Image reconstruction from multiple sensors using Stein’s principle. Application to parallel MRI
Alexandru Marin, Caroline Chaux, Jean-Christophe Pesquet, Philippe Ciuciu

To cite this version:
Alexandru Marin, Caroline Chaux, Jean-Christophe Pesquet, Philippe Ciuciu. Image reconstruction from multiple sensors using Stein’s principle. Application to parallel MRI. ISBI 2011, Mar 2011, United States. hal-00733471

HAL Id: hal-00733471
https://hal-upec-upem.archives-ouvertes.fr/hal-00733471
Submitted on 19 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
IMAGE RECONSTRUCTION FROM MULTIPLE SENSORS USING STEIN’S PRINCIPLE. APPLICATION TO PARALLEL MRI.

Alexandru Marin¹, Caroline Chaux¹, Jean-Christophe Pesquet¹ and Philippe Ciuciu²

¹ Univ. Paris-Est, LIGM UMR CNRS 8049
5 Bd Descartes, 77454 Marne-la-Vallée, France
E-mails: alex@seneka.ro, {caroline.chaux, jean-christophe.pesquet}@univ-paris-est.fr
² CEA/DSV/I2BM/Neurospin, CEA Saclay, Bat. 145, Point Courrier 156, 91191 Gif-sur-Yvette cedex, France.
E-mail: philippe.ciuciu@cea.fr

ABSTRACT

We are interested in image reconstruction when data provided by several sensors are corrupted with a linear operator and an additive white Gaussian noise. This problem is addressed by invoking Stein’s Unbiased Risk Estimate (SURE) techniques. The key advantage of SURE methods is that they do not require prior knowledge about the statistics of the unknown image, while yielding an expression of the Mean Square Error (MSE) only depending on the statistics of the observed data. Hence, they avoid the difficult problem of hyperparameter estimation related to some prior distribution, which traditionally needs to be addressed in variational or Bayesian approaches. Consequently, a SURE approach can be applied by directly parameterizing a wavelet-based estimator and finding the optimal parameters that minimize the MSE estimate in reconstruction problems. Simulations carried out on parallel Magnetic Resonance Imaging (pMRI) images show the improved performance of our method with respect to classical alternatives.

Index Terms— Reconstruction, Stein’s principle, wavelets, nonlinear estimation, multiple sensors, pMRI.

1. INTRODUCTION

Much attention has been paid to Stein’s principle [1] in the recent statistical literature in order to derive MSE estimates in statistical problems involving an additive Gaussian noise. In particular, it was successfully used in wavelet-based nonlinear denoising [2]. More recently, such an approach was extended to deconvolution problems [3, 4] and it was also employed for Poisson data denoising [5]. The key advantage of Stein’s principle [6] is that it does not require prior knowledge about the statistics of the unknown image. Hence, it allows us to circumvent the difficult problem of the estimation of the hyperparameters of a prior distribution as frequently encountered in Bayesian approaches.

In this paper, we propose to further extend the scope of Stein-based approaches by addressing image reconstruction problems. More precisely, we employ Stein’s principle in order to propose a wavelet-based estimator relying on Linear Expansion of Thresholds (LET) functions [7, 8]. This enables us to design efficient medical image reconstruction algorithms whose parameters are optimally computed to achieve the minimum MSE. The estimator operates in the wavelet transform domain and deals with multiple observations delivered by several sensors. The interest of this method is demonstrated in the parallel Magnetic Resonance Imaging (pMRI) context [9, 10]. Recall that this modality offers a means of significantly reducing the acquisition time at the expense of a degraded signal-to-noise ratio. In a recent work by Chaâri et al. [11] a novel method based on a wavelet-based variational approach was proposed, which makes use of an iterative optimization algorithm to compute the Maximum A Posteriori (MAP) estimate. We aim here at proposing a faster non-iterative method where the estimator parameters are adjusted automatically. Note also that the proposed approach is very flexible concerning the wavelet choice and the estimator form.

The remainder of the paper is organized as follows. In Section 2, we briefly introduce the notation and describe the considered inverse problem. Then, Section 3 is devoted to the proposed statistical method. In particular, we show how to address the complex-valued nature of the data and build an unbiased quadratic risk estimator, before discussing the estimator choice. Finally, in Section 4, we illustrate the effectiveness of our reconstruction algorithm on a simulated set of pMRI data.

2. PROBLEM STATEMENT

2.1. Notations

We consider here pMRI images generated by considering L acquisition coils and a reduction factor R. Let \( X \times Y \) be the dimensions of the full Field Of View (FOV) image, along the respective phase and frequency directions. The reduced FOV image size in phase encoding direction is thus given by \( \Delta X = \frac{X}{R} \). \( \mathbb{B} \) will denote the set of spatial indices in the reduced FOV image.

2.2. Observation model

At each position \( x \in \mathbb{B} \) we observe an L-dimensional data vector \( d(x) = [d_1(x), \ldots, d_L(x)]^\top \), which is obtained by applying an \( L \times R \) sensitivity matrix \( S(x) \) given by

\[
S(x) = \begin{bmatrix}
S_1(x_1, x_2) & \ldots & S_1(x_1, x_2 + (R - 1)\Delta X) \\
\vdots & \ddots & \vdots \\
S_L(x_1, x_2) & \ldots & S_L(x_1, x_2 + (R - 1)\Delta X)
\end{bmatrix},
\]

to an R-dimensional vector of pixel values in the original (unknown) image \( \rho(x) = [\rho(x_1, x_2), \rho(x_1, x_2 + \Delta X), \ldots, \rho(x_1, x_2 + (R - \Delta X))]. \)
The objective of this work is to recover $X$ so that Model (1) can be reexpressed under the form

$$\mathbf{d}(x) = \mathbf{S}(x)\rho(x) + \mathbf{n}(x), \quad \forall x \in \mathbb{B}. \quad (1)$$

The objective of this work is to recover $\rho$ from $\mathbf{d}$ knowing $\mathbf{S}$.

### 3. PROPOSED METHOD

#### 3.1. Complex nature of the data

In MRI, the acquired data are complex-valued even if the magnitude is only considered for visualization purpose. Let $\Re$ and $\Im$ be the subscripts indicating the real and imaginary parts (namely, $\Re\{\cdot\}$ and $\Im\{\cdot\}$) of the data. The observed data can thus be expressed by

$$
\begin{bmatrix}
\mathbf{d}_R(x) \\
\mathbf{d}_I(x)
\end{bmatrix} =
\begin{bmatrix}
\mathbf{S}_R(x) & -\mathbf{S}_I(x) \\
\mathbf{S}_I(x) & \mathbf{S}_R(x)
\end{bmatrix}
\begin{bmatrix}
\rho_R(x) \\
\rho_I(x)
\end{bmatrix} +
\begin{bmatrix}
\mathbf{n}_R(x) \\
\mathbf{n}_I(x)
\end{bmatrix}
$$

so that Model (1) can be reexpressed under the form

$$\mathbf{d}_C(x) = \mathbf{S}_C(x)\rho_C(x) + \mathbf{n}_C(x), \quad \forall x \in \mathbb{B}, \quad (2)$$

where the involved vectors and matrices are real-valued. Assuming that we are able to compute the pseudo-inverse $\mathbf{S}^\dagger_C(x)$ of $\mathbf{S}_C(x)$ defined by

$$\mathbf{S}^\dagger_C(x) = \left( \mathbf{S}_C(x)^\top \mathbf{S}_C(x) \right)^{-1} \mathbf{S}_C(x)^\top, \quad (3)$$

it follows that $\tilde{\rho}_C(x) = \rho_C(x) + \tilde{\mathbf{n}}_C(x)$ where

$$\tilde{\rho}_C(x) = \mathbf{S}^\dagger_C(x)\mathbf{d}_C(x) \quad \text{and} \quad \tilde{\mathbf{n}}_C(x) = \mathbf{S}^\dagger_C(x)\mathbf{n}_C(x).$$

Furthermore, the following relations regarding the second-order statistics of the random field $\tilde{\mathbf{n}}_C(x)$ will be useful for the quadratic risk computation:

$$\text{Cov}\{\tilde{\mathbf{n}}_C(x), \tilde{\mathbf{n}}_C(x')\} = \mathbf{S}^\dagger_C(x)\mathbf{S}_C(x')^\top \Psi_C \Psi_C^\top \Psi_C \mathbf{S}_C^\dagger(x') \quad (4)$$

$$\text{Cov}\{\tilde{\mathbf{n}}_C(x), \mathbf{n}_C(x')\} = \mathbf{S}^\dagger_C(x)\mathbf{S}_C(x')^\top \Psi_C, \quad (5)$$

where $\text{Cov}\{\mathbf{n}_C(x), \mathbf{n}_C(x')\} = \Psi_C \delta(x - x'), \forall (x, x') \in \mathbb{B}^2$.

#### 3.2. Derivation of the SUREST estimator

For every $x \in \mathbb{B}$ and $k \in \{1, \ldots, K\}$, let $\varphi_k(x) \in \mathbb{R}^{2L}$ be an analysis vector and let $\mathbf{\tilde{\varphi}}_k(x) \in \mathbb{R}^{2K}$ be a synthesis vector. In the case of a decomposition onto a basis of the FOV image, $K = XY$, and the need of $2K$ functions stems from the fact that the real and imaginary parts of the data are considered. We choose an estimator of the form

$$\tilde{\rho}_C(x) = \sum_{k=1}^{2K} \tilde{\theta}_k \mathbf{\tilde{\varphi}}_k(x) \quad (6)$$

where $\tilde{\mathbf{\varphi}}_k(x) \in \mathbb{R}^{2L}$ and $\tilde{\theta}_k : \mathbb{R} \to \mathbb{R}$ is some differentiable estimating function (the choice of this function will be discussed in Section 3.4). Then the estimator of $\rho_C(x)$ is given by

$$\hat{\rho}_C(x) = \sum_{k=1}^{2K} \tilde{\theta}_k \mathbf{\tilde{\varphi}}_k(x) \quad (7)$$

The objective now is to compute the estimator parameters that minimize the quadratic risk.

#### 3.3. Unbiased risk estimate

The risk corresponding to a MSE estimation is defined as: $\forall x \in \mathbb{B}$,

$$E \left\{ \left\| \tilde{\rho}_C(x) - \rho_C(x) \right\|^2 \right\} = E \left\{ \left\| \tilde{\rho}_C(x) - \rho_C(x) + \tilde{\mathbf{n}}_C(x) \right\|^2 \right\} = E \left\{ \left\| \tilde{\rho}_C(x) - \tilde{\rho}_C(x) + \tilde{\mathbf{n}}_C(x) \right\|^2 \right\} + 2E \left\{ \tilde{\rho}_C(x)^\top \tilde{\mathbf{n}}_C(x) \right\} - 2E \{ \tilde{\rho}_C(x)^\top \tilde{\mathbf{n}}_C(x) \} + E \left\{ \left\| \tilde{\mathbf{n}}_C(x) \right\|^2 \right\}. \quad (8)$$

Here, we focus on the second term which, by using (7), reads:

$$E \left\{ \tilde{\rho}_C(x)^\top \tilde{\mathbf{n}}_C(x) \right\} = \sum_{k=1}^{2K} \tilde{\varphi}_k(x)^\top E \{ \tilde{\theta}_k \tilde{\mathbf{\varphi}}_k(x) \}. \quad (9)$$

In addition, applying the analysis vectors $\varphi_k(x)$ to $\mathbf{d}_C(x)$ yields coefficients $w_k$ which can be decomposed as $w_k = u_k + v_k$, where

$$u_k = \sum_{x \in \mathbb{B}} \varphi_k(x)^\top \mathbf{S}_C(x) \rho_C(x)$$

$$v_k = \sum_{x \in \mathbb{B}} \varphi_k(x)^\top \mathbf{n}_C(x),$$

so that $\tilde{\rho}_k$ can be rewritten as

$$\tilde{\rho}_k = \theta_k(u_k) = \theta_k(u_k + v_k). \quad (10)$$

Then, by applying Stein’s principle [1], it follows that:

$$E \left\{ \tilde{\rho}_k \tilde{\mathbf{n}}_C(x) \right\} = E \left\{ \theta_k(u_k) \right\} E \{ \tilde{\mathbf{n}}_C(x) \}. \quad (11)$$

After some tedious calculations, it follows that the quadratic risk estimate reads:

$$E \left\{ \left\| \tilde{\rho}_C(x) - \rho_C(x) \right\|^2 \right\} = E \left\{ \left\| \tilde{\rho}_C(x) - \tilde{\rho}_C(x) \right\|^2 \right\} + 2 \sum_{k=1}^{2K} E \left\{ \theta_k(u_k) \right\} \tilde{\varphi}_k(x)^\top \mathbf{S}^\dagger_C(x) \Psi_C \varphi_k(x)$$

$$- \text{tr} \left( \mathbf{S}^\dagger_C(x) \Psi_C \mathbf{S}^\dagger_C(x)^\top \right).$$

An unbiased estimate of the resulting global MSE is:

$$\hat{\mathcal{E}}(\hat{\rho}_C - \rho_C) = \hat{\mathcal{E}}(\tilde{\rho}_C - \tilde{\rho}_C) + \Delta \quad (12)$$

where

$$\hat{\mathcal{E}}(\tilde{\rho}_C - \tilde{\rho}_C) = \frac{1}{XY} \sum_{x \in \mathbb{B}} \left\| \tilde{\rho}_C(x) - \tilde{\rho}_C(x) \right\|^2$$

and

$$\Delta = \frac{2}{XY} \sum_{k=1}^{2K} \theta_k(u_k) \tilde{\varphi}_k(x)^\top \mathbf{S}^\dagger_C(x) \Psi_C \varphi_k(x)$$

$$- \frac{1}{XY} \sum_{x \in \mathbb{B}} \text{tr} \left( \mathbf{S}^\dagger_C(x) \Psi_C \mathbf{S}^\dagger_C(x)^\top \right).$$

Note that the risk estimate is expressed in terms of observed data only. Now, the next step consists of specifying the form of the estimating functions $\theta_k$ in order to find the parameters that minimize the global error measure $\hat{\mathcal{E}}(\cdot)$. 


3.4. Estimating function choice

Assume that the coefficients \((w_k)_{1 \leq k \leq 2K}\) are classified according to \(M \in \mathbb{N}^+\) distinct nonempty index subsets (e.g. wavelet subbands) \(K_m, m \in \{1, \ldots, M\}\). For every \(m \in \{1, \ldots, M\}\) and \(k \in K_m\), we choose subband-dependent estimating functions of the form:

\[
\theta_k(w_k) = \sum_{i=1}^{l_m} a_{m,i} f_{m,i}(w_k)
\]

(12)

where \((a_{m,i})_{1 \leq i \leq l_m}\) are scalar real-valued weighting factors and, for every \(i \in \{1, \ldots, l_m\}\), \(f_{m,i} : \mathbb{R} \rightarrow \mathbb{R}\) is a differentiable function.

These functions are the Linear Expansion of Thresholds (LET) estimating functions introduced in [7, 8] and correspond to a linear combination of \(I_m \in \mathbb{N}^+\) given univariate functions \(f_{m,i}\) applied to \(w_k\) (note that we use the same estimating function for a given \(K_m\)).

3.5. Computation of \((a_{m,i})_{1 \leq i \leq l_m}\) parameters

As mentioned earlier, we want to compute the parameters \((a_{m,i})_{1 \leq i \leq l_m}\) that minimize the quadratic risk in (11). It can be shown that this amounts to solving the following set of linear equations:

\[
\forall m \in \{1, \ldots, M\}, \forall i \in I_m, \quad \sum_{n=1}^{M} a_{n,j} \sum_{x \in \mathbb{R}} \beta_{n,j}(x)^\top \beta_{n,j}(x) = \sum_{x \in \mathbb{R}} \beta_{m,i}(x)^\top \tilde{\rho}_c(x) - \sum_{k \in K_m} f_{m,i}(w_k) \sum_{x \in \mathbb{R}} \tilde{\varphi}_k(x)^\top S_C(x) \Psi_C \varphi_k(x)
\]

where \(\beta_m(x) = \sum_{k \in K_m} f_{m,i}(w_k) \tilde{\varphi}_k(x)\).

4. SIMULATION RESULTS

4.1. Context

In our experiments, the families of analysis/synthesis functions are chosen such that

\[
\varphi_k(x) = S_C(x) \left( S_C(x)^\top S_C(x) + \lambda I \right)^{-1} \psi_k(x)
\]

\[
\tilde{\varphi}_k(x) = \tilde{\psi}_k(x)
\]

where \(\lambda > 0\), \(I\) is the identity matrix and the components of vector \(\psi_k(x)\) (resp. \(\tilde{\psi}_k(x)\)) are obtained from shifted versions of a 2D separable wavelet analysis function (resp. synthesis function).

A scheme summarizing the proposed reconstruction approach is given in Fig. 1.

Concerning the estimating functions, we choose \(I_m = 2\). \(f_{m,1}\) is the identity function and

\[
\forall \rho \in \mathbb{R}, \quad f_{m,2}(\rho) = \left(1 - \exp \left( -\frac{\rho}{\omega \sigma_m} \right) \right) \rho
\]

where \(\omega \in ]0, \infty[\) and \(\sigma_m\) is the standard deviation of \((\nu_k)_{k \in K_m}\).

These LET estimating functions were introduced in [7].

We compare our results with SENSE reconstruction [9], which is based on a weighted least-squares estimator:

\[
\tilde{\rho}_{\text{SENSE}}(x) = \left( (S'(x))^\top \Psi^{-1} S(x) \right)^{-1} (S'(x))^\top \Psi^{-1} d(x).
\]

Fig. 1. Flowchart of the proposed reconstruction method.

4.2. Numerical results

To assess the performance of the proposed method, we simulated synthetic data from a full FOV reference image displayed in Fig. 2(a) according to model (1). Note that the dynamic range of the absolute value of the pixels is large.

The number of antennas is here equal to \(L = 8\) and the reduction factor is \(R = 4\). The covariance matrix of the noise \(\psi\) is chosen to be diagonal (no between coil correlation) and all the diagonal entries are chosen equal to \(\sigma^2\) (homoscedasticity). Next, we compared our results with those provided by the classical SENSE reconstruction. Here, an orthonormal wavelet analysis (symlets of length 8) is performed on 4 resolution levels. The regularization parameter \(\lambda\) was chosen empirically. Typical values around 0.001 were observed to provide accurate and reliable results. The performance was measured in terms of Peak Signal-to-Noise Ratio (PSNR) as reported in Table 1. In short, our method enables a significant gain in PSNR values with respect to the SENSE approach.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SENSE</td>
<td>36.29</td>
<td>35.48</td>
<td>38.75</td>
</tr>
<tr>
<td>Proposed</td>
<td>(\lambda = 10^{-4})</td>
<td>38.28</td>
<td>36.77</td>
</tr>
<tr>
<td></td>
<td>(\lambda = 10^{-5})</td>
<td>38.30</td>
<td>36.79</td>
</tr>
<tr>
<td></td>
<td>(\sigma^2 = 12.5 \times 10^{-6})</td>
<td>38.30</td>
<td>36.79</td>
</tr>
<tr>
<td>SENSE</td>
<td>42.94</td>
<td>42.40</td>
<td>45.58</td>
</tr>
<tr>
<td>Proposed</td>
<td>(\lambda = 10^{-4})</td>
<td>43.57</td>
<td>42.74</td>
</tr>
<tr>
<td></td>
<td>(\lambda = 10^{-5})</td>
<td>43.57</td>
<td>42.74</td>
</tr>
<tr>
<td></td>
<td>(\sigma^2 = 2 \times 10^{-4})</td>
<td>43.57</td>
<td>42.74</td>
</tr>
</tbody>
</table>

Interestingly, the larger the noise, the stronger the PSNR improvements we observed in favour of the proposed approach. For visualization purposes, the reconstructed images by SENSE and Stein-LET methods are shown in Fig. 2(b)-(c) when \(\sigma^2 = 5 \times 10^{-6}\). The main differences can be observed in the center of the images where more residual noise is found for SENSE reconstruction. Note that the Stein-LET reconstruction in Fig. 2(c) is achieved in about 15 s. on a Xeon(R) CPU X5450@3.00GHz.
5. CONCLUSION

We proposed a new reconstruction method based on Stein’s principle and LET functions which was applied to pMRI. In this context, our approach provides more accurate results than the classical SENSE alternative. In our approach, most of the parameters are set in an unsupervised manner. In addition, the computational complexity is quite reasonable. In the future, we plan to validate this restoration method in a more realistic context i.e. without knowing the amount of noise and using an unperfectly known sensitivity matrix. Furthermore, we will consider 3D decomposition to extend the present contribution and enable the reconstruction of all slices simultaneously. It would be also interesting to apply this approach to other reconstruction strategies [12, 13].

6. REFERENCES