Bayesian Estimation & Transforms To Denoise Astronomical Images

Albert Bijaoui
UNS/CNRS/OCA

BP 4229 06304 Nice Cedex 04
bijaoui@oca.eu

Outline

- Gaussian noise & Posterior expectation
  - The Miyasawa relation
- Mean variance reduction
  - Estimation of the corresponding index
  - Iterative k-sigma rejection & Voigt distribution
- Variance reduction & Transforms
- Karhunen-Loève transform with rejection
- Conclusion

Posterior Bayesian Expectation

- y = observation – x = true value
- *Posterior PDF=Observed PDF*

$$p(y) = \int_x p_x(x) q(y/x) dx$$

- $p_x(x)$: *prior PDF* – $q(y/x)$*:* conditional PDF
- *Posterior* conditional (Bayes's theorem):

$$p_{x/y}(x) = \frac{p_x(x) q(y/x)}{p(y)}$$

- Best quadratic x estimation :

$$\hat{x} = \int_x x \frac{p_x(x) q(y/x)}{p(y)} dx$$

The Miyasawa Relation (1961)

- Hypothesis: the noise is Gaussian

$$q(y/x) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{(y-x)^2}{2N}}$$

- We deduce

$$\hat{x} = y + \frac{1}{\sqrt{2\pi N}} \int_x (x-y) \frac{p_x(x) e^{-\frac{(y-x)^2}{2N}}}{p(y)} dx$$

- As:

$$p'(y) = \frac{1}{N} \int_x (x-y) p_x(x) q(y/x) dx$$

- The Miyasawa relation results:

$$\boxed{\hat{x} = y + N \log' p(y)}$$

- The estimation depends on the posterior PDF

Comparison with the MAP

- In the Gaussian case the MAP leads to:

$$\hat{x} = y + N \log' p_x(\hat{x})$$

- While the posterior expectation is:

$$\hat{x} = y + N \log' p(y)$$

- Equation versus functional relation
  - Uniqueness of the solution
  - Take into account the whole conditional PDF
- Case of a vector
  - Advantage for the MAP estimation

Case of a Gaussian signal

- The *prior* PDF is Gaussian with variance S
- The *posterior* PDF is Gaussian with variance $S+N$
- The Miyasawa relation leads to :

$$\hat{x} = y - \frac{N}{S+N} y = \frac{S}{S+N} y$$

- The Miyasawa relation is a generalization of the Wiener filter

Mean noise variance reduction

- For a given y:

$$E(\hat{x}-x)^2 = N^2(\log p(y))'' + N$$

- The mean variance is:

$$D = N - N^2 \int_y \frac{p'^2(y)}{p(y)} dy = N(1-R)$$

- In proportion the variance is reduced of:

$$\boxed{R = N \int_y \frac{p'^2(y)}{p(y)} dy}$$

- The SNR gain is:

$$\Delta = -10 \log_{10}(1-R)$$

Some properties

- The index is bounded: $0 \leq R \leq 1$
- It is easy to show that: $R \geq \frac{N}{S+N}$
  - Equality for a Gaussian signal
- If $N \to 0$

$$R \to N \int \frac{p'^{\,2}_x(x)}{p_x(x)} dx$$

  - $R \to 0$ Full signal = no noise reduction
- If $N \to \infty$ $R \to 1$
  - No signal can be identified
- R can be considered as a sparsity index

Index Estimation

- Direct estimation → Very bad results
- Smoothing pyramid of the histogram
- From quantiles (*Bijaoui, ADA4, 2004*)
- Parzen windows
- Histogram denoising with the DWT
- Gaussian mixture (*Bijaoui, SP 2001*)
- Truncated distributions
  - Exponential (*Raphan & Simoncelli 2007*)
  - Gaussian (*Bijaoui, GRETSI, 2009*)
- PDF fit → Voigt distributions

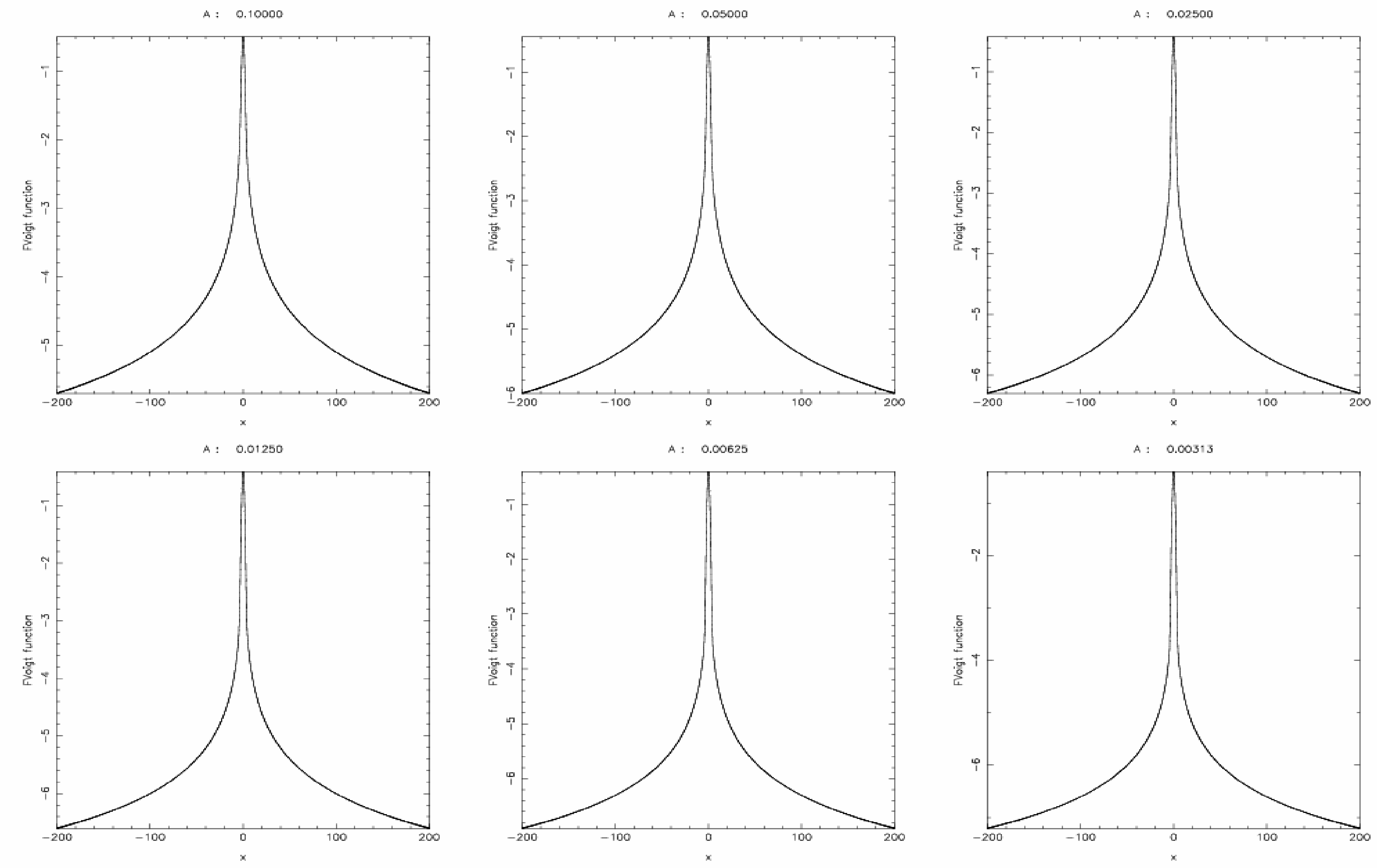
The Voigt Distribution

- Definition :

$$V(x,\sigma,a) = \frac{a}{\pi\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} \frac{e^{-\frac{y^2}{2\sigma^2}}}{a^2+(x-y)^2}dy$$

- Well adapted to astronomical images
- Approximations (*Garcia, MNRAS, 2006*):
  - Quasi Gaussian PDF
  - Close to a Cauchy(-Lorentz) PDF
- Heavy PDF distribution tails
- Parameter estimation
  - Maximum Likelihood
  - k-sigma rejection

Voigt Functions (log)

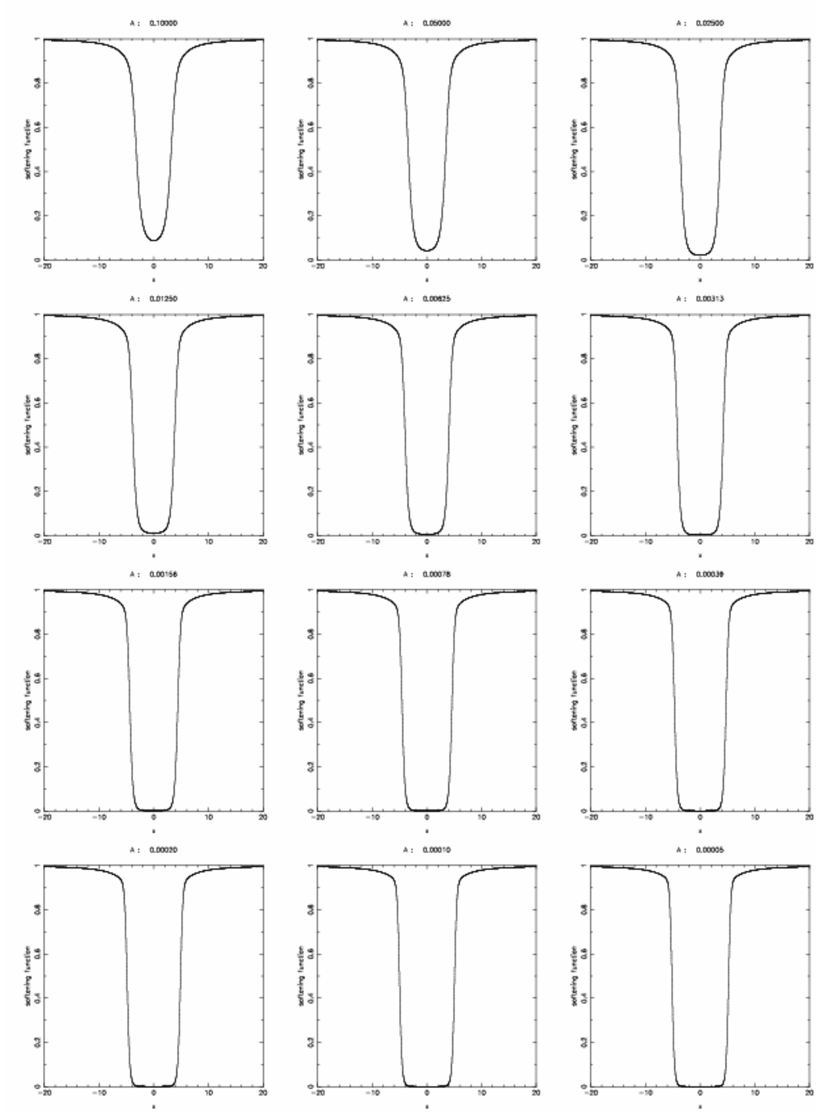


Derived filters

- Filter expression

$$W = 1 + N\frac{\log' p(y)}{y}$$

- Cauchy Signal & Gaussian noise
- Converge to a hard thresholding
- The threshold increases for a decreasing a



Iterative k-sigma Rejection

- The parameters m and s are defined as:

$$m = \frac{\int_{m-ks}^{m+ks} x p(x) dx}{\int_{m-ks}^{m+ks} p(x) dx} \qquad s^2 = \frac{\int_{m-ks}^{m+ks} (x-m)^2 p(x) dx}{\int_{m-ks}^{m+ks} p(x) dx}$$

- The parameters are estimated iteratively with the k-sigma rejection
- Proportion of $(1\text{-}r)$ values are rejected
  - Significant values
- R approximation:

$$R = r\frac{N}{s^2}$$

Scalar / vectorial Process

- Taking into account local correlations
- Vectorial Miyasawa relation
- Direct estimation is impossible
  - Too few events
- n-D Estimation with rejection
- Orthonormal transforms
  - Keep noise properties
  - PDF coefficients # PDF values → # R indices
  - The neighborhood relations can be taken into account by the transform

Index and Orthonormal Transforms

- For each coefficient → PDF Estimation?
  - Transforms with sliding windows
- Case of a Gaussian prior
  - The Karhunen-Loève transform is optimal
    - Consider a window
    - Compute the autocorrelation matrix C
    - C Eigenvectors → components
    - For each component → Wiener filtering
- Case of Voigt distributions → rejection

KL transform with Rejection (KLR)

1. Start from the identity matrix (pixels)
2. K-sigma rejection for each component
3. Autocorrelation matrix without taking into account the rejected elements
4. Eigenvectors → New components
5. Come back to point 2 up to convergence

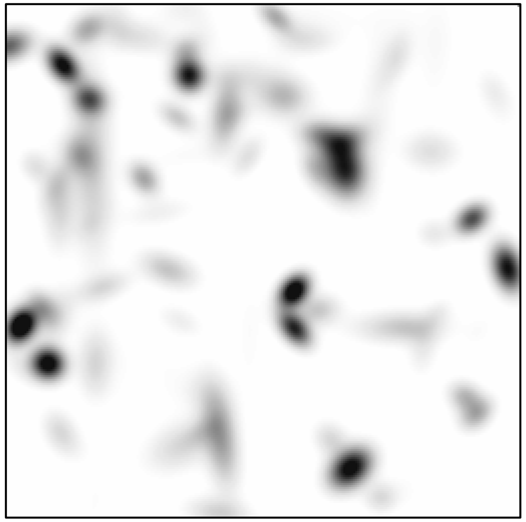
- Algorithm results:
  - Map of rejected elements
  - Clean autocorrelation matrix
  - Orthonormal transform which leads to the highest R index

KLR Application

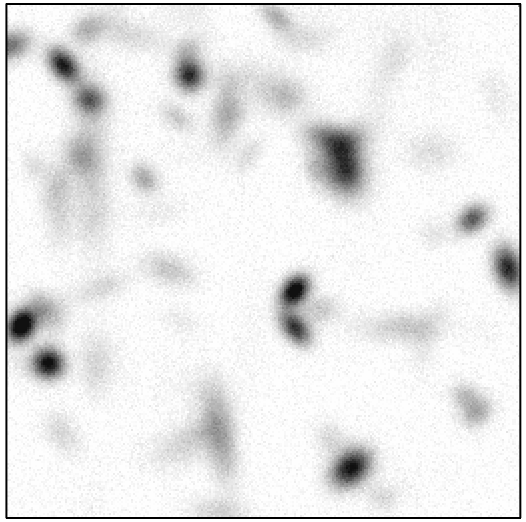
- Window size
  - Small (3x3): the noise can not be sufficiently reduced
  - Large (11x11): the autocorrelation matrix becomes too large
- Overlapping windows? → redundancy
- Multiscale KLR
  - Keep the first component and apply again KLR
  - Redundancy → À trous algorithm !
  - Similarity with the DWT with rejection

Case of a faint noise

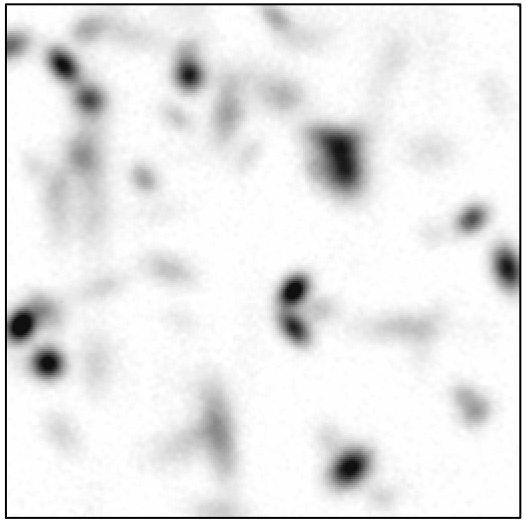
- Simulation = Sum of Gaussian patterns
- Few rejected pixels



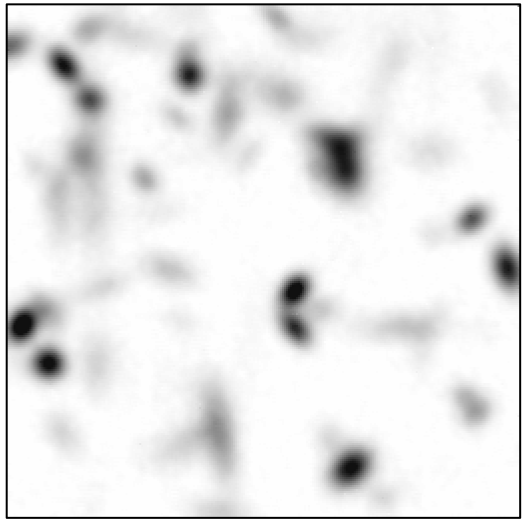
Original image



Noisy image SNR=14.73



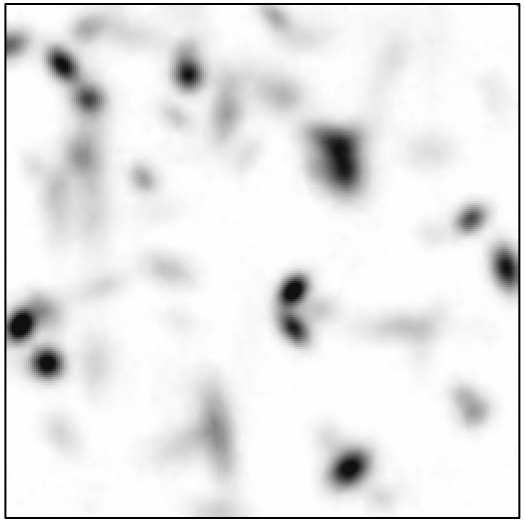
3x3 SNR= 24.11



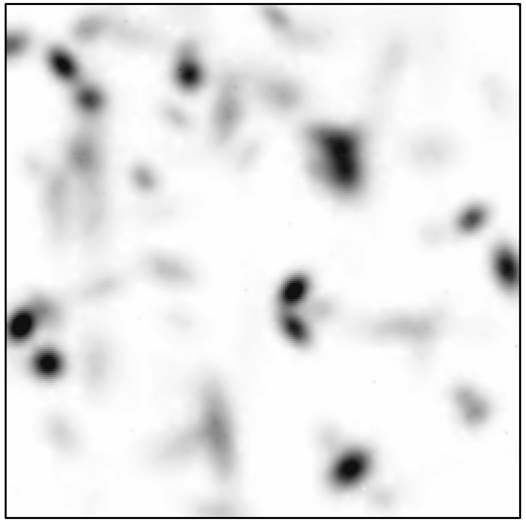
5x5 SNR=26.12

Window size & Wavelets

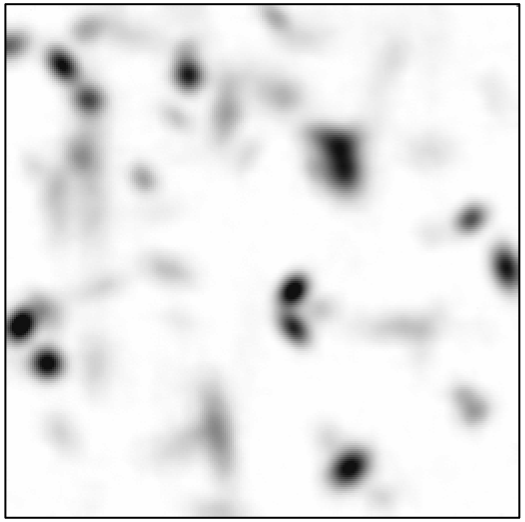
- Even with faint noise a large window is necessary
- KL instability with the size increasing
- Redundant Wavelet superiority



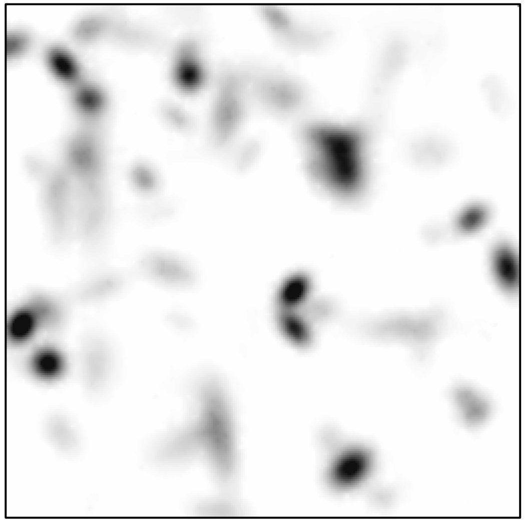
7x7 SNR=27.85



A trous Hard SNR=28.29

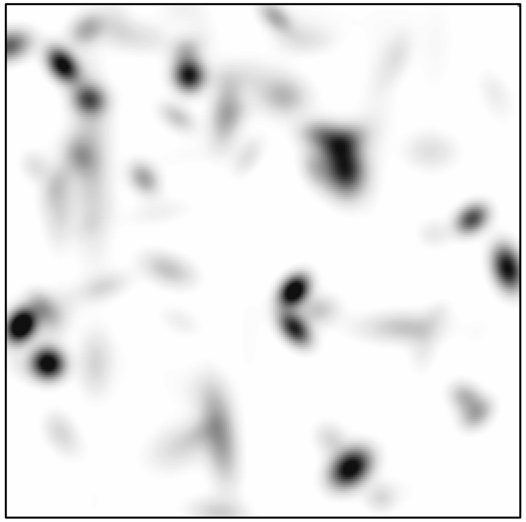


11x11 SNR=28.49

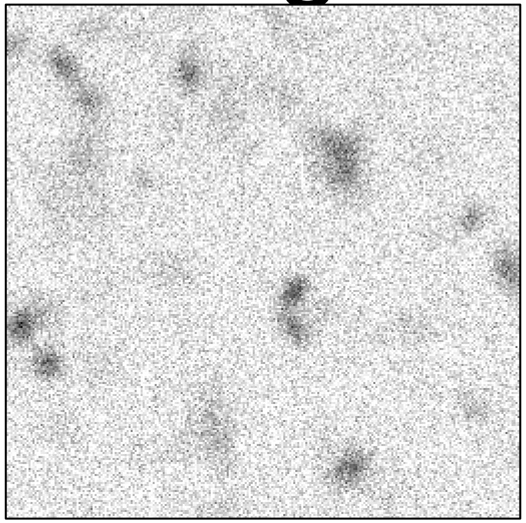


A trous Smooth SNR=28.74

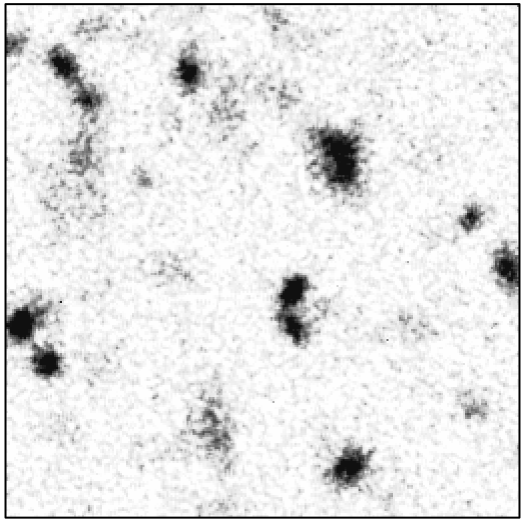
With a larger Noise



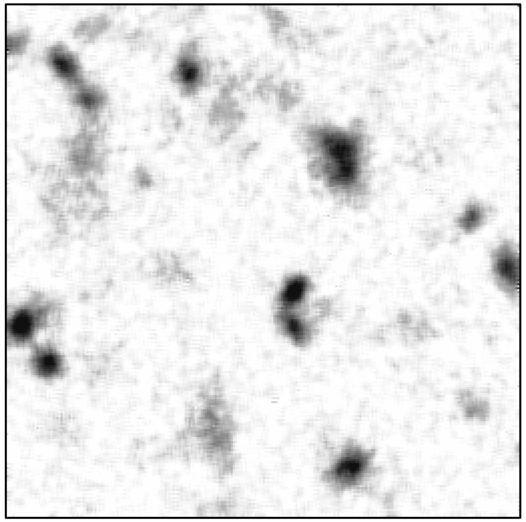
Original



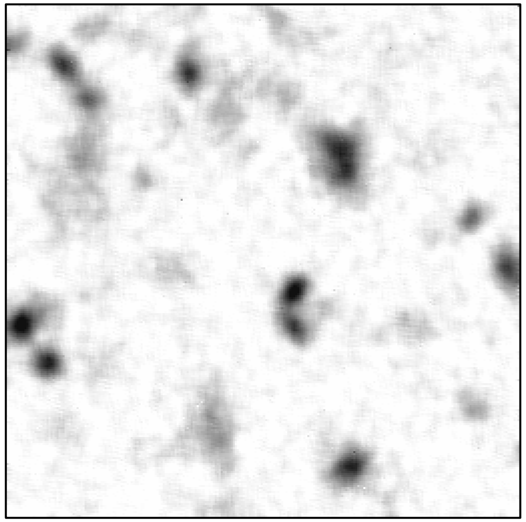
Noisy Image SNR=-5.27



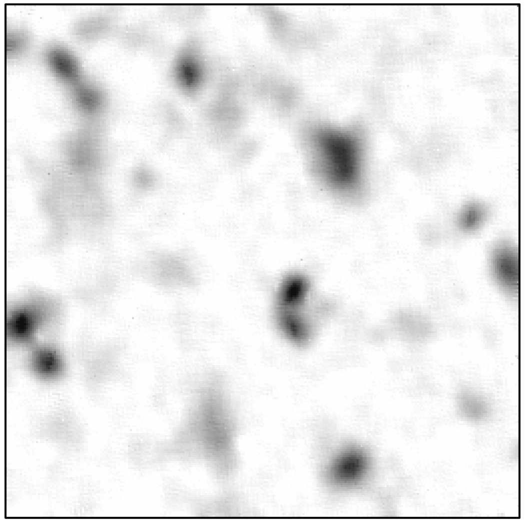
3x3 SNR=7.34



5x5 SNR=10.38



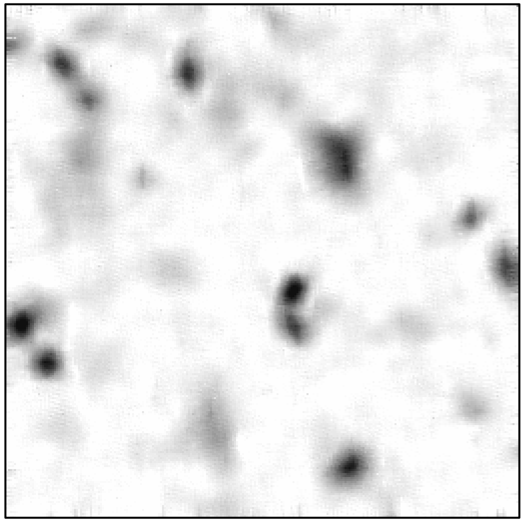
7x7 SNR=12.08



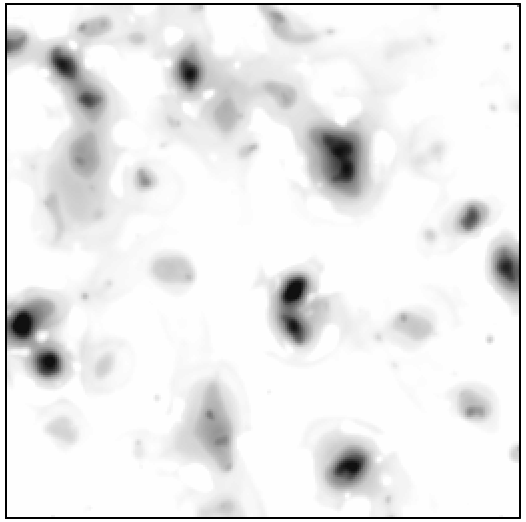
11x11 SNR=12.76

Comparison with wavelets

→ A multiscale approach is essential for a large denoising



KL 15x15
SNR=12.54



A trous wavelet transform
Bayesian Softening
SNR=14.54

Conclusion

- The posterior Bayesian expectation can be directly obtained from the data
- The reduction factor of the noise variance R is a sparsity index
- A robust estimation results from an algorithm with a k-sigma iterative rejection
- The maximum index value would be reached by the best transformation
- A Karhunen-Loève transform with rejection would be the best solution
  - Window size → Multiscale approach