A photograph of a desk with a calculator, a pen, and a mouse. The calculator is a TI-84 Plus CE, black with a silver face, positioned on the left. A blue and white pen lies horizontally across the top of the papers. A silver and black mouse is on the right. The background is a wooden desk surface.

Bivariate Shrinkage  
Functions for Wavelet-Based  
Denoising Exploiting  
Interscale Dependency

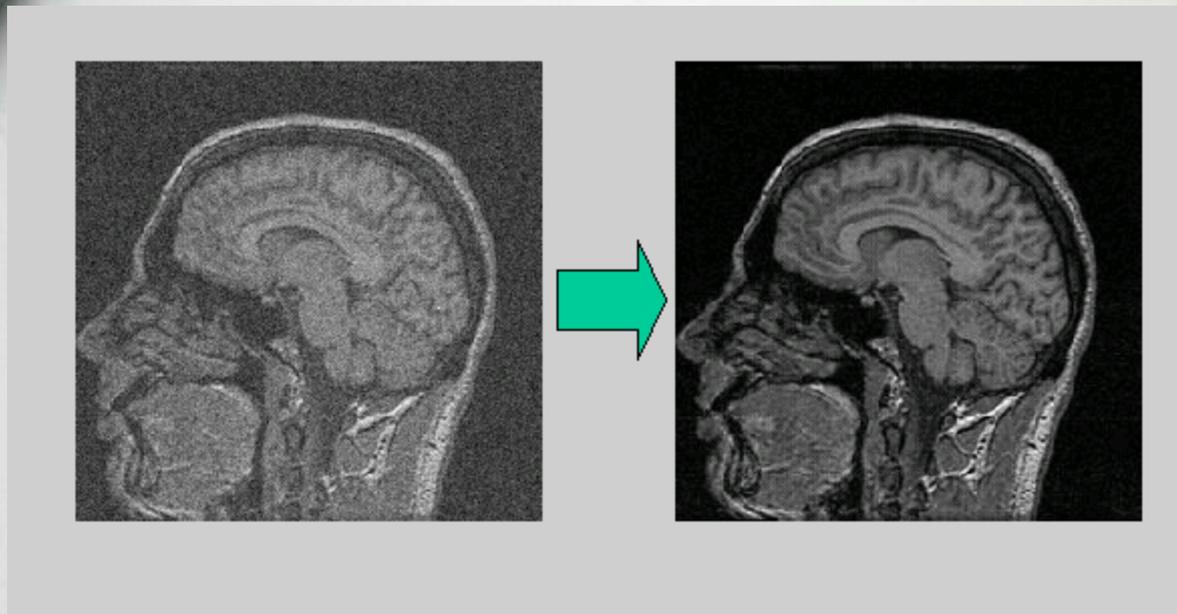
Student: Andreea Teodorescu  
Coodonator: Corina Nafoarnita

# Denoising

## What does it mean?

- Denoising (noise reduction) is to remove noise as much as possible while preserving useful information. This is the first step in many applications
- Applications:
  - Medical signal/image analysis (ECG, CT, MRI etc.)
  - Data mining
  - Radio astronomy image analysis

The left is a noisy MR brain image. Before the analysis, we would like to clean it up



- Multiscale decompositions have shown significant advantages in the representation of signals, and they are used extensively in image compression, segmentation and denoising
- The denoising of a natural image corrupted by Gaussian noise is a classic problem in signal processing.

## Denoising algorithm that use the wavelet transform - Steps

- 1) Calculate the wavelet transform of the noisy signal.
- 2) Modify the noisy wavelet coefficients according to some rule.
- 3) Compute the inverse transform using the modified coefficients.

- Denoising by thresholding in the wavelet domain has been developed principally by Donoho who introduced RiskShrink with the minimax threshold, VisuShrink with the universal threshold, and discussed both hard and soft thresholds in a general context that included ideal denoising in both the wavelet and Fourier domains.

# Dave Donoho

- David Leigh Donoho, born on March 5, 1957 in Los Angeles, is a professor of mathematics and statistics at Stanford University
- His work includes multiscale geometric analysis, developments of wavelets for denoising and compressed sensing



- The main idea is to report coefficients to the threshold value  $T$
- Statistical approaches: The basic idea is to model wavelet transform coefficients with prior probability distributions.
- MAP estimator - requires a priori knowledge about the distribution of wavelet coefficients.

# BAYESIAN DENOISING

- In this section, the denoising of an image corrupted by white Gaussian noise will be considered:

$$g = x + n$$

- In the wavelet domain, if we use an orthogonal wavelet transform, the problem can be formulated as

$$y = w + n$$

# Marginal Model

- The classical MAP estimator

$$\hat{w}(y) = \arg \max_w p_{wy}(wy)$$

- Using Bayes rule we can write this estimation

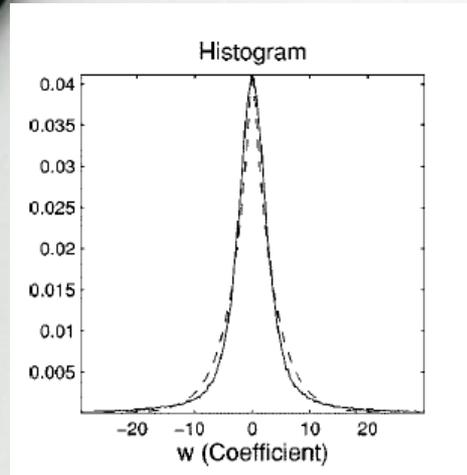
$$\hat{w}(y) = \arg \max_w [p_{y|w}(y|w) \cdot p_w(w)]$$

$$= \arg \max_w [p_n(y-w) \cdot p_w(w)]$$

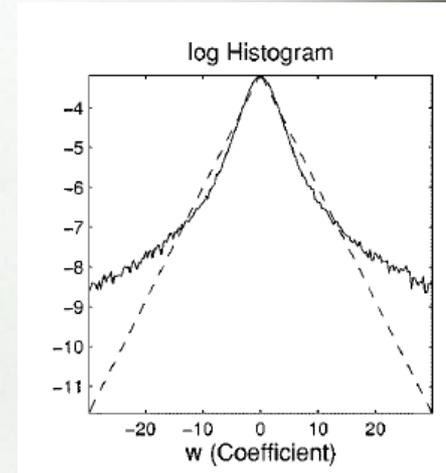
# Problems

- It can be difficult to estimate the parameters for a specific image, especially from noisy data
- the estimators for these models may not have simple closed form solution and can be difficult to obtain. The solution for these problems usually requires numerical techniques

# Examples



Illustrates the histogram of the wavelet coefficients computed from several natural images.



Same data is illustrated in log domain in order to emphasize the tail difference.

# Developing MAP estimator

The formula for MAP estimator is also equivalent to:

$$\hat{w}(y) = \arg \max_w [\log(p_n(y - w)) + \log(p_w(w))]$$

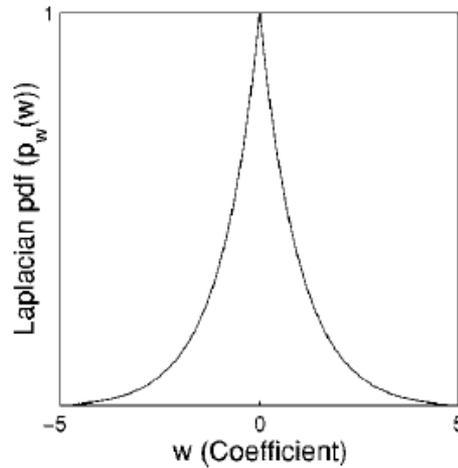
If  $p_w$  is Laplacian

then

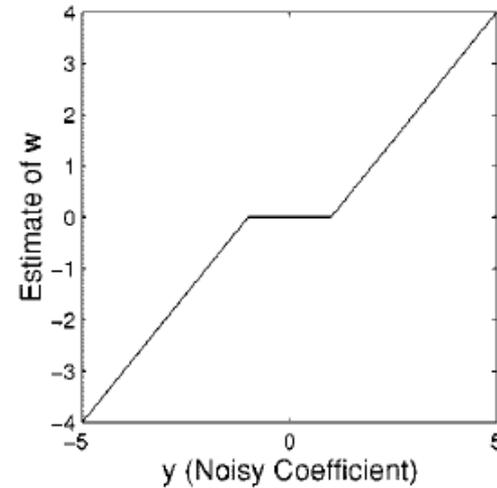
$$\hat{w}(y) = \text{sign}(y) \left( \left| y - \frac{\sqrt{2} \sigma_n^2}{\sigma} \right|_+ \right)$$

The equation above is the classical soft shrinkage function.

# The Laplacian pdf and corresponding shrinkage function



Laplacian pdf



Corresponding shrinkage function

# Bivariate Models

There are strong dependencies between neighbor coefficients such as between a coefficient, its parent (adjacent coarser scale locations), and their siblings (adjacent spatial locations)

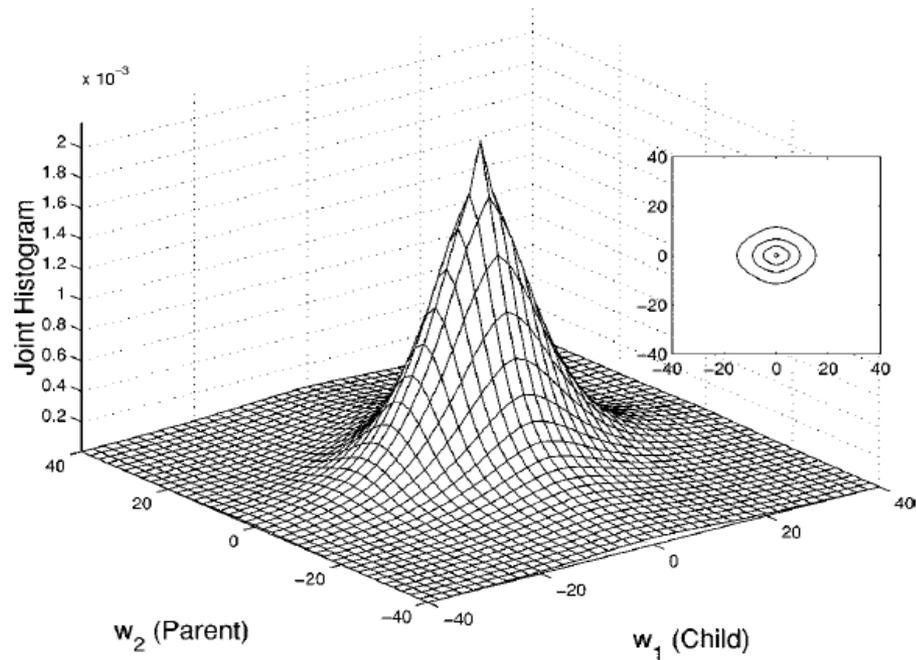
- By modifying the Bayesian estimator to take into account the statistical dependency between a coefficient and its parent, the equation can be written as follows:

$$\hat{w}(y) = \arg \max_w [p_{y|w}(y | w) \cdot p_w(w)]$$

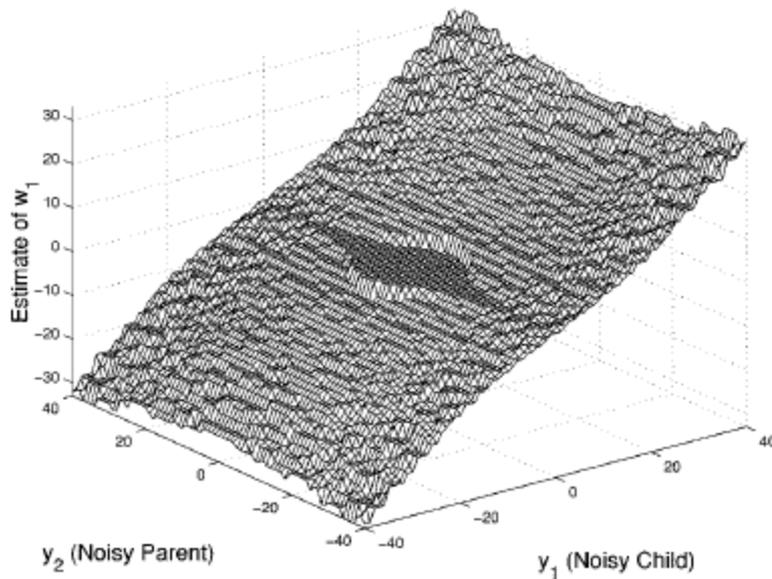
$$\hat{w}(y) = \arg \max_w [p_n(y - w) \cdot p_w(w)]$$

- The same problem as in marginal case appears. What kind of joint pdf models the wavelet coefficients? The joint empirical coefficient-parent histogram can be used to observe  $p_w(w)$
- For this purpose, we used 200 512 x 512 images from the Corel image database in order to stabilize the corresponding statistic.
- We used Daubechies length-8 filter to compute the wavelet transform.

Empirical joint parent-child histogram of wavelet coefficients (computed from the Corel image database).



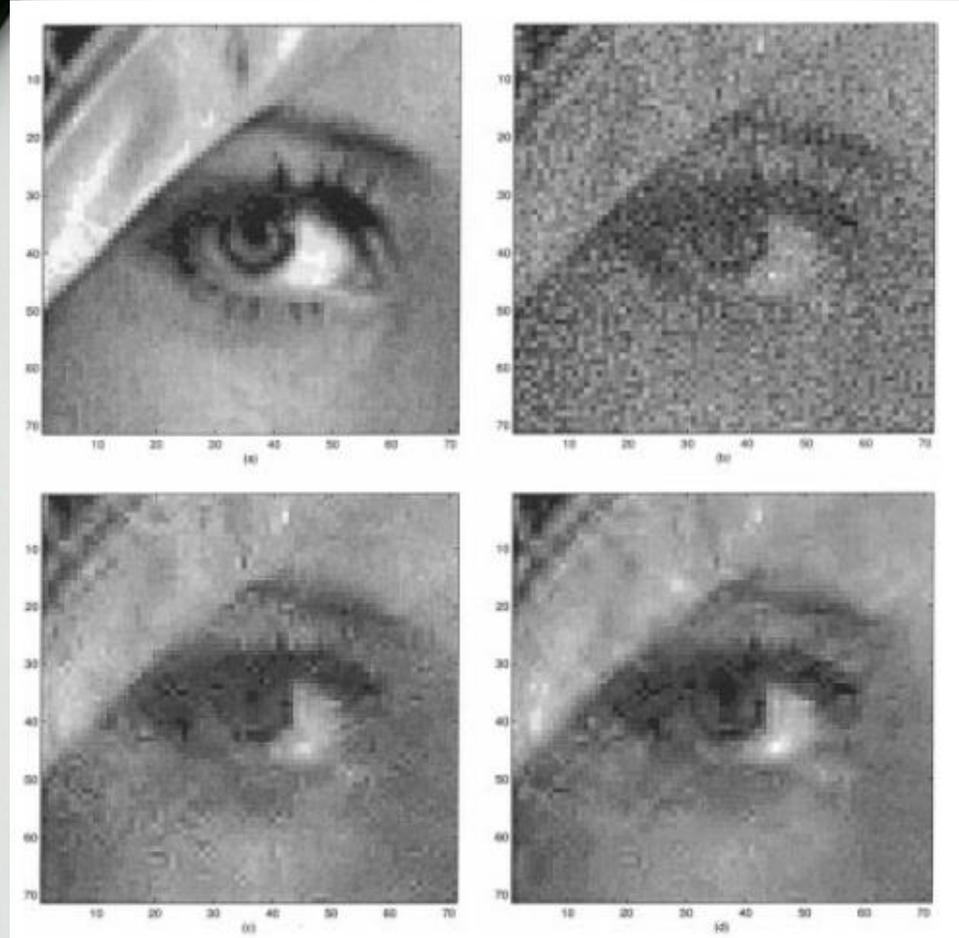
Joint shrinkage function  
derived numerically from  
the empirical joint  
parent-child histogram



## APPLICATION TO IMAGE DENOISING

- a critically sampled orthogonal discrete wavelet transform, with Daubechies length-8 filter, is used. We have compared our bivariate shrinkage function with the classical soft thresholding for image denoising.
- The  $512 \times 512$  Lena image is used for this purpose.
- Zero mean white Gaussian noise is added to the original image  $\sigma_n = 25.5$

1. Original image
2. Noisy image with PSNR = 20:02 dB
3. Denoised image using soft thresholding; PSNR = 27:73 dB
4. Denoised image using new bivariate shrinkage function given in (30); PSNR = 28:83 dB.



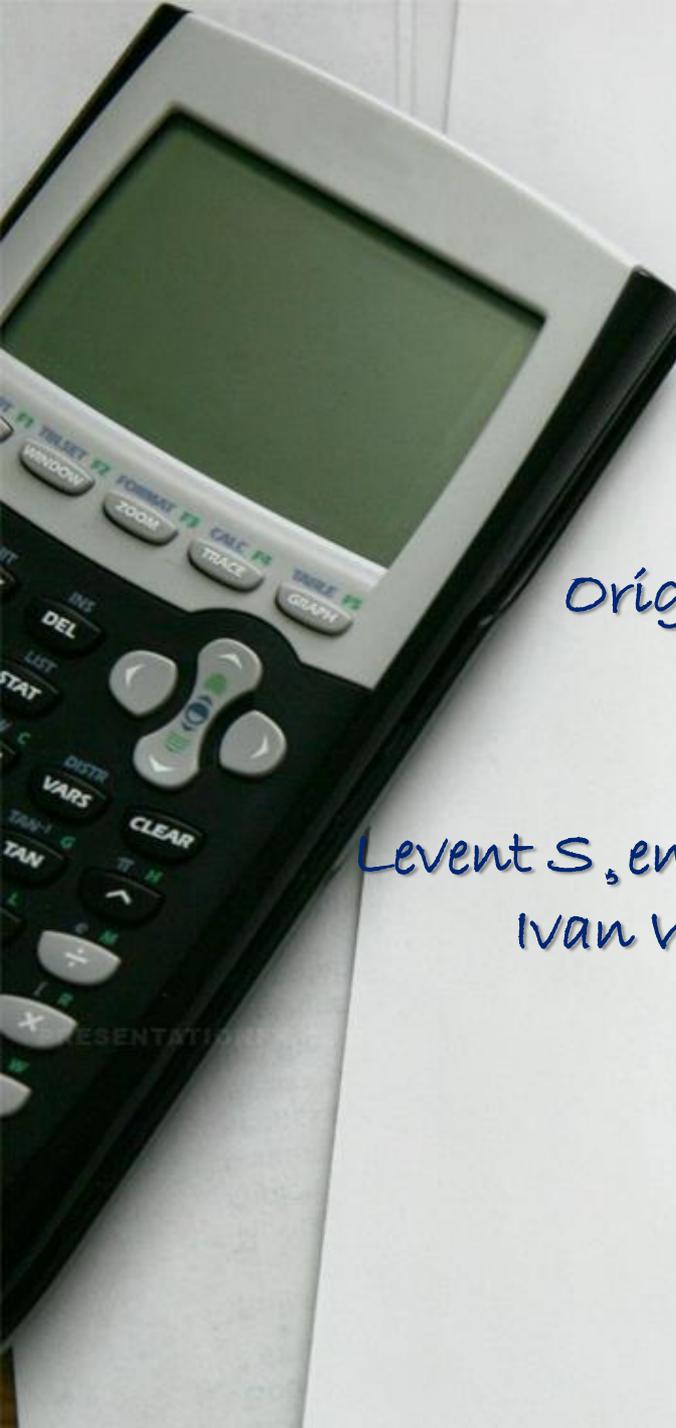
## Conclusions

- new bivariate distributions are proposed for wavelet coefficients of natural images in order to characterize the dependencies between a coefficient and its parent
- the corresponding bivariate shrinkage functions are derived from them using Bayesian estimation, in particular, the MAP estimator.



Thank you for your attention





Original authors of the article

Levent S. endur, Student Member, IEEE, and  
Ivan W. Selesnick, Member, IEEE

