The SURE-LET Methodology

A Prior-Free Approach to Signal and Image Denoising

Thierry Blu

Department of Electronic Engineering
The Chinese University of Hong Kong

May 2009

Joint work with **Florian Luisier** *(EPFL)*

Outline

1. Image denoising
   - The problem
   - Prior-based approaches for image denoising
2. The SURE-LET Approach
   - Stein's Unbiased Risk Estimate
   - A Linear Expansion of Thresholds (LET)
3. SURE-LET algorithms in image denoising
   - Orthogonal representations
   - Non-Orthogonal/Redundant Representations
4. Possible extensions
   - Other MSE estimates
   - PURE-LET Haar denoising

Noisy data

Usual acquisition devices provide signals

$$\mathrm{Y} = [y_1, y_2, \dots, y_N]^{\mathrm{T}}$$

that are corrupted with noise.

Frequent modelization using an **additive white Gaussian noise** hypothesis

$$\underbrace{\mathrm{Y}}_{\text{"noisy" signal}} = \underbrace{\mathrm{X}}_{\text{"original" signal}} + \underbrace{\mathrm{B}}_{\text{"noise"}}$$

where $\mathscr{E}\{\mathrm{B}\mathrm{B}^{\mathrm{T}}\} = \sigma^2 \mathbf{Id}$.

Signal denoising consists in fiding a "good" candidate $\hat{\mathrm{X}}$ of X *using only the noisy signal* Y: i.e., find the algorithm $\mathbf{F}$ such that

$$\hat{\mathrm{X}} = \mathbf{F}(\mathrm{Y})$$

An Abundant Literature

Many approaches available

1. *Explicit hypotheses on the signal*
   - Statistics-based (Bayesian)
   - Regularization
   - Model fitting
2. *Heuristic approaches*
   - Filtering
   - Non-Local Means
   - Any combination of approaches 1 when the hypotheses are not satisfied/checked

In the details, algorithms also differ whether they operate in the *signal domain* directly, or in a *transformed domain*.

NOTE: Most approaches involve parameters which are often set empirically.

The goal of this talk is not to compare algorithms,
but to propose a *method* to obtain (fast) algorithms.

Statistical approaches

Based on an explicit knowledge of the prior probability density of the signal to restore. Various objectives are possible, among which

- Maximum A Posteriori (MAP)
- Minimum MSE (e.g., Wiener)

This means that these methods assume that the following are given

- The probability density of the noise $q(\mathrm{B}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\|\mathrm{B}\|^2}{2\sigma^2}\right)$;
- The probability density of the original signal $p(\mathrm{X})$.

Goals of this talk

Show that it is possible to

- *avoid statistical assumptions* on the original signal (SURE)
- devise *non-iterative* algorithms (LET) that are optimal

Maximum a Posteriori

The MAP consists in choosing the estimate $\hat{\mathrm{X}}$ that maximizes the *posterior probability density*

$$p(\hat{\mathrm{X}}|\mathrm{Y}) = \max_{\mathrm{X}} p(\mathrm{X}|\mathrm{Y})$$

which in this case amounts to maximize $q(\mathrm{Y}-\mathrm{X})p(\mathrm{X})$.

Optimal detector: given noisy measurements of a signal X having a finite number of values X_1, X_2, ..., X_K occurring with probabilities p_1, p_2,..., p_K, the MAP minimizes the error probability

$$\mathscr{P}\left\{\hat{\mathrm{X}} \neq \mathrm{X}\right\}$$

NOTE: Description of the prior $p(\mathrm{X})$ may require many parameters.

For signals with large, or infinite number of levels, the probabilistic optimality of the MAP becomes irrelevant $\rightsquigarrow$ MSE.

Minimum MSE: Wiener

The Wiener "filter" consists in finding the linear[1] estimate, $\hat{\mathrm{X}} = \hat{\mathbf{A}}\mathrm{Y}$, that minimizes the *Mean Square Error* (MSE)

$$\underbrace{\mathscr{E}\left\{\frac{1}{N}\|\hat{\mathbf{A}}\mathrm{Y}-\mathrm{X}\|^2\right\}}_{\text{MSE between } \hat{\mathrm{X}} \text{ and } \mathrm{X}} = \min_{\mathbf{A}} \mathscr{E}\left\{\frac{1}{N}\|\mathbf{A}\mathrm{Y}-\mathrm{X}\|^2\right\}$$

Solution: Requires only the knowledge of the covariance matrix $\mathbf{R} = \mathscr{E}\{\mathrm{X}\mathrm{X}^{\mathrm{T}}\}$ of the original signal

$$\hat{\mathrm{X}} = \mathbf{R}\left(\mathbf{R}+\sigma^2\mathbf{Id}\right)^{-1}\mathrm{Y}$$

NOTE: Although very popular, linear processing is not well-adapted to the processing of transient signals.

[1] if $\mathscr{E}\{\mathrm{X}\} = \mathbf{0}$ — an affine estimate is used, otherwise.

Minimum MSE: Non-linear case

It is possible to solve Wiener's problem *without the linear processing hypothesis* (see e.g., Raphan/Simoncelli); i.e., find the optimal processing $\mathbf{F}(\cdot)$ that yields the estimate $\hat{\mathrm{X}} = \mathbf{F}(\mathrm{Y})$ such that

$$\mathscr{E}\left\{\frac{1}{N}\|\mathbf{F}(\mathrm{Y})-\mathrm{X}\|^2\right\} \text{ is minimized.}$$

Solution: $\hat{\mathrm{X}} = \mathscr{E}\{\mathrm{X}|\mathrm{Y}\}$, the posterior expectation. This expression can be simplified to

$$\hat{\mathrm{X}} = \mathrm{Y} + \sigma^2 \frac{\nabla r(\mathrm{Y})}{r(\mathrm{Y})}$$

where $r(\mathrm{Y}) = (p * q)(\mathrm{Y})$ is the (marginal) probability density of Y.

NOTE: The optimal MSE processing is infinitely differentiable.

The optimal algorithm only requires the knowledge of the *pdf of the noisy signal* $\rightsquigarrow$ no prior information is needed!

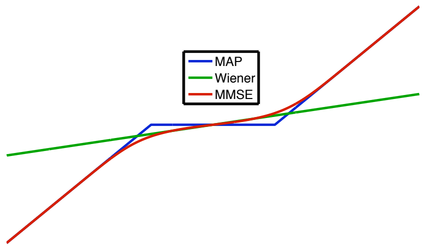
Example

Assuming a Laplace prior, $p(\mathrm{X}) = \prod_{n=1}^{N} \frac{\lambda}{2} \mathrm{e}^{-\lambda |x_n|}$, these statistical approaches yield a pointwise thresholding involving $T = \lambda\sigma^2$:

MAP $\hat{x}_n = \mathrm{soft}_T(y_n)$

Wiener $\hat{x}_n = \dfrac{y_n}{1 + \frac{T^2}{2\sigma^2}}$

MMSE $\hat{x}_n = y_n - T \dfrac{\mathrm{e}^{-\lambda y_n} \operatorname{erfc}\left(\frac{-y_n + T}{\sigma\sqrt{2}}\right) - \mathrm{e}^{\lambda y_n} \operatorname{erfc}\left(\frac{y_n + T}{\sigma\sqrt{2}}\right)}{\mathrm{e}^{-\lambda y_n} \operatorname{erfc}\left(\frac{-y_n + T}{\sigma\sqrt{2}}\right) + \mathrm{e}^{\lambda y_n} \operatorname{erfc}\left(\frac{y_n + T}{\sigma\sqrt{2}}\right)}$



Regularization approaches

Choice of a functional $J(\mathrm{X})$ that is known to be small when applied to the original signal. Typical choices are

- Tikhonov (e.g., smoothness prior): $J(\mathrm{X}) = \|\mathbf{R}\mathrm{X}\|^2$
- Sparsity prior: $J(\mathrm{X}) = \|\mathrm{X}\|_{\ell^0} \rightsquigarrow J(\mathrm{X}) = \|\mathrm{X}\|_{\ell^1}$
- Total variation (edge prior): $J(\mathrm{X}) = \sum_n |x_n - x_{n-1}|$

The signal estimate $\hat{X}$ is then selected as the solution of

$$\min_{\mathrm{X}} J(\mathrm{X}) \text{ such that } \|\mathrm{Y} - \mathrm{X}\|^2 \leq N\sigma^2$$

Note: Using Lagrange's multipliers method, $J(\mathrm{X})$ can be re-interpreted as a statistical prior and the optimization equivalent to a MAP.

No explicit distance minimization between original and denoised signal.

Estimation of the MSE without signal prior

Thanks to the *white Gaussian noise* hypothesis, Stein's estimate

$$\mathrm{SURE}(\mathrm{Y}) = \frac{1}{N}\|\mathbf{F}(\mathrm{Y}) - \mathrm{Y}\|^2 + \frac{2\sigma^2}{N}\operatorname{div}\big(\mathbf{F}(\mathrm{Y})\big) - \sigma^2$$

satisfies[2] $\mathscr{E}\{\mathrm{SURE}(\mathrm{Y})\} = \mathscr{E}\left\{\|\hat{\mathrm{X}} - \mathrm{X}\|^2/N\right\}$.

Moreover, $\mathrm{SURE}(\mathrm{Y})$ has a small variance ($\propto 1/N$), thus

$$\frac{1}{N}\|\hat{\mathrm{X}} - \mathrm{X}\|^2 \approx \mathrm{SURE}(\mathrm{Y})$$

Note: Particularly adapted for large data sizes (e.g., images).

No assumptions on the original signal X, no statistical characterization.

[2] Expectation taken over all possible realizations of the noise.

A simple proof

On the one hand

$$\mathscr{E}\left\{\|\mathbf{F}(\mathrm{Y}) - \mathrm{X}\|^2\right\} = \mathscr{E}\left\{\|\mathbf{F}(\mathrm{Y})\|^2\right\} - 2\underbrace{\mathscr{E}\{\mathrm{X}^{\mathrm{T}}\mathbf{F}(\mathrm{Y})\}}_{\mathscr{E}\{(\mathrm{Y}-\mathrm{B})^{\mathrm{T}}\mathbf{F}(\mathrm{Y})\}} + \underbrace{\|\mathrm{X}\|^2}_{\mathscr{E}\{\|\mathrm{Y}\|^2\} - N\sigma^2}$$
$$= \mathscr{E}\left\{\|\mathbf{F}(\mathrm{Y}) - \mathrm{Y}\|^2\right\} + 2\,\mathscr{E}\{\mathrm{B}^{\mathrm{T}}\mathbf{F}(\mathrm{Y})\} - N\sigma^2$$

and on the other hand (*Stein's Lemma*)

$$\mathscr{E}\{\mathrm{B}^{\mathrm{T}}\mathbf{F}(\mathrm{Y})\} = \int \underbrace{q(\mathrm{B})\,\mathrm{B}^{\mathrm{T}}}_{-\sigma^2\nabla q(\mathrm{B})^{\mathrm{T}}} \mathbf{F}(\mathrm{X} + \mathrm{B})\,\mathrm{d}^N\mathrm{B}$$
$$= \int \sigma^2 q(\mathrm{B}) \operatorname{div}\big(\mathbf{F}(\mathrm{X} + \mathrm{B})\big)\,\mathrm{d}^N\mathrm{B} \qquad \text{(by parts)}$$
$$= \mathscr{E}\left\{\sigma^2 \operatorname{div}\big(\mathbf{F}(\mathrm{Y})\big)\right\}$$

SURE minimization

Because it is an estimate of the MSE of a processing, it is natural to minimize the SURE for finding good estimates of the parameters that define the processing.

Example: Donoho's *SureShrink*; find the optimal threshold T such that $\text{SURE}_{\text{soft}(.,T)}$ is minimal[3].

$$N.\text{SURE}_{\text{soft}(.,T)} = \underbrace{\sum_n \left|\text{soft}(y_n,T) - y_n\right|^2}_{\left(\sum_{|y_n|<T} y_n^2\right) + T^2 \#_{|y_n|\geq T}} + \underbrace{\sum_n 2\sigma^2 \frac{\text{d soft}}{\text{d}y}(y_n,T)}_{2\sigma^2 \#_{|y_n|\geq T}} - N\sigma^2$$

NOTE: Very few other examples in the SP literature (Pesquet et al.).

[3] $\#_{|y_n|\geq T}$ is the number of coefficients y_n such that $|y_n| \geq T$.

Prior-free parametric processing

A change of emphasis

Standard: Choice of a *parametric prior*, find the parameters from the noisy data, then derive the optimal processing (e.g., MAP)

Proposed: *Parametrize the processing* directly, then find the optimal parameters (SURE minimization)

In the SURE-based approach, the *signal estimation* problem is replaced by a *processing approximation* problem — i.e., approximation of a *functional*, not a signal:

$$\underbrace{\mathrm{Y} \longmapsto \hat{\mathrm{X}}}_{\text{standard}} \quad \text{replaced by} \quad \underbrace{\mathrm{Y} \longmapsto \mathbf{F}(\cdot)}_{\text{proposed}}$$

Optimization over a class of processings
vs. optimization over a class of signals

Linear approximation

It is particularly attractive to perform a *linear* decomposition of the processing onto a basis of *elementary* processings

Linear Expansion of Thresholds (LET)

$$\underbrace{\mathbf{F}(\cdot)}_{\hat{\mathrm{X}}=\mathbf{F}(\mathrm{Y})} = \overbrace{\sum_{k=1}^{K} a_k \underbrace{\mathbf{F}_k(\cdot)}_{\text{elementary "thresholds"}}}^{\text{Linear Expansion of Thresholds (LET)}}$$

Advantages

- *Explicit* description of the processing;
- Using enough (reasonable) basis elements, it is possible to *approximate most non-linear parametric* processing;
- Minimization of a quadratic objective (e.g., SURE) yields a *linear system of equations* (non-iterative solution).

SURE-LET processing

Minimization of the SURE for processings described as a LET: the coefficients a_k of the linear combination are obtained as

$$\{a_k\}_{k=1\ldots K} = \arg\min_{\{a_k\}_{k=1\ldots K}} \frac{1}{N}\Big\|\sum_{k=1}^{K} a_k \mathbf{F}_k(\mathrm{Y}) - \mathrm{Y}\Big\|^2 + \frac{2\sigma^2}{N}\sum_{k=1}^{K} a_k \operatorname{div}\big(\mathbf{F}_k(\mathrm{Y})\big) - \sigma^2$$

i.e., by solving a *linear system of equations*:

$$\sum_{k=1}^{K} a_k \mathbf{F}_l(\mathrm{Y})^{\mathrm{T}} \mathbf{F}_k(\mathrm{Y}) = \mathbf{F}_l(\mathrm{Y})^{\mathrm{T}}\mathrm{Y} - \sigma^2 \operatorname{div} \mathbf{F}_l(\mathrm{Y}) \qquad \text{for } l = 1, 2, \ldots K$$

NOTE: When model order K increases, the variance of SURE increases $\rightsquigarrow$ MSE estimation quality decreases.

Non-iterative optimization, naturally fast.

Transformed domain denoising

It is frequent to use linear transformations (wavelets, DCT) to represent signals/images better: e.g., to "decorrelate" them, or to sparsify them:

$$\underbrace{\mathrm{W} = \mathbf{D}\mathrm{Y}}_{\text{analysis}} \rightsquigarrow \underbrace{\mathrm{Y} = \mathbf{R}\mathrm{W}}_{\text{synthesis}}$$

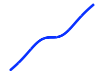
where $\mathbf{RD} = \mathbf{Id}$. Typical transformations may be

- **orthogonal** — useful because of *MSE preservation* $\rightsquigarrow$ *separate* processing of transformed coefficients;
- **redundant** — useful because *simple (coefficientwise) processing* of transformed coefficients is sufficient to produce high-quality results.

Transformed domain LET processing: $\mathbf{F}(\mathrm{Y}) = \sum_{k=1}^{K} a_k \mathbf{R}\mathbf{\Gamma}_k(\mathrm{W})$

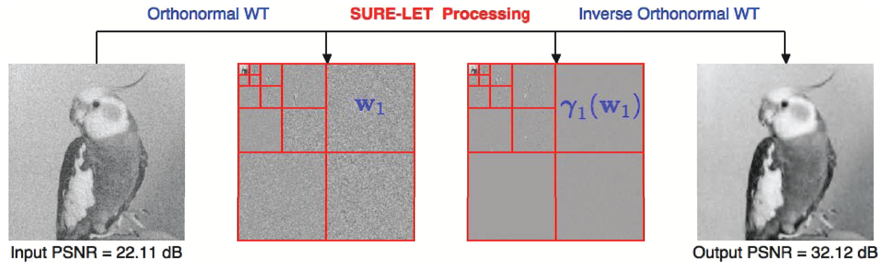
Pointwise wavelet thresholding

Principle: use an orthogonal (non-redundant) wavelet representation (e.g., symlet 8) and threshold each wavelet band using



$$\gamma_{a,b}(w) = aw + bwe^{-\frac{w^2}{12\sigma^2}}$$

where a, b minimize the SURE in each subband.



Example of result

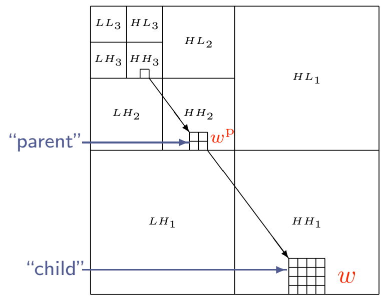
Noisy | SureShrink | SURE-LET pointwise

PSNR=15 dB | PSNR=28.08 dB | PSNR=28.33 dB

NOTE: Adding more parameters brings almost no improvement. Better denoising efficiency requires *multivariate* thresholding rules.

InterScale wavelet thresholding

The relative locality of the DWT implies that there may be a *spatial correlation* between different wavelet scales: three potential *tree*-structures — LH, HH and HL



Interscale thresholding consists in expressing the denoised estimate as

$$\hat{x}_w[n] = \gamma(w[n], w^{\mathrm{P}}[n])$$

InterScale wavelet thresholding

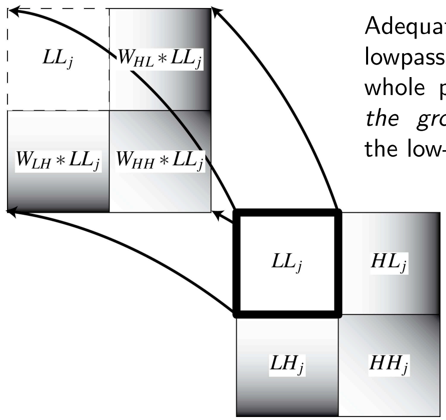
Principle: separate the parent into *large* and *small* coefficients, and within each zone so defined, apply a pointwise thresholding function:

$$\gamma(w, w^{\mathrm{p}}) = \underbrace{\mathrm{e}^{-\frac{(w^{\mathrm{p}})^2}{12\sigma^2}}\left(aw + bw\mathrm{e}^{-\frac{w^2}{12\sigma^2}}\right)}_{\text{small parents}} + \underbrace{\left(1 - \mathrm{e}^{-\frac{(w^{\mathrm{p}})^2}{12\sigma^2}}\right)\left(a'w + b'w\mathrm{e}^{-\frac{w^2}{12\sigma^2}}\right)}_{\text{large parents}}$$

NOTE: DWT is orthogonal, hence w and w^p are *statistically independent* $\rightsquigarrow$ same SURE formula as for the pointwise case.

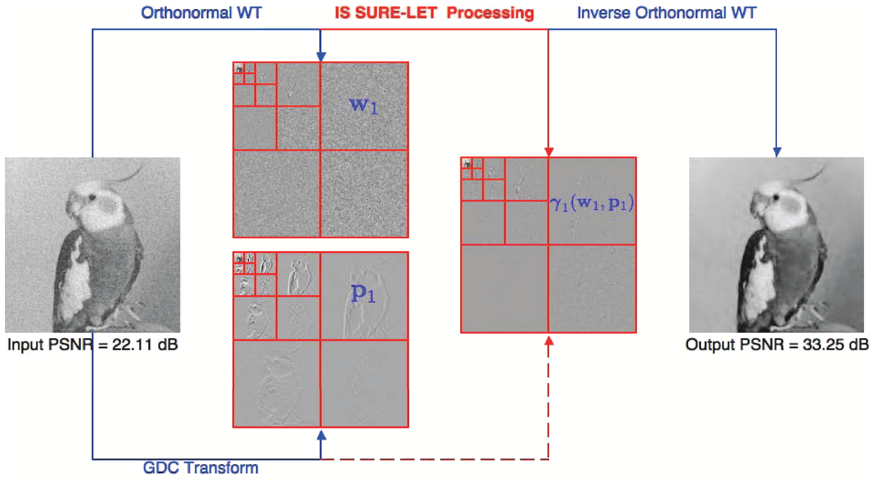
PROBLEM: the wavelet coefficients are not exactly aligned from band to band (filtering and downsampling effect). How to obtain a parent aligned exactly with his child?

Parent/child alignment: Group-Delay Compensation



Adequate high-pass filtering of the lowpass LL_j — which contains the whole parent tree: W *compensates the group-delay* difference between the low-pass and the high-pass band.

Overview of the interscale SURE-LET denoising



Example of result

Noisy	SureShrink	SURE-LET interscale

PSNR=15 dB	PSNR=28.08 dB	PSNR=29.29 dB

Best non-redundant transform-domain algorithm.

Extension to multichannel denoising

Direct generalization by replacing:

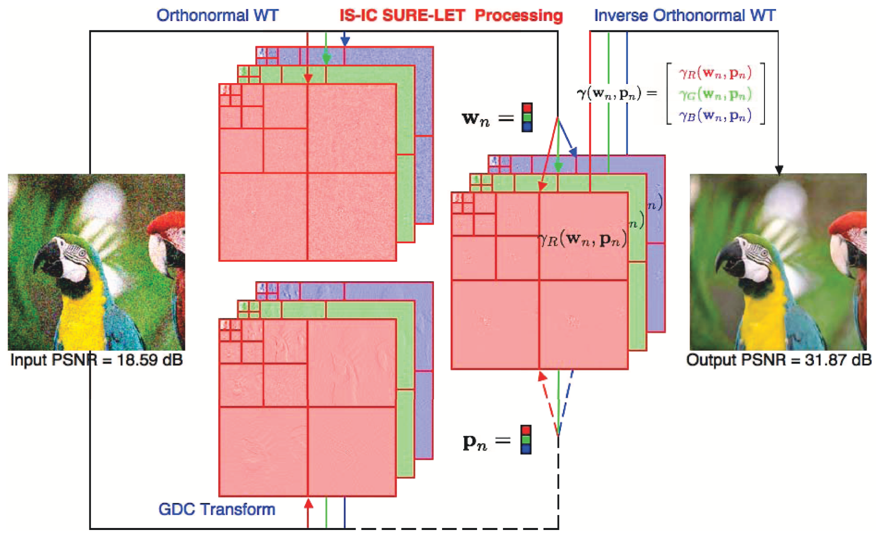
- *scalar*-valued by *vector*-valued wavelet coefficients;
- *scalar*-valued by *matrix*-valued LET parameters.

Assuming $\mathbf{R}$=covariance matrix of the noise, and $g(x) = \exp(-x/12)$

$$\begin{aligned}\gamma(\mathbf{w}_n,\mathbf{p}_n) = & \underbrace{g(\mathbf{p}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{p}_n)g(\mathbf{w}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_n)}_{\text{small parents and small coefficients}}\mathbf{a}_1^{\mathrm{T}}\mathbf{w}_n \\ & + \underbrace{(1-g(\mathbf{p}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{p}_n))g(\mathbf{w}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_n)}_{\text{large parents and small coefficients}}\mathbf{a}_2^{\mathrm{T}}\mathbf{w}_n \\ & + \underbrace{g(\mathbf{p}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{p}_n)(1-g(\mathbf{w}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_n))}_{\text{small parents and large coefficients}}\mathbf{a}_3^{\mathrm{T}}\mathbf{w}_n \\ & + \underbrace{(1-g(\mathbf{p}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{p}_n))(1-g(\mathbf{w}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{w}_n))}_{\text{large parents and large coefficients}}\mathbf{a}_4^{\mathrm{T}}\mathbf{w}_n\end{aligned}$$

NOTE: Automatically selects the best color space.

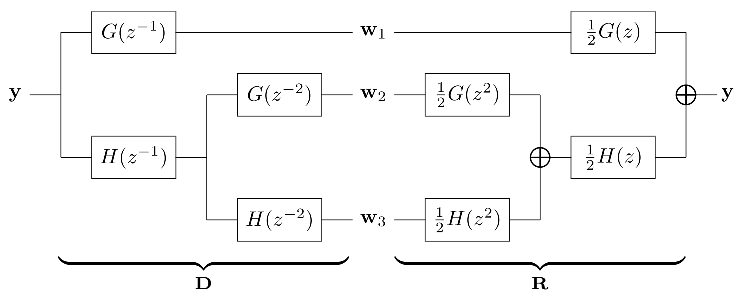
Overview of the Multichannel SURE-LET denoising



Undecimated pointwise wavelet thresholding

It has been observed 10 years ago (Coifman, Guo *et al.*) that redundant DWT are substantially more efficient for image denoising.

Two iterations of a 1D UDWT

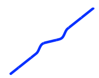


Perfect reconstruction condition: $\mathbf{R}.\mathbf{D} = \mathbf{Id}$

Undecimated pointwise wavelet thresholding

Thresholding rule

Defining $\mathbf{\Gamma}_{\mathrm{a,b}}(\mathrm{W}) = [\gamma_{a_1,b_1}(w_1), \gamma_{a_2,b_2}(w_2), \dots \gamma_{a_N,b_N}(w_N)]$, the processing takes the form $\mathbf{F}(\mathrm{Y}) = \mathbf{R}.\mathbf{\Gamma}_{\mathrm{a,b}}(\mathbf{D}.\mathrm{Y})$ where



$$\gamma_{a,b}(w) = aw + bw\left(1 - \mathrm{e}^{-(\frac{w}{3\sigma})^8}\right)$$

and where the (a_k, b_k) are all identical within the same wavelet subband — i.e., two parameters per subband.

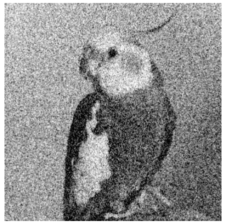
The optimal set of parameters $\{\mathrm{a},\mathrm{b}\}$ is then found by minimizing the global image-domain SURE.

NOTE: Contrary to the nonredundant case, a hard-like threshold works better than a softer version.

Undecimated pointwise wavelet thresholding

Undecimated discrete symlet transform

Noisy



PSNR=15 dB

SureShrink

PSNR=28.08 dB

SURE-LET

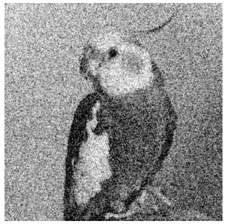
PSNR=29.49 dB

NOTE: Surprisingly, it is the simplest wavelet type (Haar) that works best. Smallest support?

Undecimated pointwise wavelet thresholding

Undecimated discrete Haar wavelet transform

Noisy



PSNR=15 dB

SureShrink

PSNR=28.08 dB

SURE-LET

PSNR=30.28 dB

NOTE: Surprisingly, it is the simplest wavelet type (Haar) that works best. Smallest support?

Linear modifications

It is possible to adapt the SURE so as to take into account

1. An arbitrary noise covariance: $\mathscr{E}\{\mathrm{B}\mathrm{B}^{\mathrm{T}}\} = \mathbf{R}$;
2. A distortion: $\mathrm{Y} = \mathbf{A}\mathrm{X} + \mathrm{B}$;
3. A non-Euclidian, but quadratic quality measure: $\mathscr{E}\left\{\|\mathbf{Q}(\hat{\mathrm{X}} - \mathrm{X})\|^2\right\}$.

Given all these linear modifications, the SURE formula has to be modified

$$\mathsf{SURE}(\mathrm{Y}) = \frac{1}{N}\|\mathbf{Q}(\mathbf{F}(\mathrm{Y}) - \mathbf{A}^{-1}\mathrm{Y})\|^2 + \frac{2}{N}\,\mathrm{div}\big(\mathbf{R}\mathbf{A}^{-\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{F}(\mathrm{Y})\big) - \frac{\mathrm{Tr}(\mathbf{Q}\mathbf{A}^{-1}\mathbf{R}\mathbf{A}^{-\mathrm{T}}\mathbf{Q}^{\mathrm{T}})}{N}$$

NOTE: Prior information on X may be needed when matrices involved are singular. Application to deconvolution (Vonesch,Pesquet/Benazza/Chaux).

Other noise statistics

It is possible to obtain unbiased estimate of the MSE for non Gaussian statistics. Typically (Raphan/Simoncelli, Eldar) for

- Additive arbitrary pdf
- Exponential families of pdf

Example of the Poisson Unbiased Risk Estimate (PURE)

- Estimate x from noisy Poisson measurements y
$$\mathscr{P}\{y = n\} = x^n e^{-x}/n!$$
- Processing on y to obtain an estimate $\hat{x}$ of x: $\hat{x} = f(y)$
- $\mathsf{PURE} = f(y)^2 - 2yf(y-1) + y(y-1)$ is such that
$$\mathscr{E}\{\mathsf{PURE}\} = \mathscr{E}\{|\hat{x} - x|^2\}$$

NOTE: All these estimates are quadratic in $\mathbf{F}(\cdot) \rightsquigarrow$ LET parametrization.

Haar and Poisson

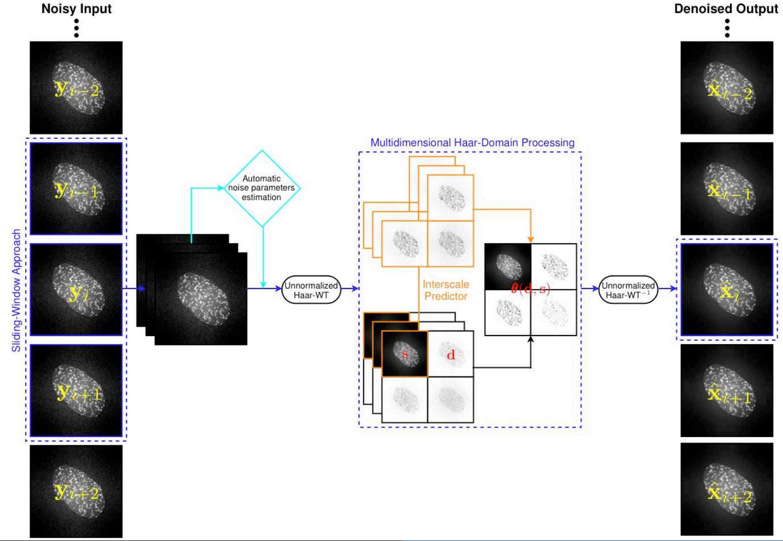
The Haar wavelet transform has two important properties

- Orthogonality, i.e., preservation of the MSE in the wavelet transform
- "Propagation" of the Poisson statistics at coarser scales.

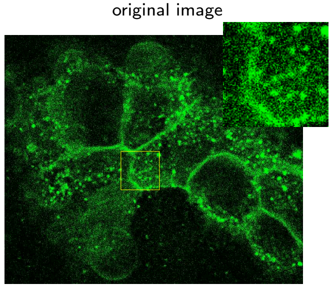
⤳ PURE involving neighboring scales.
⤳ thresholding function involving interscale dependencies.
⤳ application to fluorescence microscopy images.

Natural extension (with Florian Luisier and Cédric Vonesch) of the interscale SURE-LET approach to Haar PURE-LET.

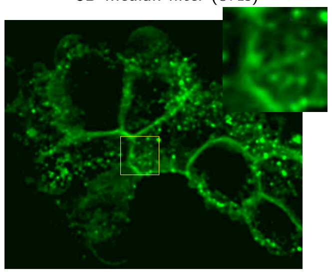
Overview of the multi-frame algorithm



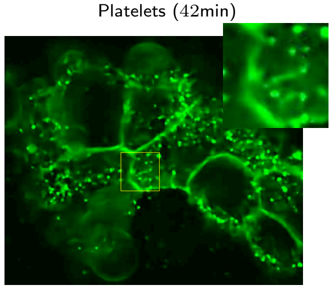
original image



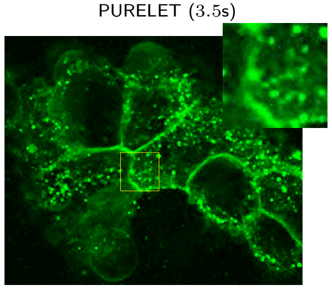
3D median filter (8.4s)



Platelets (42min)



PURELET (3.5s)



Conclusion

Presentation of a generic framework for signal/image denoising.

Advantages:

- Does not require hypotheses on the signal, only on the noise (SURE)
- Linear approximation of the denoising process on a basis of "thresholds" (LET)
- Fast, non-iterative (SURE + LET)
- Natural construction of multivariate thresholding rules.
- Extensions to non-Gaussian noise corruptions.

Papers available at `http://www.ee.cuhk.edu.hk/~tblu/`