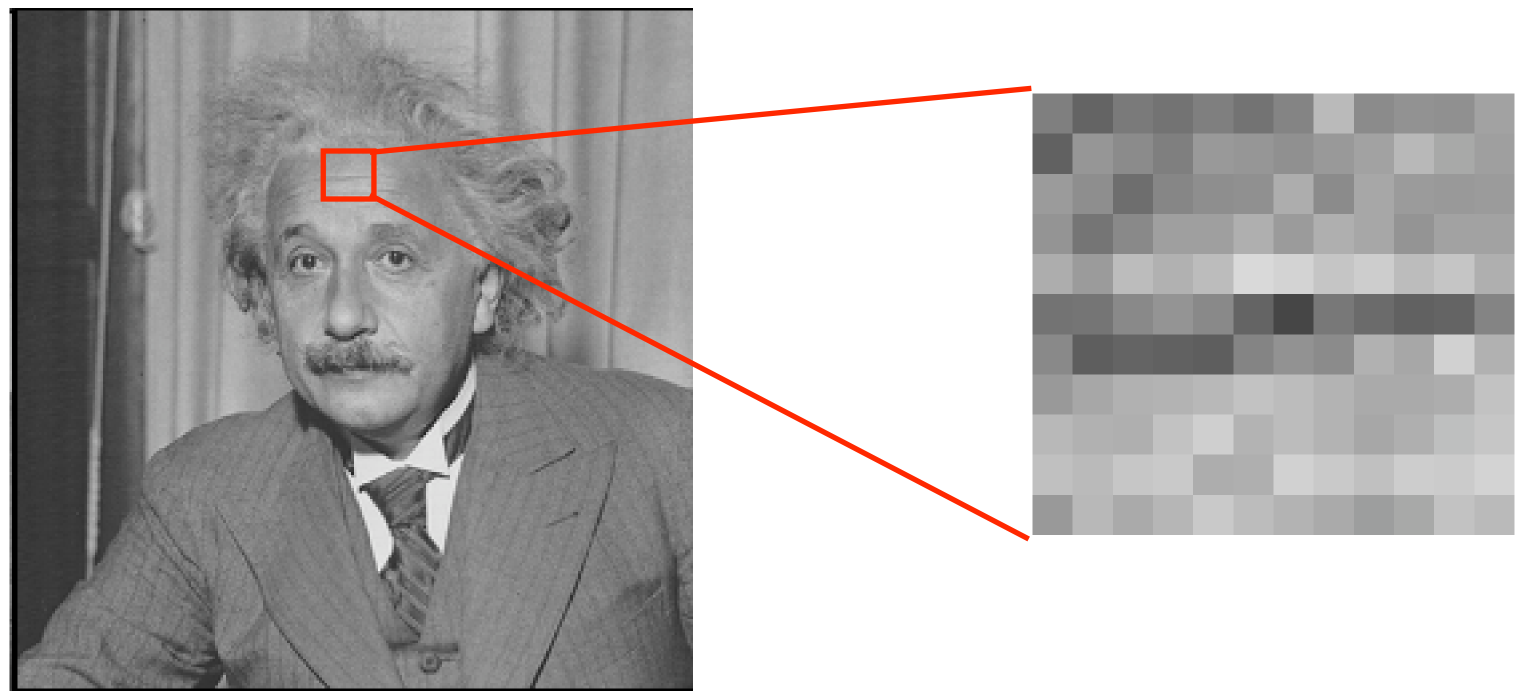
Image representation with multi-scale gradients

Eero P Simoncelli
Center for Neural Science, and
Courant Institute of Mathematical Sciences
New York University

http://www.cns.nyu.edu/~eero

Visual image representation



- Clean edges are rare
- One scale's texture is another scale's edge
- Need seamless transitions from isolated features to dense textures
- ~~Sparseness?~~ → explicit geometric quantities

Sparseness

Seems intuitively appealing, but in practice:

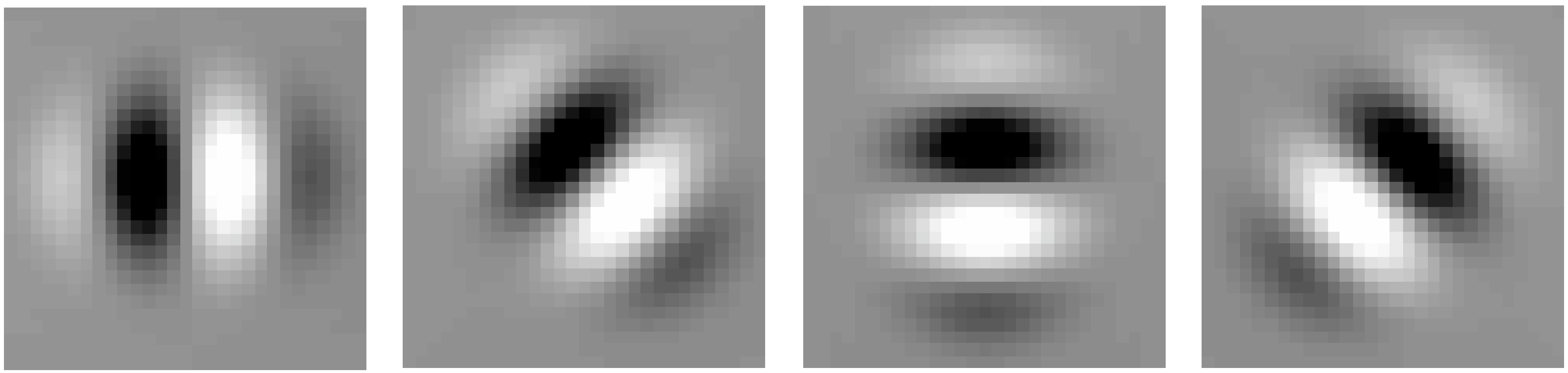
- Hard to compute
- Unlikely to benefit compression (must code addresses!)
- Discontinuous under geometric distortion of image
- Relevance to neuroscience is questionable

Dissent (at the Republican National Convention)

Multi-scale gradient basis

- Multi-scale bases: efficient representation
- Derivatives: good for analysis
  - Local Taylor expansion of image structures
  - Explicit geometry (orientation)
- Combination:
  - Explicit incorporation of geometry in basis
  - Bridge between PDE / harmonic analysis approaches

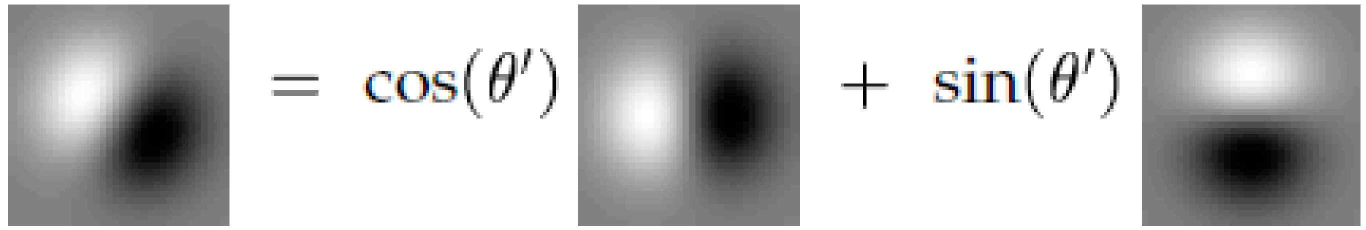
Steerable pyramid



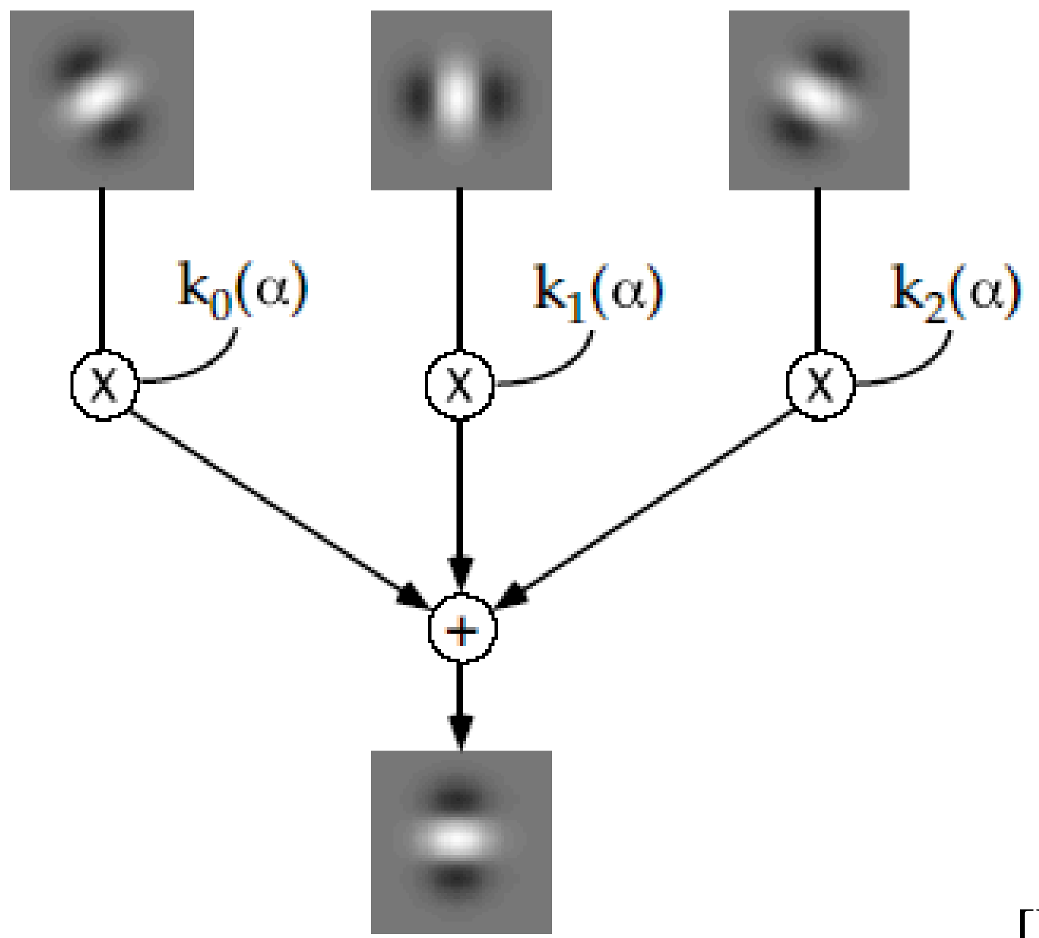
- Basis functions are *K*th derivative operators, related by translation/dilation/rotation
- Tight frame (*4(K-1)*/*3* overcomplete)
- Translation-invariance, rotation-invariance

[Simoncelli et.al., 1992; Simoncelli & Freeman 1995]

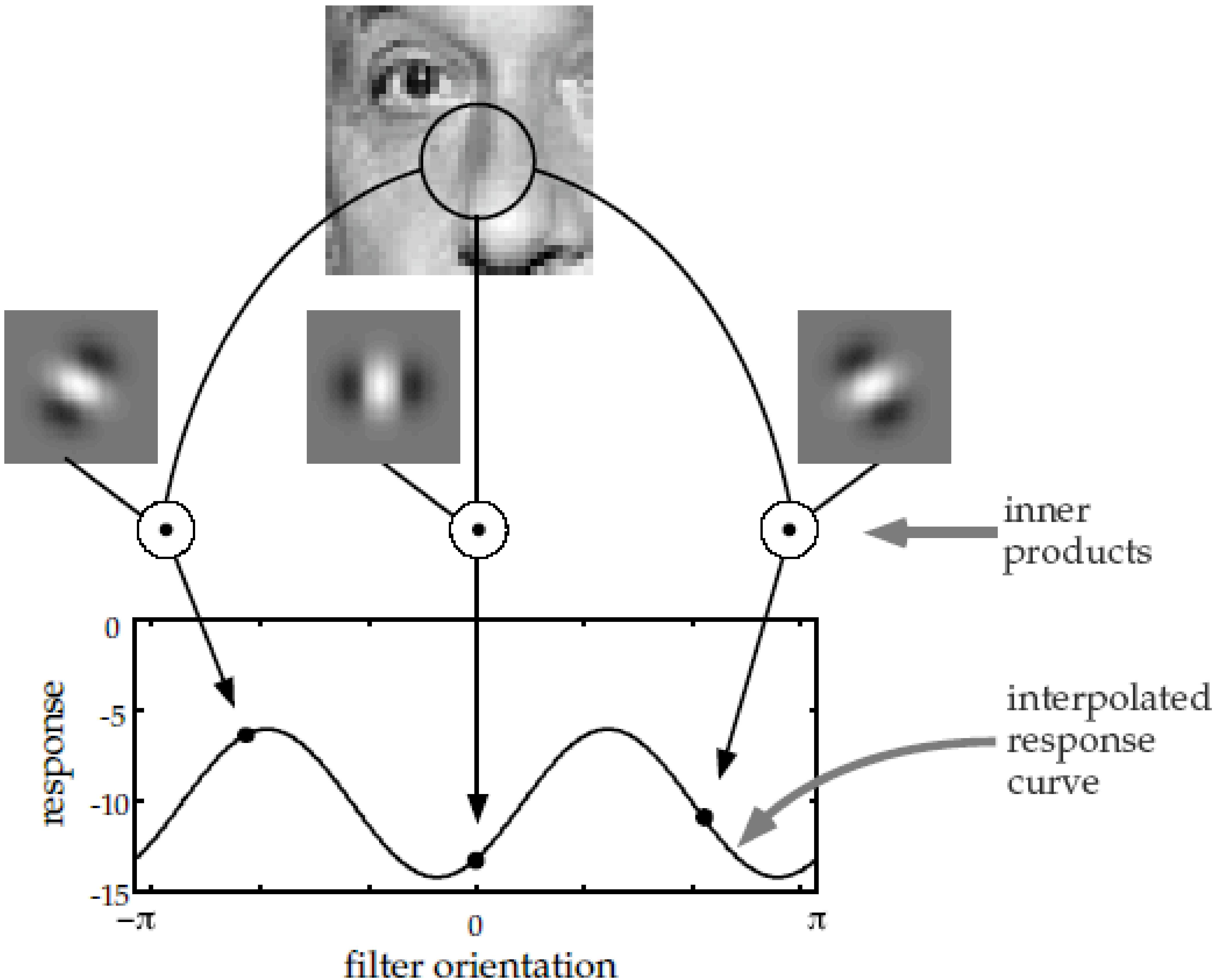
- First derivative of radial function:



- *k+1* *k*th-order directional derivatives:

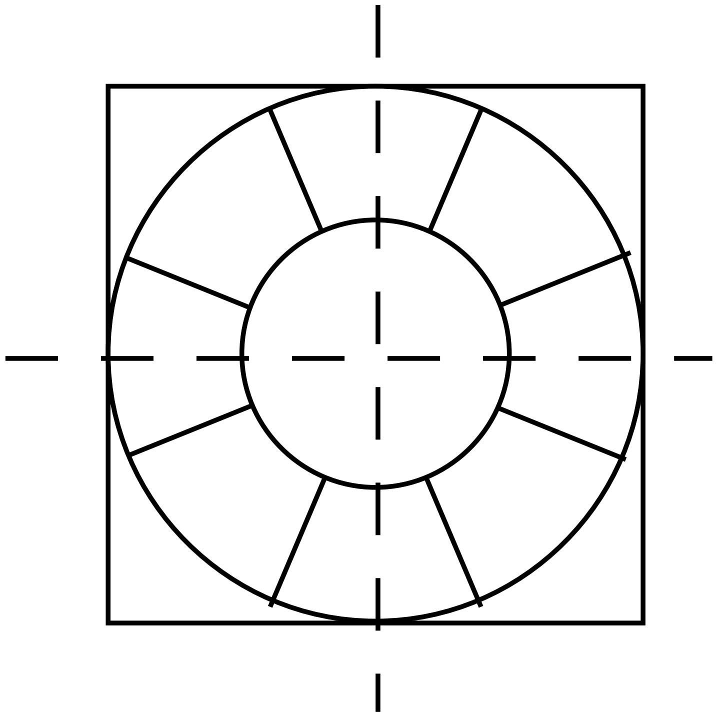


[Freeman & Adelson, ‘91]



inner
products
0
-5
-10
-15
response
interpolated
response
curve
−π
0
π
filter orientation

Filters



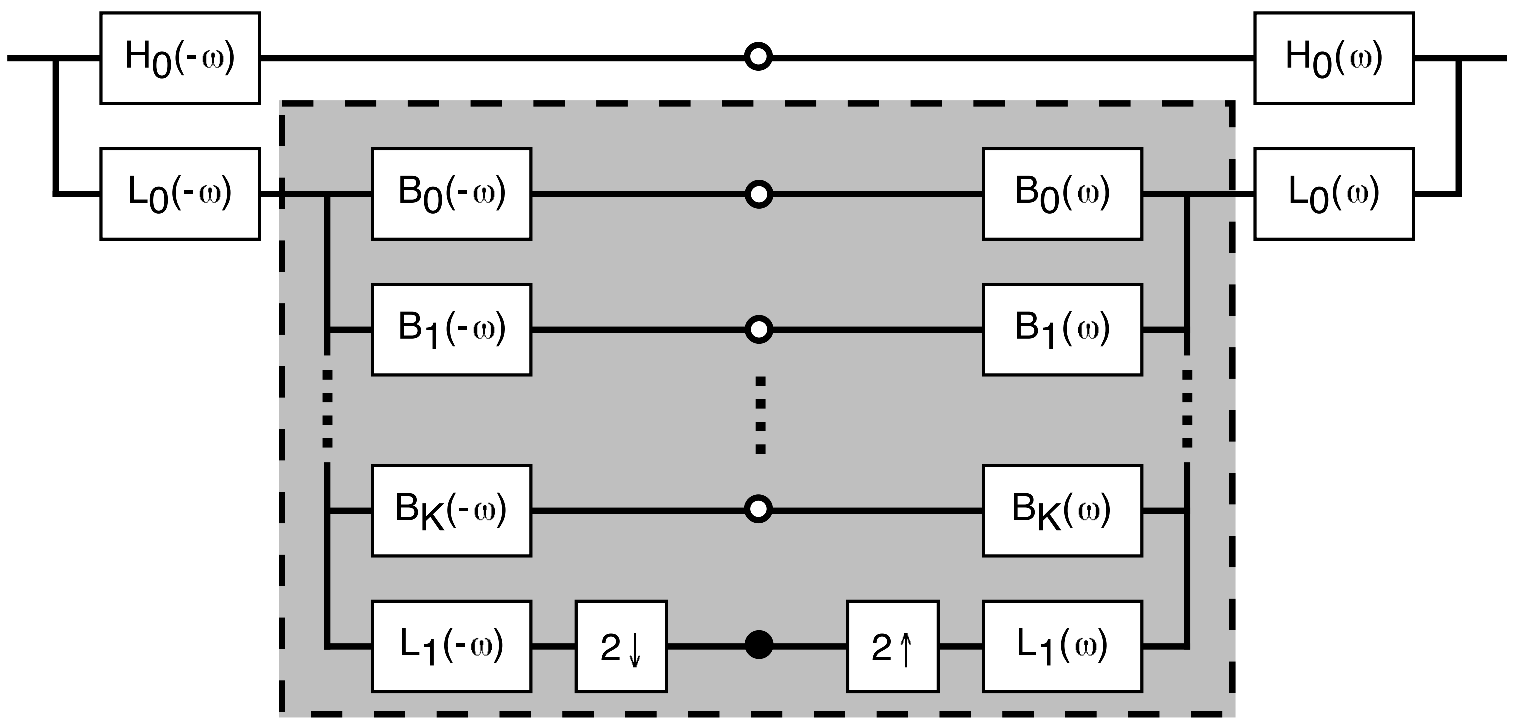
- Polar-separable:

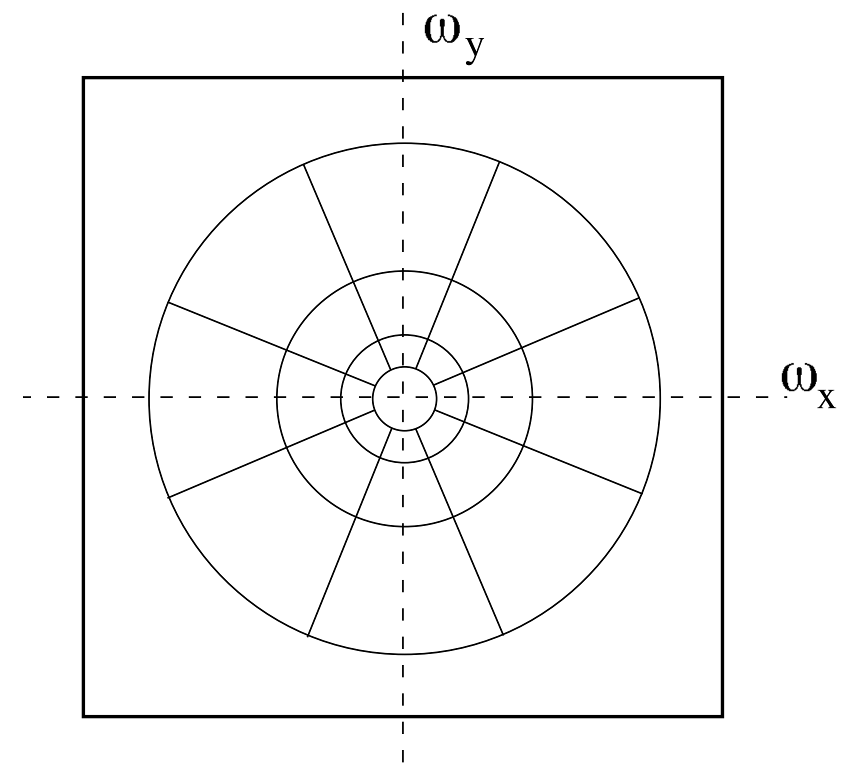
$$B_k(r,\theta) = H(r)G_k(\theta), \quad k \in [0, K-1],$$

$$H(r) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2\left(\frac{2r}{\pi}\right)\right), & \frac{\pi}{4} < r < \frac{\pi}{2} \\ 1, & r \geq \frac{\pi}{2} \\ 0, & r \leq \frac{\pi}{4} \end{cases}$$

$$G_k(\theta) = \begin{cases} \alpha_K \left[\cos(\theta - \frac{\pi k}{K})\right]^{K-1}, & |\theta - \frac{\pi k}{K}| < \frac{\pi}{2} \\ 0, & \text{otherwise}, \end{cases}$$

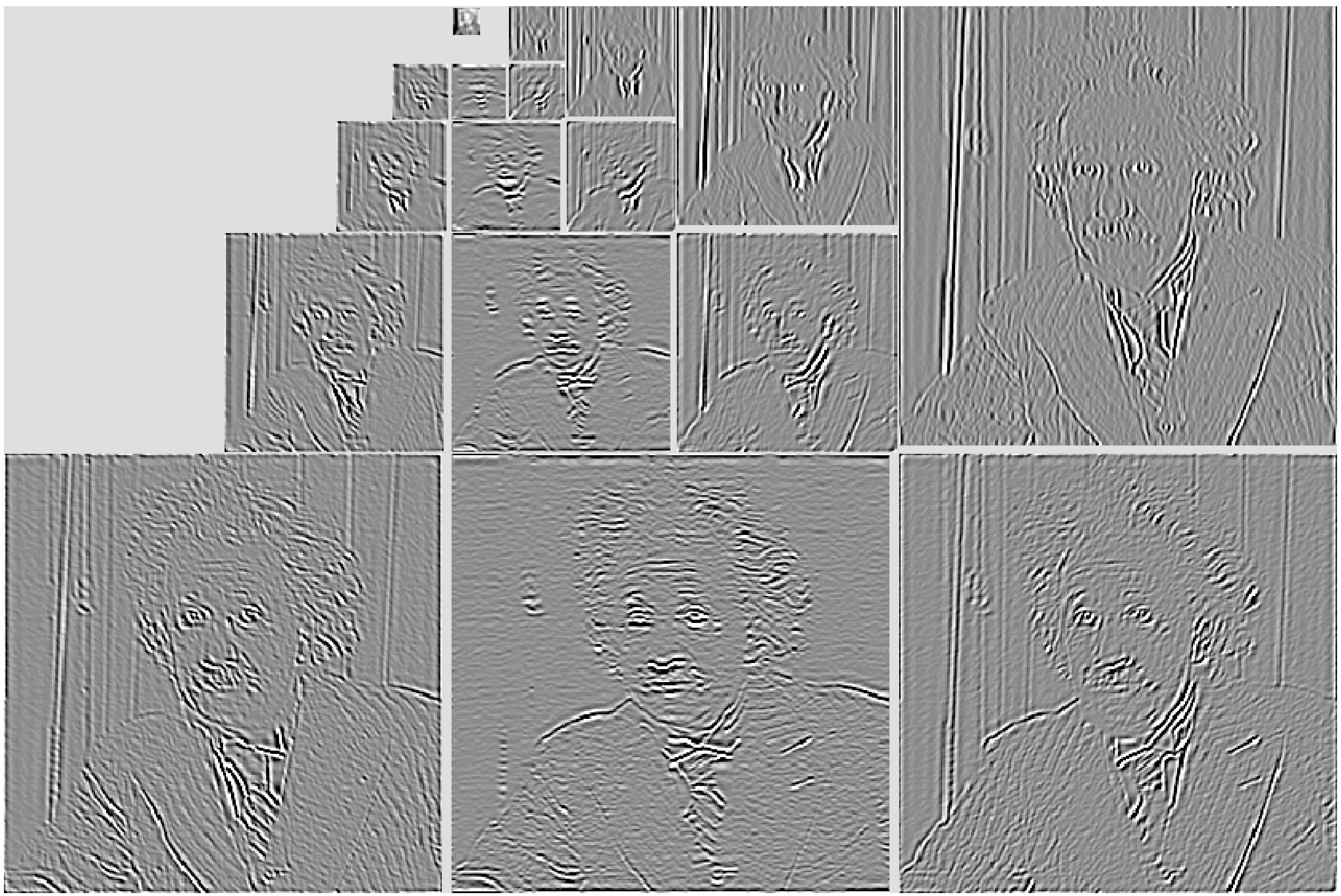
Recursive computation





Example decomposition

(3rd derivatives)



Bayes denoising

- Additive Gaussian noise:

$$y = x + w$$

$$P(y|x) \propto \exp[-(y-x)^2/2\sigma_w^2]$$

- Bayes' least squares solution is conditional mean:

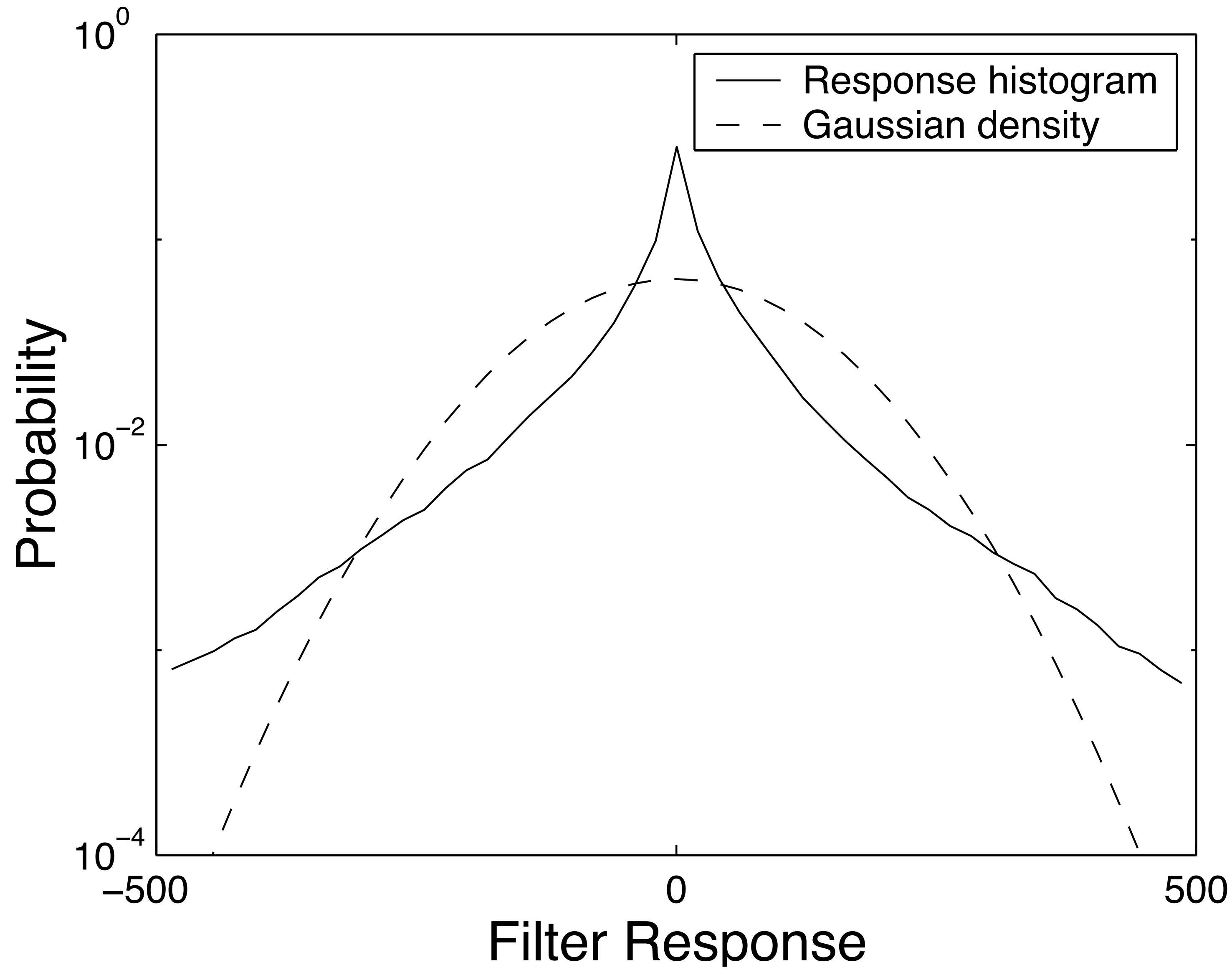
$$\begin{aligned}\hat{x}(y) &= \mathbb{E}(x|y) \\ &= \int dx \mathcal{P}(y|x)\mathcal{P}(x)x/\mathcal{P}(y)\end{aligned}$$

Denoising: classical

- Assume signal is Gaussian
- Then estimator is linear:

$$\mathbb{E}(\vec{x}|\vec{y}) = C_x(C_x + C_w)^{-1}\vec{y}$$

Wavelet statistics: Marginal



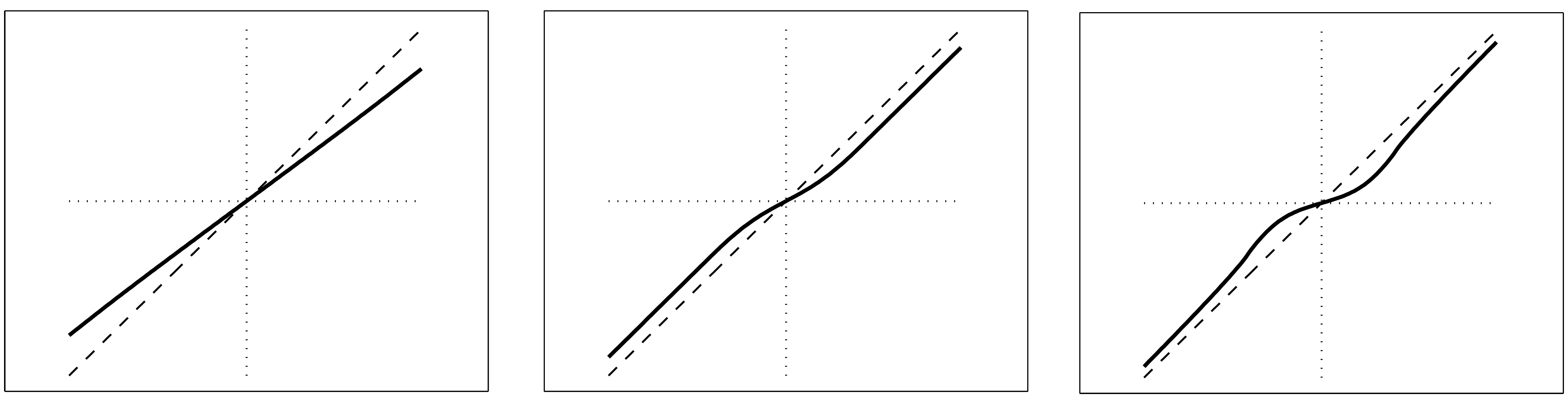
[Field ‘87; Mallat ‘89; Daugman ‘89; etc]

Denoising: shrinkage

- Assume marginal distribution [Mallat, '89]

$$P(x) \propto \exp -|x/s|^p$$

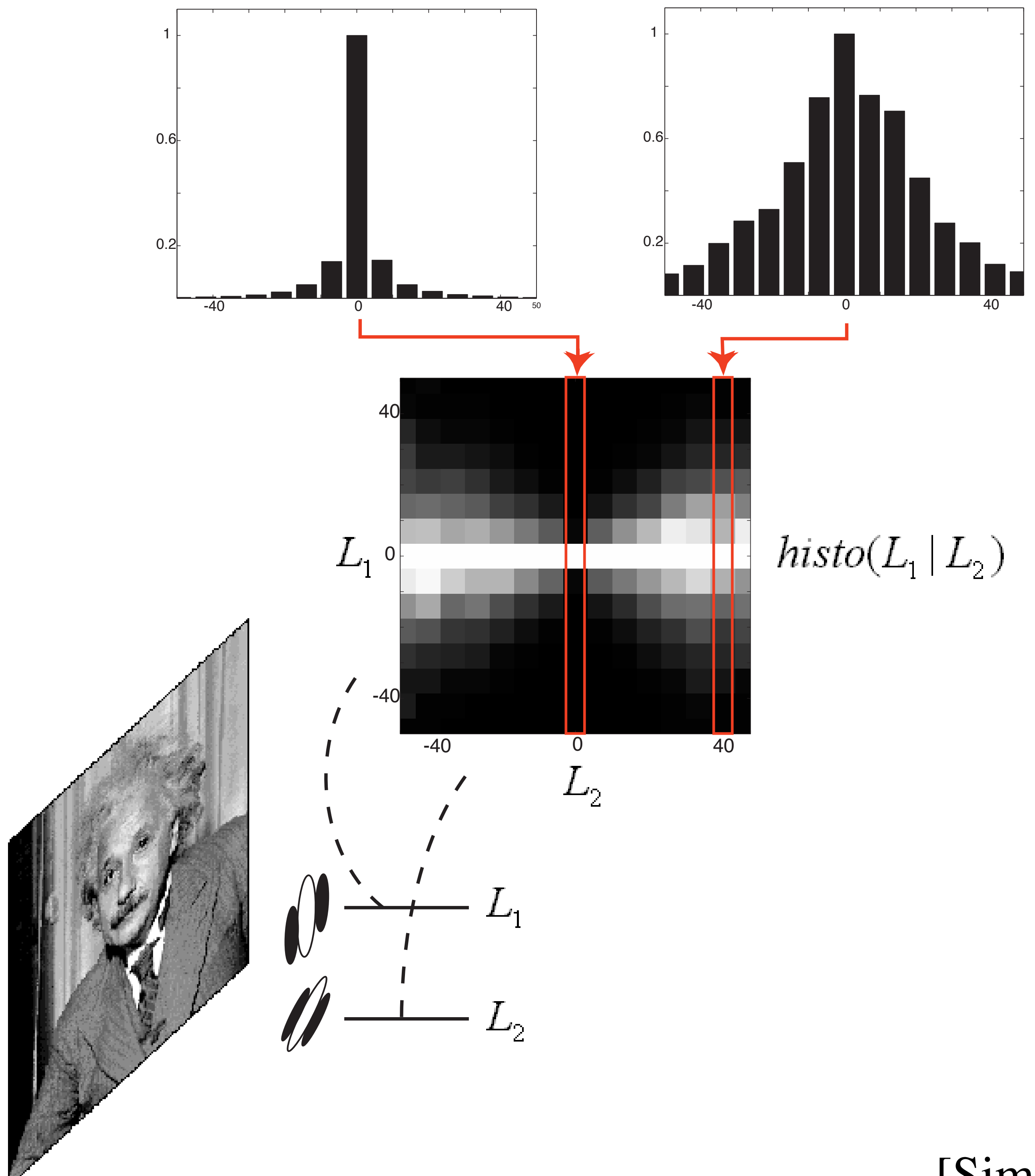
- Then estimator is generally nonlinear:



p = 2.0   *p = 1.0*   *p = 0.5*

[Simoncelli & Adelson, '96]

Statistics: Joint

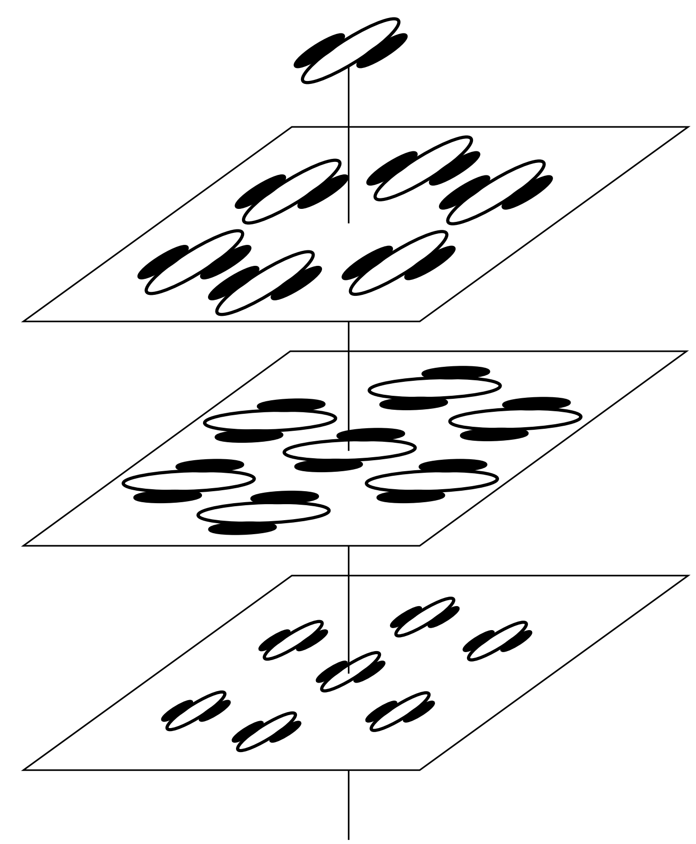


[Simoncelli, '97]

GSM model

Model generalized neighborhood of coefficients as a Gaussian Scale Mixture (GSM) [Andrews & Mallows '74]:

$\vec{x} = \sqrt{z}\,\vec{u}$, where



- z and $\vec{u}$ are independent

- $\vec{x}|z$ is Gaussian, with covariance zC_u

- marginals are always leptokurtotic

[Wainwright & Simoncelli, '99]

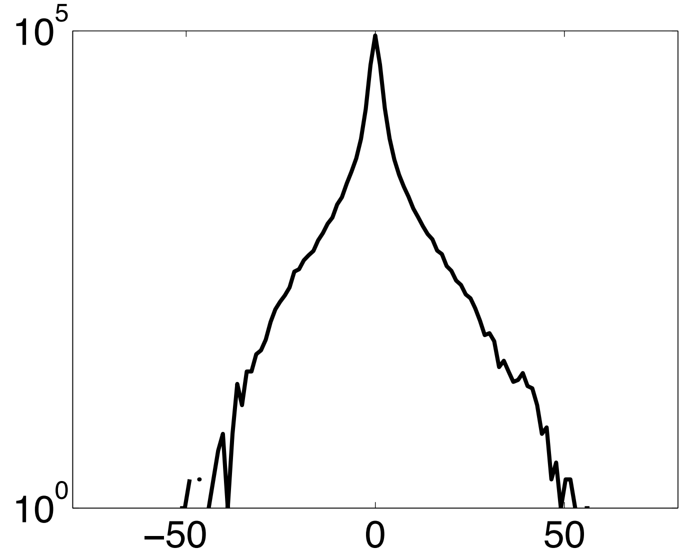
GSM - prior on z

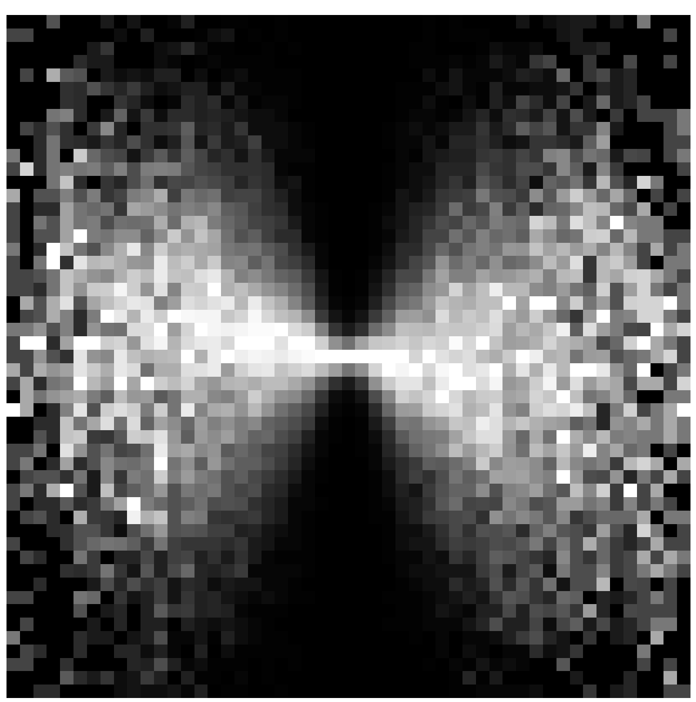
- Empirically, z is approximately lognormal
- Alternatively, can use Jeffrey's noninformative prior:

$$P(z) \propto 1/z$$

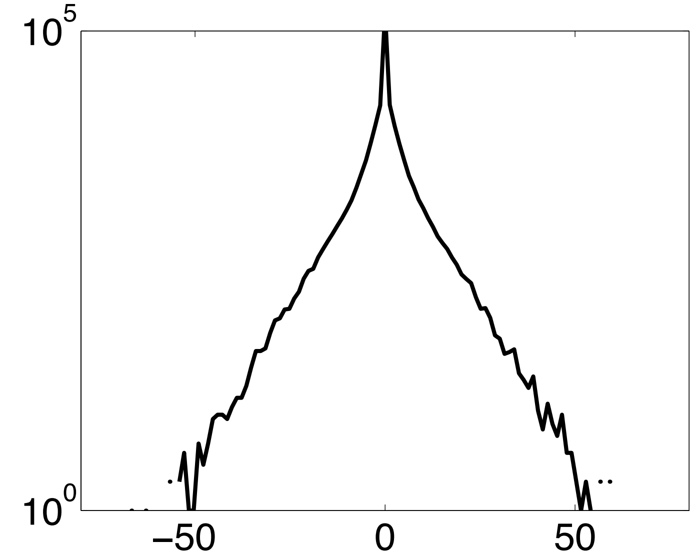
Simulation

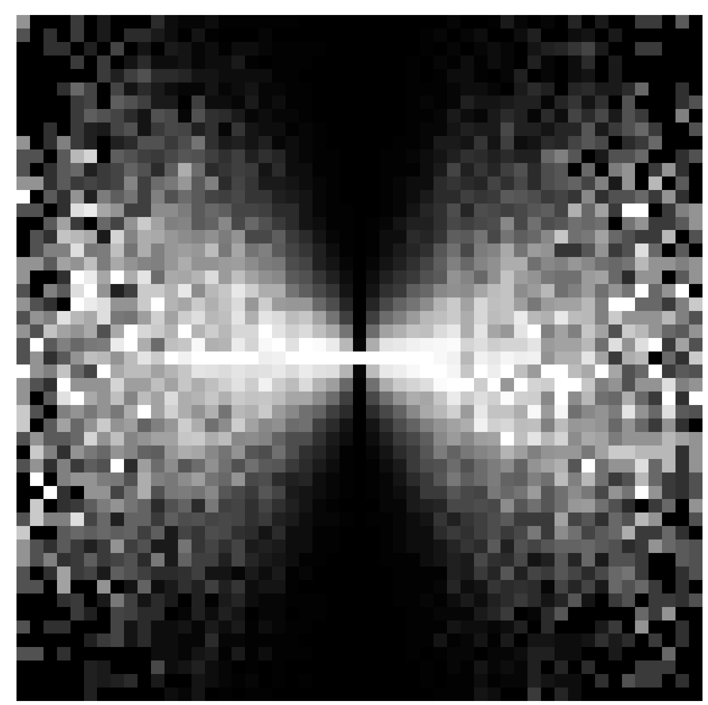
Image data





GSM simulation





Denoising: Joint

$$\mathbb{E}(x|\vec{y}) = \int dz\, \mathcal{P}(z|\vec{y})\ \mathbb{E}(x|\vec{y}, z)$$
$$= \int dz\, \mathcal{P}(z|\vec{y})\ \left[zC_u(zC_u + C_w)^{-1}\vec{y}\right]_{\mathrm{ctr}}$$

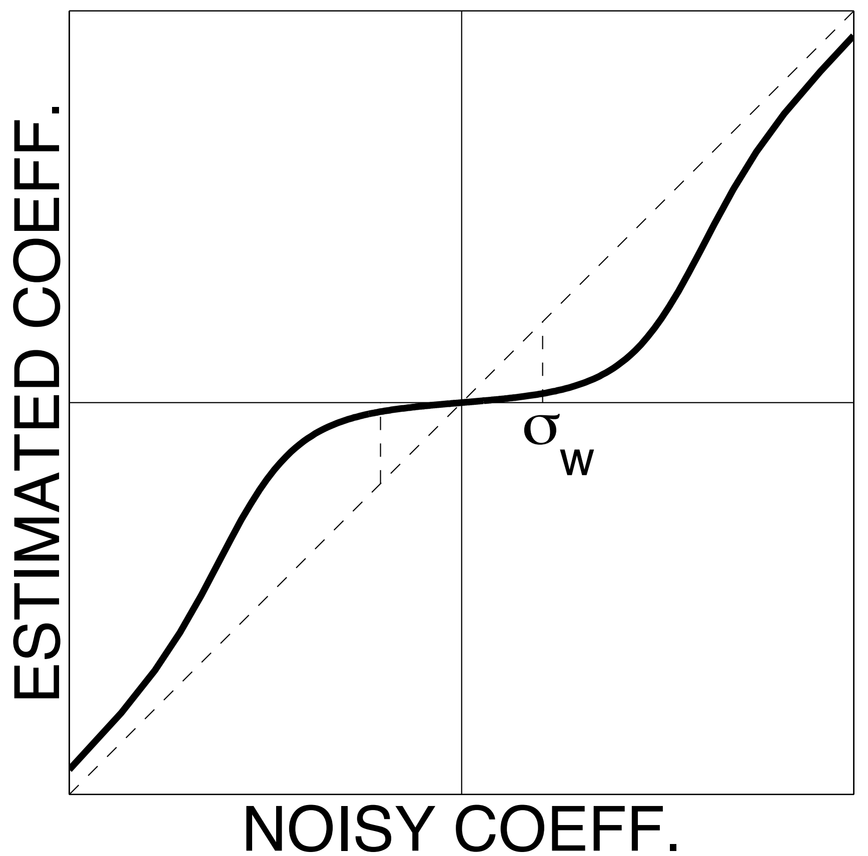
where

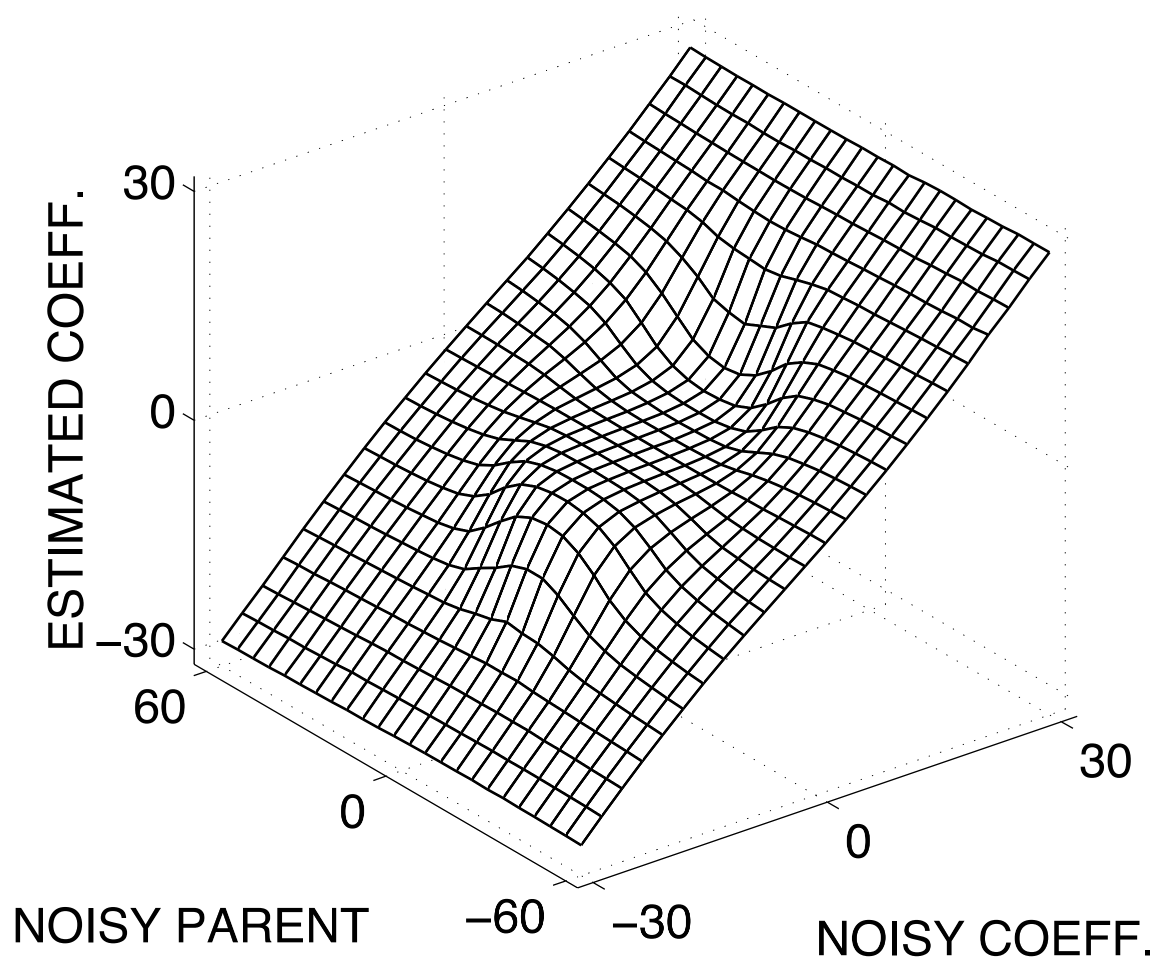
$$\mathcal{P}(z|\vec{y}) = \frac{\mathcal{P}(\vec{y}|z)\ \mathcal{P}(z)}{\mathcal{P}\vec{y}}, \quad \mathcal{P}(\vec{y}|z) = \frac{\exp(-\vec{y}^T(zC_u + C_w)^{-1}\vec{y}/2)}{\sqrt{(2\pi)^N |zC_u + C_w|}}$$

Numerical computation of solution is reasonably efficient if one jointly diagonalizes C_u and C_w ...

[Portilla, Strela, Wainwright, Simoncelli, '03]

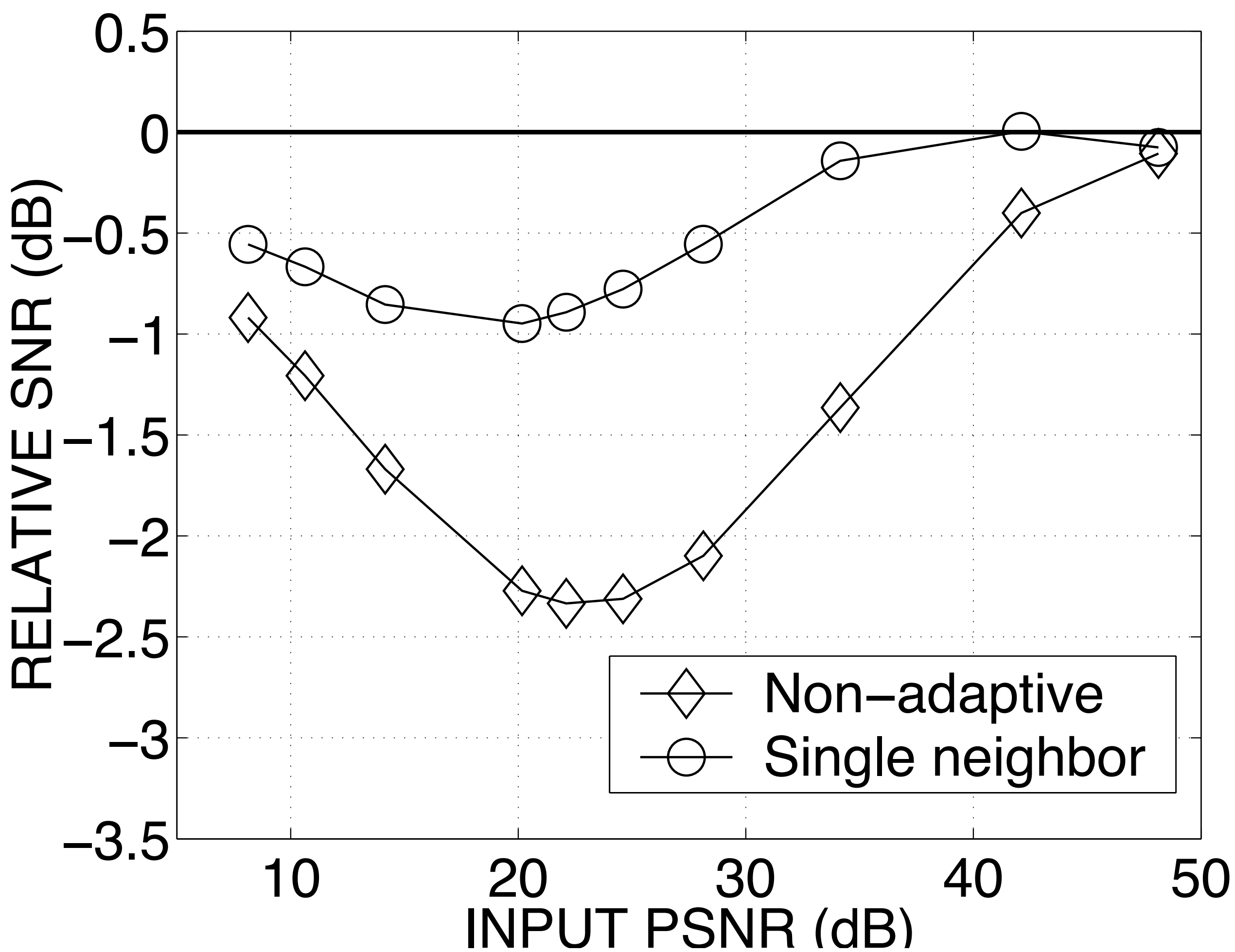
Example estimators

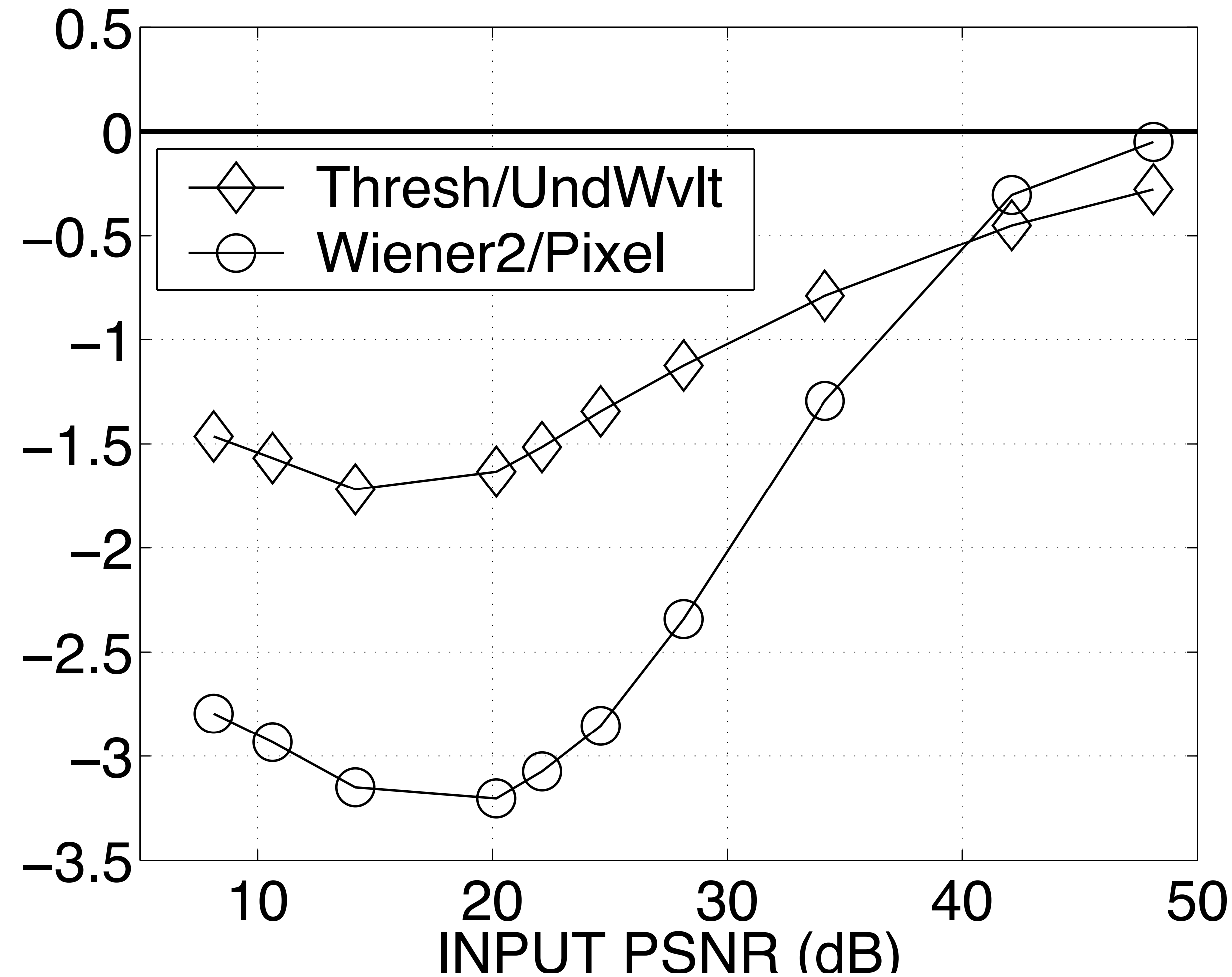




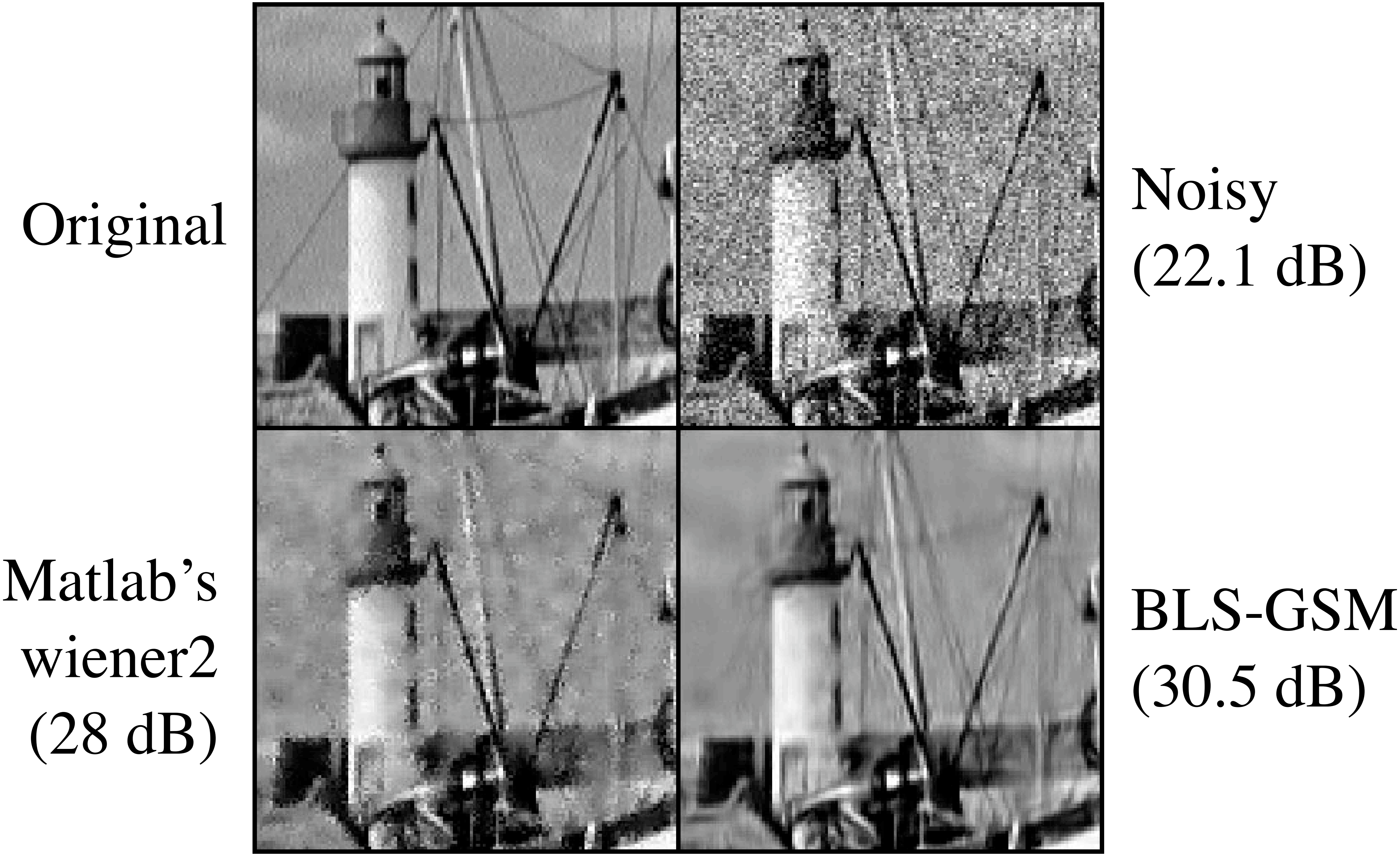
Estimators for the scalar and single-neighbor cases

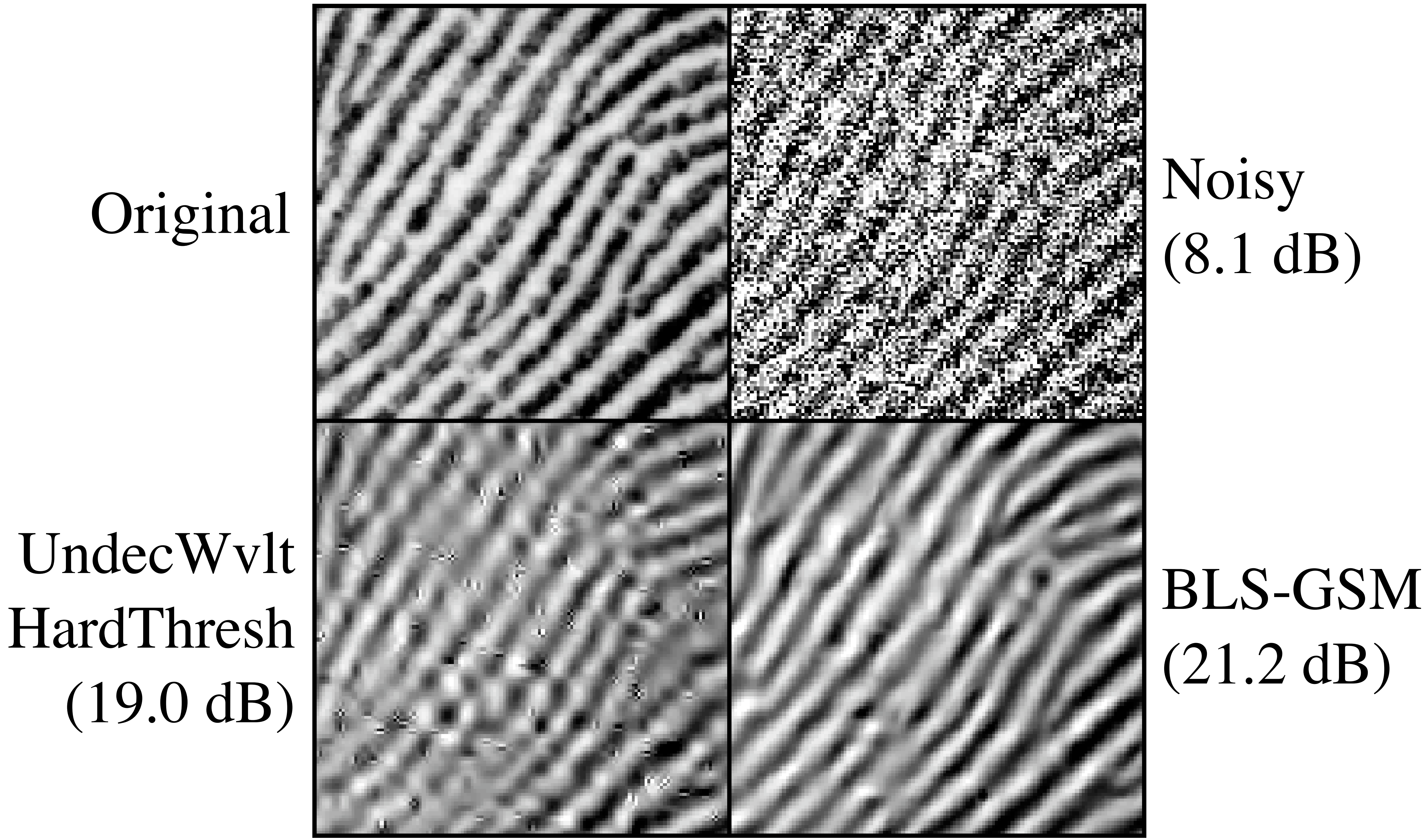
Comparison to other methods





Results averaged over 3 images


Original
Noisy
(22.1 dB)
Matlab's
wiener2
(28 dB)
BLS-GSM
(30.5 dB)


Original
Noisy
(8.1 dB)
UndecWvlt
HardThresh
(19.0 dB)
BLS-GSM
(21.2 dB)

Real sensor noise

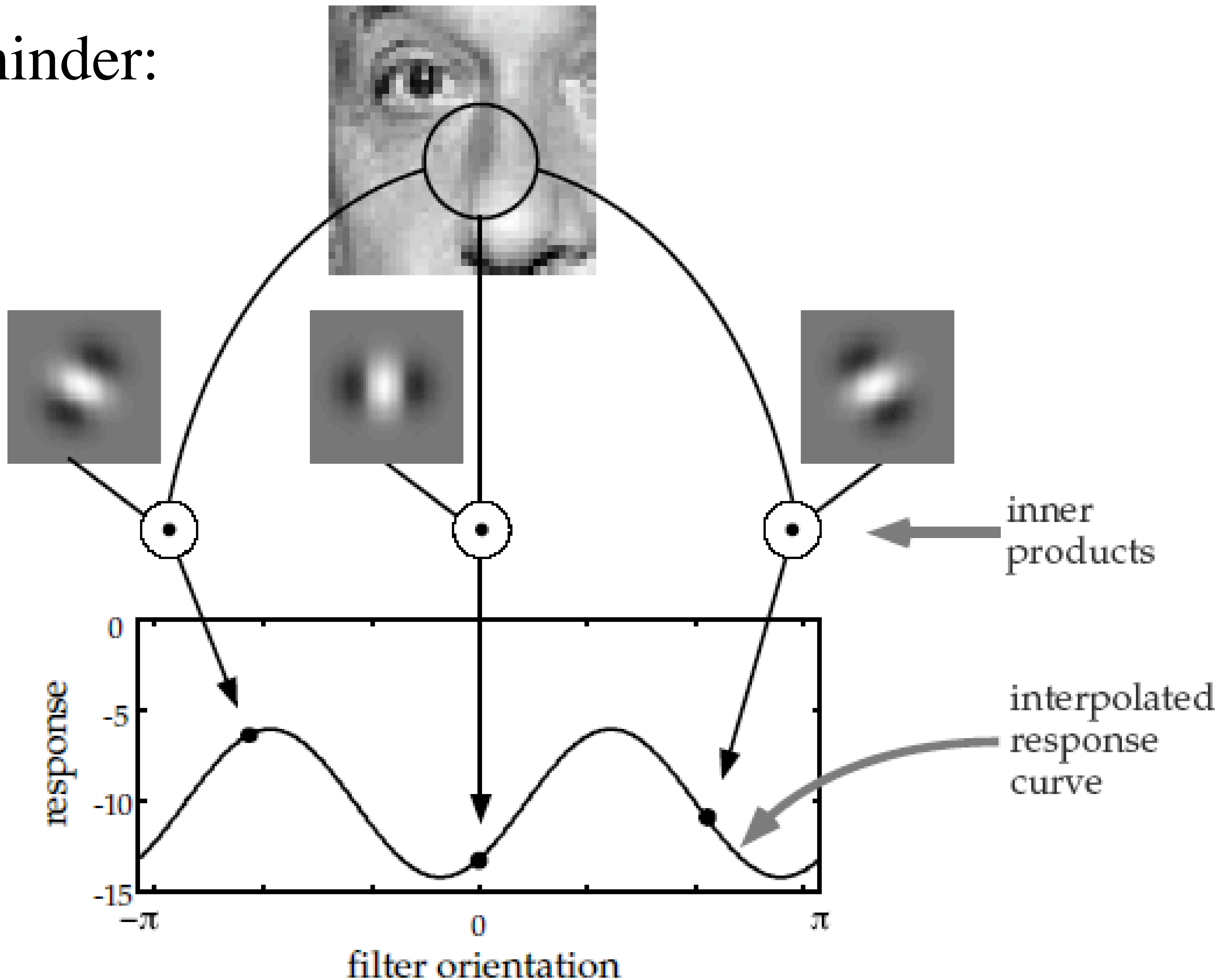
400 ISO

denoised

GSM summary

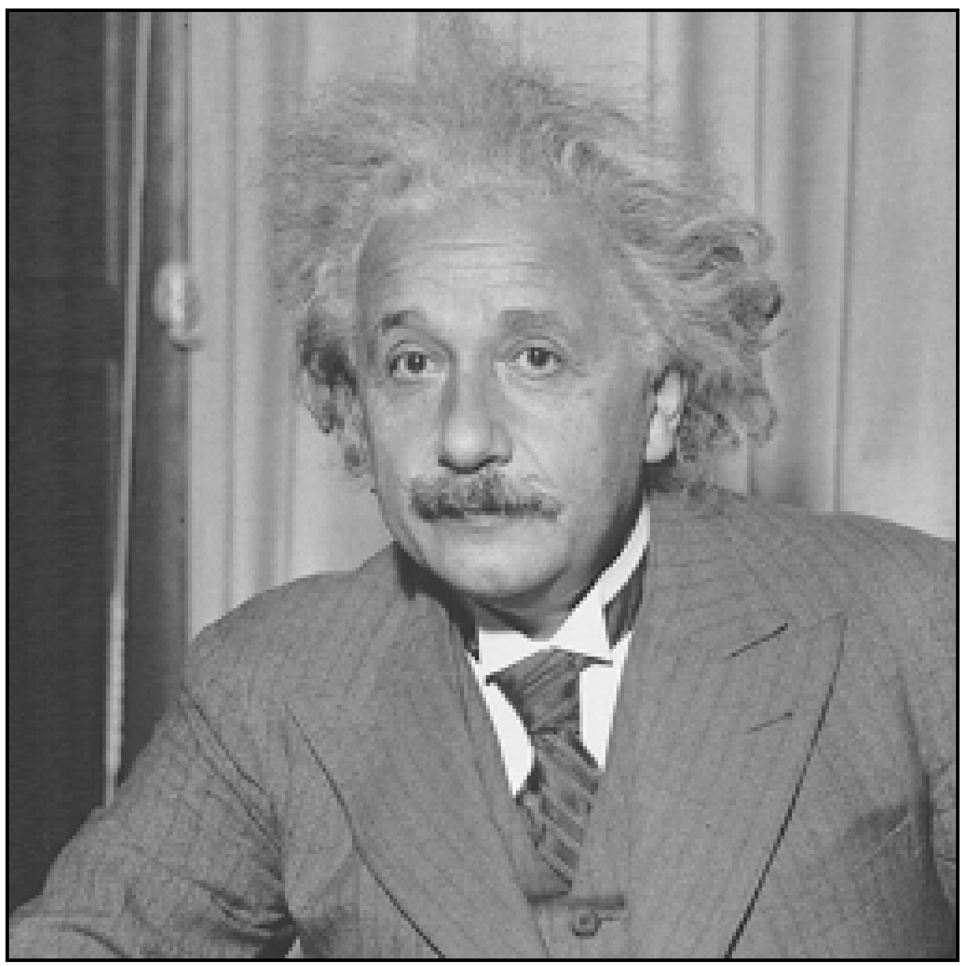
- GSM captures local variance
- Underlying Gaussian leads to simple computation
- Excellent denoising results
- What's missing?
  - Joint model of z variables [Wainwright etal '99; Romberg etal '99; Hyvarinen/Hoyer '02; Karklin/Lewicki '02; etc.]
  - Explicit geometry...

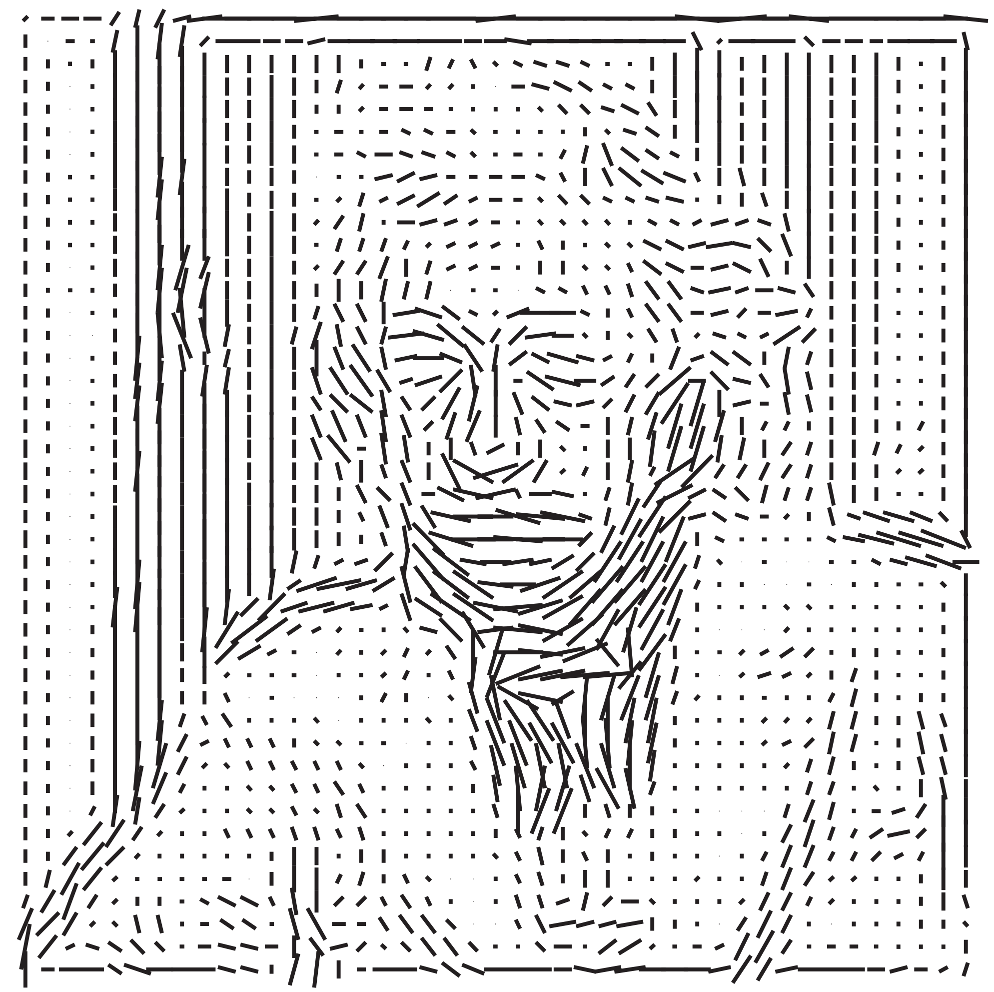
Reminder:



inner products
interpolated response curve
0
-5
-10
-15
response
−π
0
π
filter orientation

Local orientation

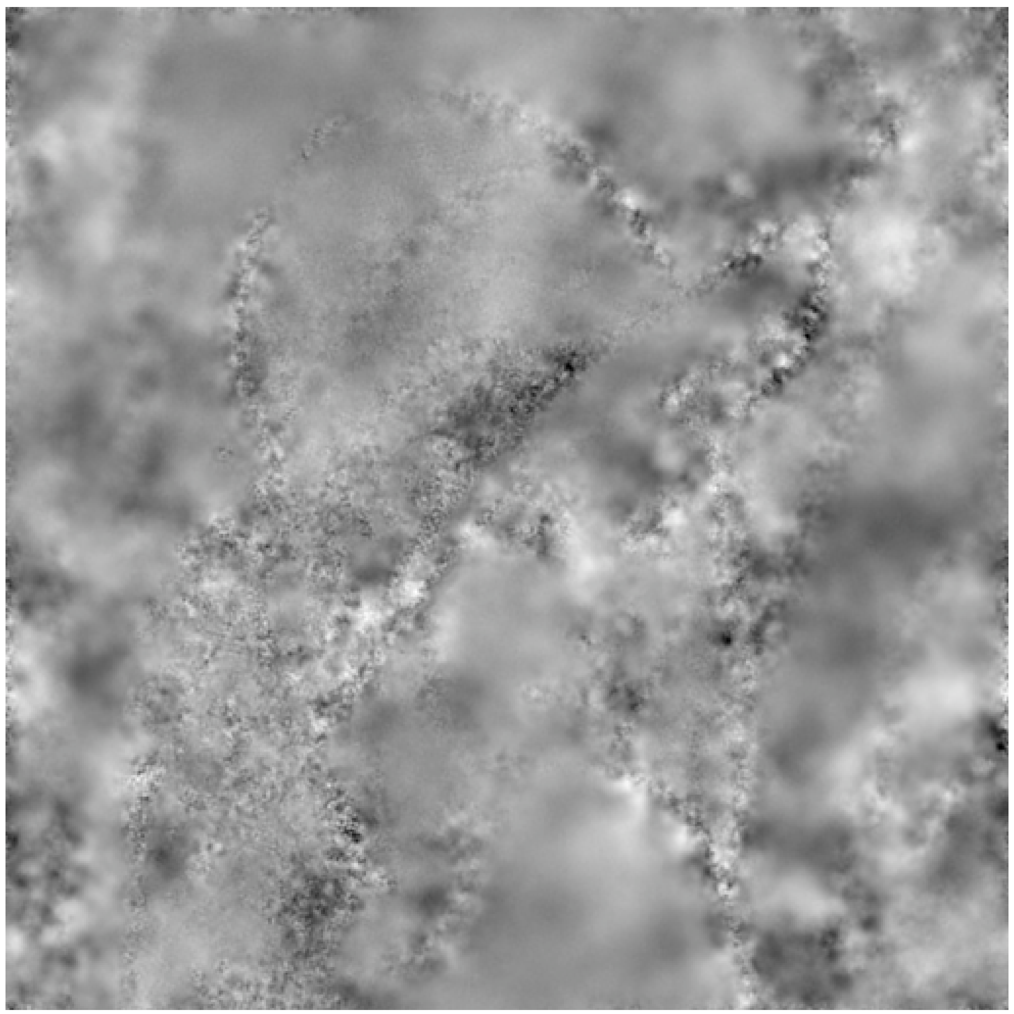




Lines normal to gradient, length proportional to magnitude

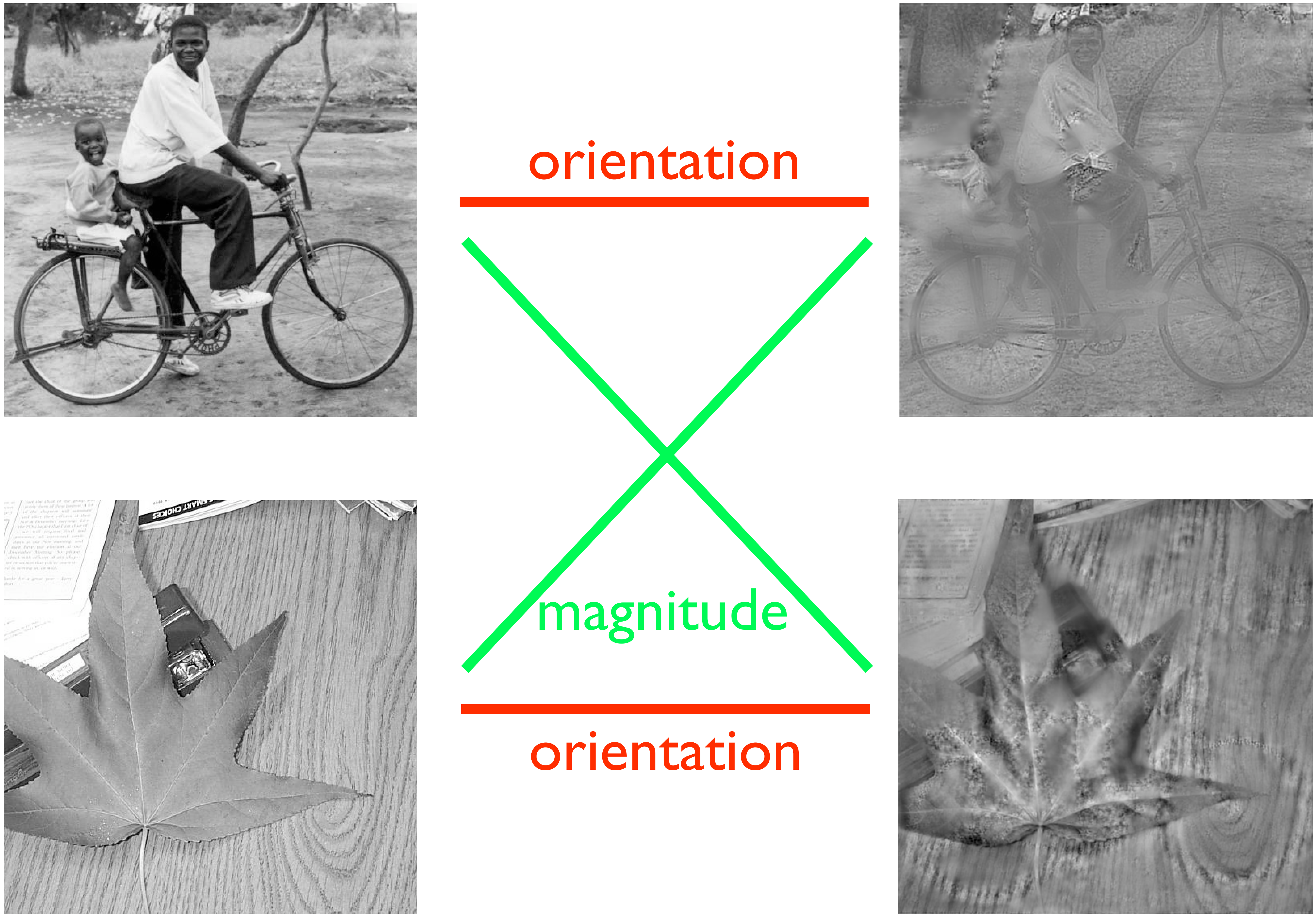
Importance of local orientation

Randomized orientation



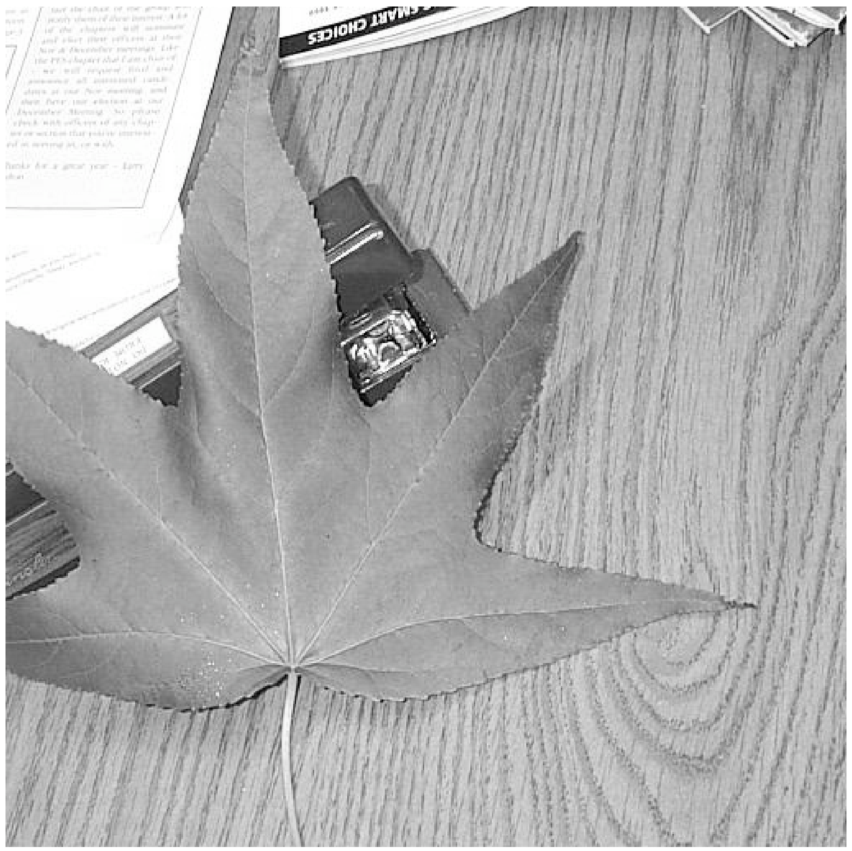
Randomized magnitude

Two-band, 6-level steerable pyramid


orientation
magnitude
orientation

Reconstruction from orientation

Original

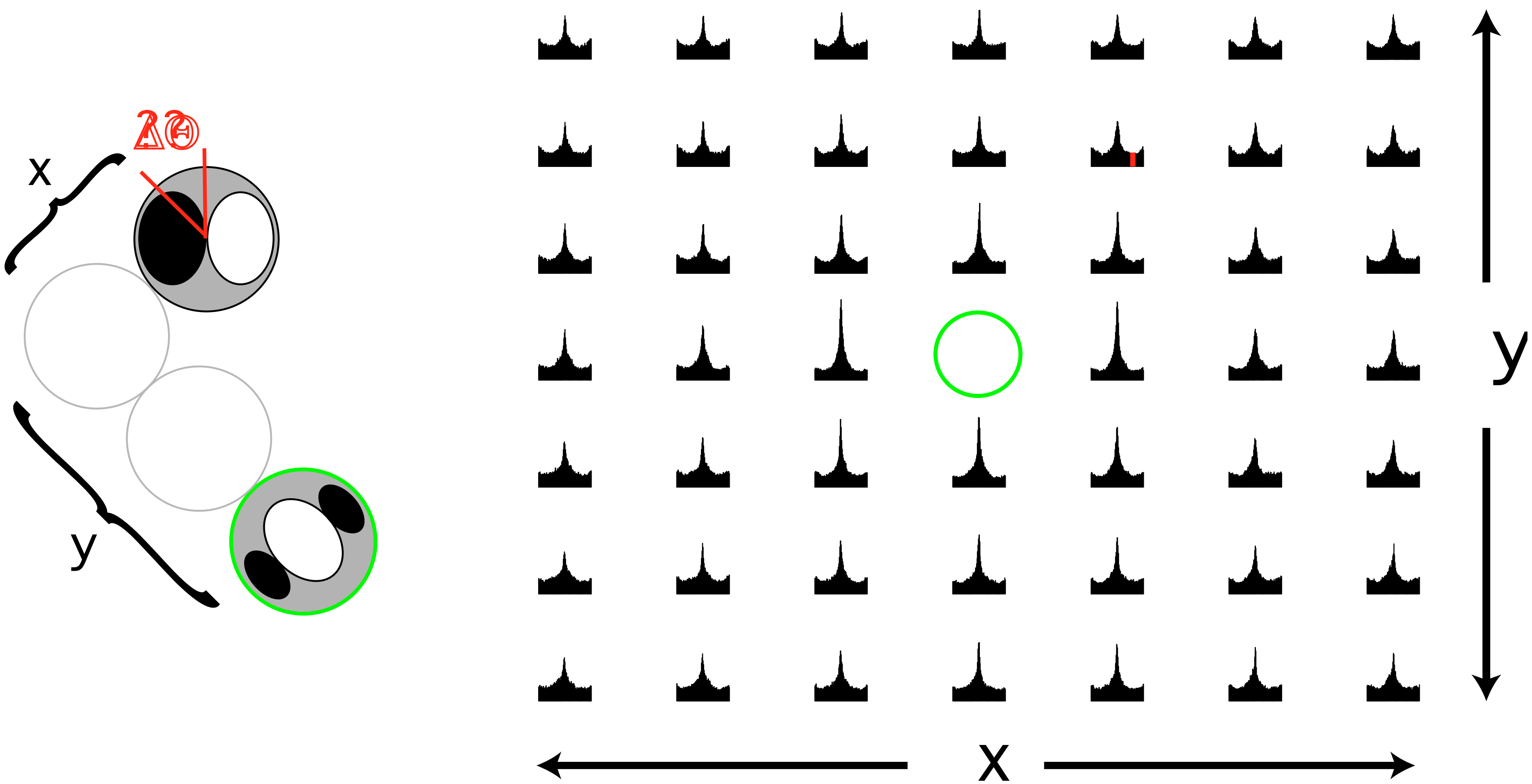


Quantized to 2 bits

- Converges: projections onto convex sets
- Resilient to quantization
- Highly redundant, across both spatial position and scale

[with David Hammond]

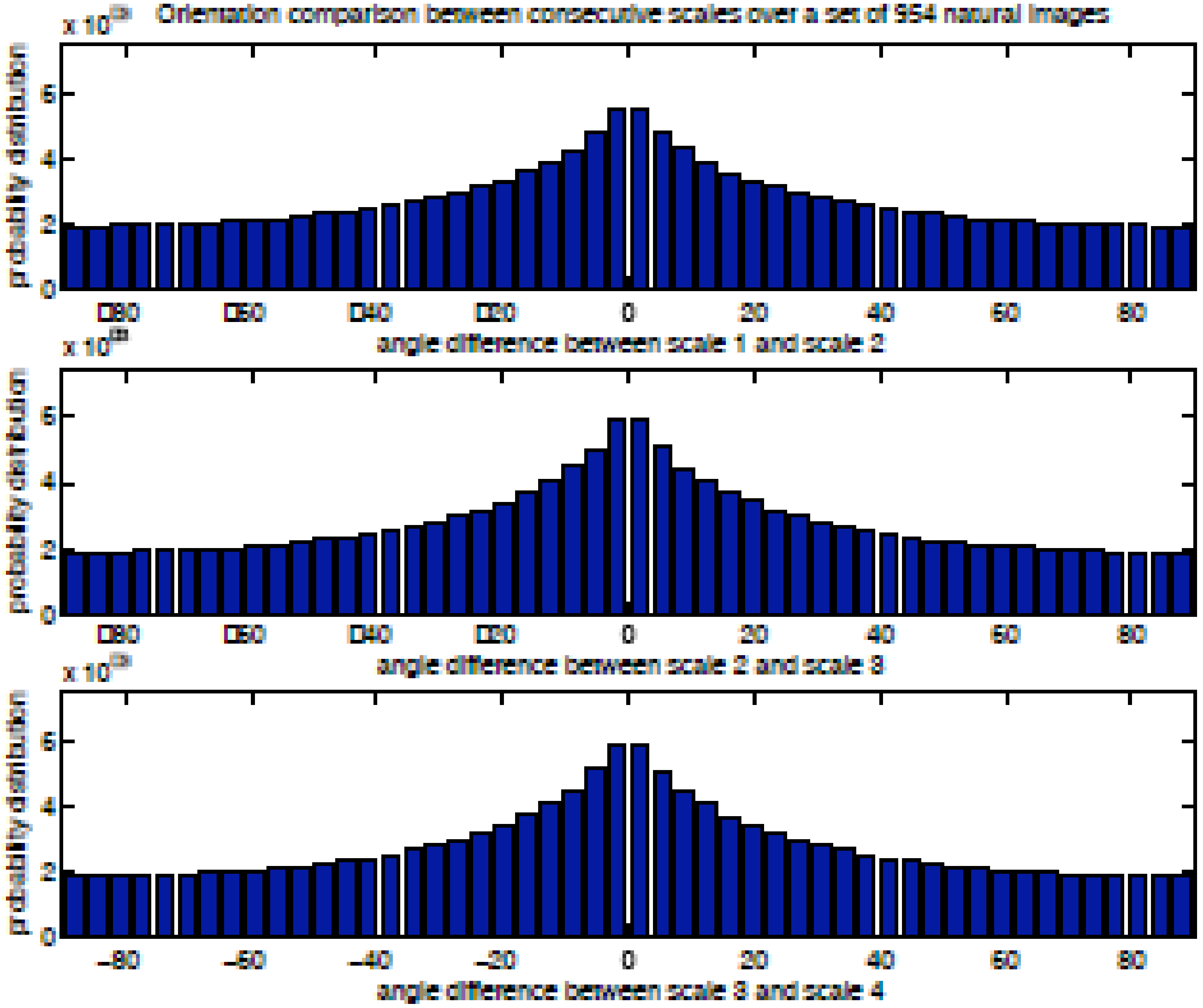
Spatial redundancy



- Relative orientation histograms, at different locations
- See also: Geisler, Elder

[with Patrik Hoyer & Shani Offen]

Scale redundancy



[with Clementine Marcovici]

Cast

- Local GSM model: Martin Wainwright, Javier Portilla, Vasily Strela
- Denoising: Javier Portilla, Martin Wainwright, Vasily Strela
- GSM tree model: Martin Wainwright, Alan Willsky
- Orientation: David Hammond, Clementine Marcovici
- Compression: Robert Buccigrossi
- Texture representation/synthesis: Javier Portilla
- Local phase: Zhou Wang