Lecture 10

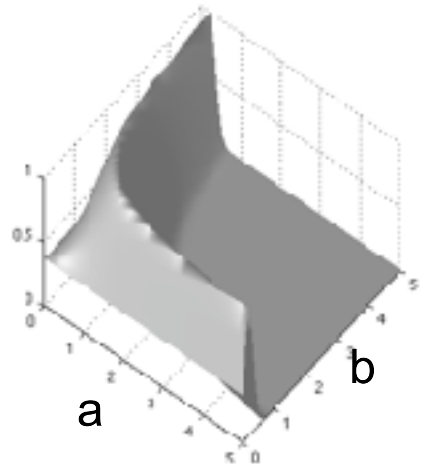
Statistical Image Models, part II

Bayesian approach

Use P(a, b | y = 1) = k P(y=1|a, b) P(a, b)

Likelihood function

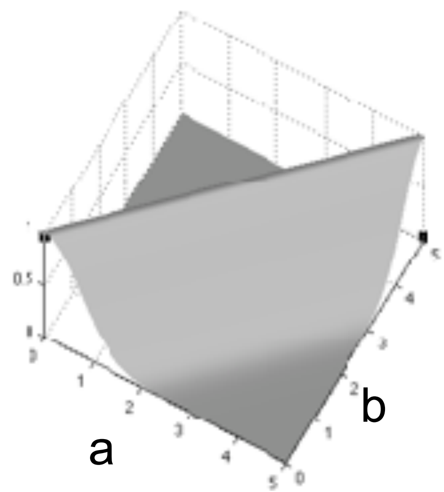
$$P(y=1|a,b) = ke^{-\frac{(1-ab)^2}{2\sigma^2}}$$

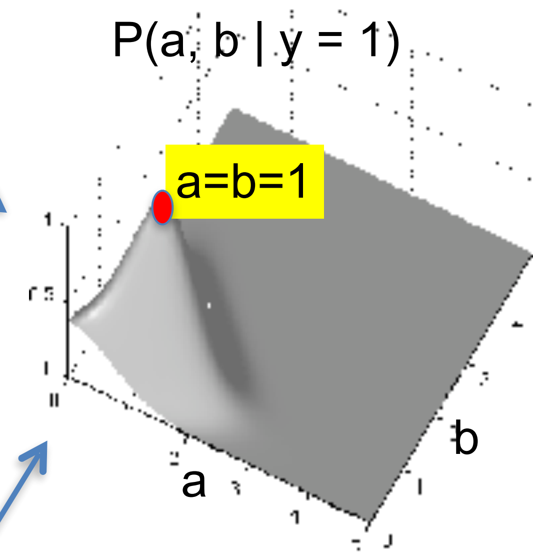


Prior probability

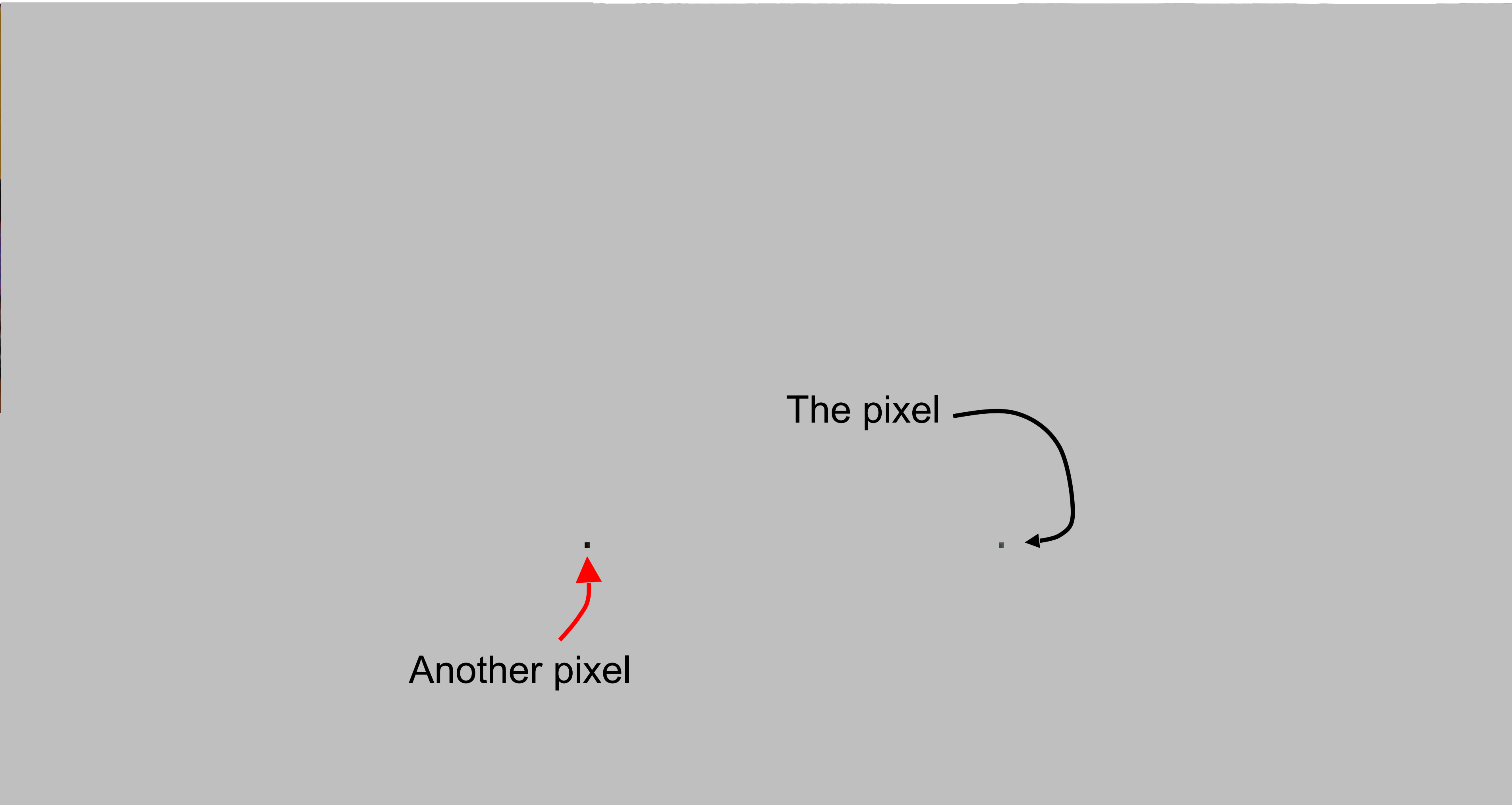
$$P(a,b) = ke^{-\frac{(a-b)^2}{2\sigma^2}} \quad \text{If } a>0,\ b>0$$

$= 0$ otherwise





Statistical modeling of images



$$C(\Delta x, \Delta y) = E\left[\mathbf{I}(x + \Delta x, y + \Delta y), \mathbf{I}(x, y)\right]$$

Gaussian model

We want a distribution that captures the correlation structure typical of natural images.

Let **C** be the covariance matrix of the image:

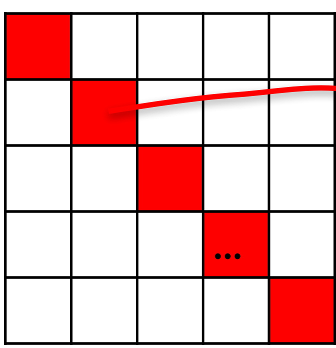
$$p(\mathbf{I}) = \exp\left(-\frac{1}{2}\mathbf{I}^T\mathbf{C}^{-1}\mathbf{I}\right) \qquad C = \begin{bmatrix} c_0 & c_1 & c_2 & & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & c_2 & & \vdots \\ & c_{n-1} & c_0 & c_1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & & c_2 \\ & & & & & c_1 \\ c_1 & \cdots & & c_{n-1} & & c_0 \end{bmatrix}$$

Stationarity assumption: Symmetrical circulant matrix

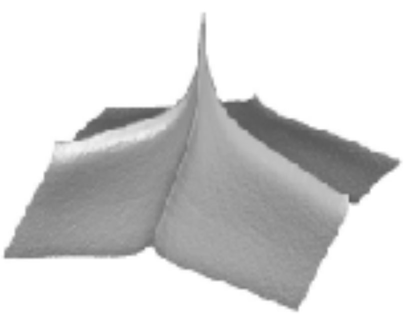
Diagonalization of circulant matrices: C = EDE^T

The eigenvectors are the Fourier basis

The eigenvalues are the squared magnitude of the Fourier coefficients

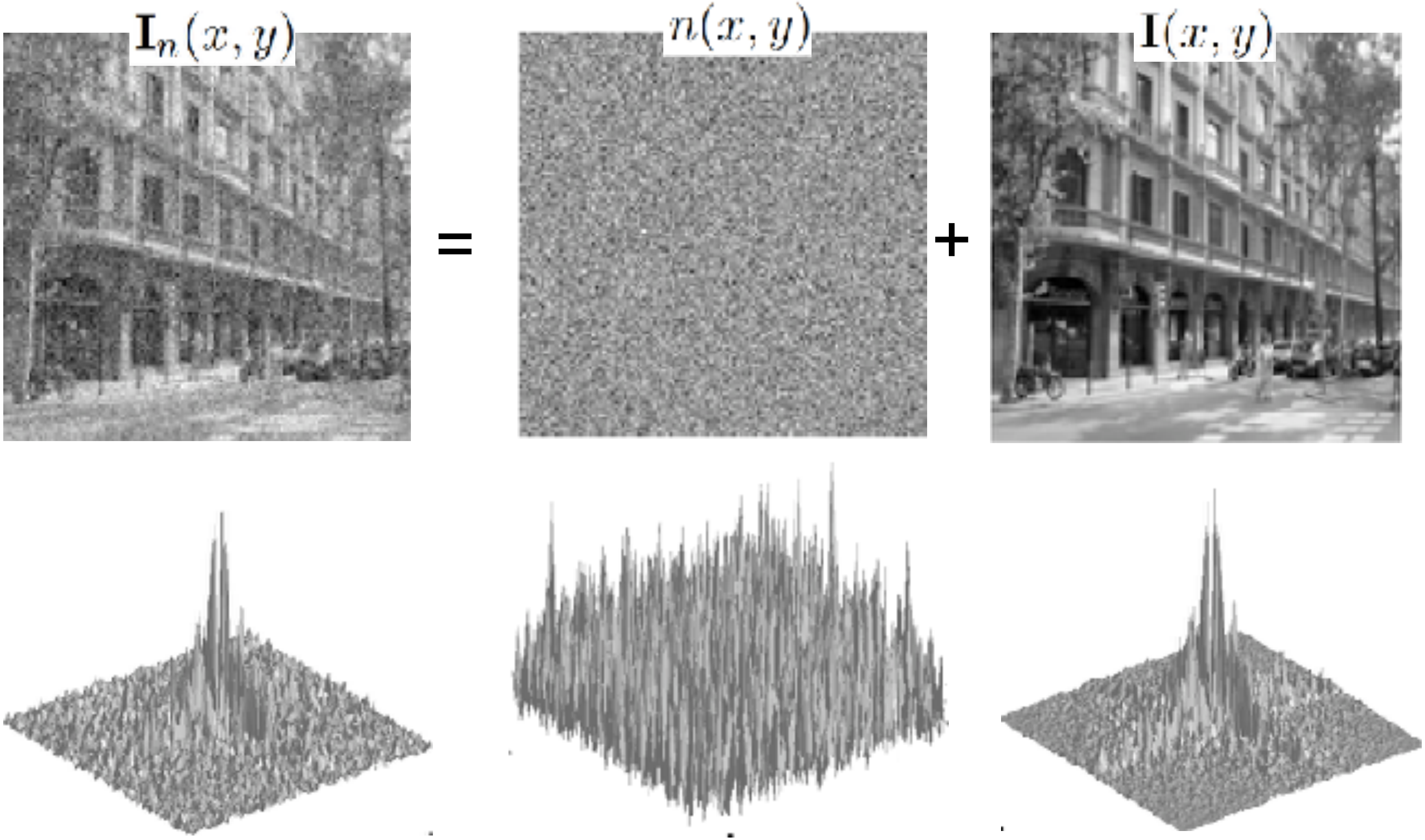
D= 

$$\left|\hat{\mathbf{I}}(v)\right|^2 \simeq \frac{1}{|v|^{2\alpha}}$$



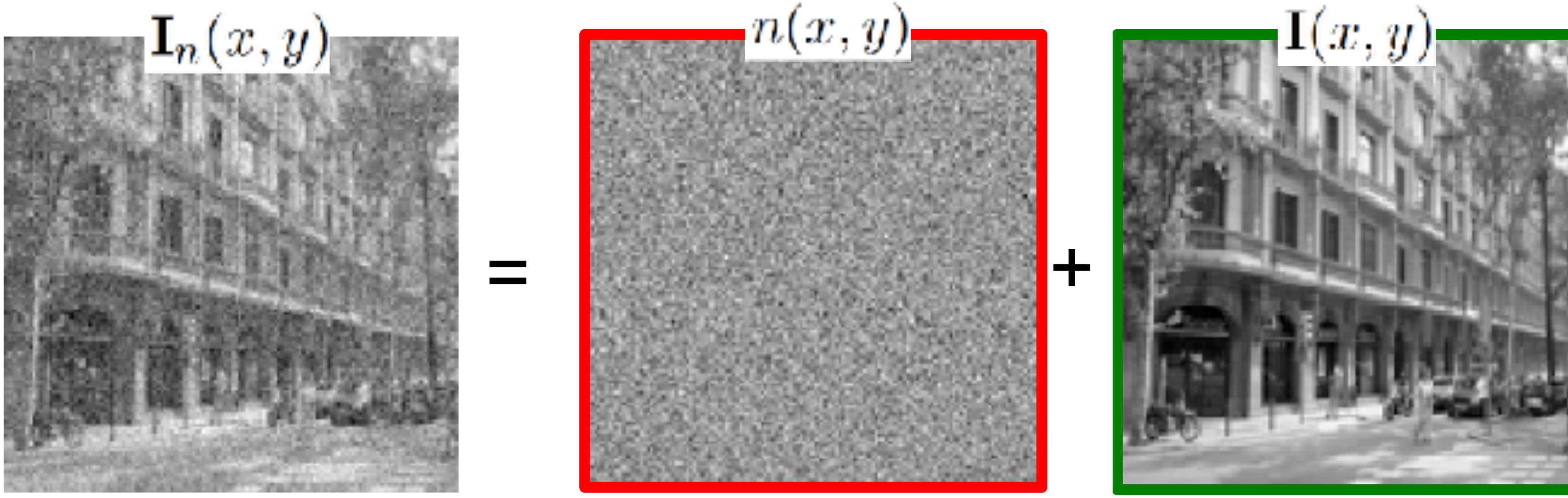
Denoising

Decomposition of a noisy image



Denoising

Decomposition of a noisy image



White Gaussian noise: $N(0, \sigma_n^2)$ Natural image

Find I(x,y) that maximizes the posterior (maximum a posteriory, MAP):

$$\max_{\mathbf{I}} p(\mathbf{I}|\mathbf{I}_n) = \max_{\mathbf{I}} \underbrace{p(\mathbf{I}_n|\mathbf{I})}_{\text{likelihood}} \times \underbrace{p(\mathbf{I})}_{\text{prior}}$$

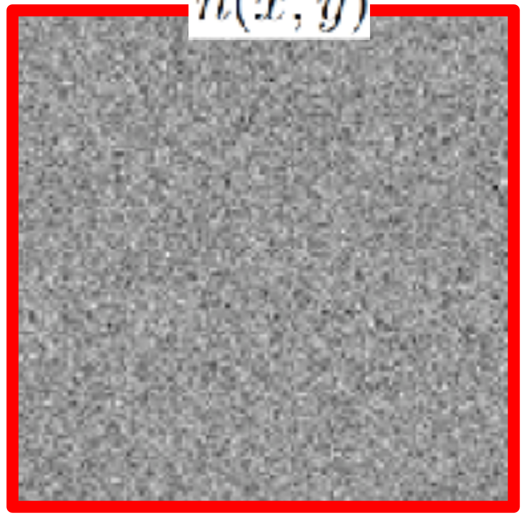
Denoising

Decomposition of a noisy image

$\mathbf{I}_n(x,y)$

=

$n(x,y)$



White Gaussian noise: $N(0, \sigma_n^2)$

+

$\mathbf{I}(x,y)$

Natural image

Find I(x,y) that maximizes the posterior (maximum a posteriory, MAP):

$$\max_{\mathbf{I}} p(\mathbf{I}|\mathbf{I}_n) = \max_{\mathbf{I}} \underbrace{p(\mathbf{I}_n|\mathbf{I})}_{\text{likelihood}} \times \underbrace{p(\mathbf{I})}_{\text{prior}}$$

$$= \max_{\mathbf{I}} \exp(-|\mathbf{I}_n - \mathbf{I}|^2/\sigma_n^2) \times \exp\left(-\frac{1}{2}\mathbf{I}^T\mathbf{C}^{-1}\mathbf{I}\right)$$

Denoising

$$\max_{\mathbf{I}} p(\mathbf{I}|\mathbf{I}_n) = \max_{\mathbf{I}} \quad p(\mathbf{I}_n|\mathbf{I}) \quad \text{x} \quad p(\mathbf{I})$$

likelihood prior

$$= \max_{\mathbf{I}} \exp(-|\mathbf{I}_n - \mathbf{I}|^2/\sigma_n^2) \text{ x } \exp\left(-\frac{1}{2}\mathbf{I}^T\mathbf{C}^{-1}\mathbf{I}\right)$$

The solution is:

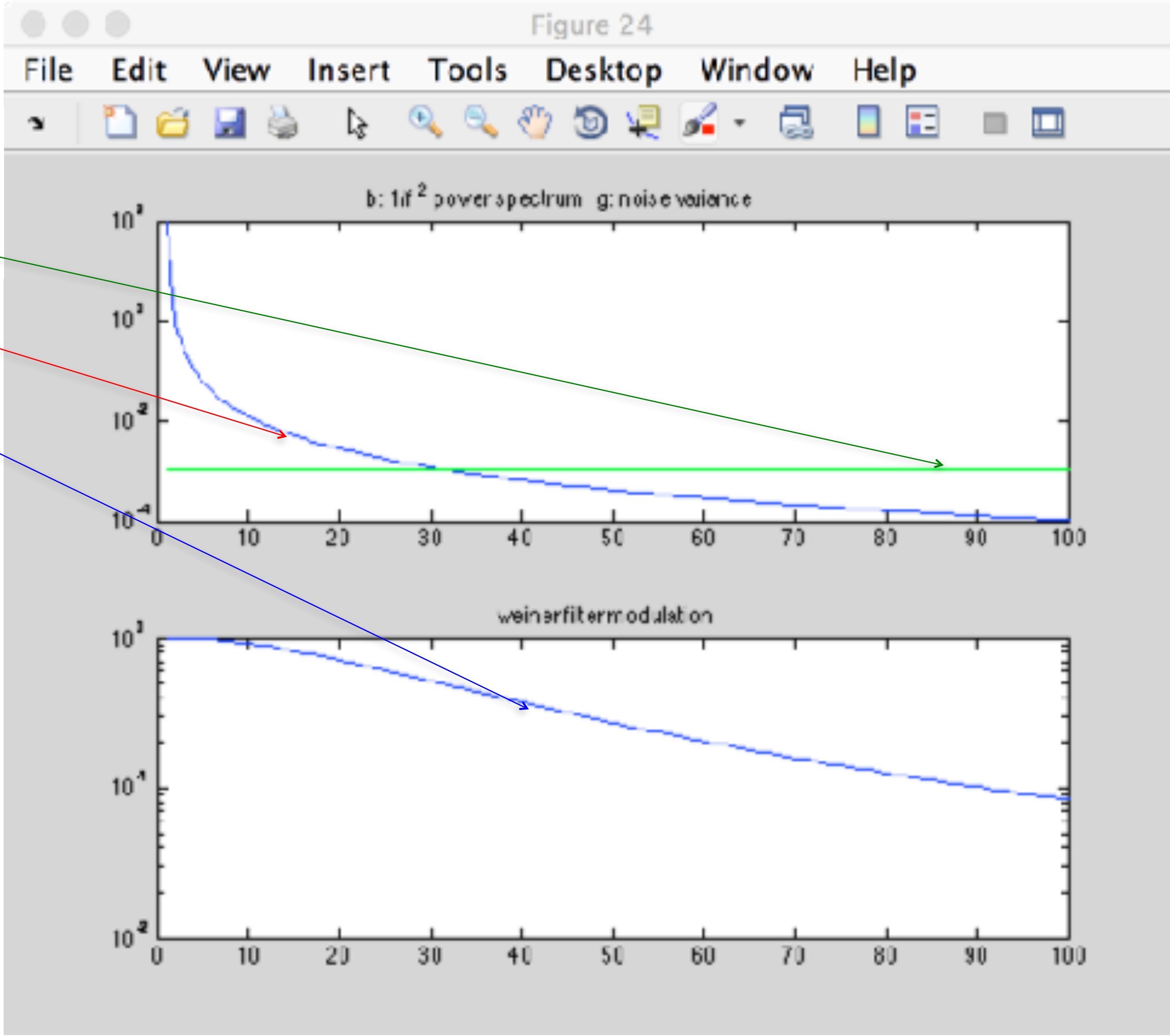
$$\mathbf{I} = \mathbf{C}\left(\mathbf{C} + \sigma_n^2\mathbb{I}\right)^{-1}\mathbf{I}_n$$ (note this is a linear operation

This can also be written in the Fourier domain, with C = EDE^T:

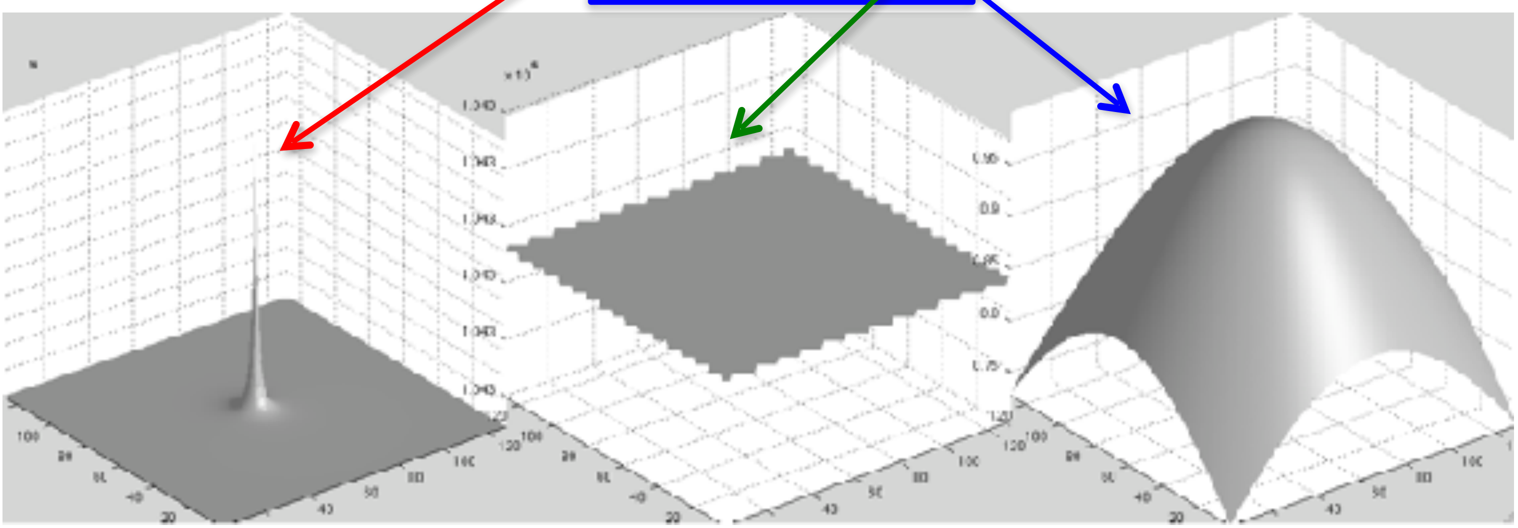
$$\widetilde{\mathbf{I}}(v) = \frac{A/|v|^{2\alpha}}{A/|v|^{2\alpha} + \sigma_n^2}\widetilde{\mathbf{I}}_n(v)$$

Wiener filter
(optimal filter for Gaussian image, additive Gaussian noise)

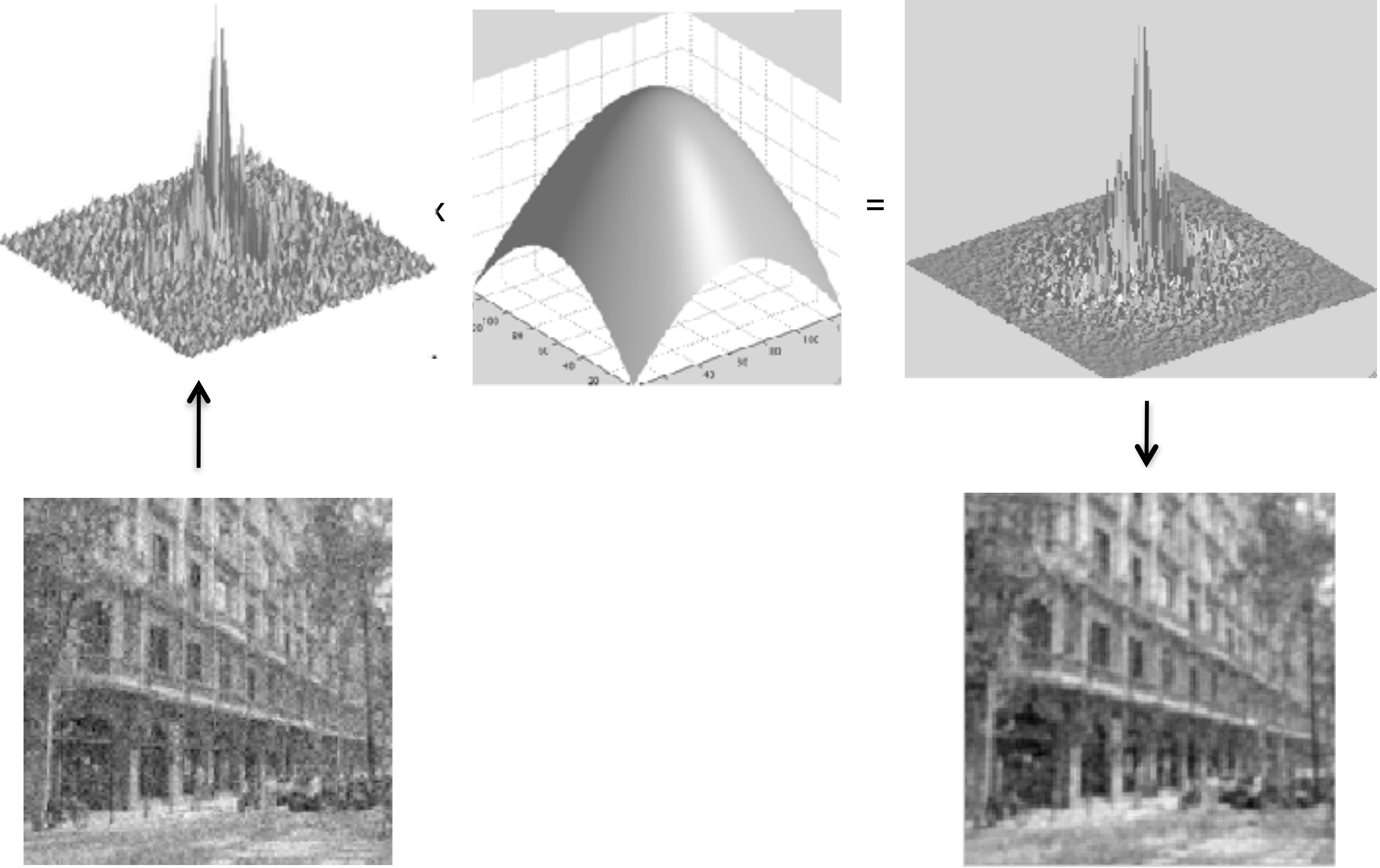
$$\widetilde{\mathbf{I}}(v) = \frac{A/|v|^{2\alpha}}{A/|v|^{2\alpha} + \sigma_n^2} \widetilde{\mathbf{I}}_n(v)$$



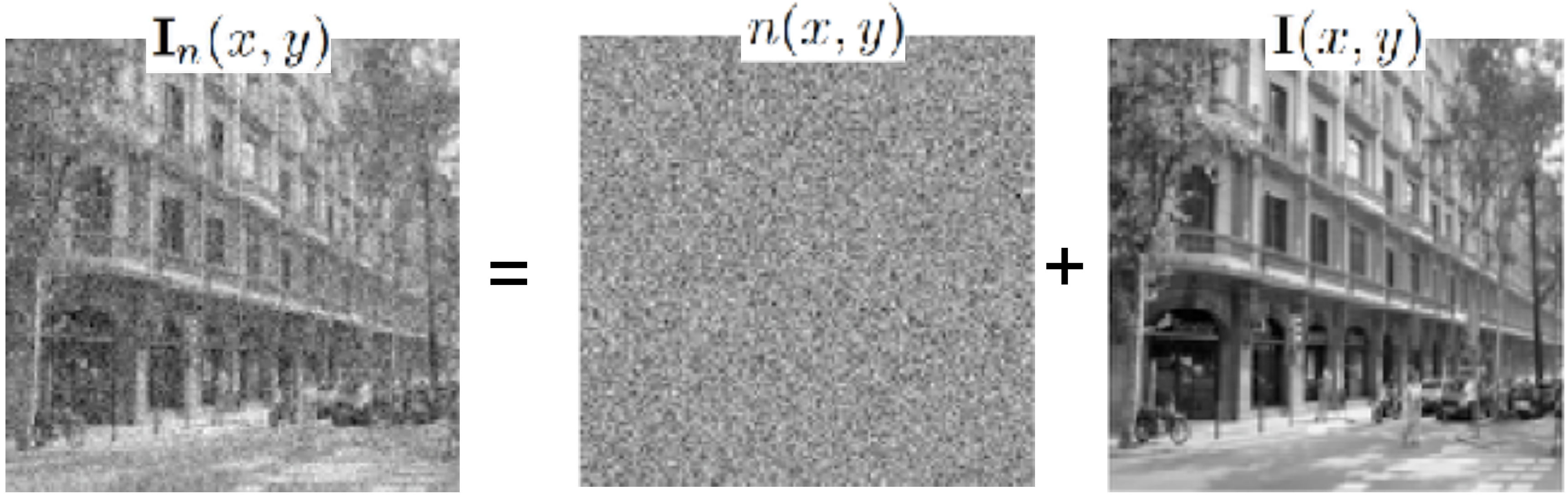
$$\widetilde{\mathbf{I}}(v) = \frac{A/|v|^{2\alpha}}{A/|v|^{2\alpha} + \sigma_n^2} \widetilde{\mathbf{I}}_n(v)$$



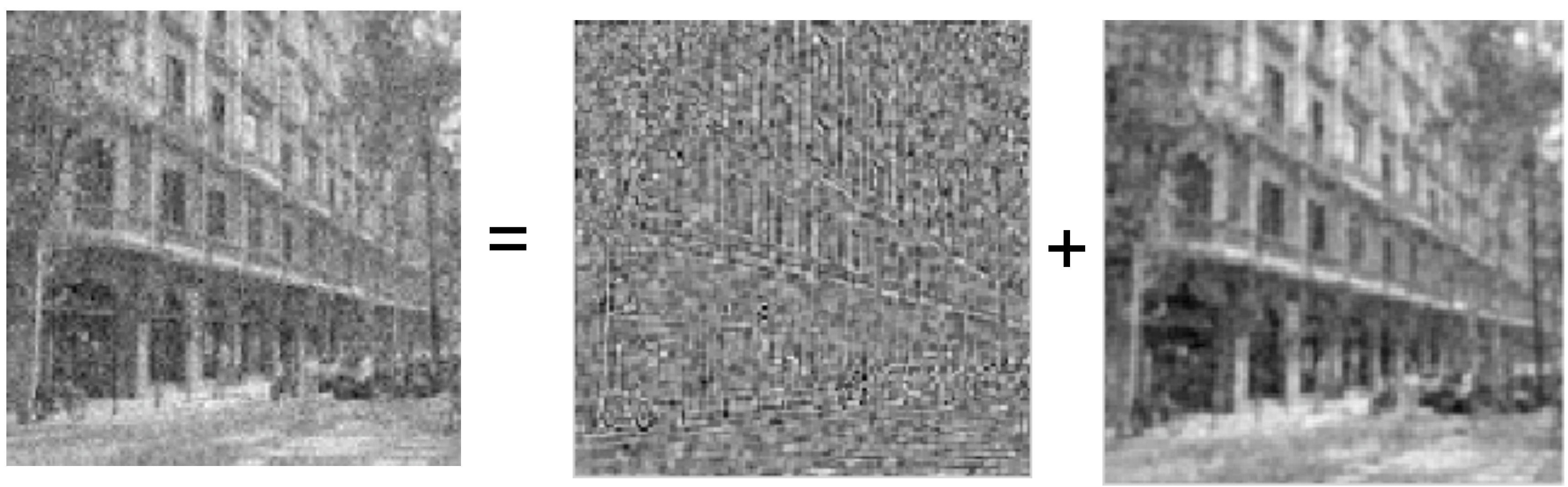
$$\frac{A/|v|^{2\alpha}}{A/|v|^{2\alpha} + \sigma_n^2}$$



The truth:

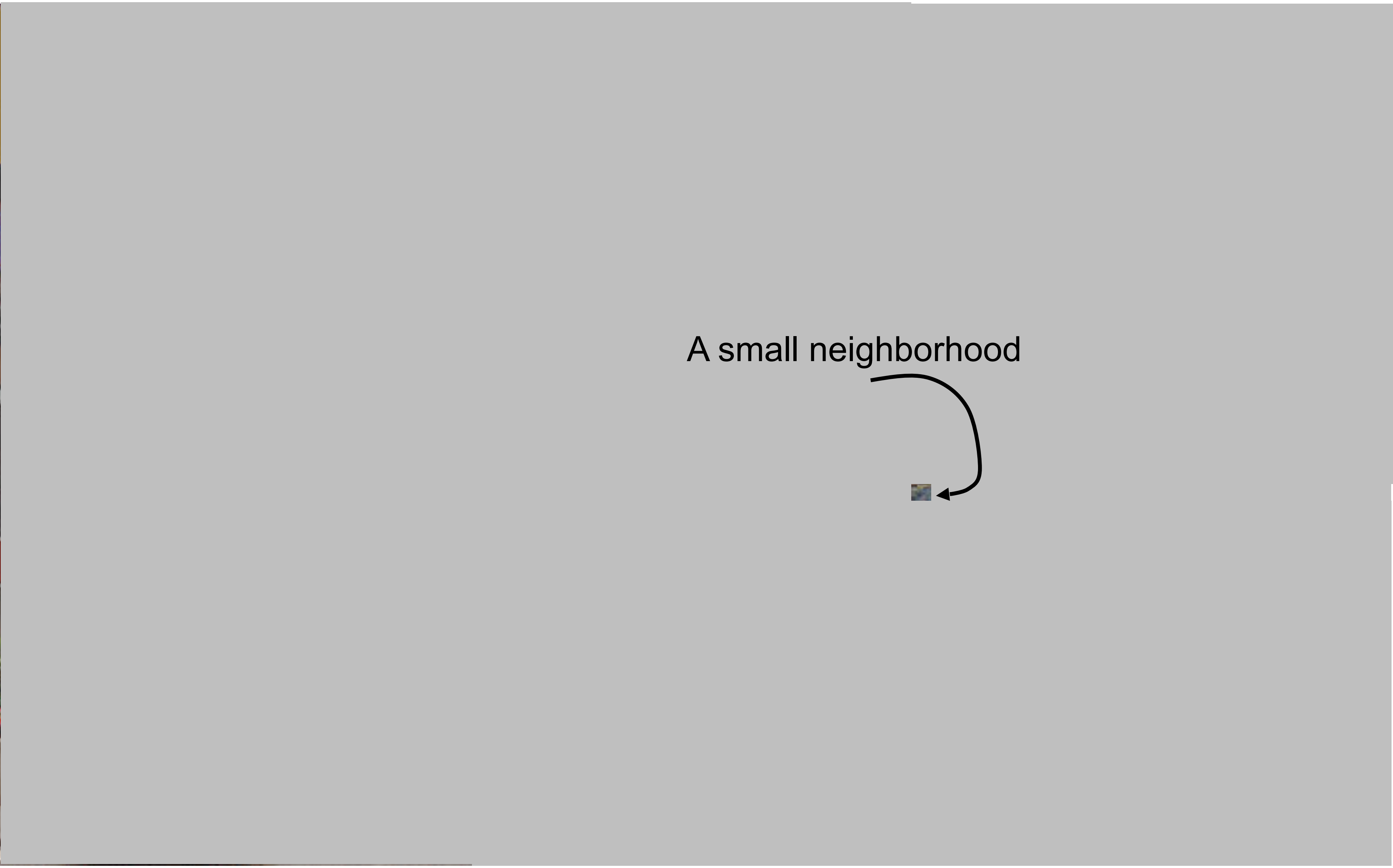


The estimated decomposition:



And we got all this from just modeling the correlation between pairs of pixels!

Statistical modeling of images



Edge responses, and responses to other bandpass filters

[-1 1]

g[m,n]

$\otimes$

[-1, 1]

h[m,n]

=

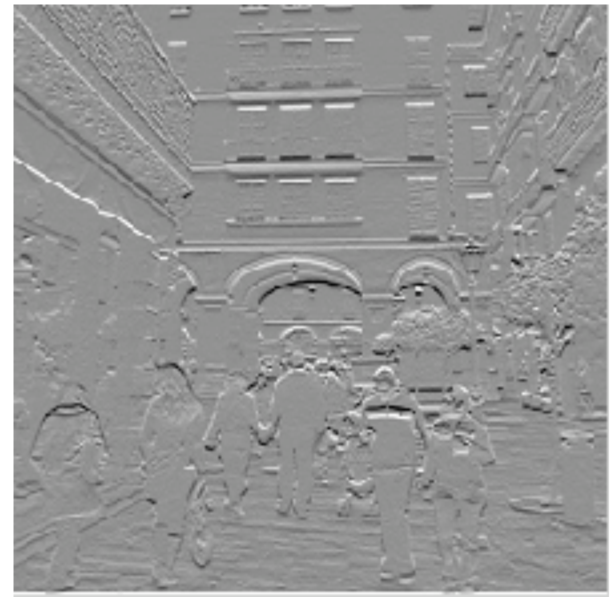
f[m,n]

$[-1\ 1]^T$

$\otimes$

$[-1, 1]^T$ =

h[m,n]



g[m,n]

f[m,n]

Observation: Sparse filter response

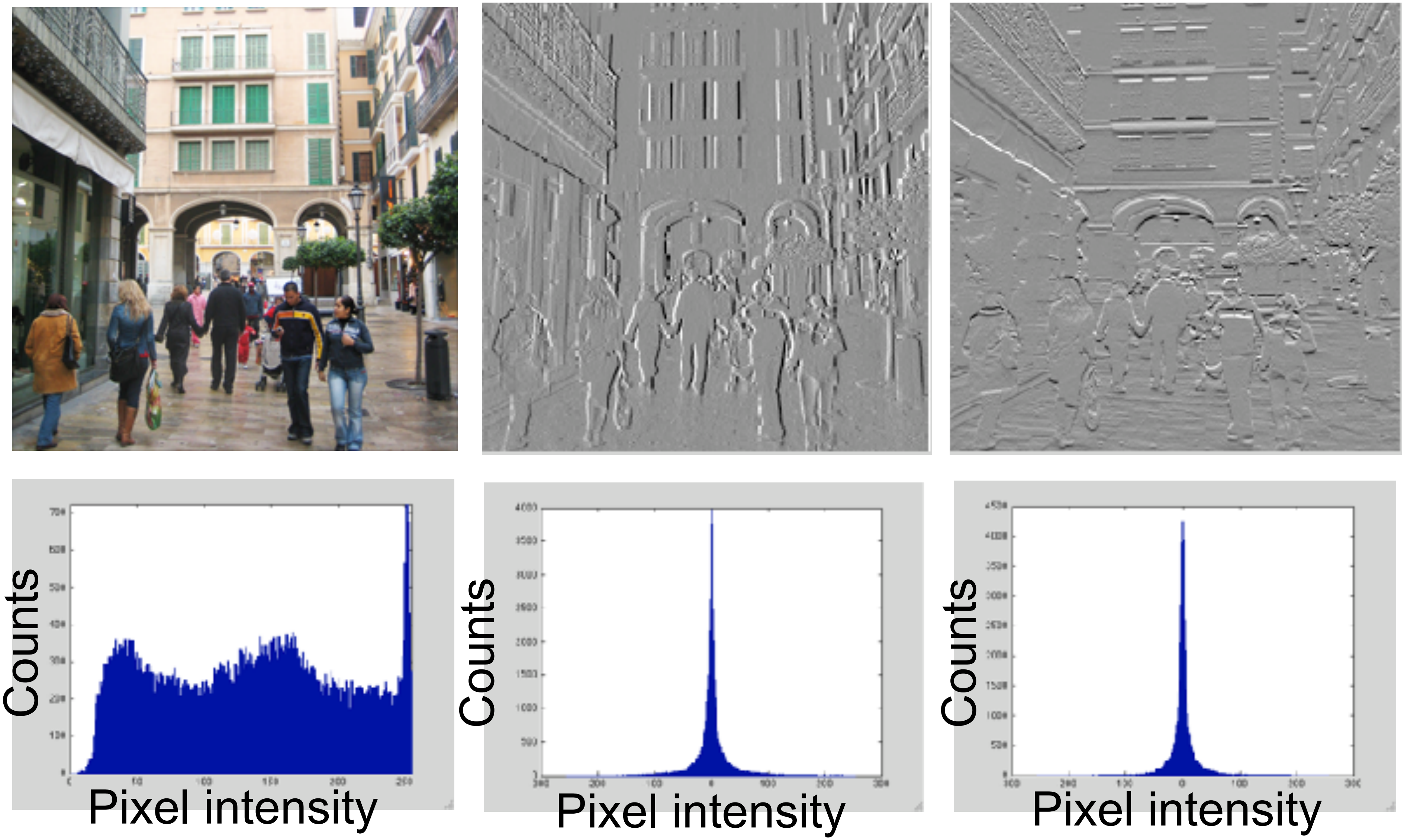
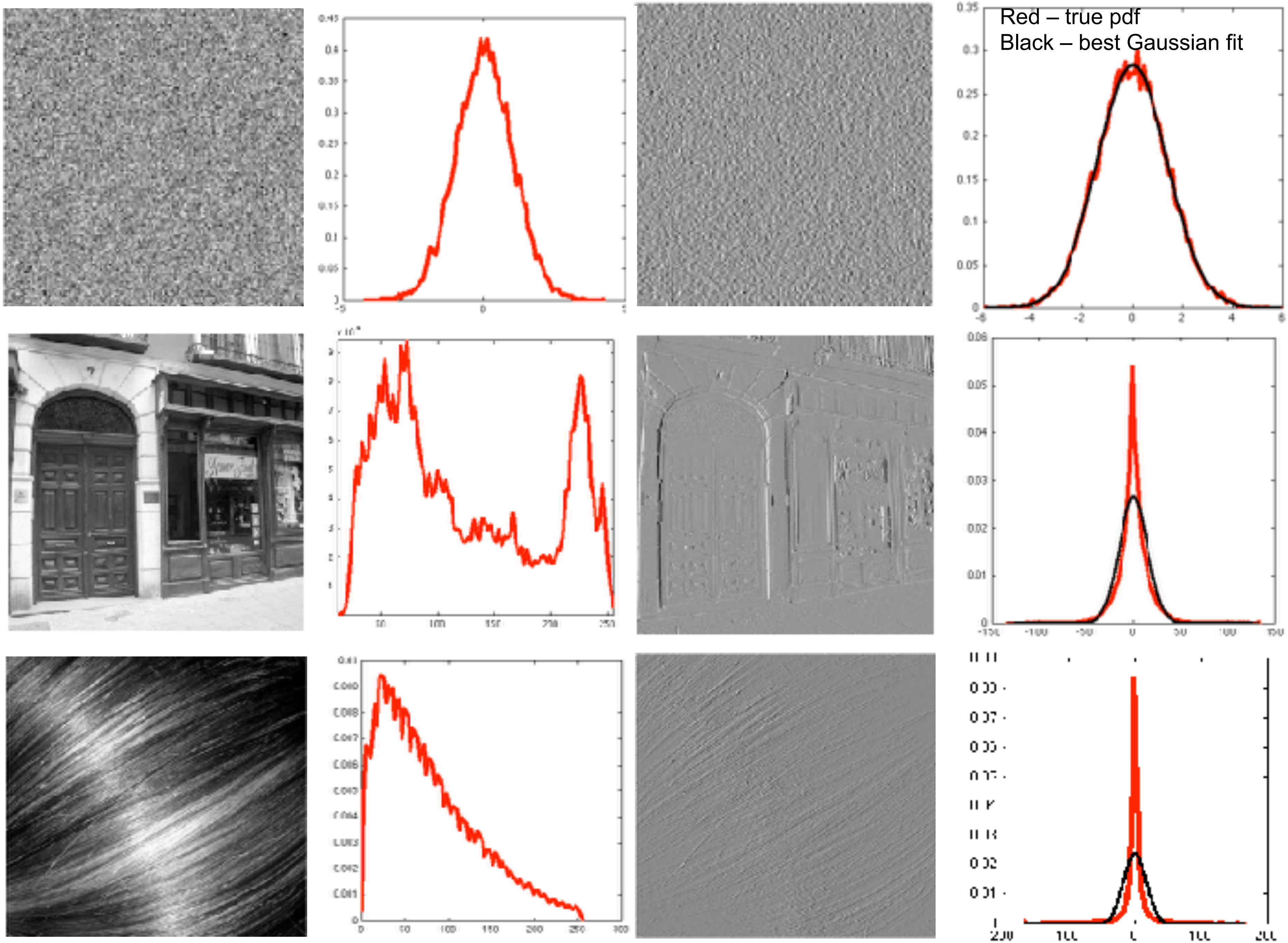
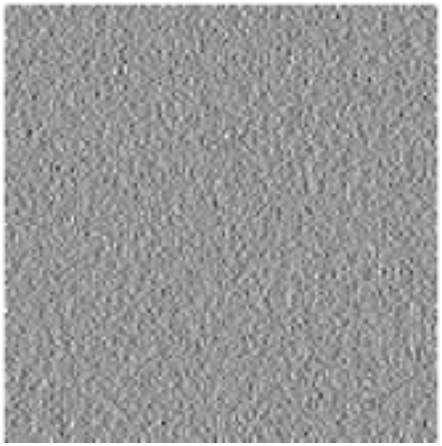
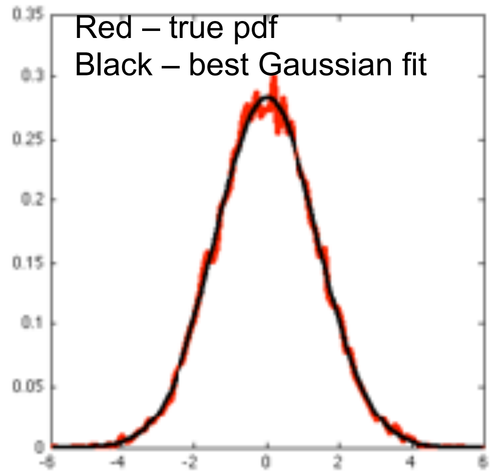


Image | Intensity histogram | [1 -1] filter output | [1 -1] output histogram

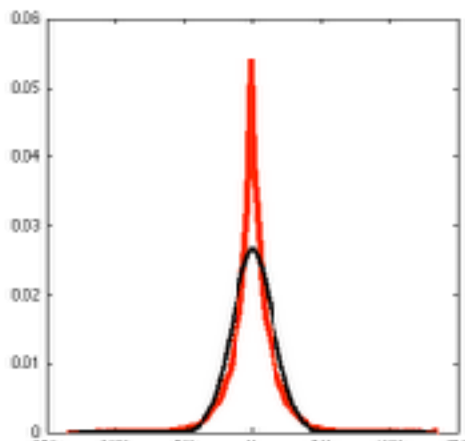


A model for the distribution of filter outputs





$$p(x) = \frac{\exp(-x^2/2\sigma^2)}{\sqrt{2\pi\sigma^2}}$$



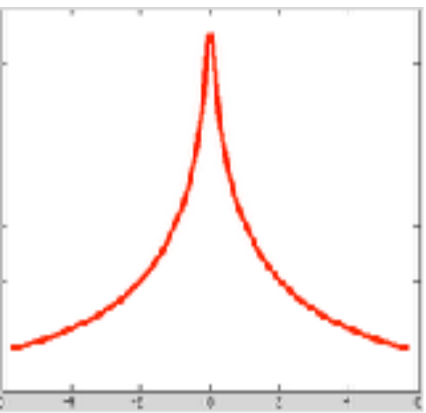
$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$

$r \sim 0.8 \ \ (< 2)$

Generalized Gaussian

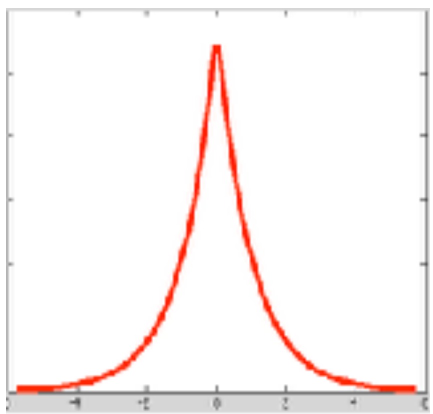
$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$

r = 0.5



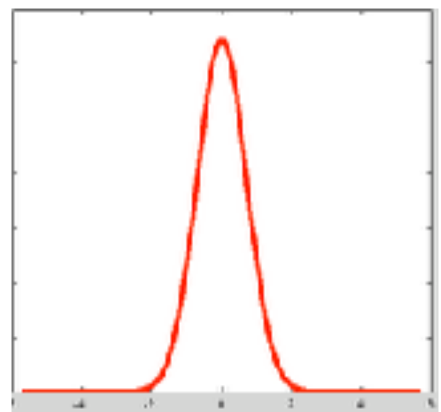
r = 1

Laplacian distribution

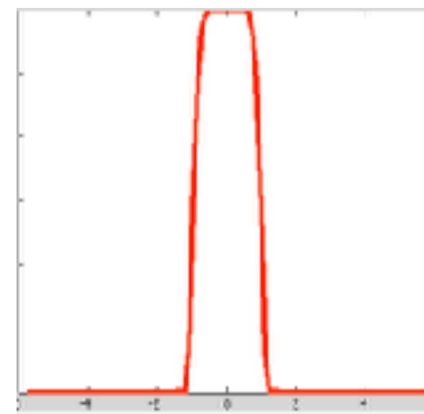


r = 2

Gaussian distribution



r = 10



Uniform distribution
r -> infinite

The wavelet marginal model

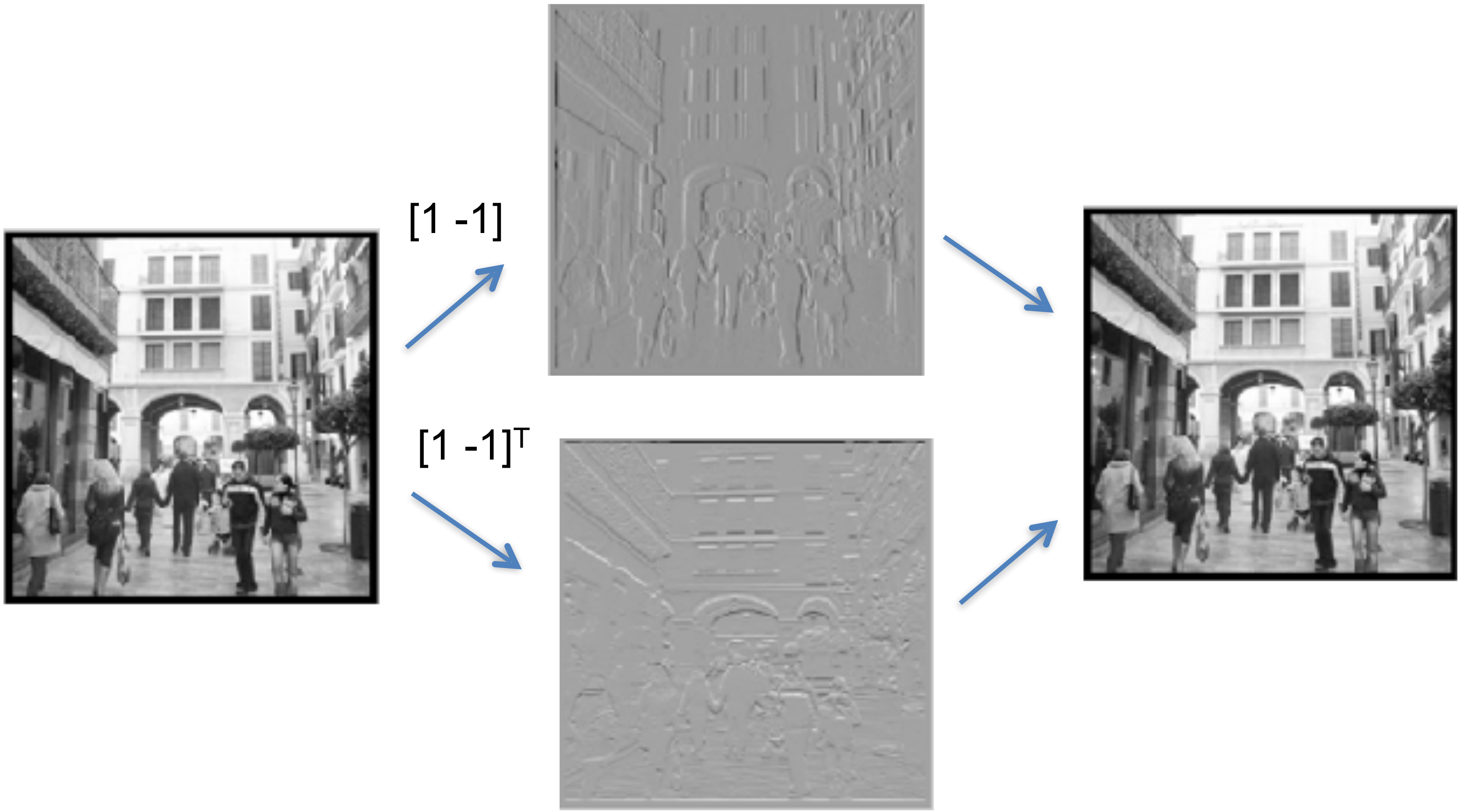
A small neighborhood

$$p(\mathbf{I}) = \prod_{k} \prod_{x,y} p(h_k(x,y))$$

All pixels and all outputs are independent

Filter outputs

The wavelet marginal model



$$p(\mathbf{I}) = \prod_{k} \prod_{x,y} p(h_k(x,y))$$

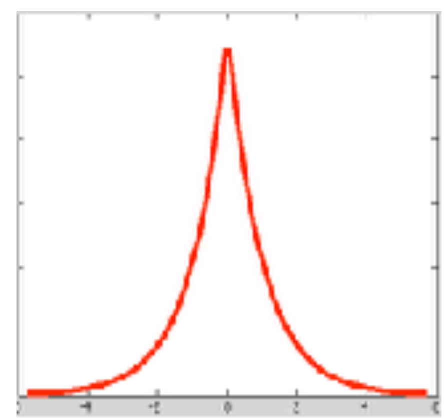
What is the most probable image under the wavelet marginal model?

[1 -1]

[1 -1]^T

$$p(\mathbf{I}) = \prod_k \prod_{x,y} p(h_k(x,y))$$

$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$



Sampling typical images

Gaussian model

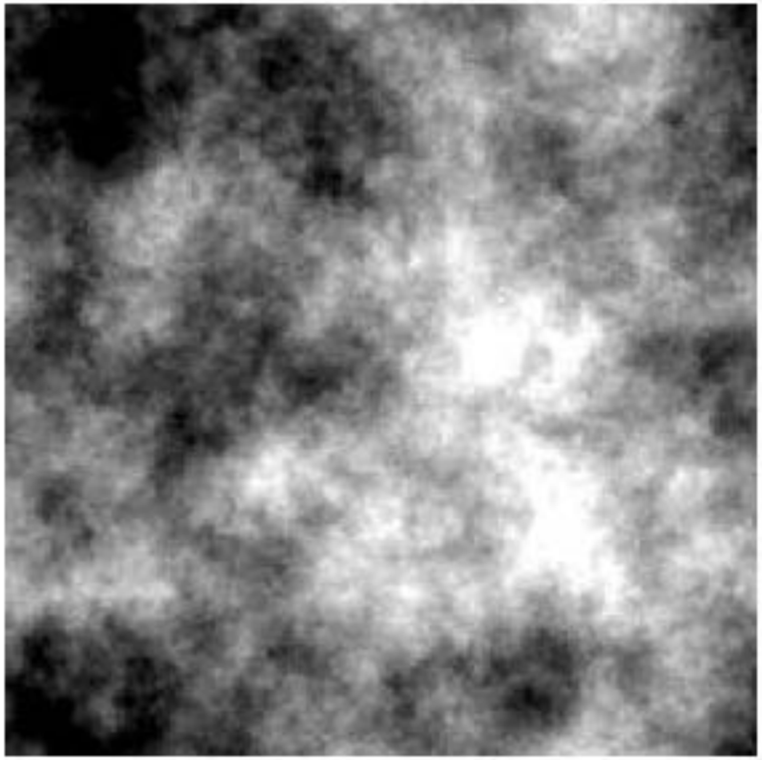


Fig. 3. Example image randomly drawn from the Gaussian spectral model, with $\gamma = 2.0$.

Wavelet marginal model

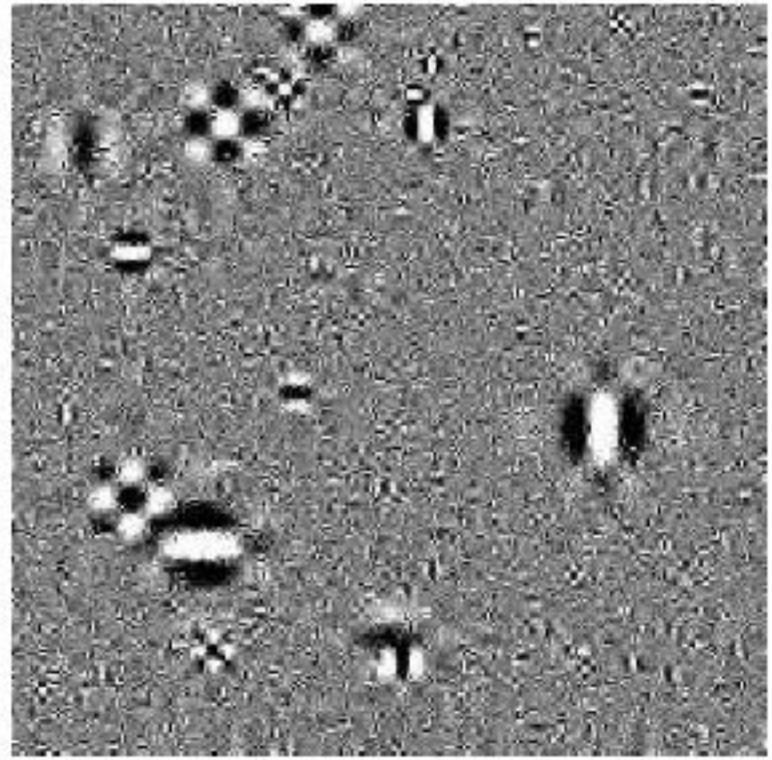
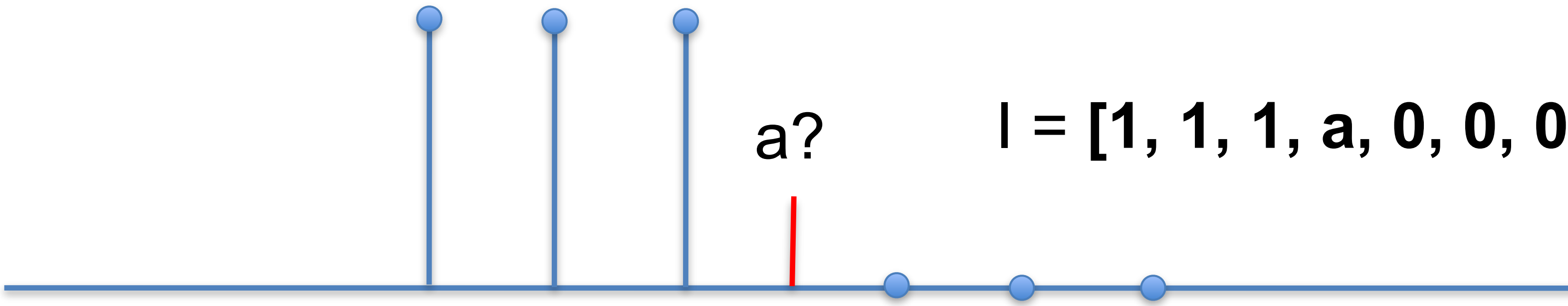


Fig. 6. A sample image drawn from the wavelet marginal model, with subband density parameters chosen to fit the image of Fig. 7.

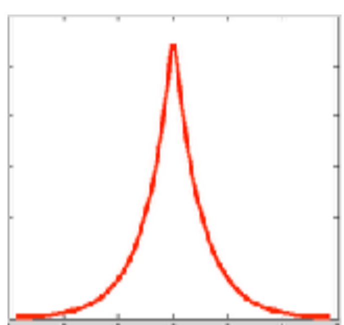
1D example



I = [1, 1, 1, a, 0, 0, 0]

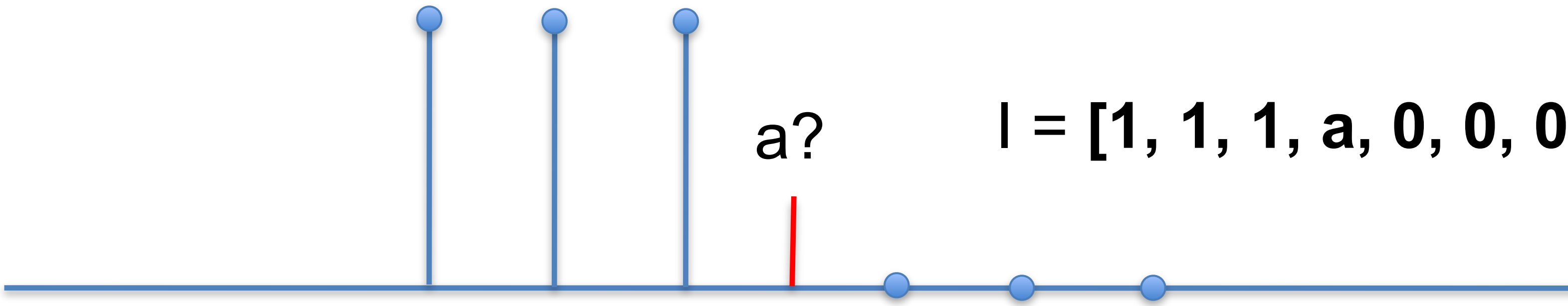
The prior model:

The output to the filter [-1,1] produces independent values each following the distribution:



$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$

1D example



I = [1, 1, 1, a, 0, 0, 0]

1. First, let's compute the [-1,1] filter output:

$$h = I * [-1,1] = \mathbf{[0, 0, 0, 1\text{-}a, a, 0, 0]}$$

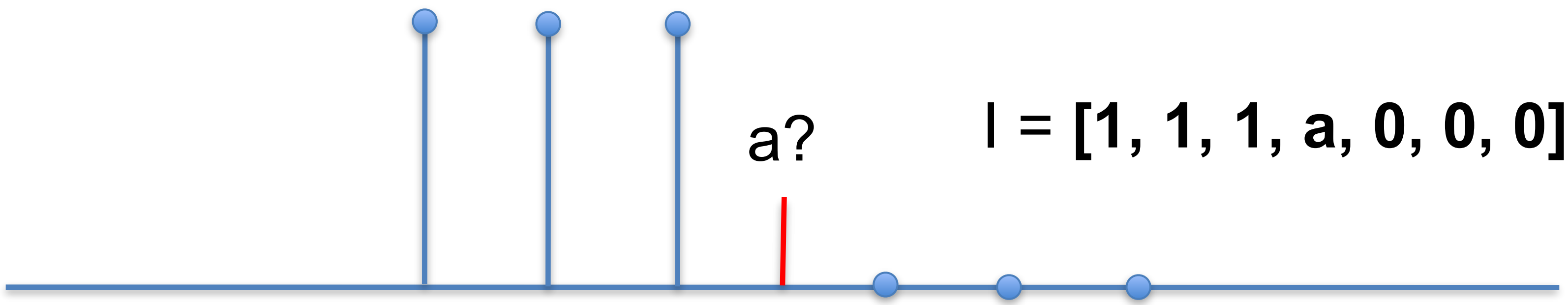
2. Let's write the probability of the output as a function of **a**

$$p(\mathbf{I}) = \prod_x p(h(x)) = \frac{\exp\left(-|(1-a)/s|^r\right)\exp\left(-|a/s|^r\right)}{(2s/r\Gamma(1/r))^7} =$$

$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$

$$= \frac{1}{(2s/r\Gamma(1/r))^7} \exp -\frac{|1-a|^r + |a|^r}{s^r}$$

What is the preferred value of **a** as we change **r**?



I = [1, 1, 1, a, 0, 0, 0]

$$p(\mathbf{I}) = \frac{1}{(2s/r\Gamma(1/r))^7}\exp -\frac{|1-a|^r+|a|^r}{s^r}$$

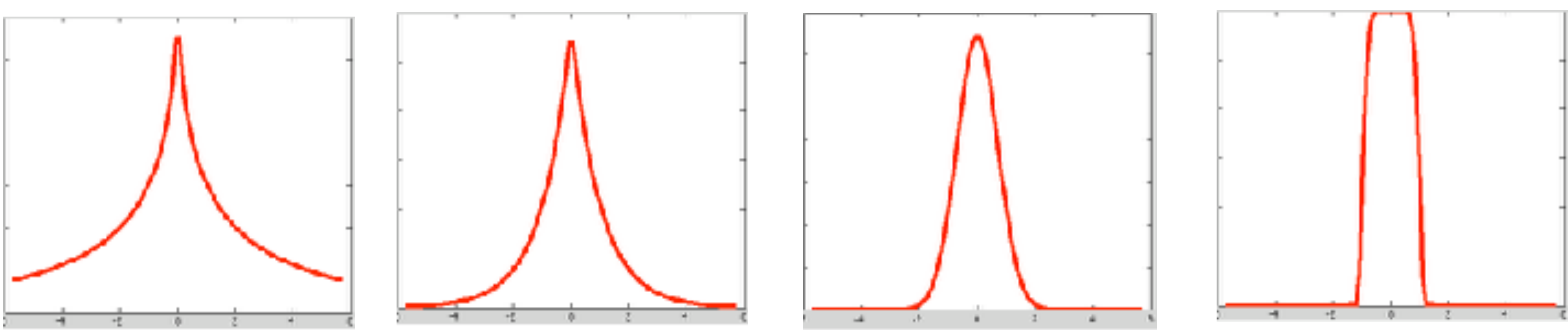
$$p(x) = \frac{\exp(-|x/s|^r)}{2s/r\Gamma(1/r)}$$

r = 0.5

r = 1 Laplacian distribution

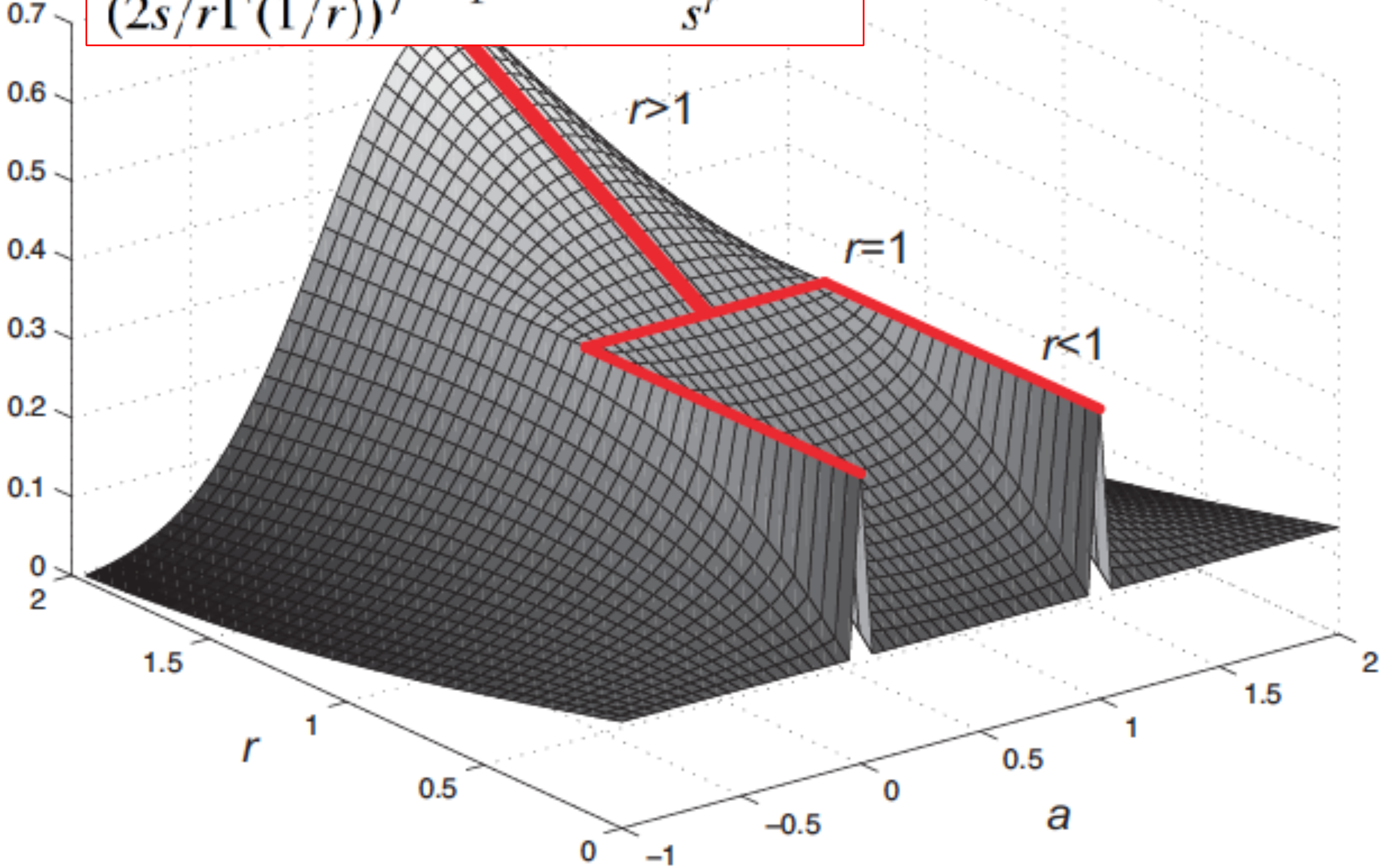
r = 2 Gaussian distribution

r = 10

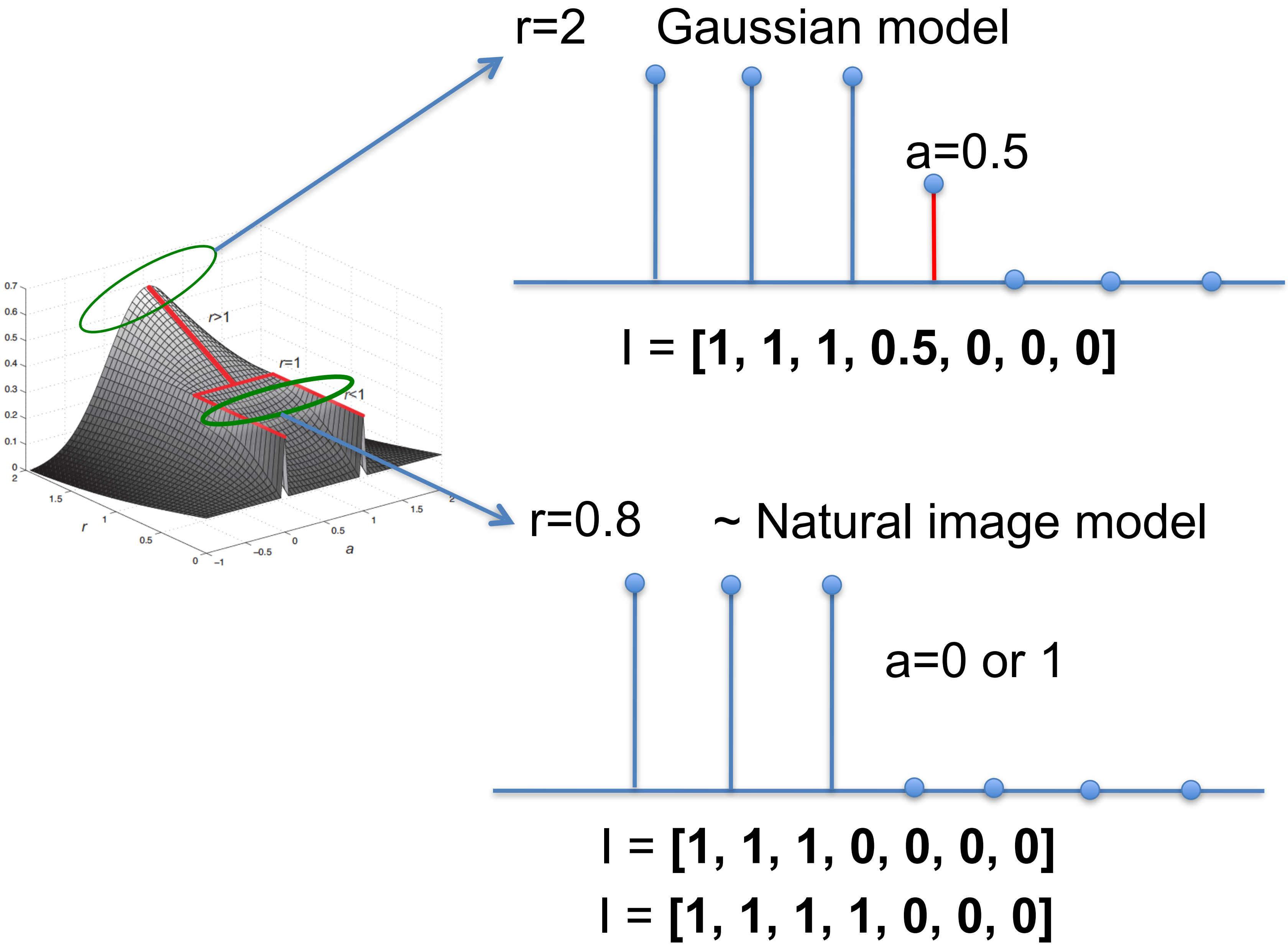


1D example

$$\frac{1}{(2s/r\Gamma(1/r))^7}\exp -\frac{|1-a|^r+|a|^r}{s^r}$$



For different values of **r** we get different maximum

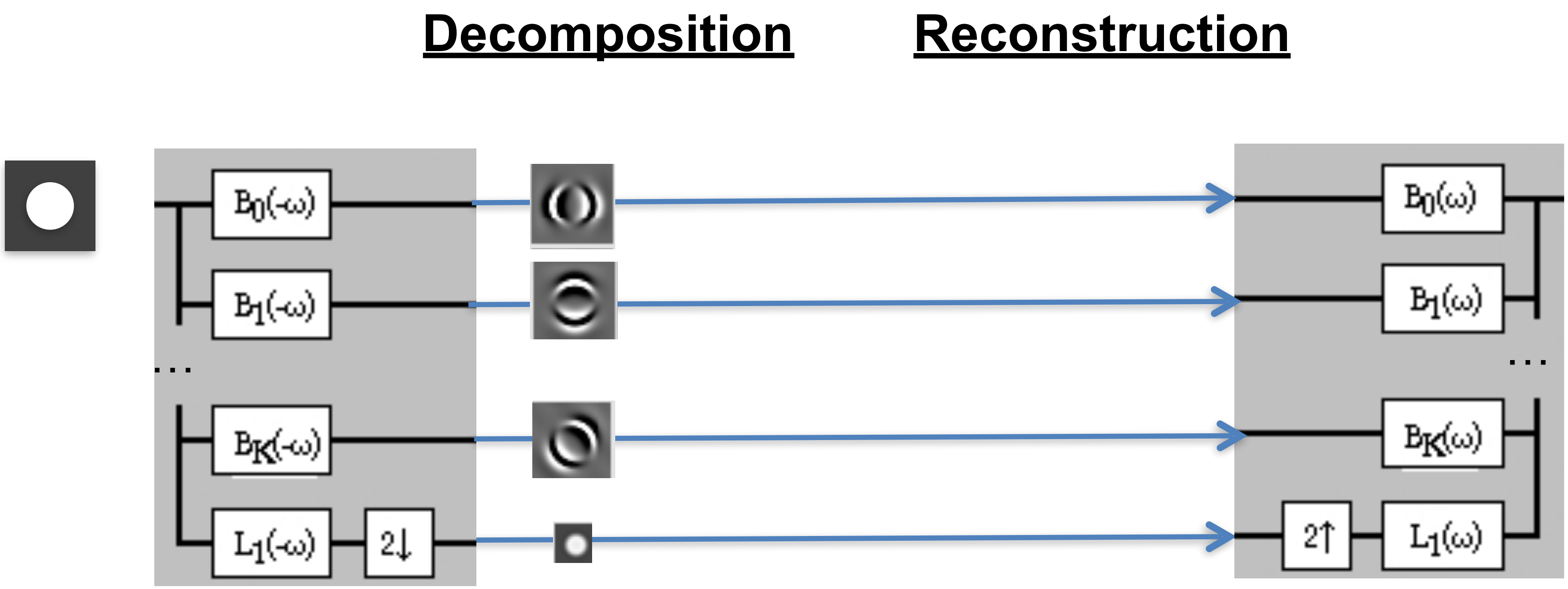


r=2
Gaussian model
a=0.5
I = [1, 1, 1, 0.5, 0, 0, 0]
r>1
r=1
r<1
r
a
r=0.8
~ Natural image model
a=0 or 1
I = [1, 1, 1, 0, 0, 0, 0]
I = [1, 1, 1, 1, 0, 0, 0]

The sparse responses for image subbands is a useful image prior. But what should the subbands be?

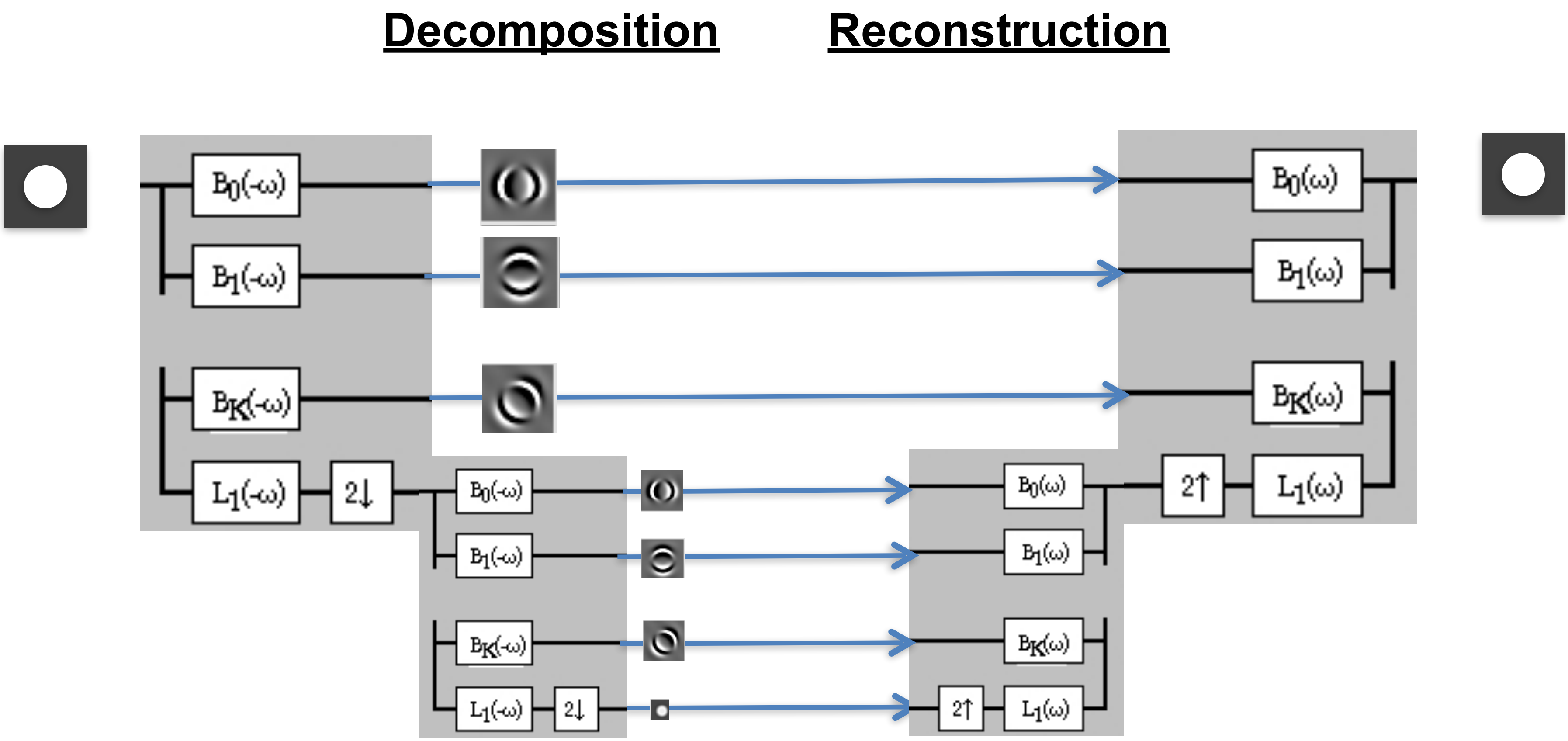
Following are 3 different subband image representations. Each derived differently, each arriving at approximately the same representation

Steerable Pyramid

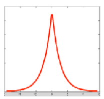


Images from: http://www.cis.upenn.edu/~eero/steerpyr.html

Steerable Pyramid



$$p(\mathbf{I}) = \prod_{k} \prod_{x,y} p(h_k(x,y))$$



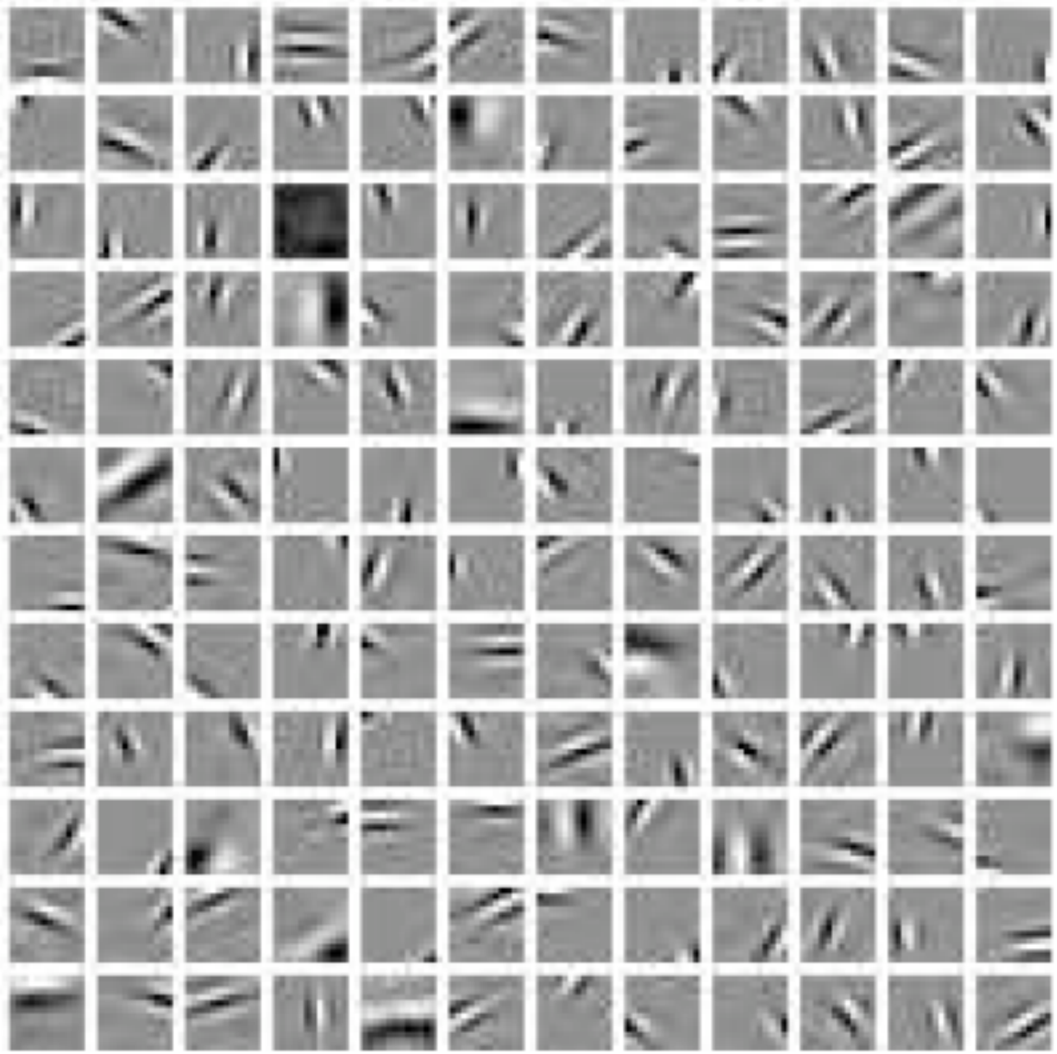
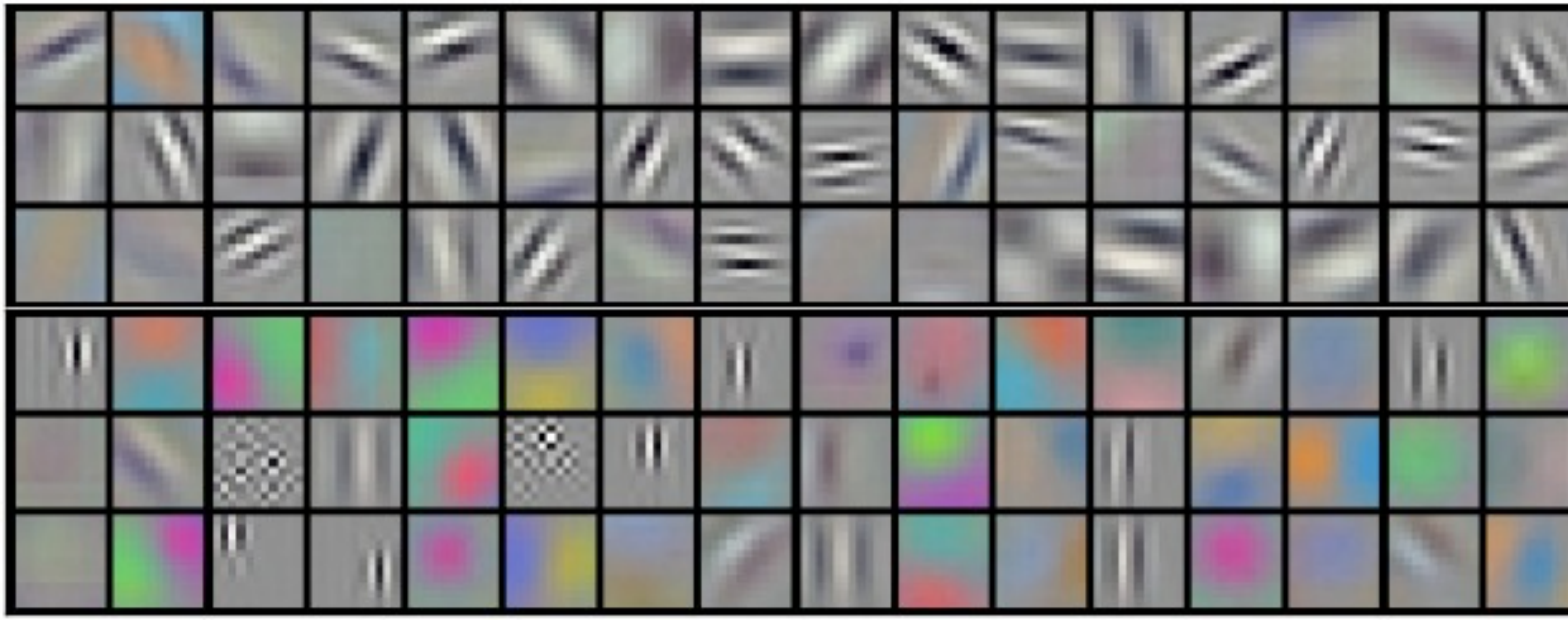


Fig. 5. Example basis functions derived by optimizing a marginal kurtosis criterion [see 35].

Olshausen BA, and Field DJ. (1996). "Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images." Nature, 381: 607-609.

Learned with a convNet



1st layer

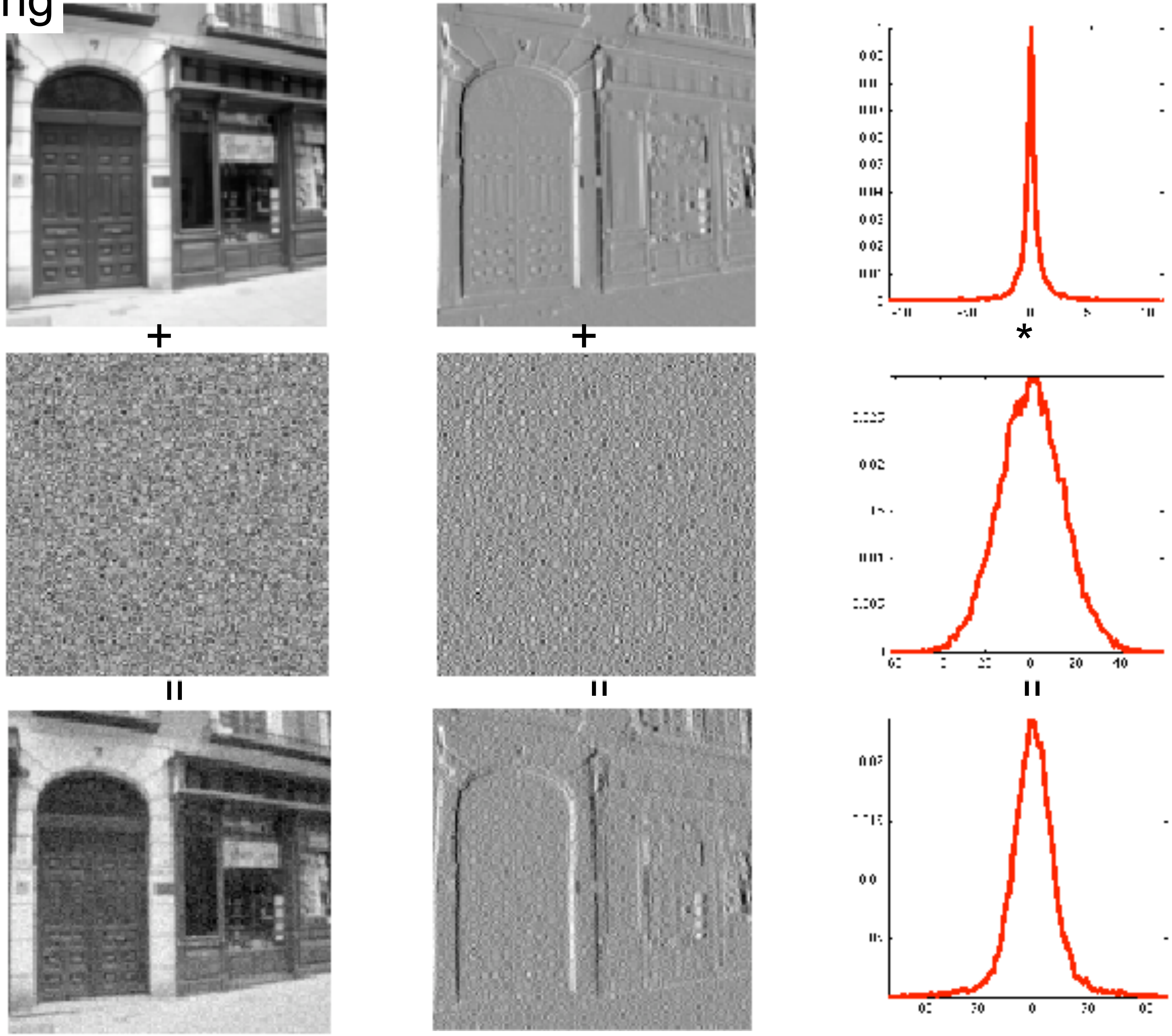
Krizhevsky, A., Sutskever, I. and Hinton, G. E.
ImageNet Classification with Deep Convolutional Neural Networks
NIPS 2012: Neural Information Processing Systems

http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks

Denoising

White Gaussian noise

Noisy image



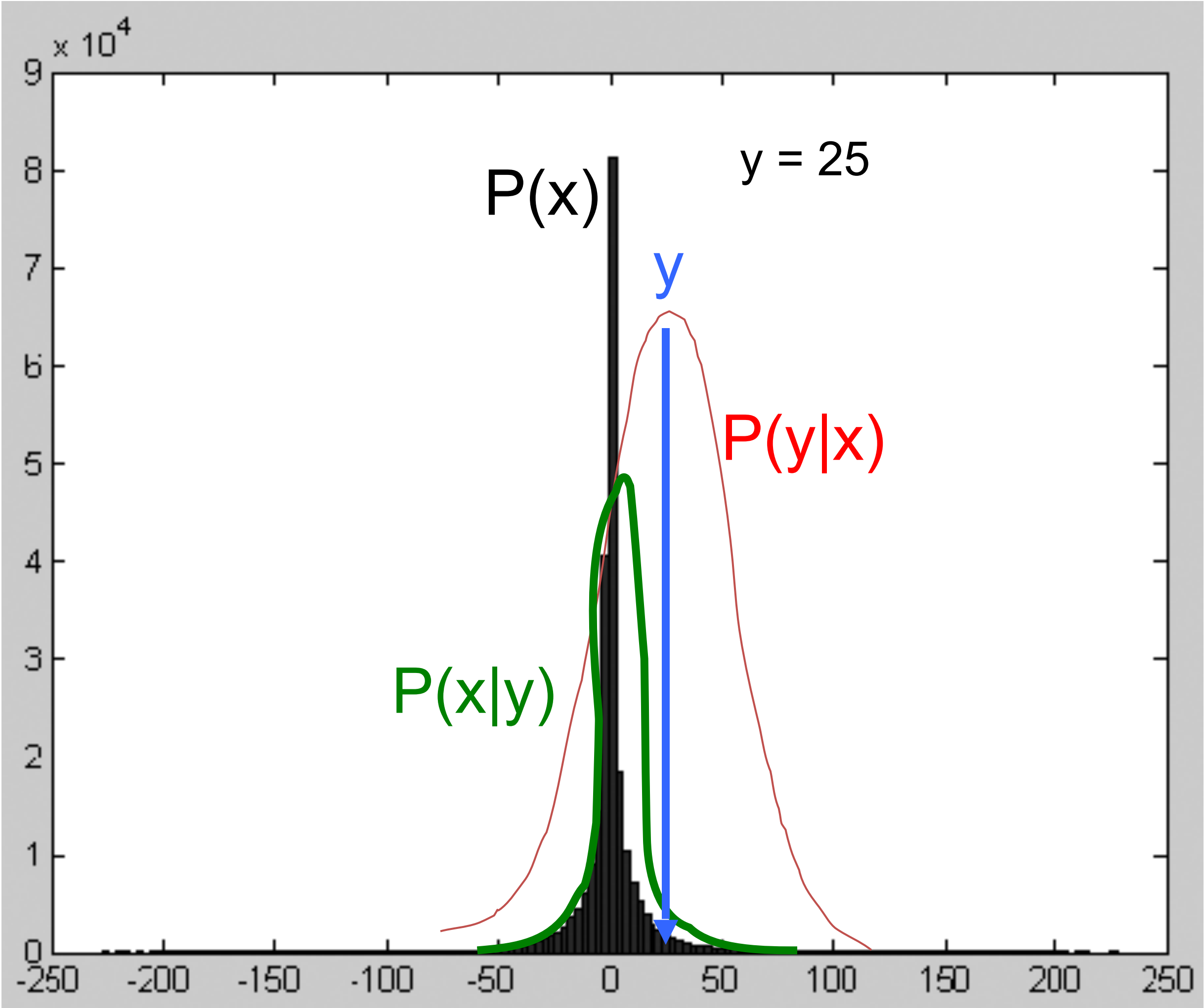
Denoising with the marginal wavelet model

Let y = noise-corrupted observation: y = x+n, with n ~ gaussian.

Let x = bandpassed image value before adding noise.

By Bayes theorem

P(x|y) ~ P(y|x) P(x)

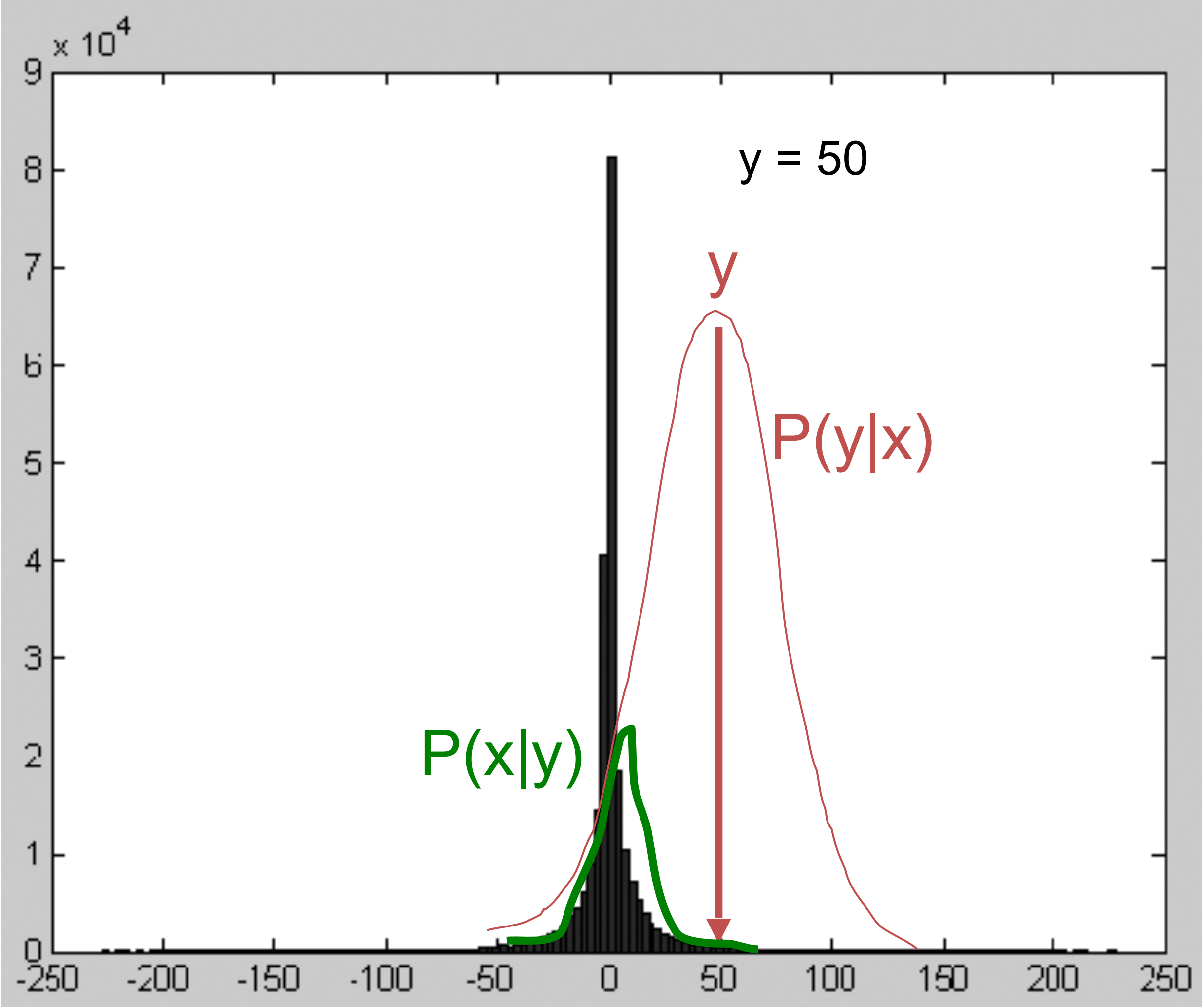


Denoising with the marginal wavelet model

Let x = bandpassed image value before adding noise.
Let y = noise-corrupted observation.

By Bayes theorem

P(x|y) ~ P(y|x) P(x)

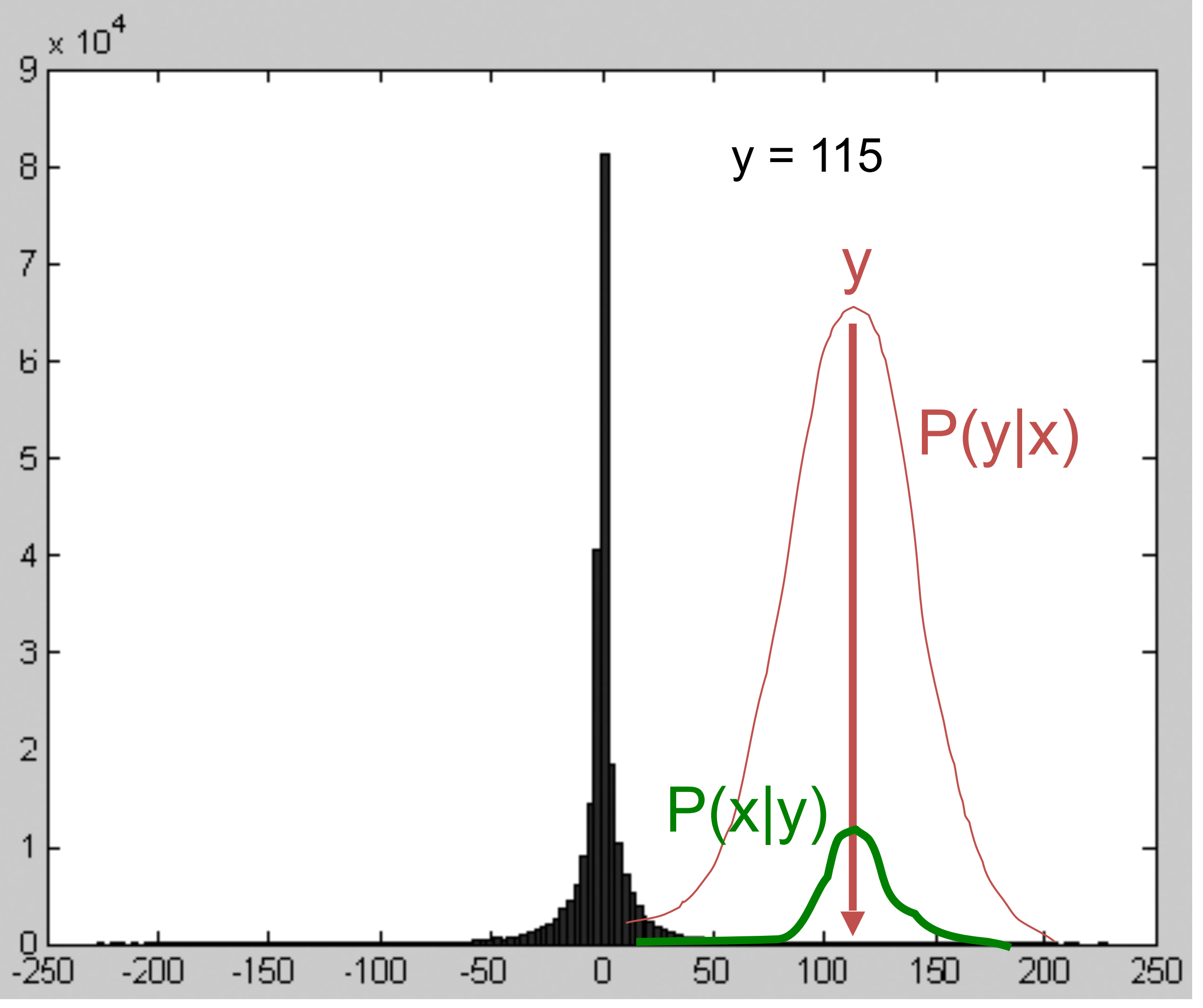


Denoising with the marginal wavelet model

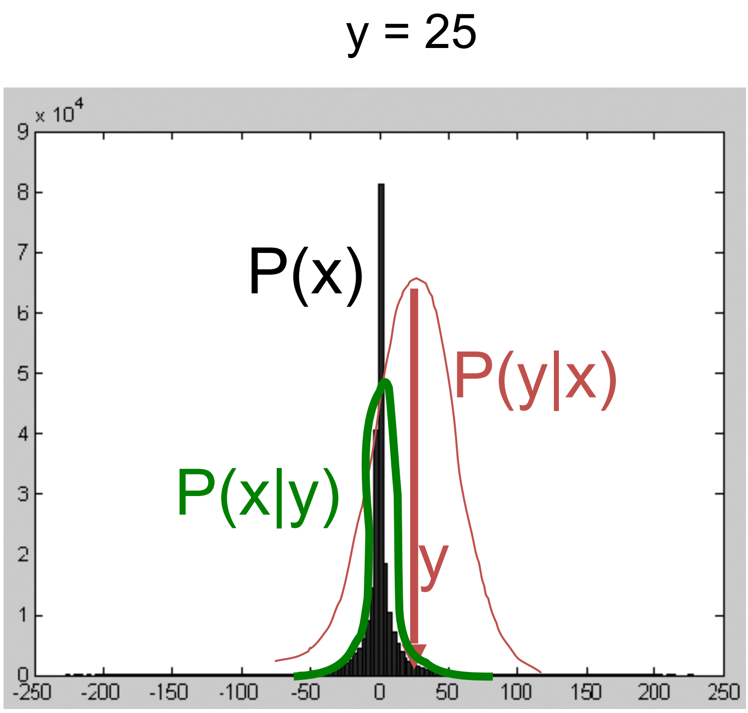
Let x = bandpassed image value before adding noise.
Let y = noise-corrupted observation.

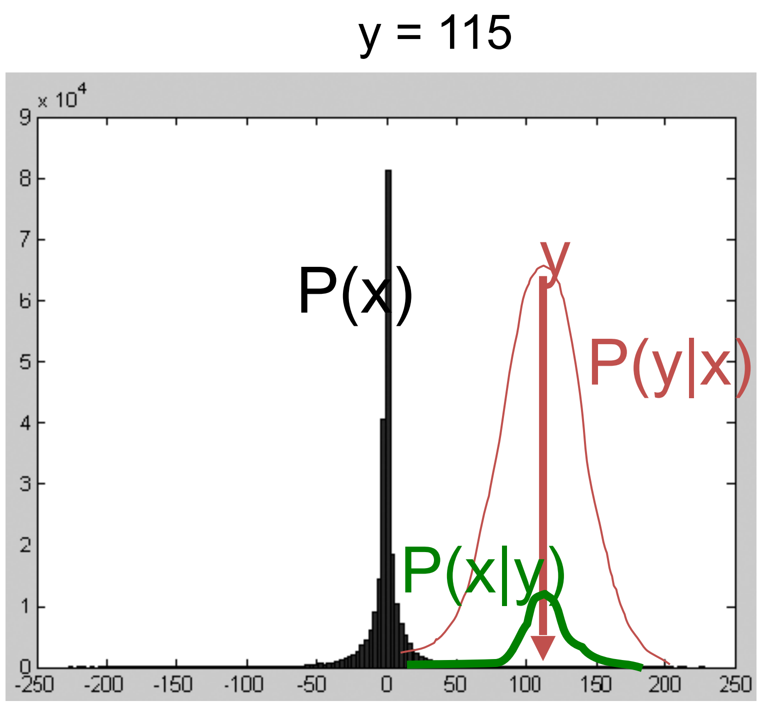
By Bayes theorem

P(x|y) ~ P(y|x) P(x)



Denoising with the marginal wavelet model





For small y: probably it is due to noise and y should be set to 0
For large y: probably it is due to an image edge and it should be kept untouched

MAP estimate, $\hat{x}$, as function of observed coefficient value, y

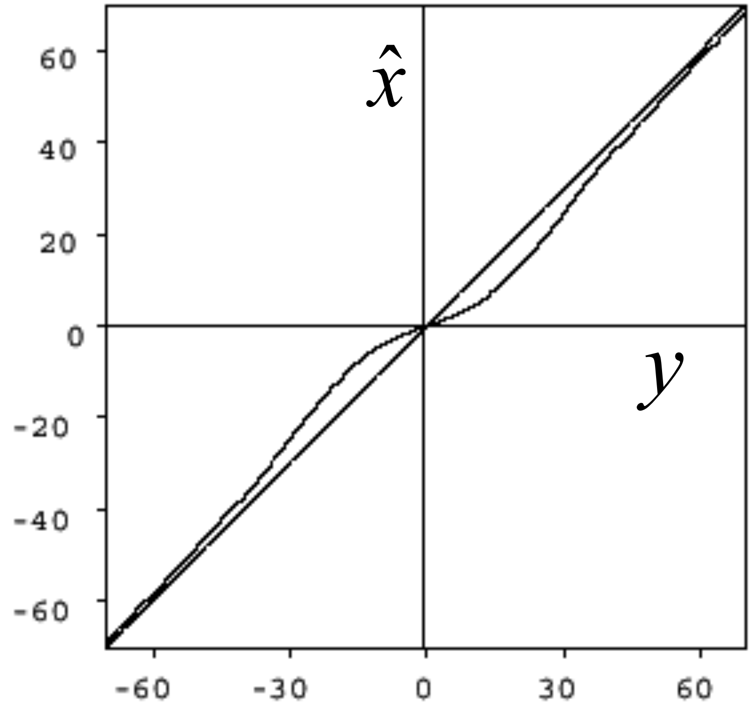
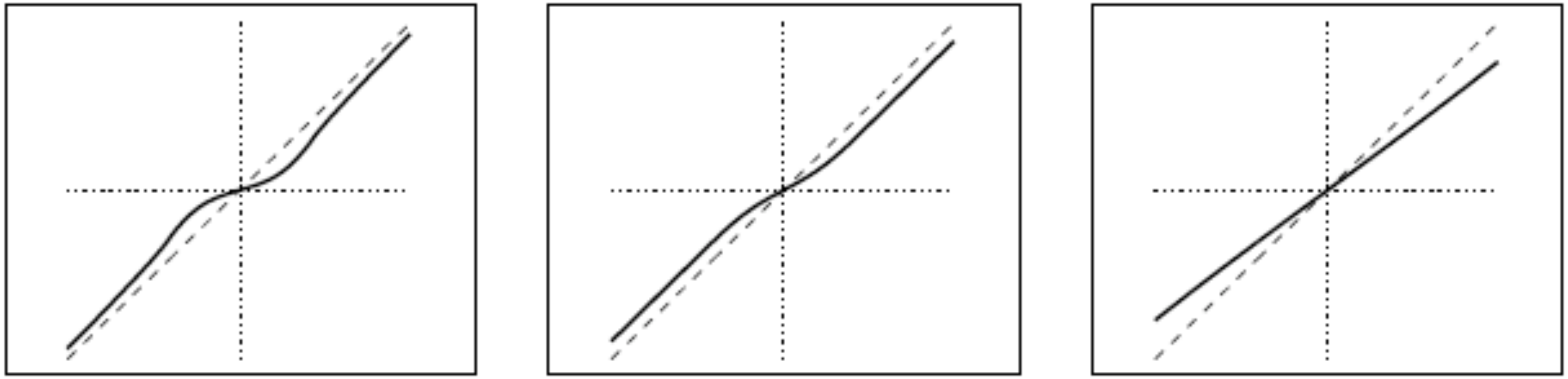


Figure 2: Bayesian estimator (symmetrized) for the signal and noise histograms shown in figure 1. Superimposed on the plot is a straight line indicating the identity function.

http://www-bcs.mit.edu/people/adelson/pub_pdfs/simoncelli_noise.pdf

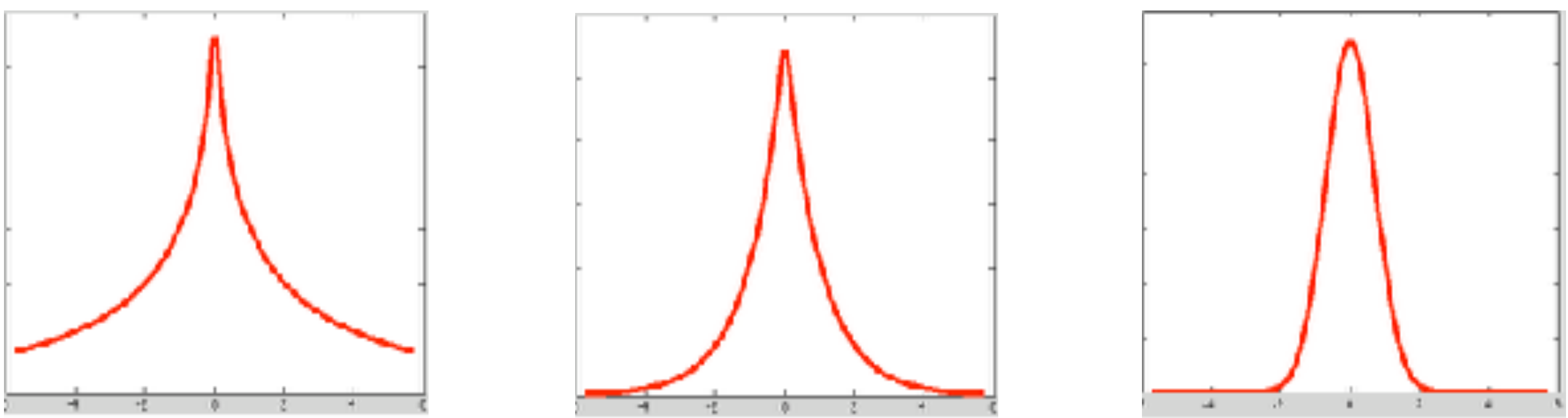
Simoncelli and Adelson, Noise Removal via Bayesian Wavelet Coring



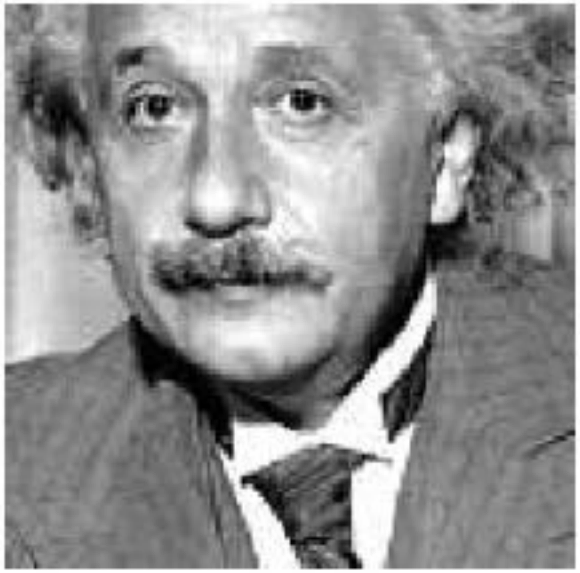
r = 0.5

r = 1
Laplacian distribution

r = 2
Gaussian distribution

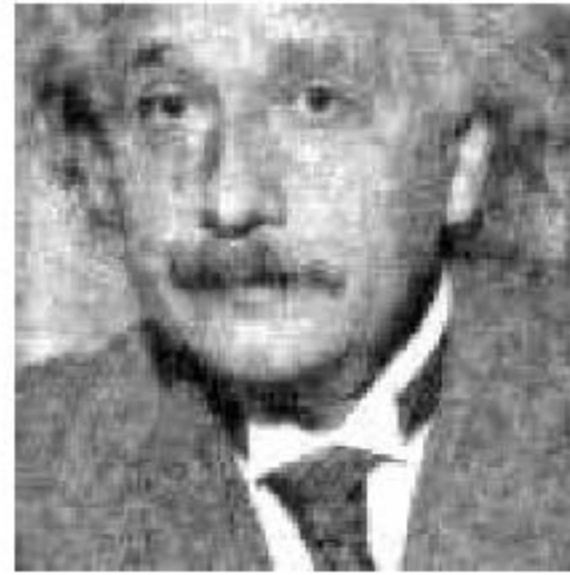


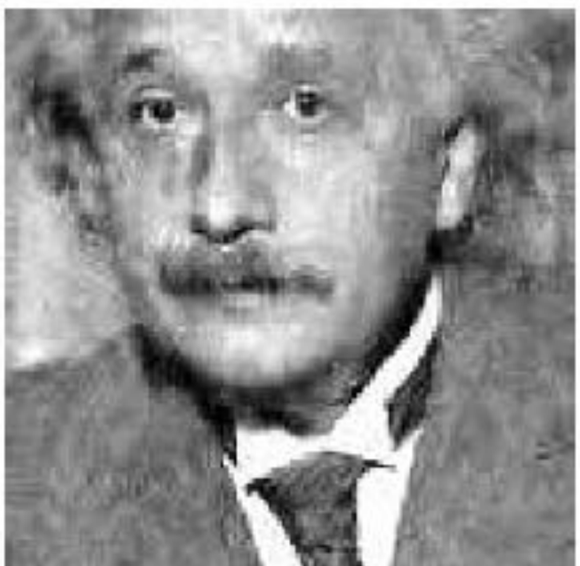
original



With Gaussian noise of std. dev. 21.4 added, giving PSNR=22.06

(1) Denoised with Gaussian model, PSNR=27.87





(2) Denoised with wavelet marginal model, PSNR=29.24

http://www.cns.nyu.edu/pub/eero/simoncelli05a-preprint.pdf

Gaussian scale mixtures

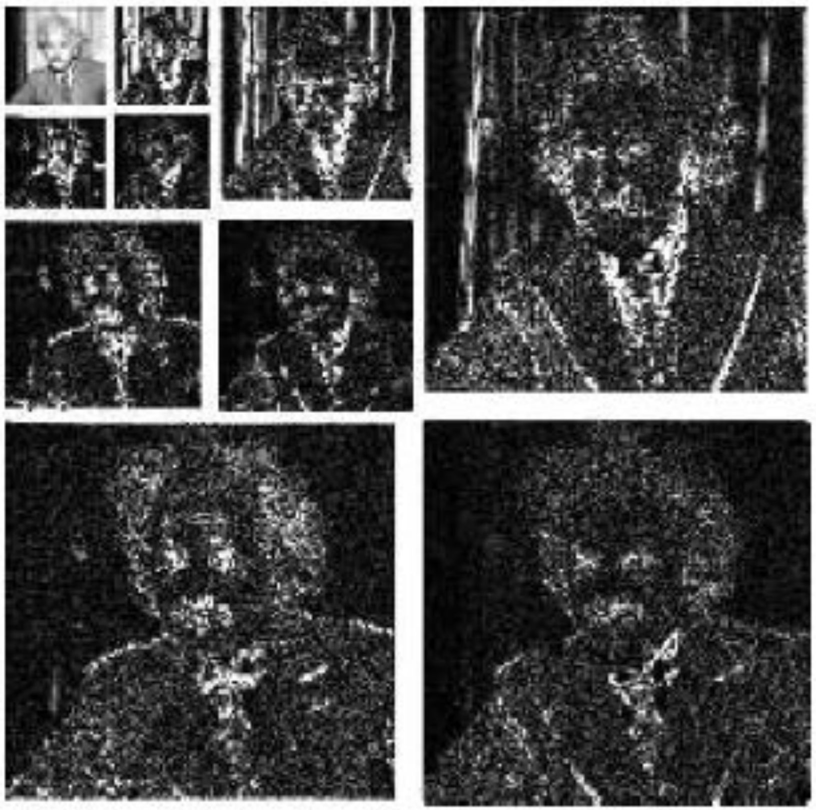
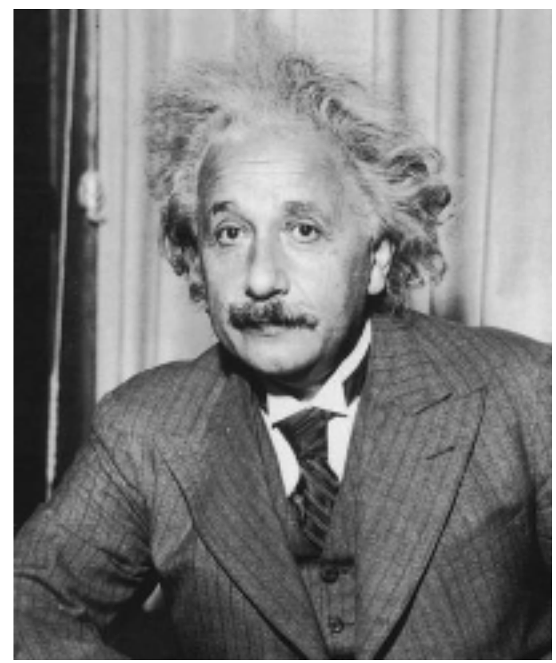


Fig. 7. Amplitudes of multi-scale wavelet coefficients for the "Einstein" image. Each subimage shows coefficient amplitudes of a subband obtained by convolution with a filter of a different scale and orientation, and subsampled by an appropriate factor. Coefficients that are spatially near each other within a band tend to have similar amplitudes. In addition, coefficients at different orientations or scales but in nearby (relative) spatial positions tend to have similar amplitudes.

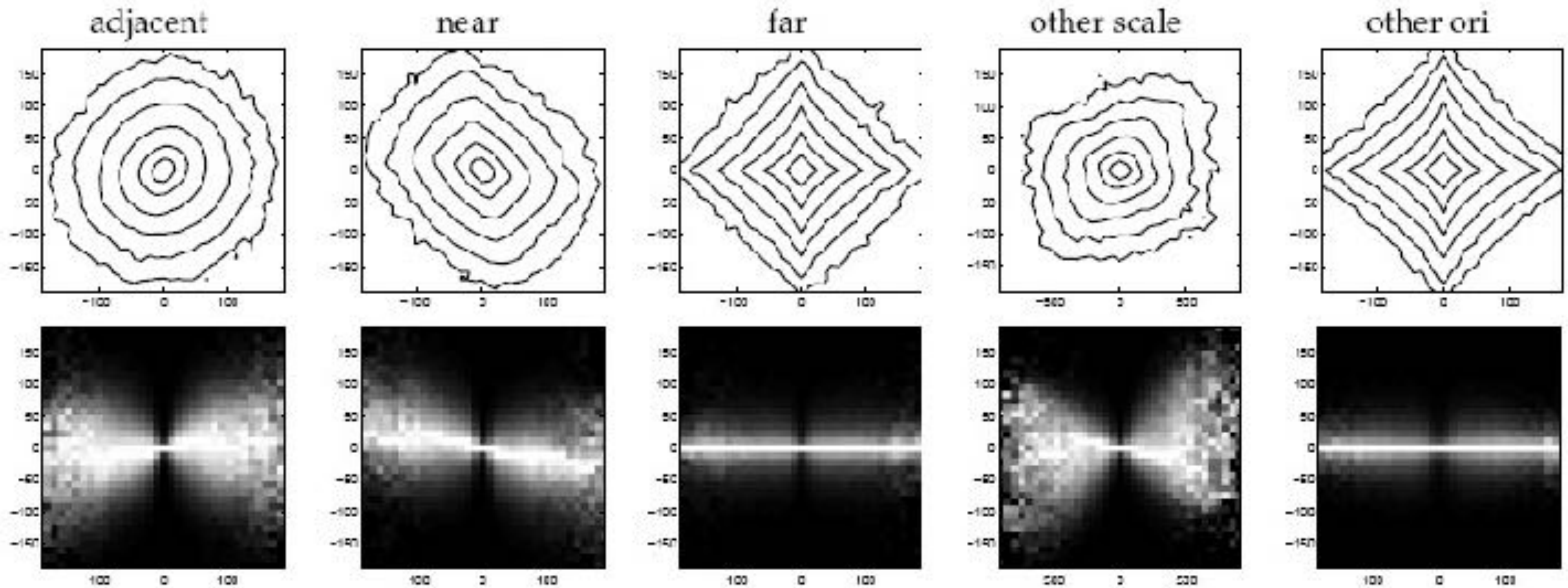
Note correlations between the amplitudes of each wavelet subband.



http://www.cns.nyu.edu/pub/eero/simoncelli05a-preprint.pdf

Statistics of pairs of wavelet coefficients

Contour plots of the joint histogram of various wavelet coefficient pairs



Conditional distributions of the corresponding wavelet pairs

Fig. 8. Empirical joint distributions of wavelet coefficients associated with different pairs of basis functions, for a single image of a New York City street scene (see Fig. 1 for image description). The top row shows joint distributions as contour plots, with lines drawn at equal intervals of log probability. The three leftmost examples correspond to pairs of basis functions at the same scale and orientation, but separated by different spatial offsets. The next corresponds to a pair at adjacent scales (but the same orientation, and nearly the same position), and the rightmost corresponds to a pair at orthogonal orientations (but the same scale and nearly the same position). The bottom row shows corresponding conditional distributions: brightness corresponds to frequency of occurance, except that each column has been independently rescaled to fill the full range of intensities.

http://www.cns.nyu.edu/pub/eero/simoncelli05a-preprint.pdf

Gaussian scale mixtures

$$P(\vec{x}) = \int \frac{\exp(-\frac{1}{2}\vec{x}^T (z\Lambda)^{-1}\vec{x})}{(2\pi)^{N/2}\,|z\Lambda|^{1/2}} P_z(z)\,dz$$

Wavelet coefficient probability

A mixture of Gaussians of scaled covariances

observed

Gaussian scale mixture model simulation

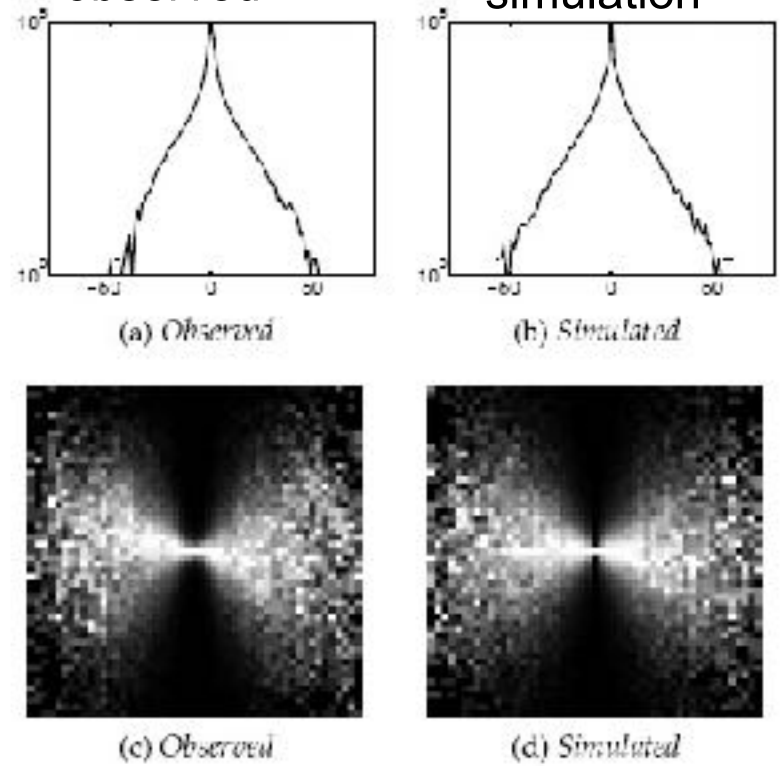
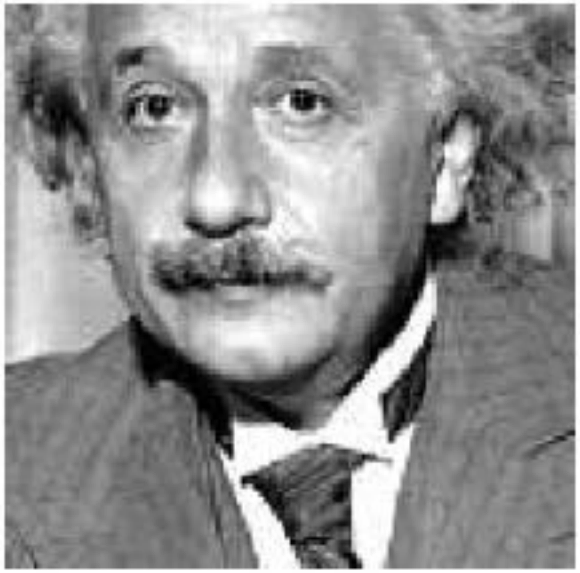


Fig. 9. Comparison of statistics of coefficients from an example image subband (left panels) with those generated by simulation of a local GSM model (right panels).

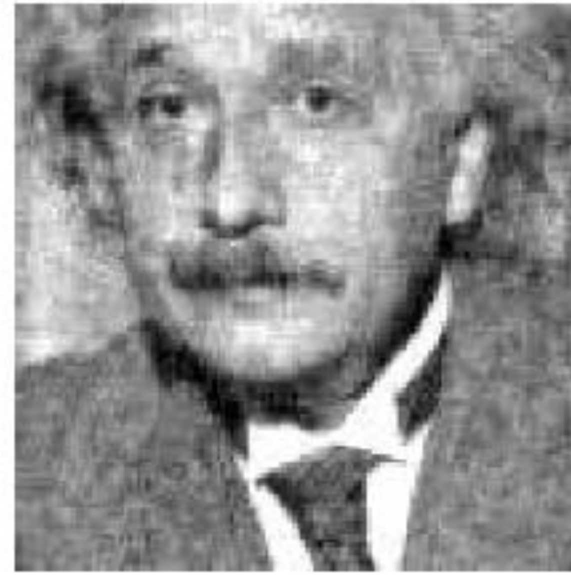
z is a spatially varying hidden variable that can be used to
(a) Create the non-gaussian histograms from a mixture of Gaussian densities, and
(b) model correlations between the neighboring wavelet coefficients.

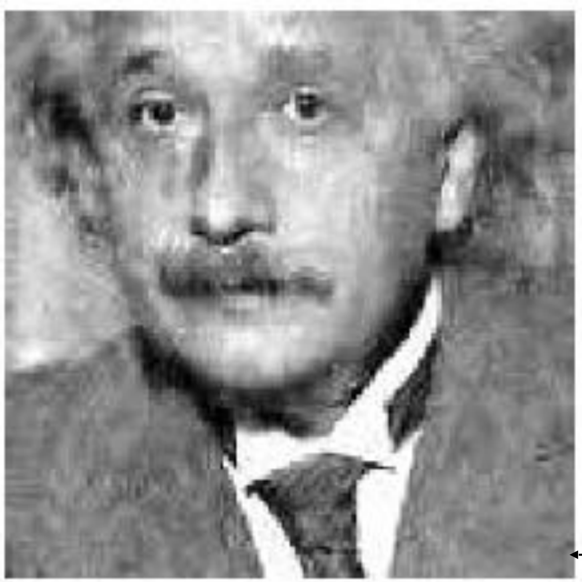
original



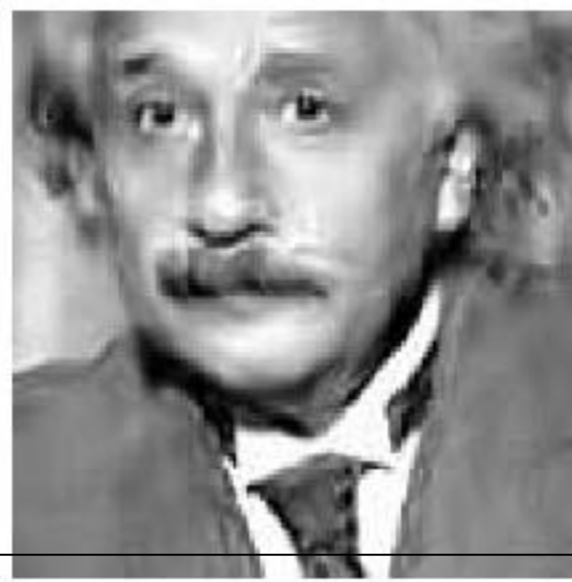
With Gaussian noise of std. dev. 21.4 added, giving PSNR=22.06

(1) Denoised with Gaussian model, PSNR=27.87





(2) Denoised with wavelet marginal model, PSNR=29.24



(3) Denoised with Gaussian scale mixture model, PSNR=30.86

http://www.cns.nyu.edu/pub/eero/simoncelli05a-preprint.pdf

Applications

- Detecting fake images

- Camera shake removal

Can we tell if a photograph is real?

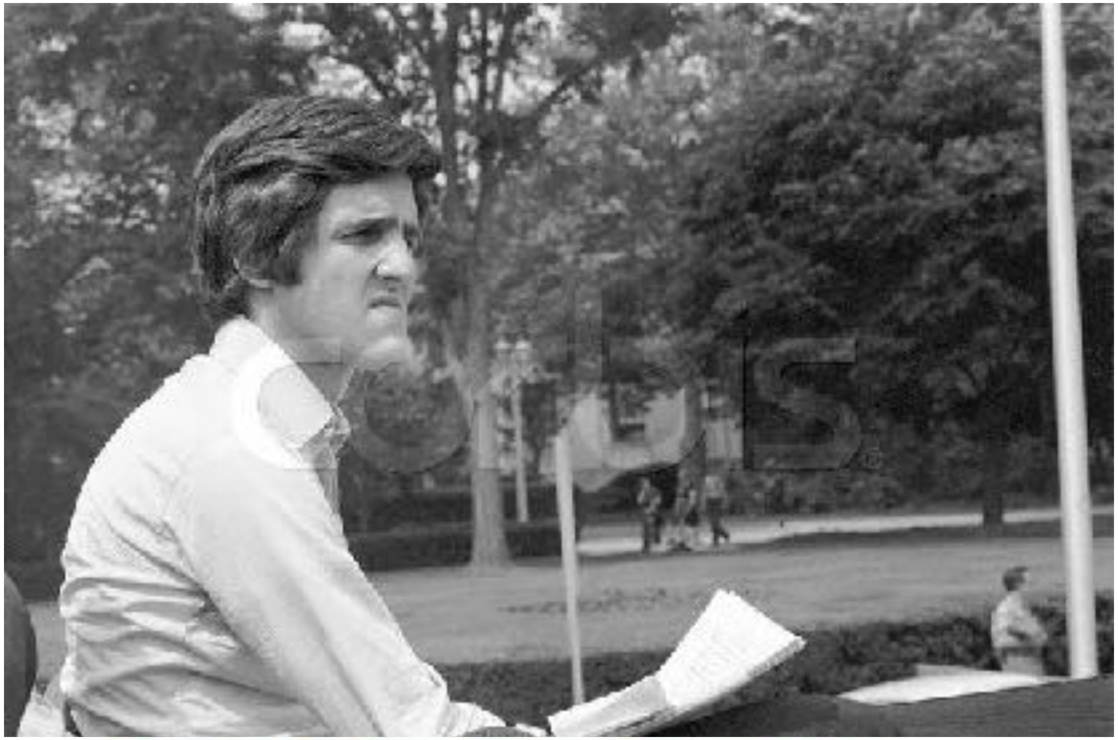
slides from Prof. Hany Farid,
Dartmouth College

Image circulated on internet

Fonda Speaks To Vietnam Veterans At Anti-War Rally

Actress And Anti-War Activist Jane Fonda Speaks to a crowd of Vietnam Veterans as Activist and former Vietnam Vet John Kerry (LEFT) listens and prepares to speak next concerning the war in Vietnam (AP Photo)

The source images



Update: Fonda, Kerry and Photo Fakery (free reg. required) - Photographer Ken Light describes the experience of discovering his 1970 photograph of John Kerry circulating in altered form on the Internet. "As far as I know, John Kerry never shared a demonstration podium with Jane Fonda, and the fact that a widely circulated photo showed him doing so — until it was exposed in recent weeks as a hoax — tells us more about the troublesome combination of Photoshop and the Internet than it does about the prospective Democratic candidate for president." (*Washington Post*)

Some applications requiring the ability to verify a photograph's validity

- Verify news photographs
- Child pornography prosecution.
  - A defense is to argue that the image is computer generated, thus there is a need to verify that an image is a photograph.

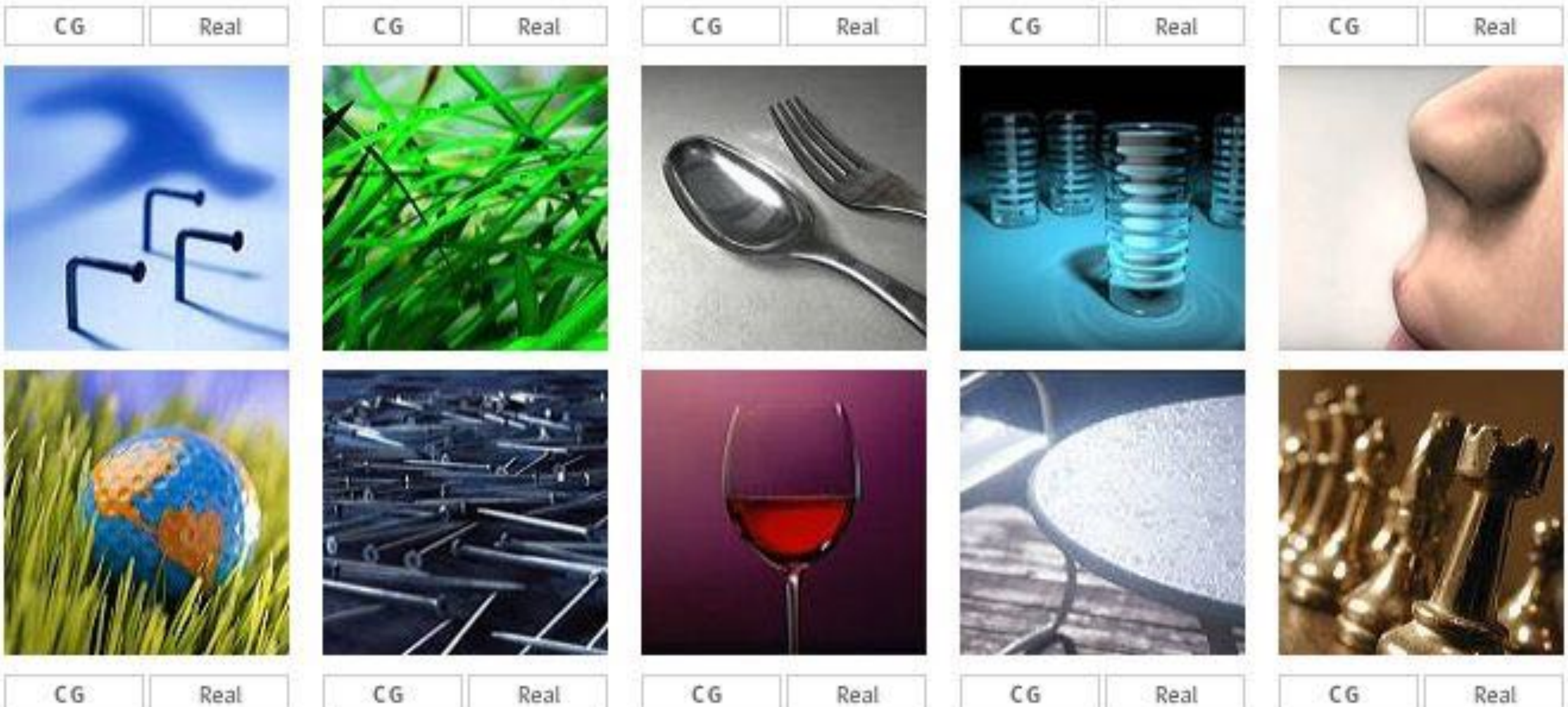
Which is the photograph?

Which is the photograph?

Which is the photograph?

Which is the photograph?

Some harder ones, from www.fakeorfoto.com



Answers

Real


©SUPERSTOCK/Four by Five Photography Inc. .


©SUPERSTOCK/Four by Five Photography Inc


©SUPERSTOCK/Four by Five Photography Inc.


©Heather Woodcock Somewhere in Time Photographics Images


©George Ihring Jüüce Interactive

Computer Graphics

Artist: Stefaan Contreras

Artist: Brad Clarkson

Artist:Kevin Mannens

Artist: Israel Yang

Artist: Matt Dougan

The CG images were created with Maya® software and the Maya® software renderer.

How can we determine which is real?

Image statistics can help us…

IEEE Transactions on Signal Processing, 53(2):845-850, 2005

How Realistic is Photorealistic?

Siwei Lyu and Hany Farid
Department of Computer Science
Dartmouth College
Hanover, NH 03755
Email: {lyu,farid}@cs.dartmouth.edu

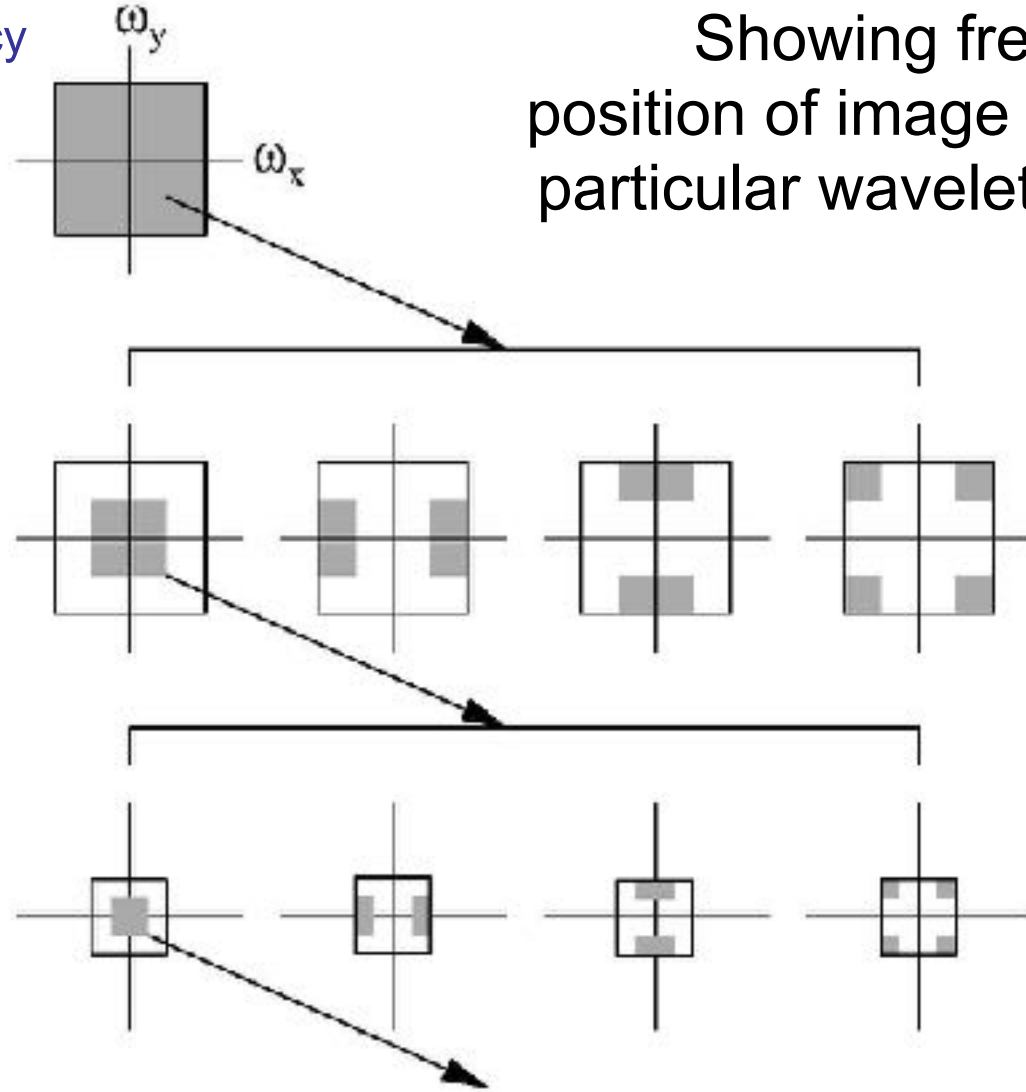
Abstract—**Computer graphics rendering software is capable of generating highly photorealistic images that can be impossible to differentiate from photographic images. As a result, the unique stature of photographs as a definitive recording of events is being diminished (the ease with which digital images can be manipulated is, of course,**

There has been some work in evaluating the photorealism of computer graphics rendered images from a human perception point of view (e.g., [10], [9], [11]). To our knowledge, however, no computational techniques exist to differentiate between photographic and photorealis-

How do statistics of photographic and photorealistic images differ?

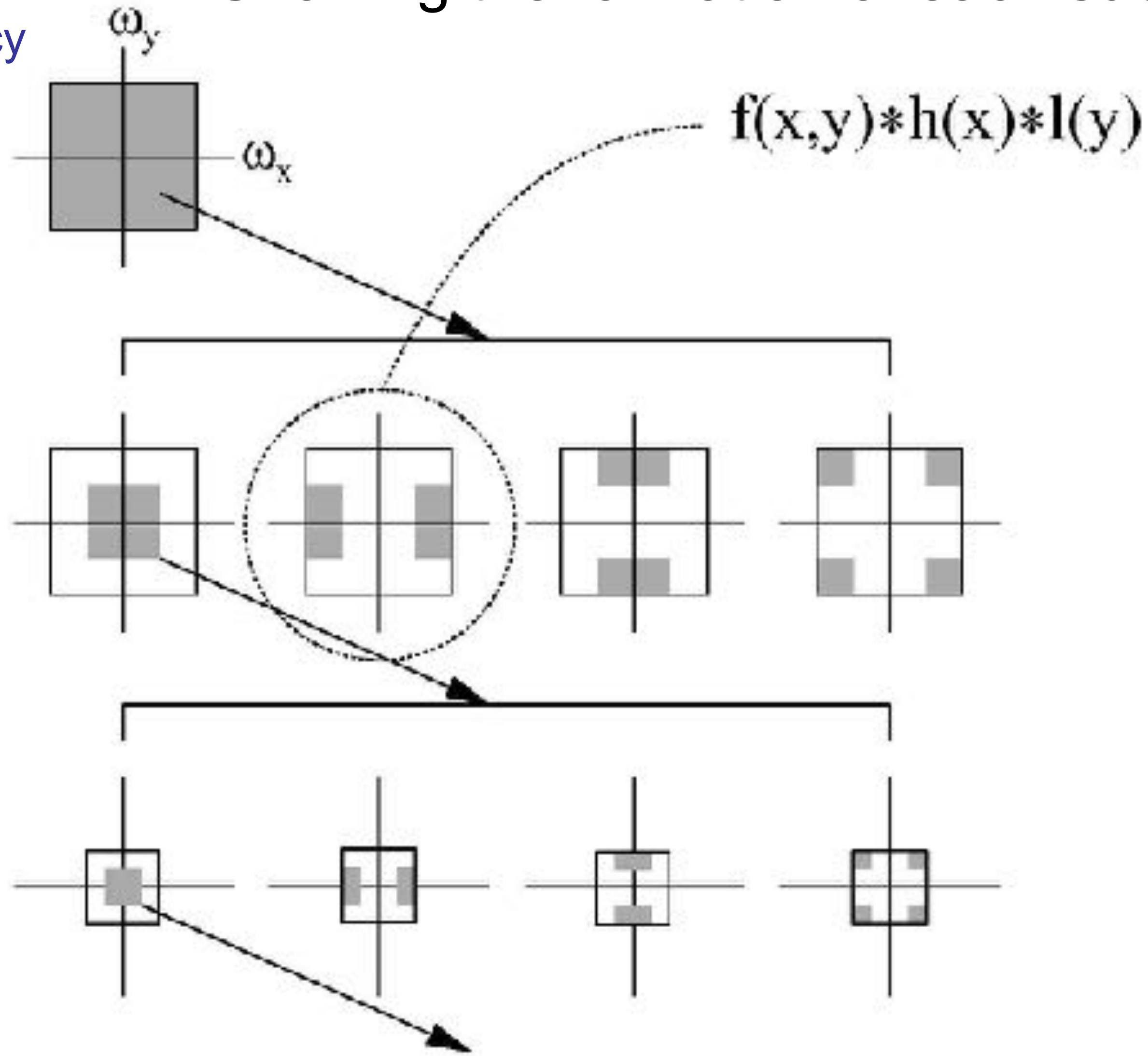
Showing frequency domain position of image subbands the a particular wavelet decomposition

Frequency domain



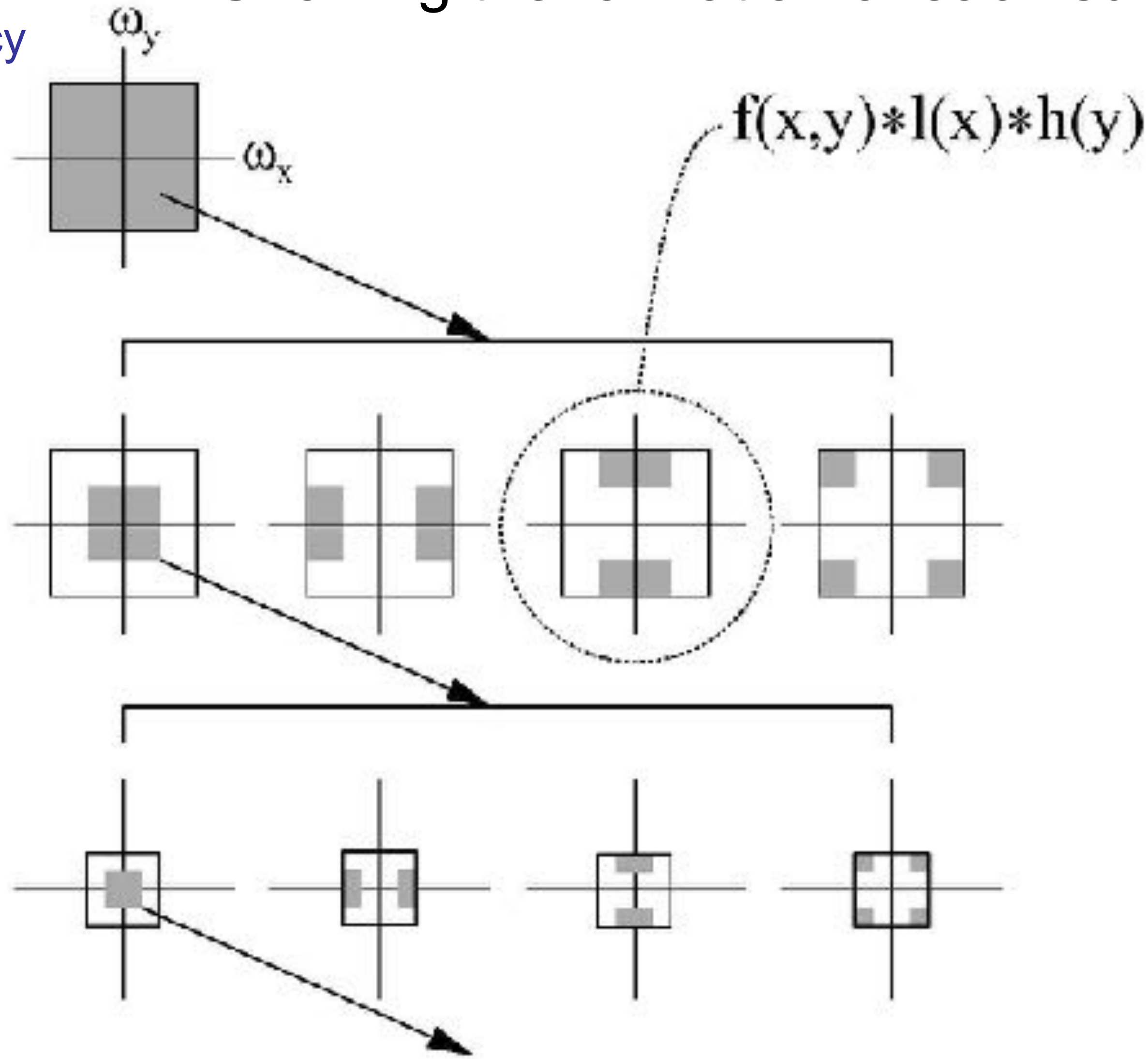
Showing the formation of each subband image

Frequency domain

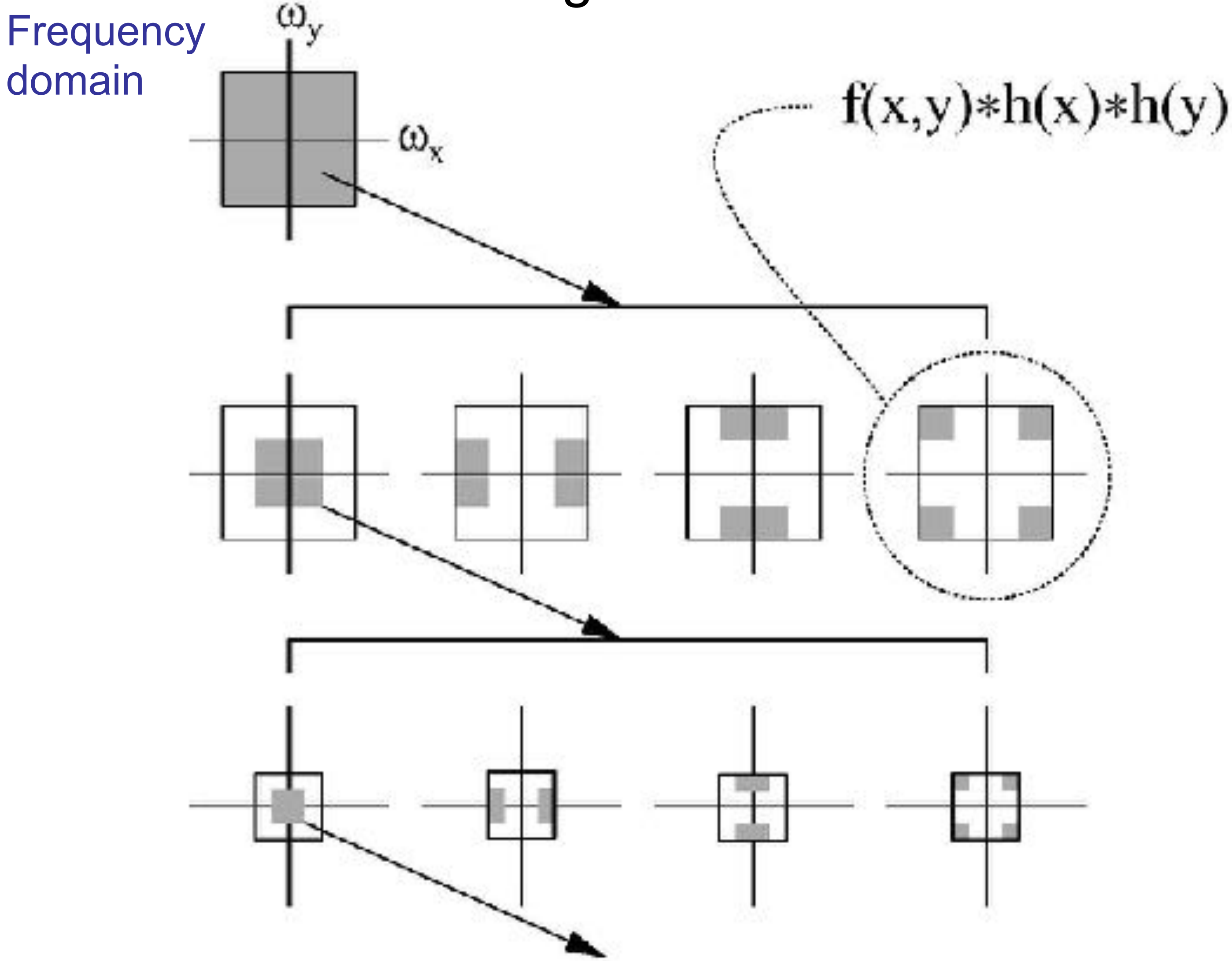


Showing the formation of each subband image

Frequency domain

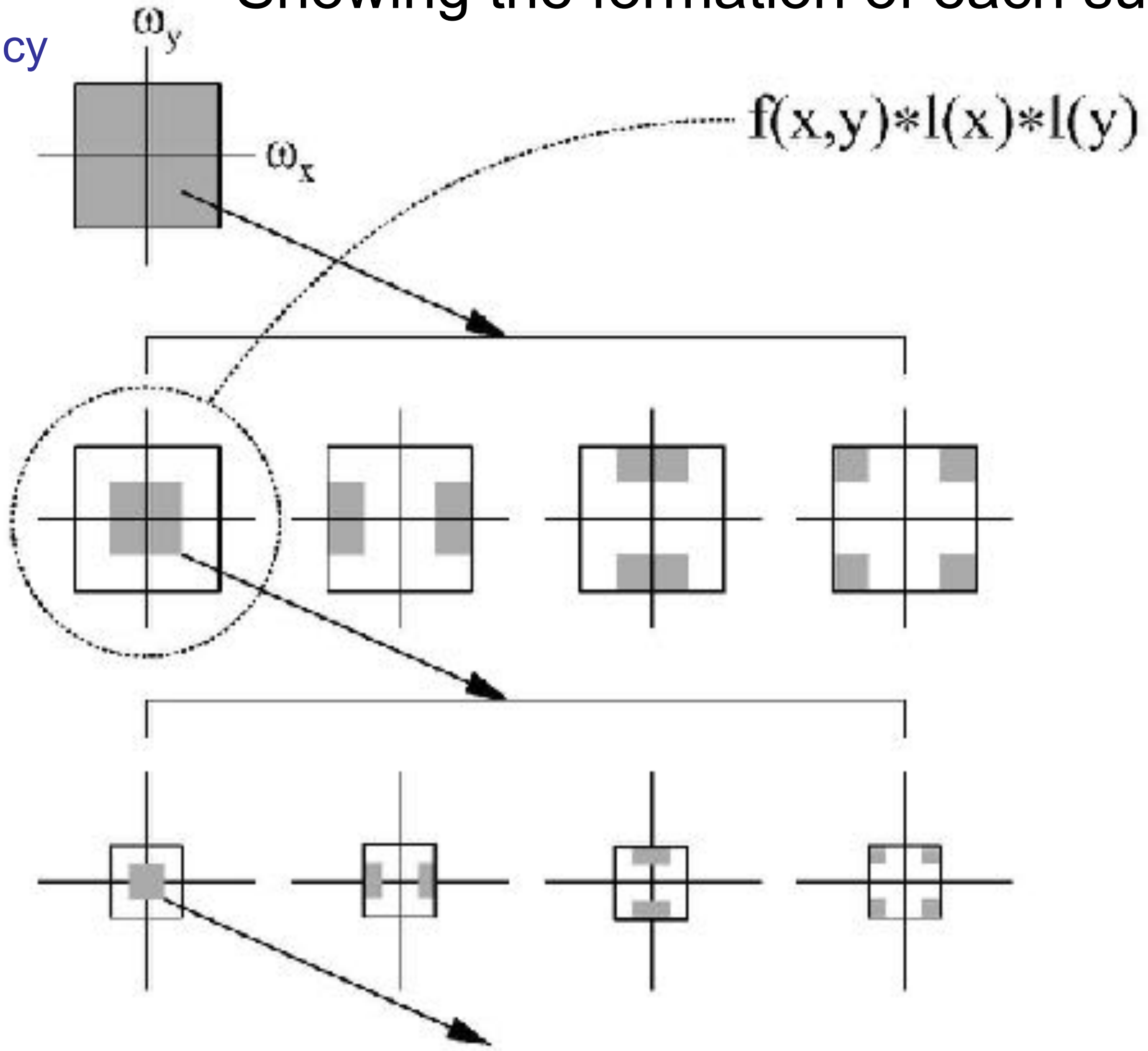


Showing the formation of each subband image



Showing the formation of each subband image

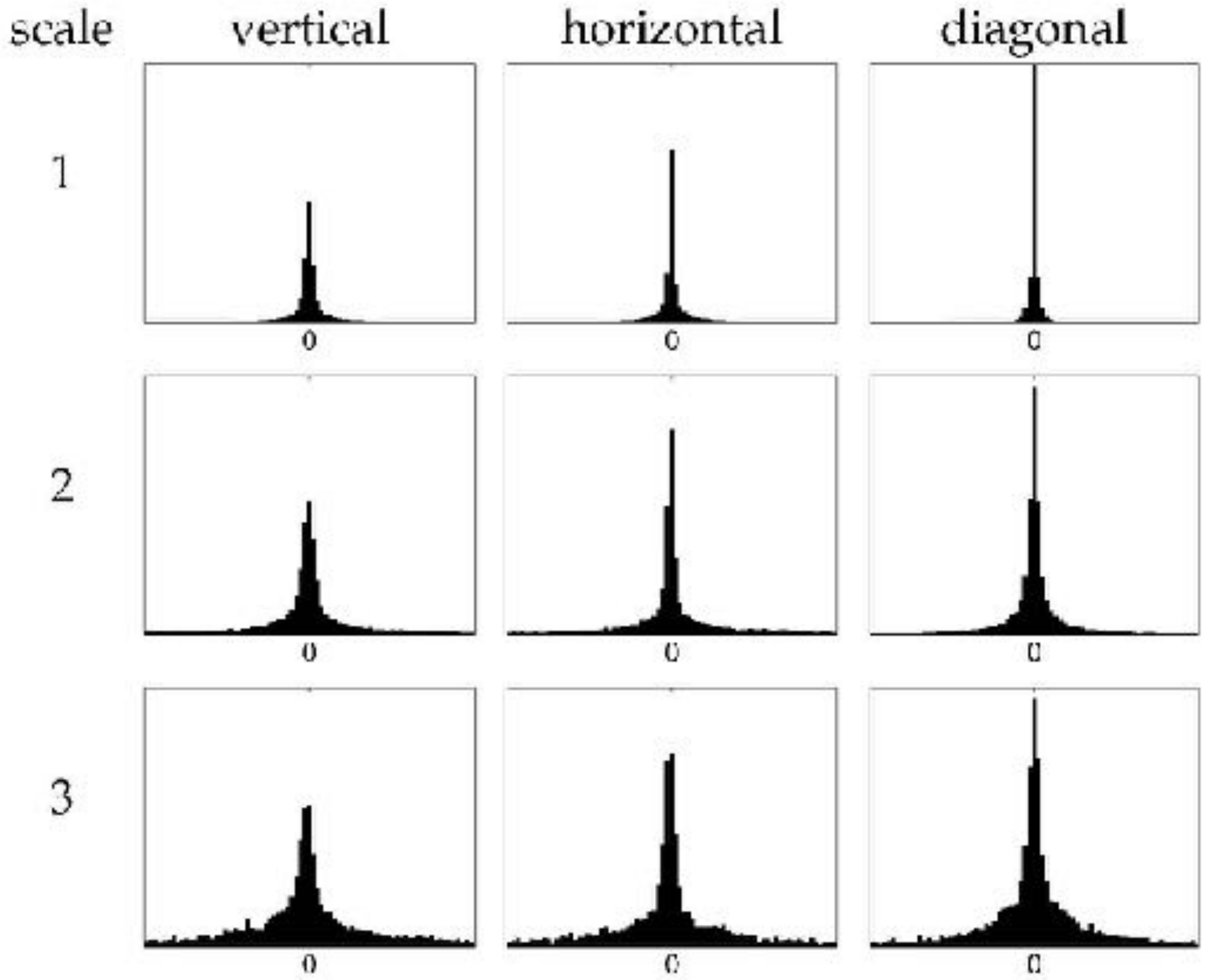
Frequency domain



Input image

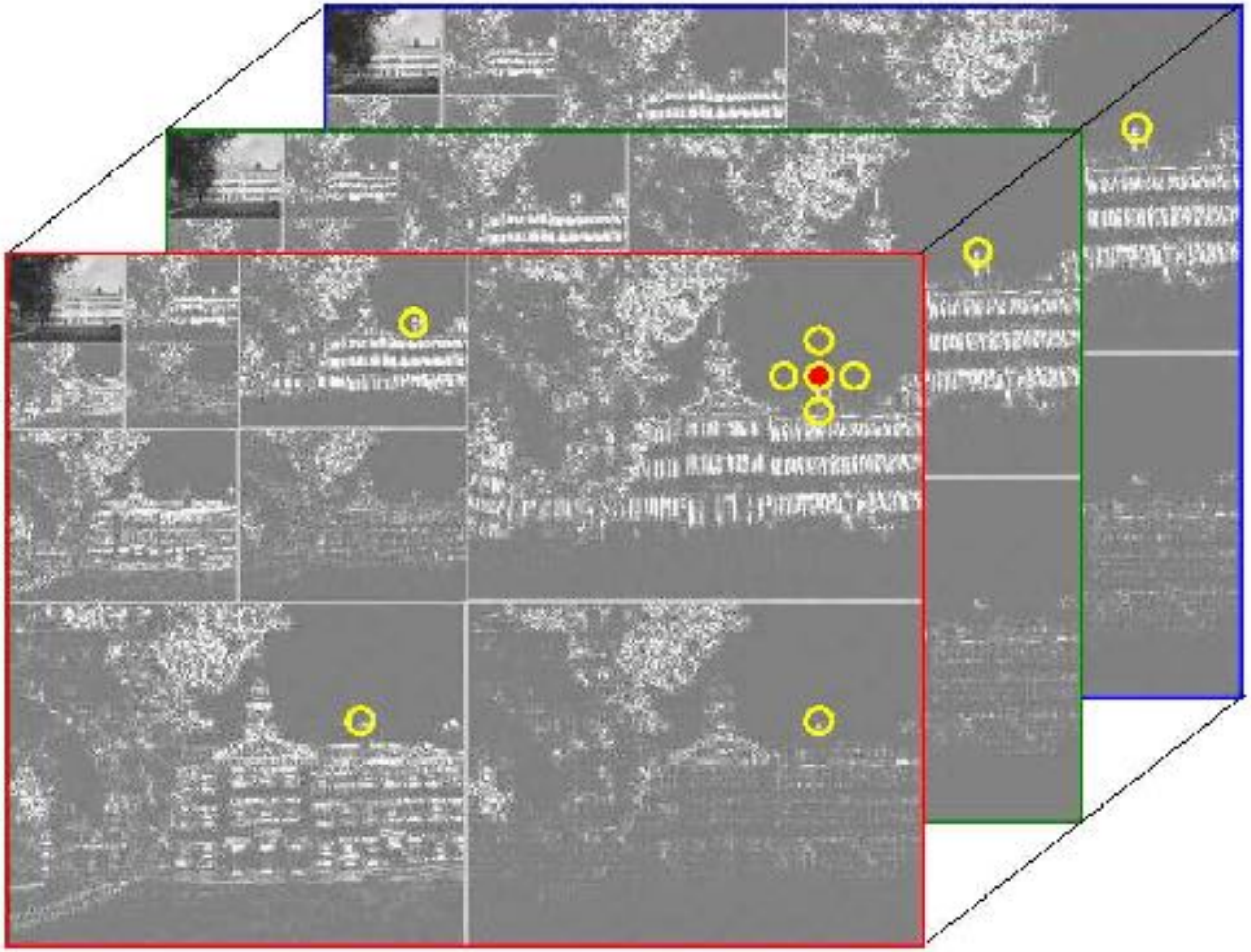
Representation of color input image in wavelet subbands

Histograms of wavelet subband coefficients



mean (μ), variance (μ_2), skewness (μ_3/σ^3), kurtosis (μ_4/σ^4)

There are correlations between subband coefficients



Predict coefficients as a linear combination of its neighbors

$$
\begin{aligned}
V_i(x,y) &= w_1 V_i(x-1,y) + w_2 V_i(x+1,y) + w_3 V_i(x,y-1) \\
&+ w_4 V_i(x,y+1) + w_5 V_{i-1}(x/2,y/2) + w_6 H_i(x,y) \\
&+ w_7 D_i(x,y) + w_8 V_i(x,y) + w_9 V_i(x,y)
\end{aligned}
$$

$$
\begin{pmatrix} | \\ | \\ V_i(x,y) \\ | \\ | \end{pmatrix} = \begin{pmatrix} | & & | \\ | & & | \\ V_i(x-1,y) & \dots & V_i(x,y) \\ | & & | \\ | & & | \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ \vdots \\ w_9 \end{pmatrix}
$$

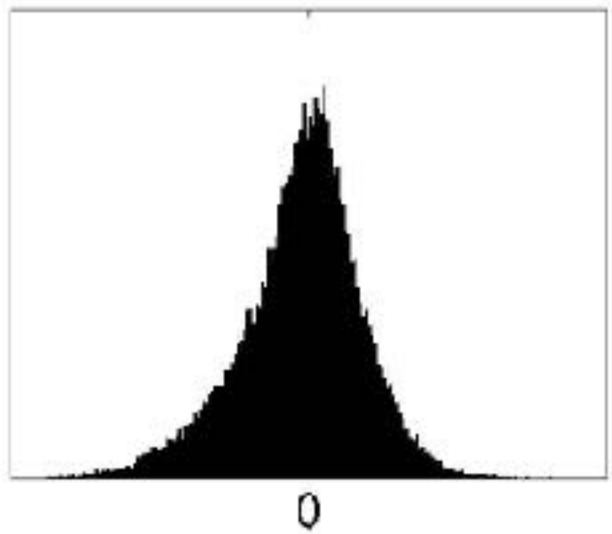
$$
\vec{V} = Q\vec{w}
$$

Straightforward to find optimal predictor of subband coefficients from neighbors

$\vec{V} = Q\vec{w}$	linear predictor
$E(\vec{w}) = [\vec{V} - Q\vec{w}]^2$	quadratic error
$\frac{dE(\vec{w})}{d\vec{w}} = 2Q^T[\vec{V} - Q\vec{w}]$	differentiate
$\frac{dE(\vec{w})}{d\vec{w}} = 0$	minimize
$\vec{w} = (Q^TQ)^{-1}Q^T\vec{V}$	optimal predictor

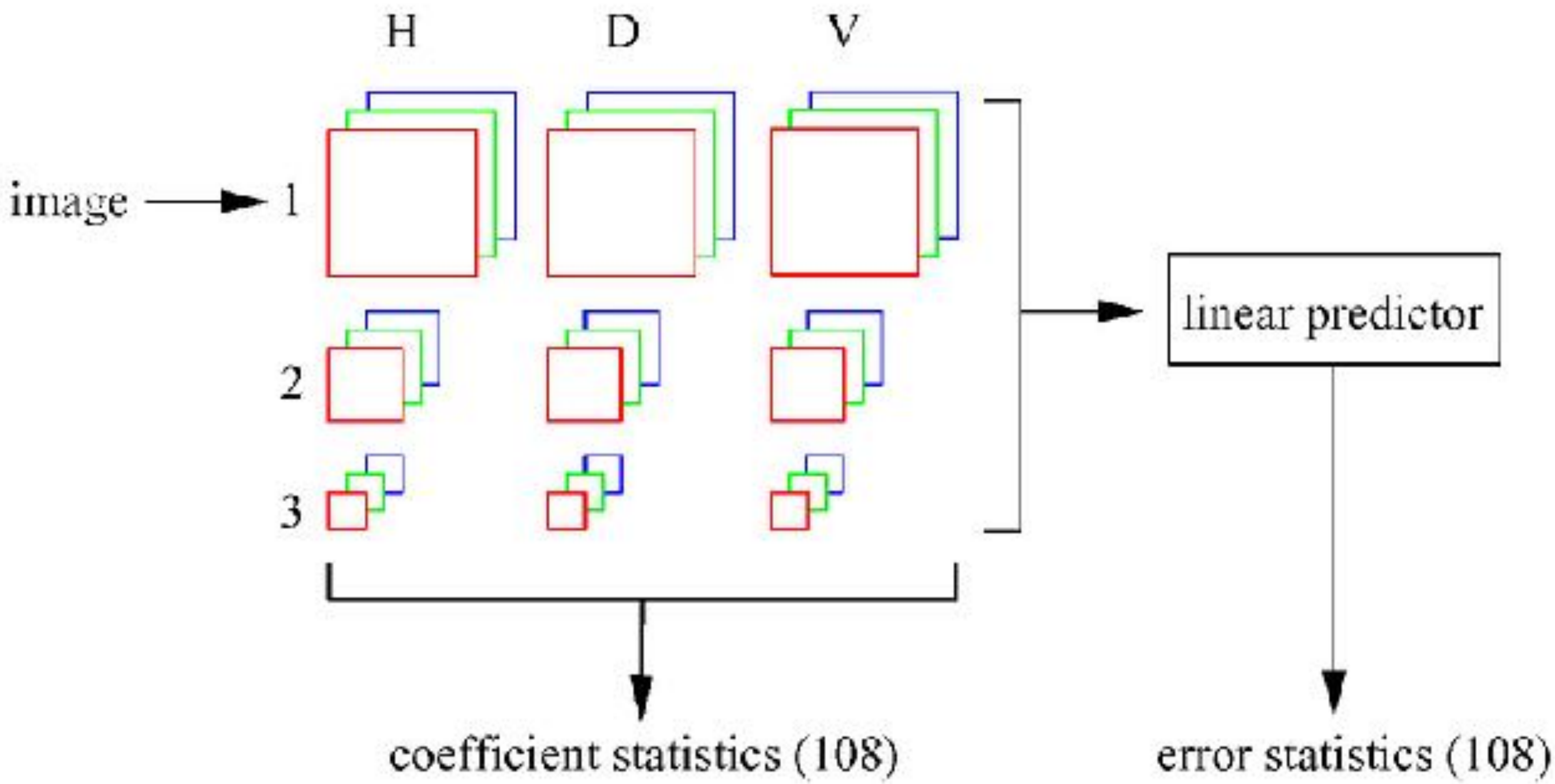
$\vec{w} = (Q^T Q)^{-1} Q^T \vec{V}$	optimal predictor
$\log_2(\vec{V}) - \log_2(\lvert Q\vec{w} \rvert)$	prediction error

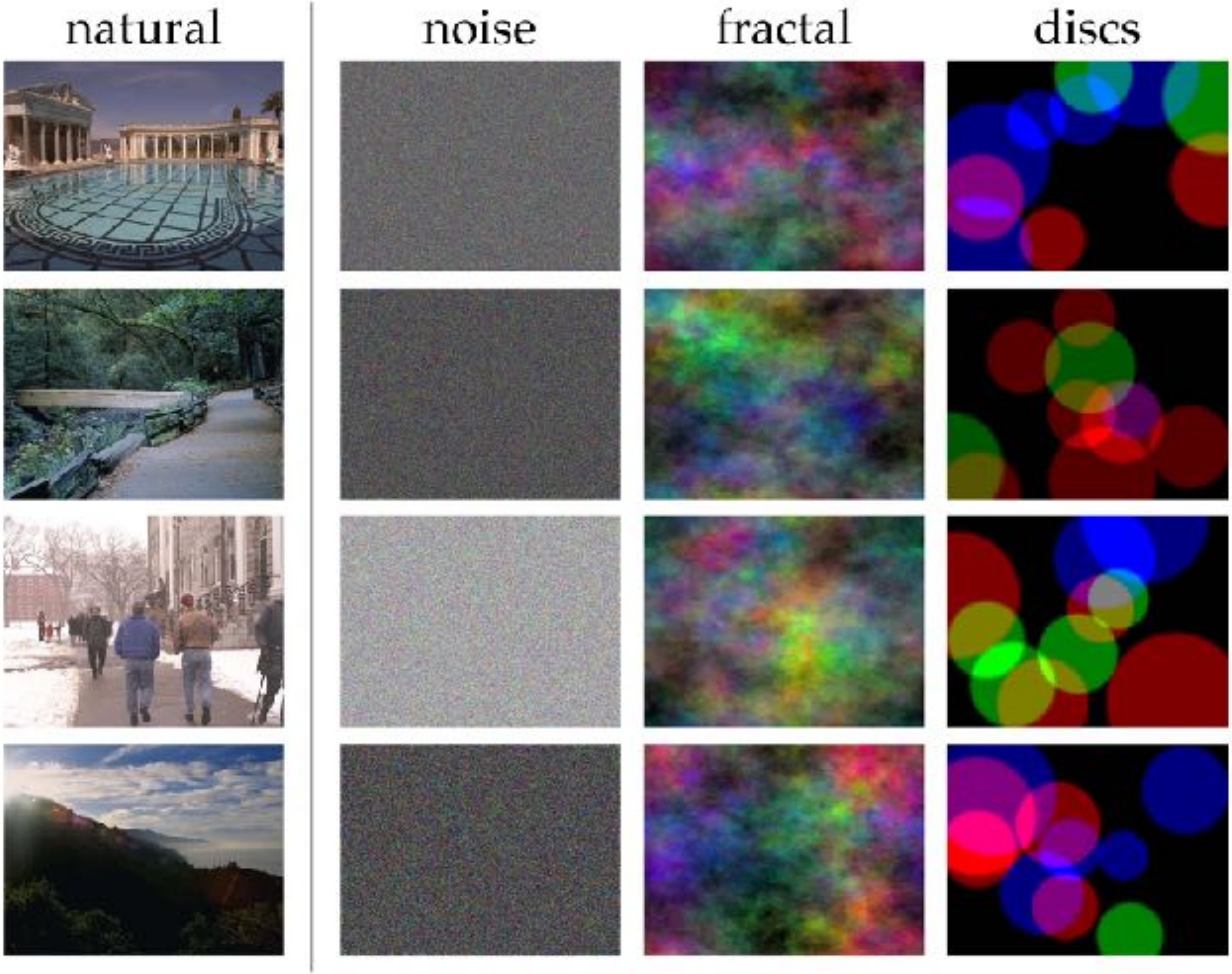
Another feature: the histogram of the prediction errors



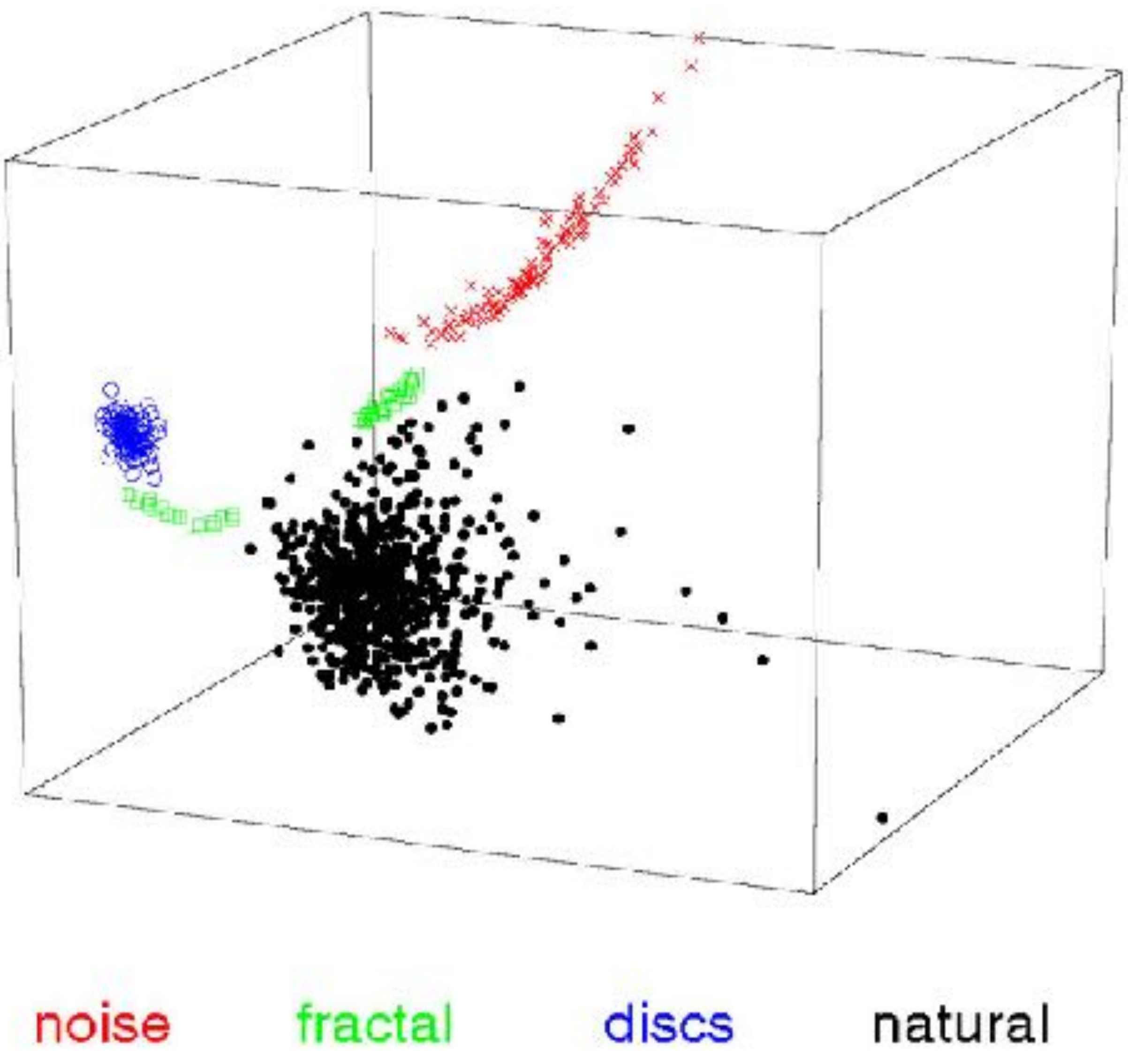
mean (μ), variance (μ_2), skewness (μ_3/σ^3), kurtosis (μ_4/σ^4)

Summary of features used for image classification




natural
noise
fractal
discs

Projection of measured features into a 3-d space: well separated even in that low-dimensional space



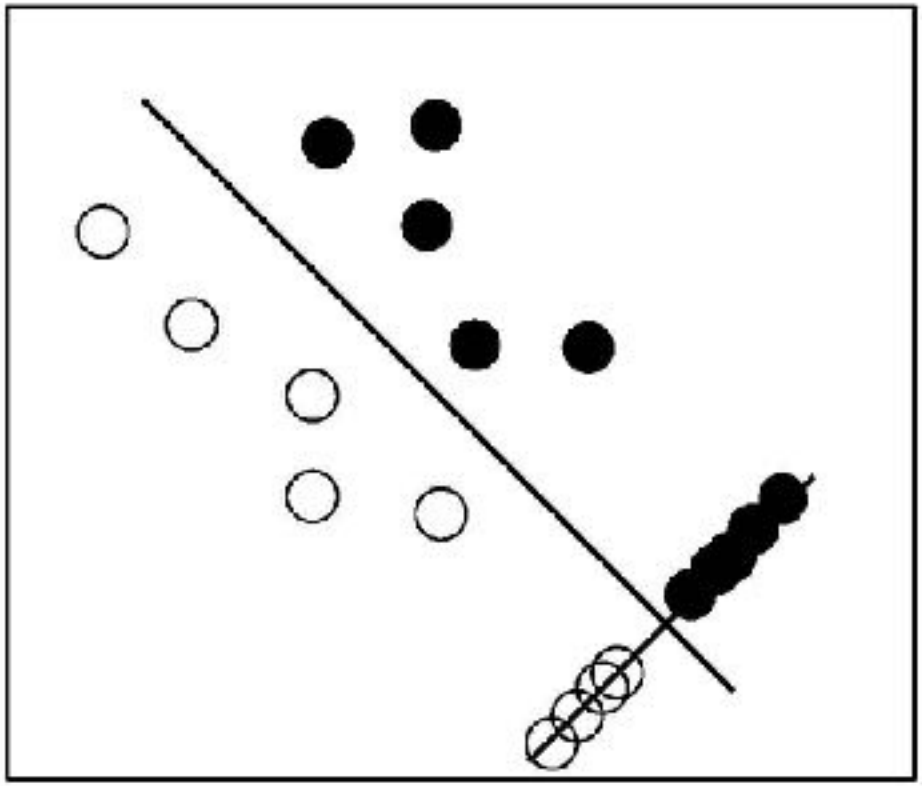
Photographic training set: downloaded from www.freefoto.com

photographic (40,000)

Photorealistic training set:
downloaded from www.raph.com and www.irtc.org

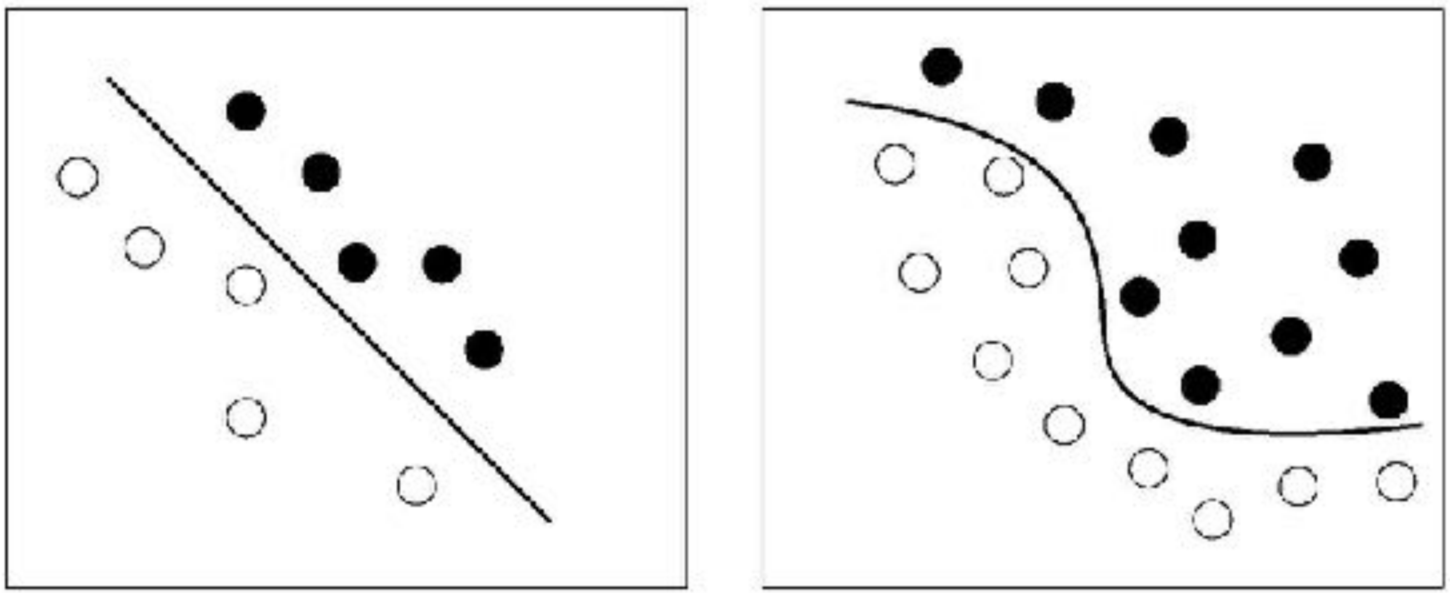
photorealistic (6,000)

Classifier 1: LDA. Simple, amenable to analysis



linear discriminant analysis (LDA)

Classifier 2: Support Vector Machine (SVM).



linear SVM — non-linear SVM

Learn a classifier from training data

Thresholds set favoring correct identification of photographs:

	training		testing	
	LDA	SVM	LDA	SVM
photographic	99.1	99.8	99.0	98.6
photorealistic	57.8	76.0	56.4	72.1

Thresholds set favoring correct identification of computer graphic images:

	training		testing	
	LDA	SVM	LDA	SVM
photographic	58.7	70.9	54.6	66.8
photorealistic	99.4	99.1	99.2	98.8

Allowing only 1 in 100 chances of a mistake against the defendant, correctly identify photos as being photographs 67% of the time.

Easily classified photographic images

Fig. 4: Easily classified photographic images.

Easily classified photorealistic images

Fig. 5: Easily classified photorealistic images.

Incorrectly classified photographic images

Fig. 6: Incorrectly classified photographic images.

Incorrectly classified photorealistic images

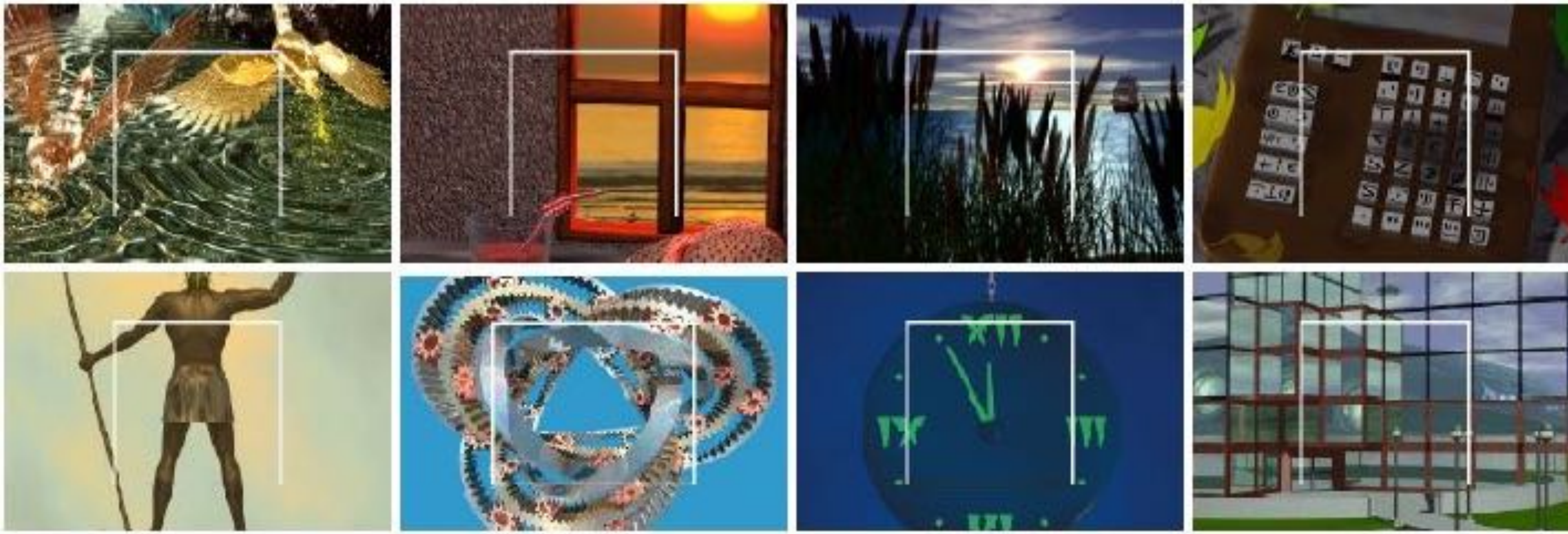
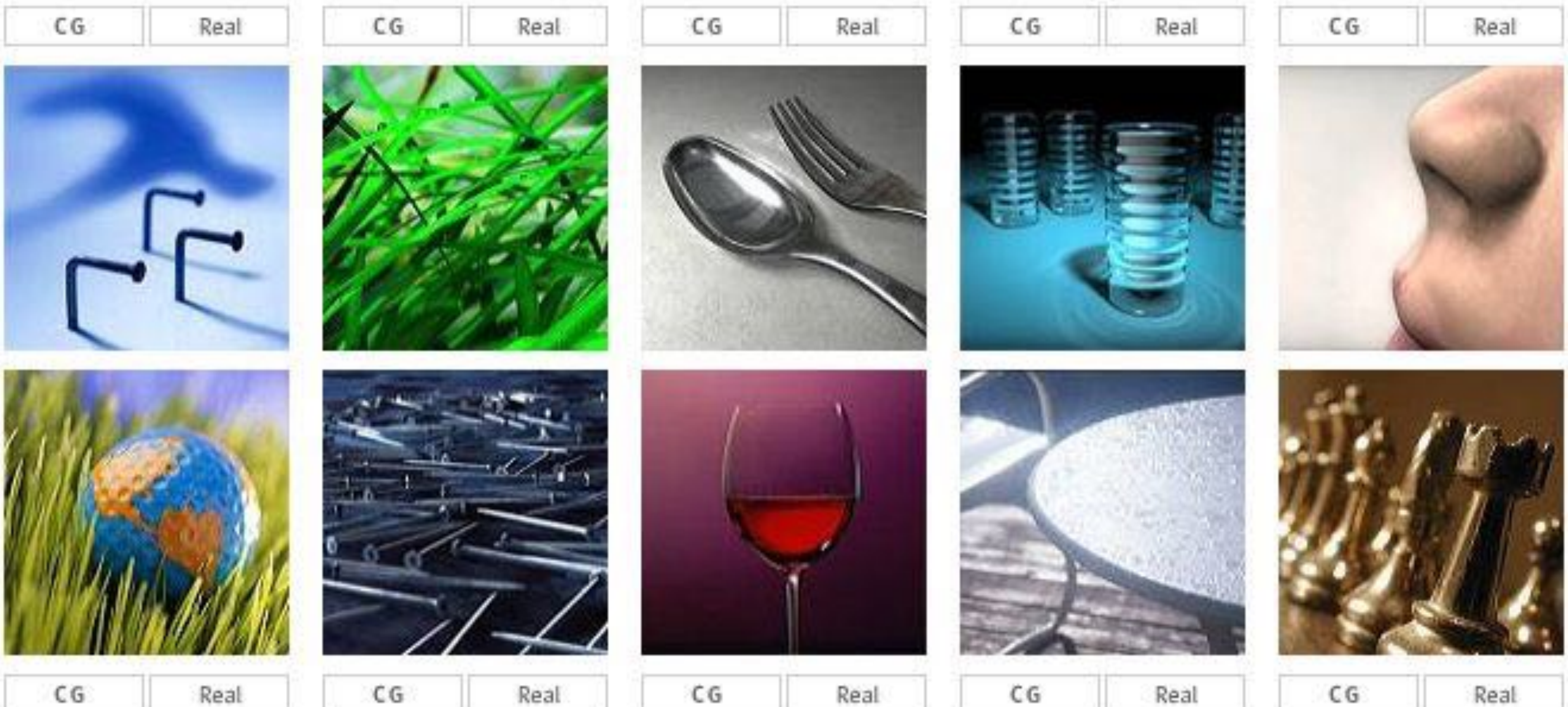


Fig. 7: Incorrectly classified photorealistic images.

www.fakeorfoto.com



Alias

StudioTools Maya

Results of algorithm

Photographic images | Photorealistic images

Correct

Incorrect

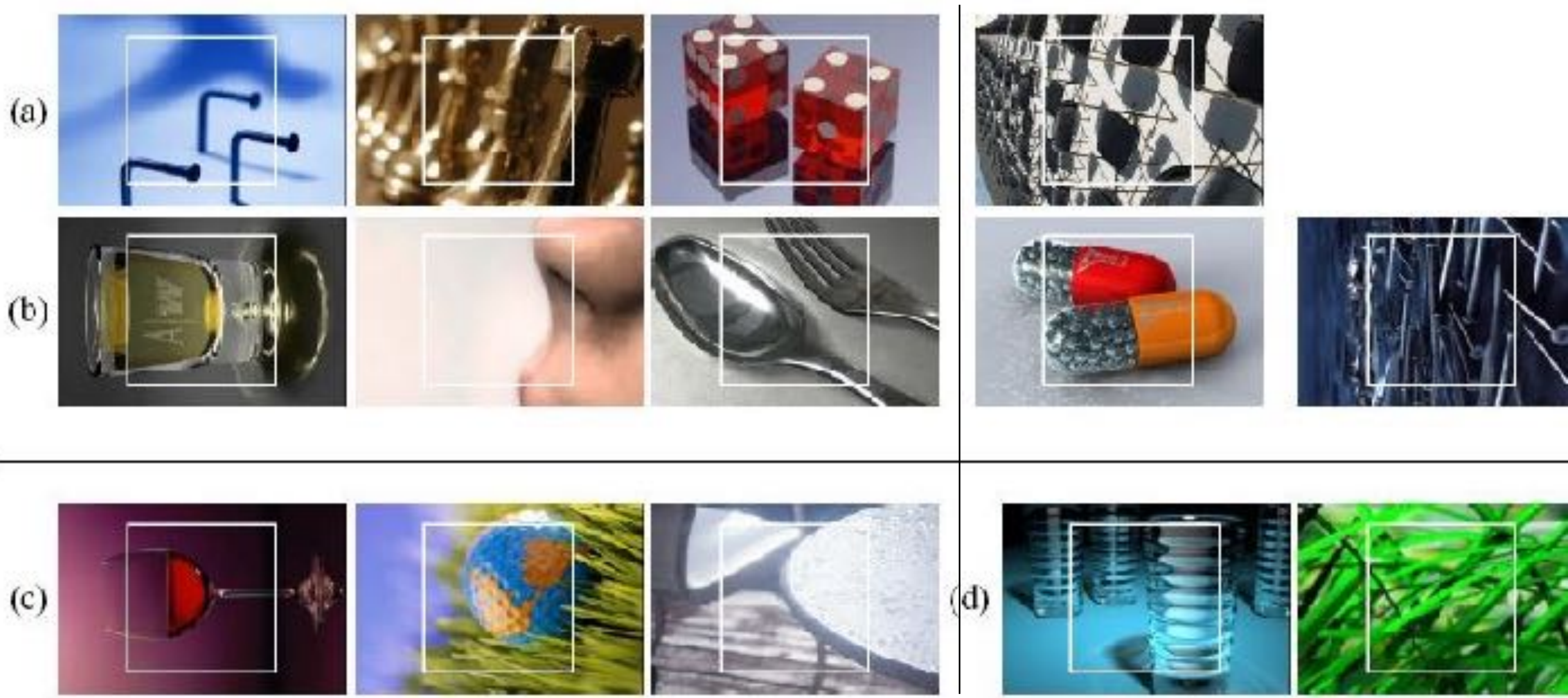


Fig. 9: Images from www.fakeorfoto.com. Shown in (a) and (c) are correctly and incorrectly classified photographic images, respectively. Shown in (b) and (d) are correctly and incorrectly classified photorealistic images, respectively.

Hany Farid's book

The MIT Press

Books | Journals | Open | About | Giving

My Account | Cart (0) | Help

Search this site... GO

COMPUTER SCIENCE AND INTELLIGENT SYSTEMS → PHOTO FORENSICS

Photo Forensics

By Hany Farid

Overview

Photographs have been doctored since photography was invented. Dictators have erased people from photographs and from history. Politicians have manipulated photos for short-term political gain. Altering photographs in the predigital era required time-consuming darkroom work. Today, powerful and low-cost digital technology makes it relatively easy to alter digital images, and the resulting fakes are difficult to detect. The field of photo forensics—pioneered in Hany Farid's lab at Dartmouth College—restores some trust to photography. In this book, Farid describes techniques that can be used to authenticate photos. He provides the intuition and background as well as the mathematical and algorithmic details needed to understand, implement, and utilize a variety of photo forensic techniques.

Farid traces the entire imaging pipeline. He begins with the physics and geometry of the interaction of light with the physical world, proceeds through the way light passes through a camera lens, the conversion of light to pixel values in the electronic sensor, the packaging of the pixel values into a digital

Rent eTextbook

Buying Options