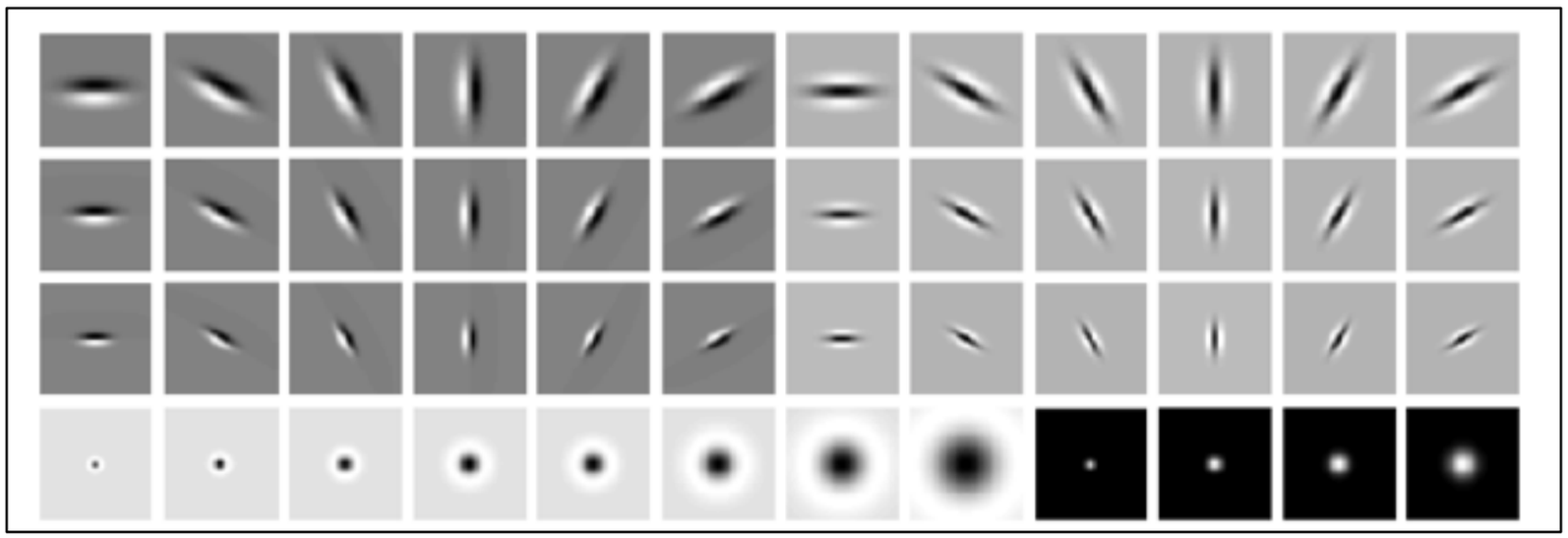
Example of Filter Banks

Isotropic Gabor

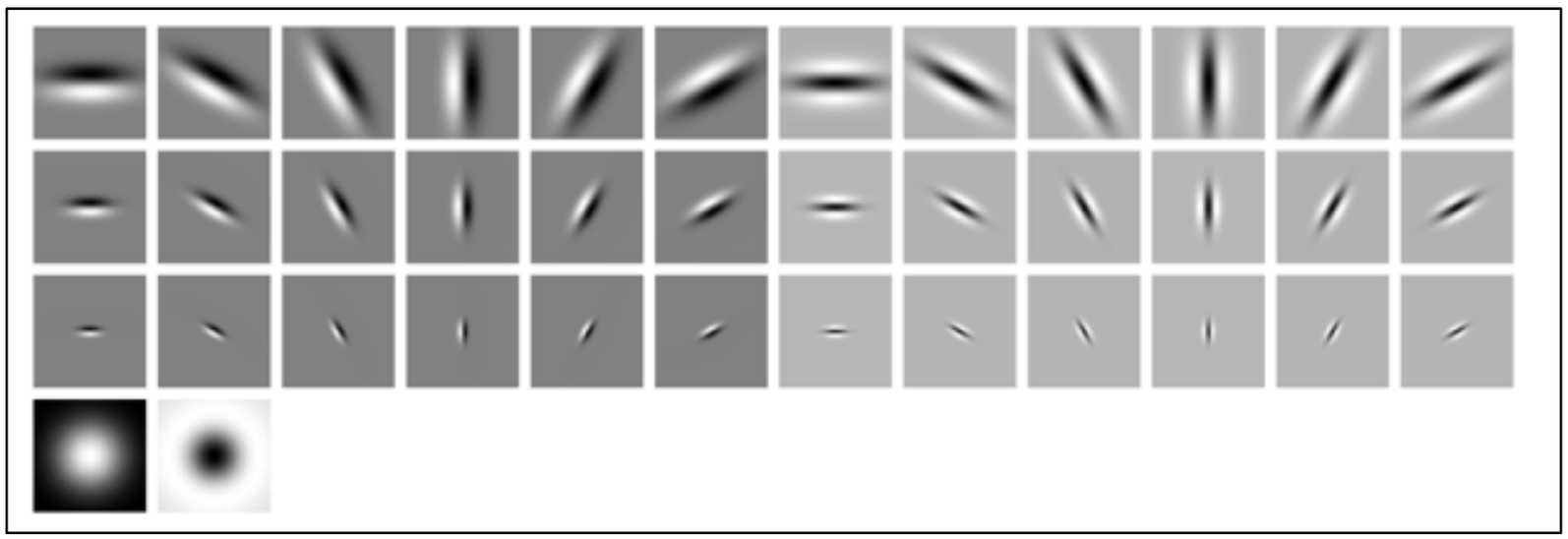


'S'

Gaussian derivatives at different scales and orientations

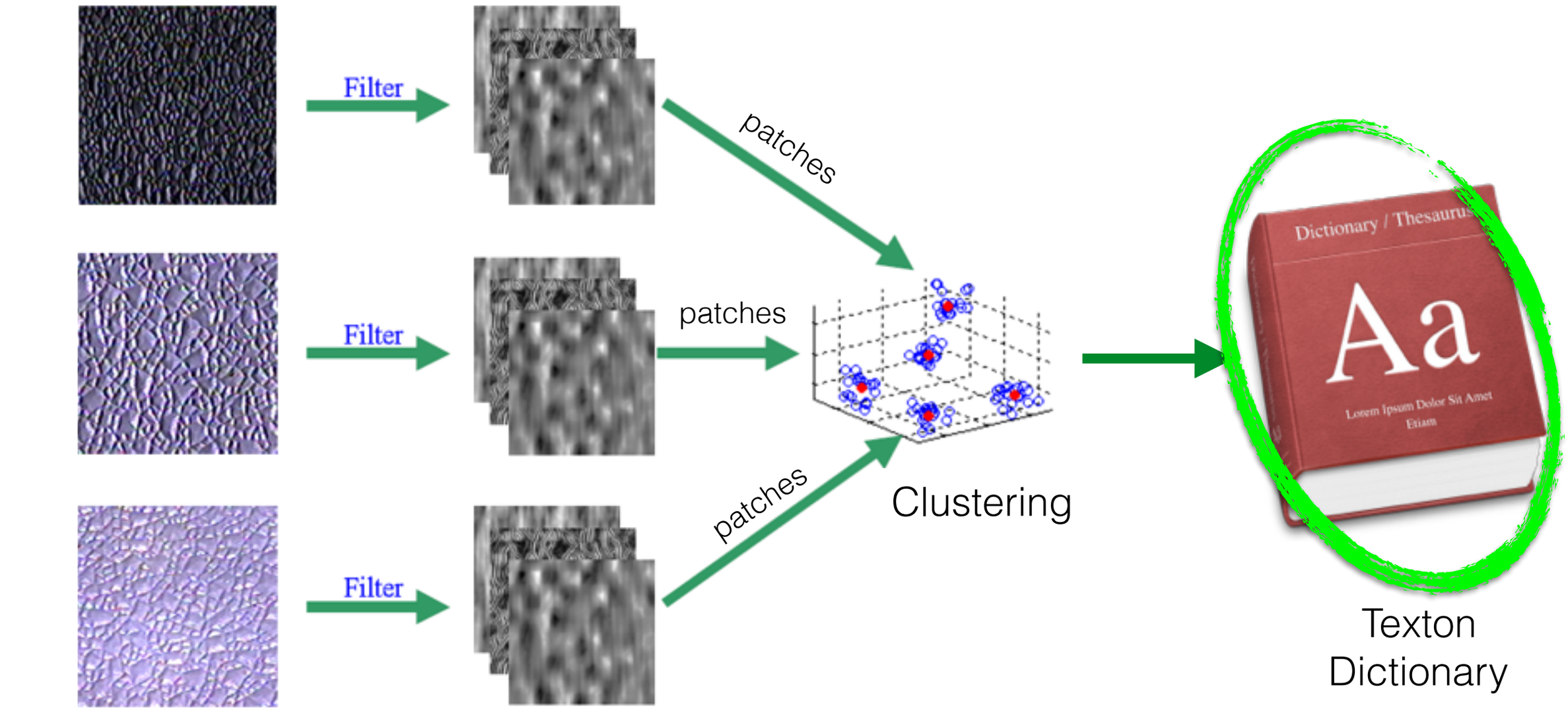


'LM'

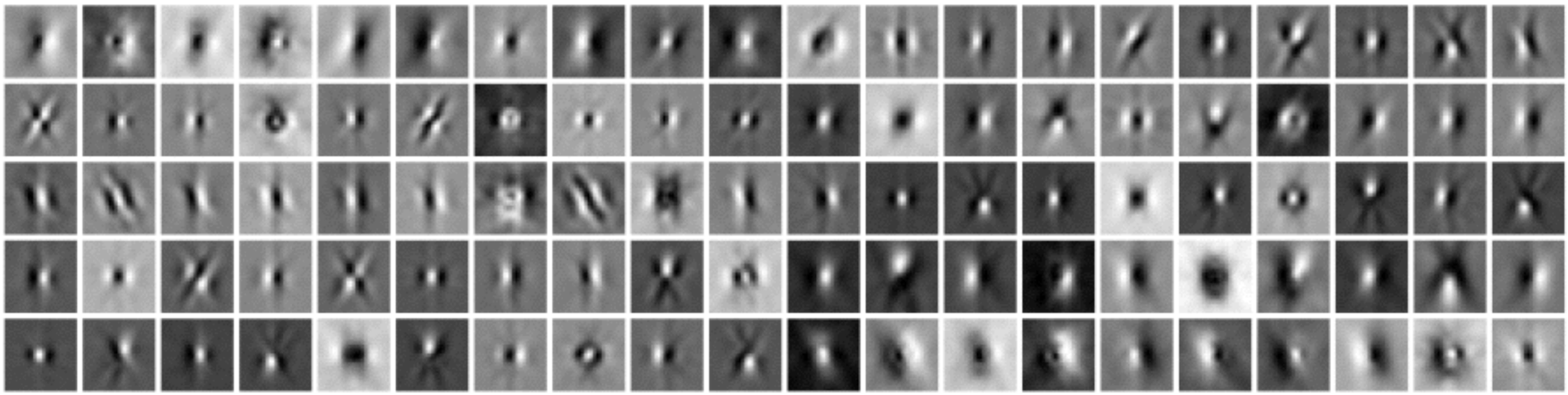


'MR8'

Learning Textons from data

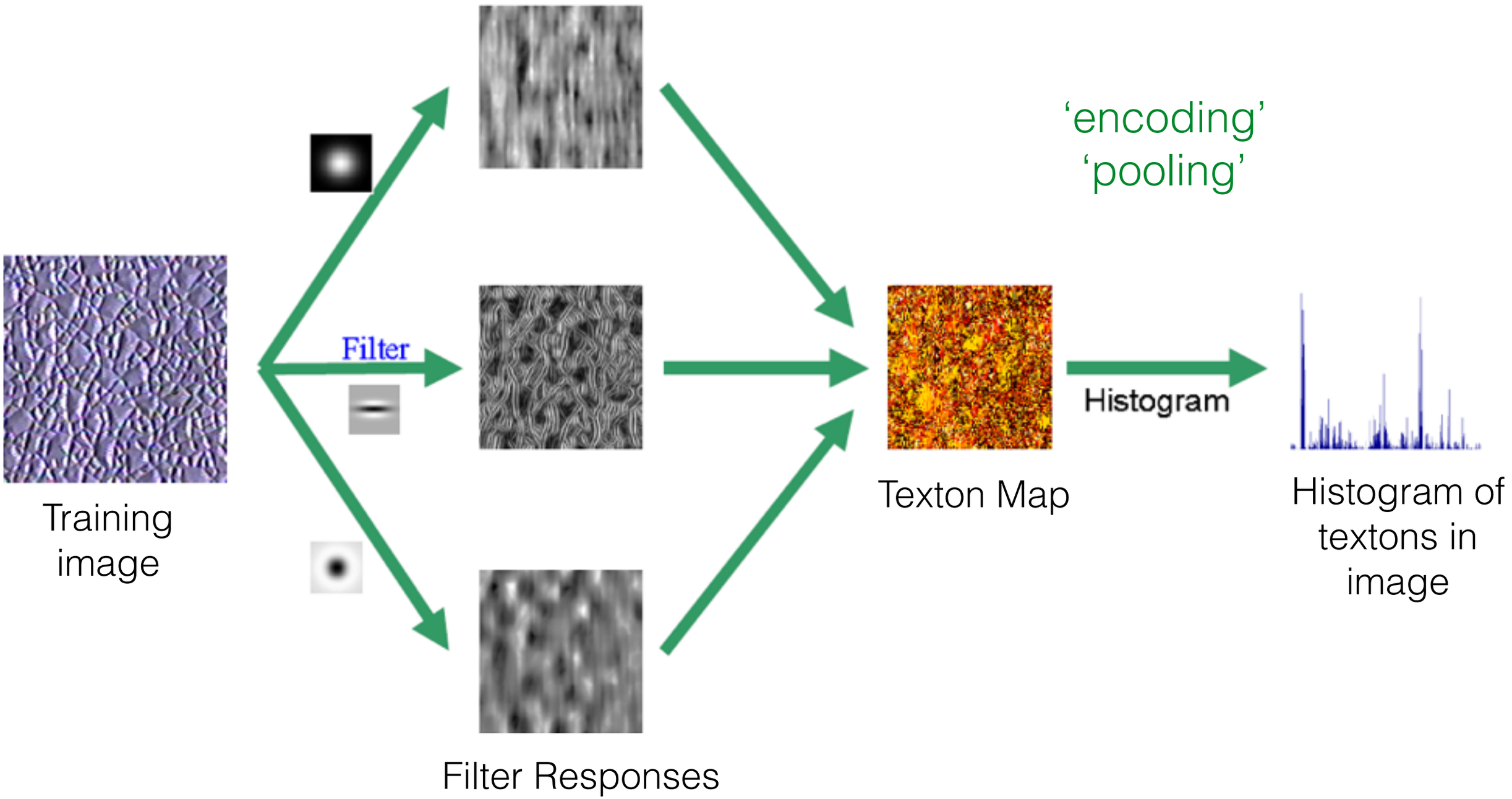


Texton Dictionary

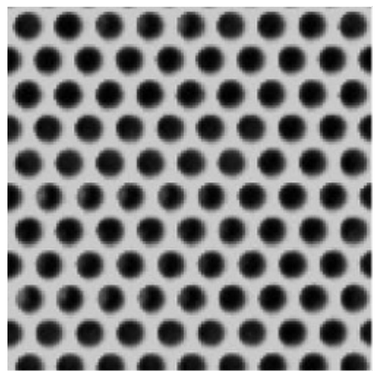


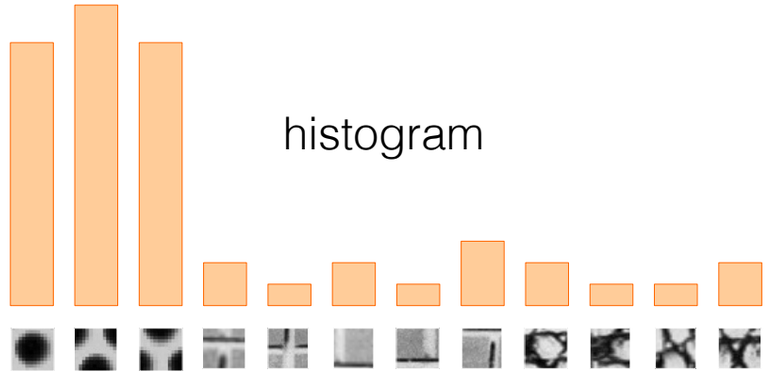
Malik, Belongie, Shi, Leung. Textons, Contours and Regions: Cue Integration in Image Segmentation. ICCV 1999.

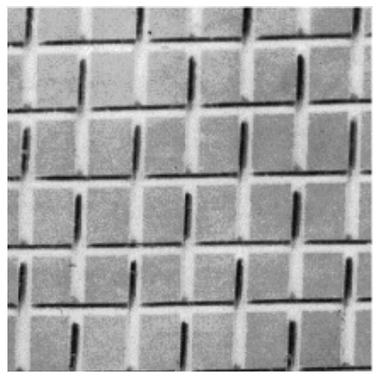
Histogram of Textons descriptor

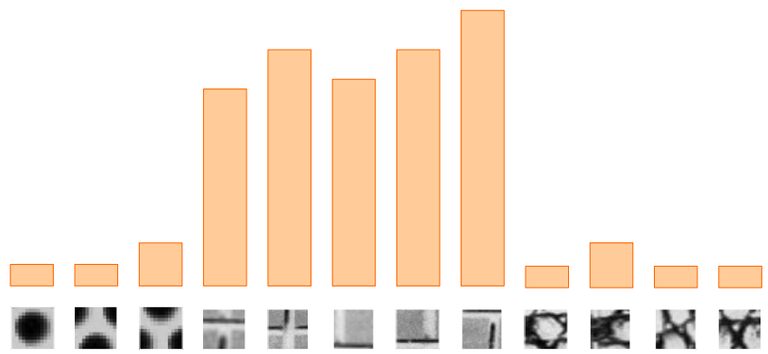


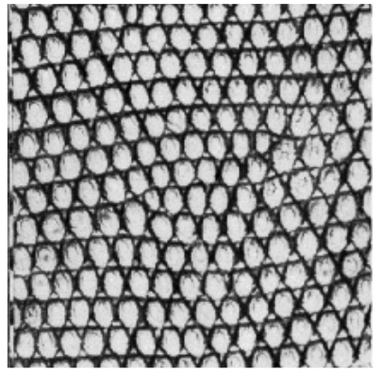
‘encoding’
‘pooling’

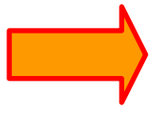


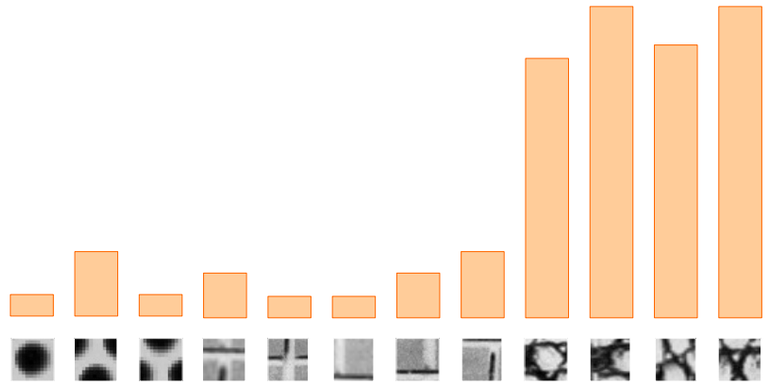










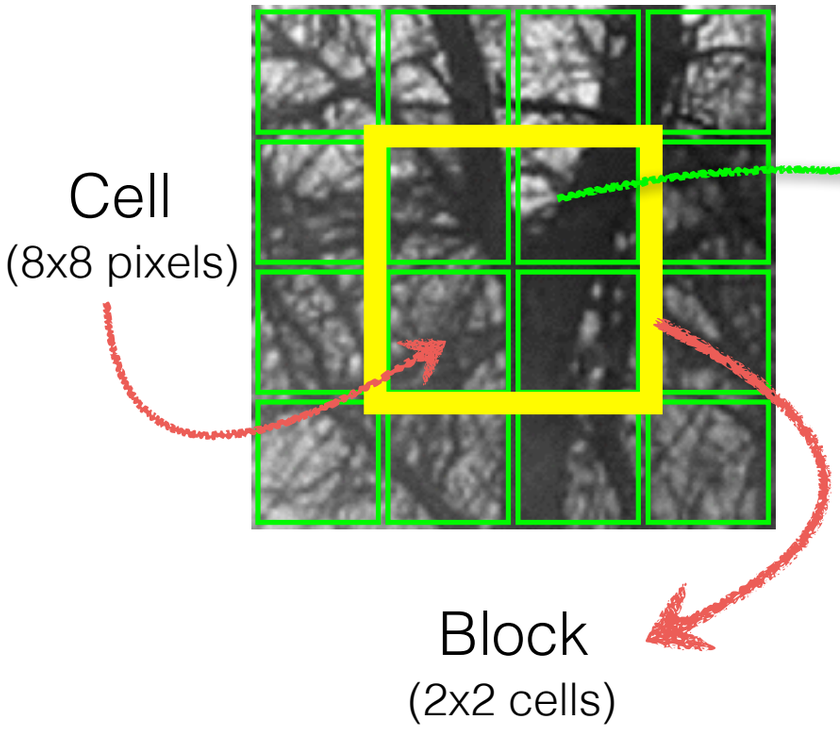


Julesz, 1981; Cula & Dana, 2001; Leung & Malik 2001; Mori, Belongie & Malik, 2001; Schmid 2001; Varma & Zisserman, 2002, 2003; Lazebnik, Schmid & Ponce, 2003

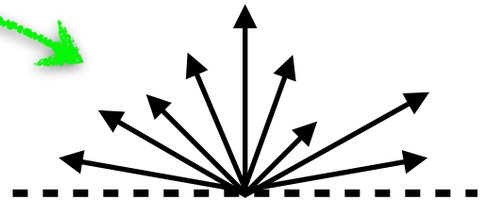
HOG descriptor

HOG

Dalal, Triggs. **Histograms of Oriented Gradients** for Human Detection. CVPR, 2005



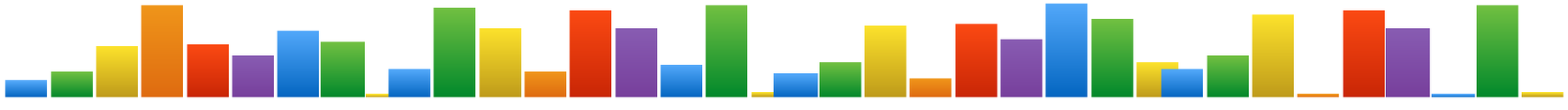
histogram of
'unsigned' gradients



soft binning

gradient magnitude
(one for each cell)

Concatenate and L-2 normalization

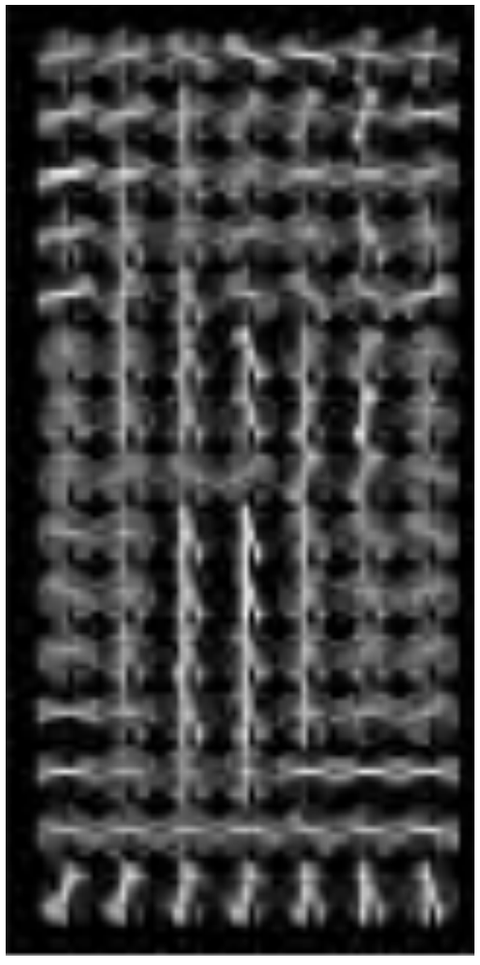


Single scale, no dominant orientation

Pedestrian detection

1 cell step size

visualization



128 pixels
16 cells
15 blocks

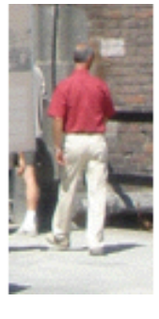
15 x 7 x 4 x 36 =
3780

64 pixels
8 cells
7 blocks

Redundant representation due to overlapping blocks

How many times is each inner cell encoded?



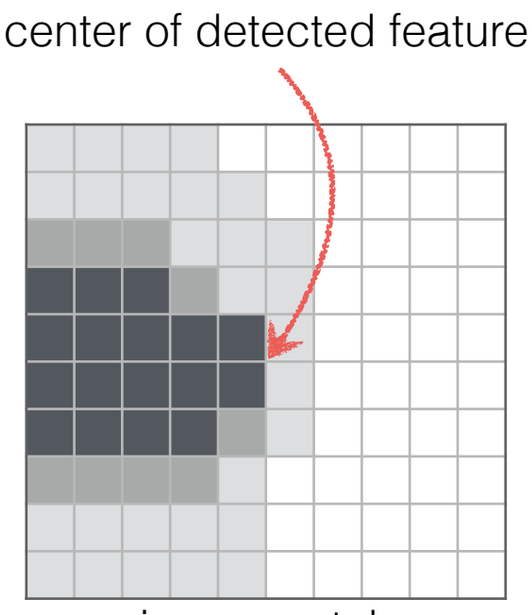


http://chrisjmccormick.wordpress.com/2013/05/09/hog-person-detector-tutorial/

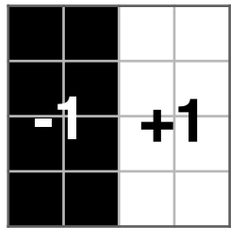
SURF

(‘Speeded’ Up Robust Features)

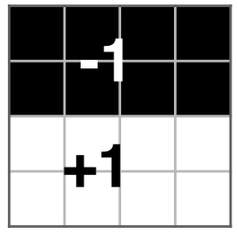
Compute Haar wavelet response at each pixel in patch



Haar wavelets filters



d_x



d_y

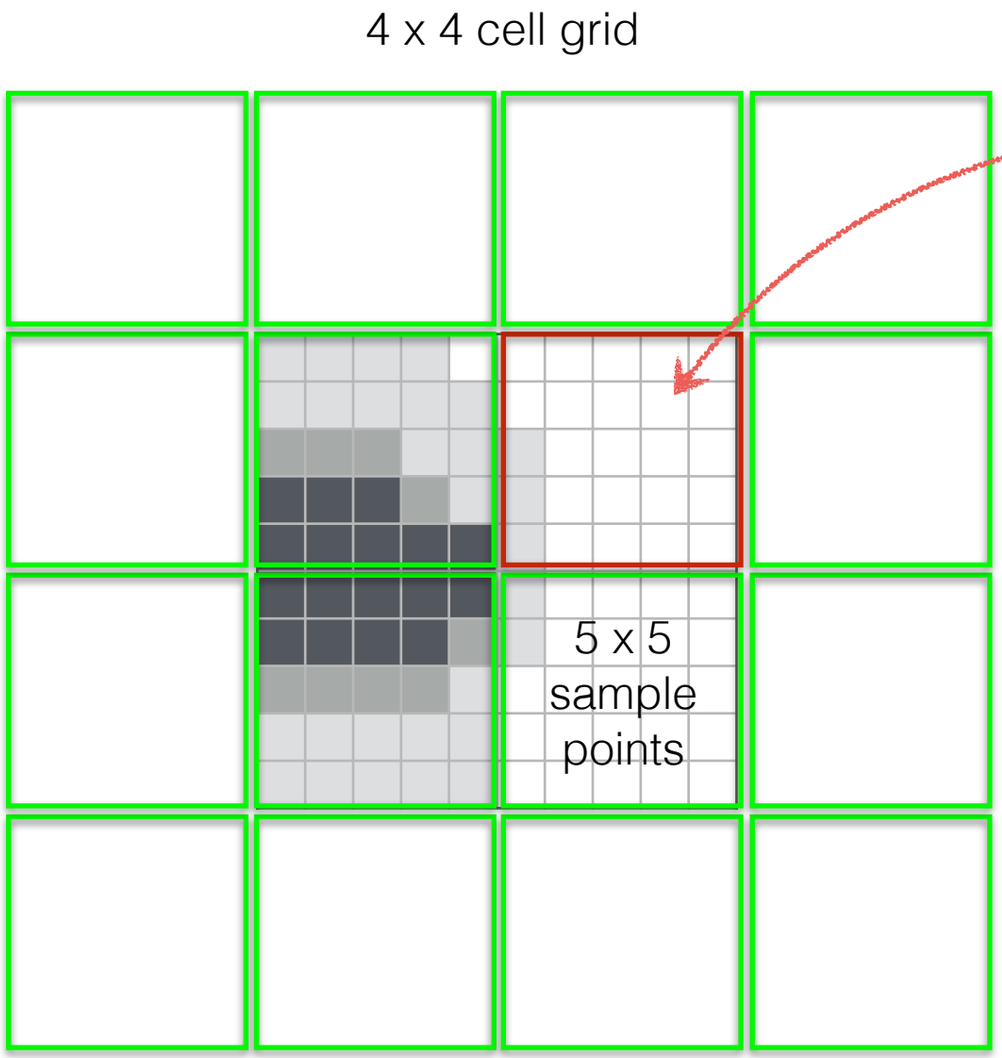
(Gaussian weighted from center)

How would do you compute the filter response?

Filtering using a sliding window can be slow
Haar wavelets are just sums over blocks
Use integral images for efficiency (6 operations)

SURF

(‘Speeded’ Up Robust Features)

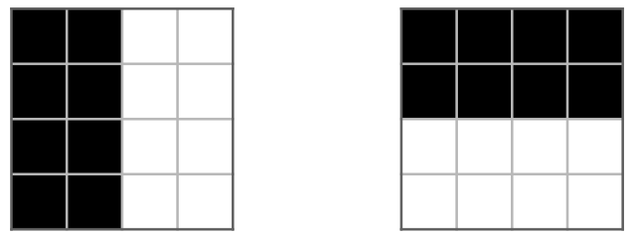


Each cell is represented by 4 values:

$$\left[\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right]$$

Haar wavelets filters

(Gaussian weighted from center)

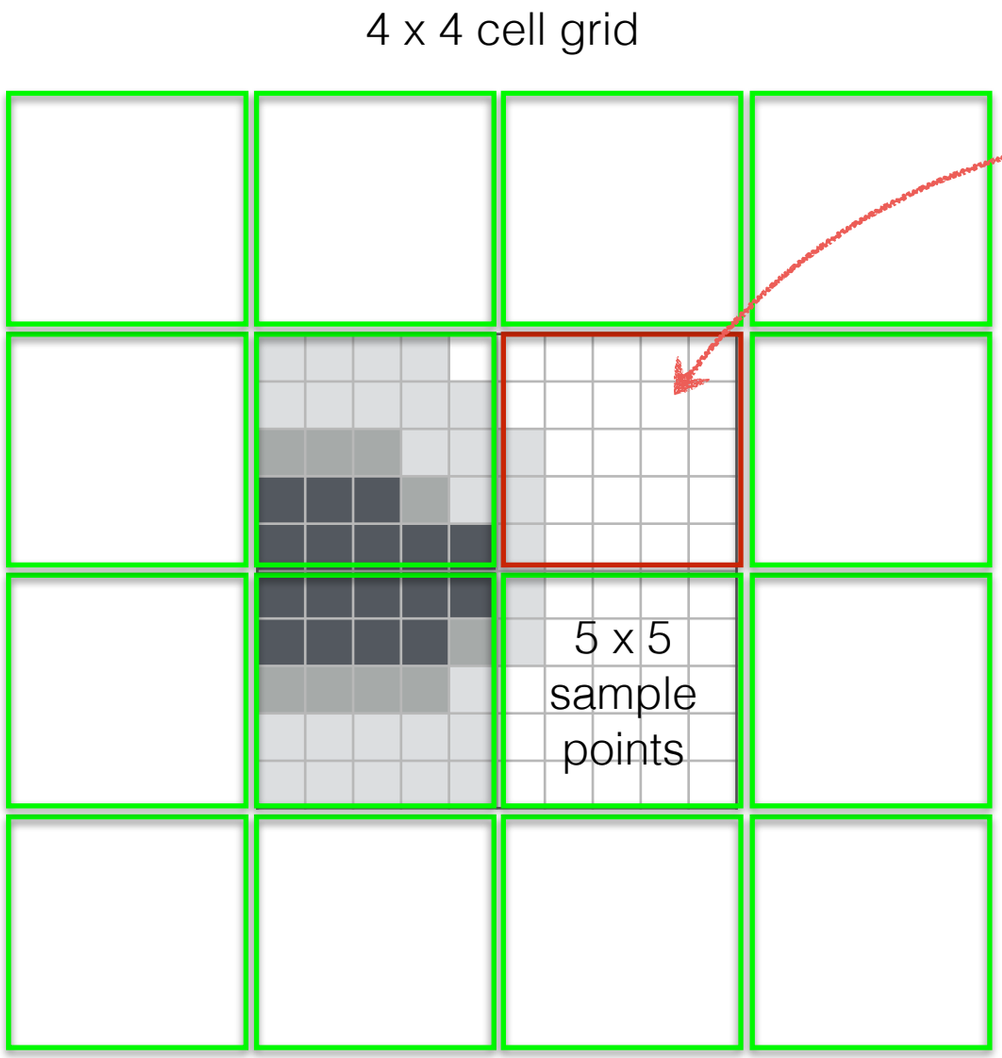


d_x d_y

How big is the SURF descriptor?

SURF

(‘Speeded’ Up Robust Features)

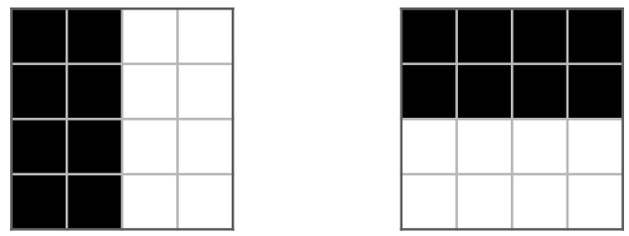


Each cell is represented by 4 values:

$$\left[\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right]$$

Haar wavelets filters

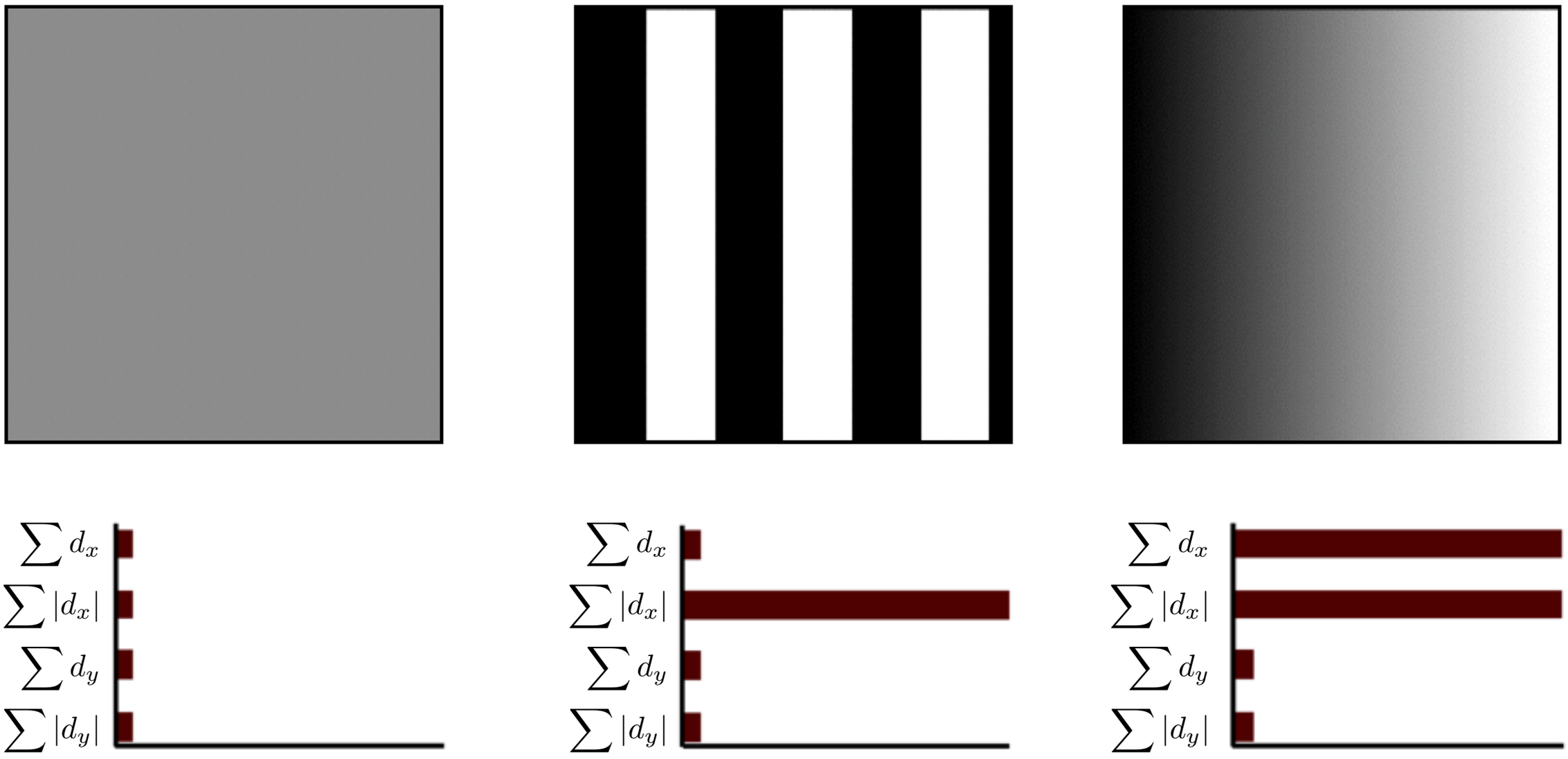
(Gaussian weighted from center)



d_x d_y

How big is the SURF descriptor?

64 dimensions


$\sum d_x$
$\sum |d_x|$
$\sum d_y$
$\sum |d_y|$
$\sum d_x$
$\sum |d_x|$
$\sum d_y$
$\sum |d_y|$
$\sum d_x$
$\sum |d_x|$
$\sum d_y$
$\sum |d_y|$

BRIEF

BRIEF: binary robust independent elementary features, Calonder, V Lepetit, C Strecha, ECCV 2010

Randomly sample pair of pixels a and b.
1 if a > b, else 0. Store binary vector.

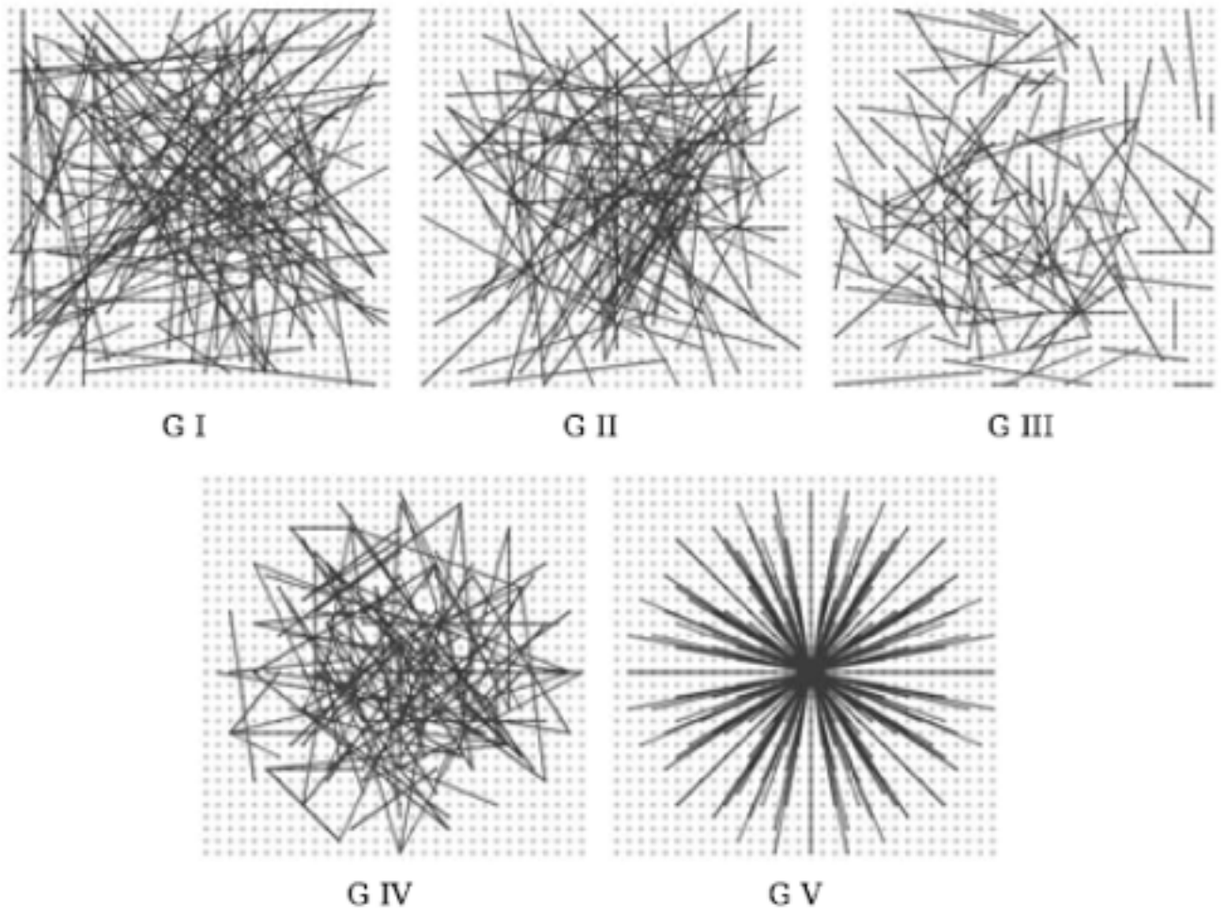


Fig. 2. Different approaches to choosing the test locations. All except the righmost one are selected by random sampling. Showing 128 tests in every image.