

# EXTENDED FORMS OF GEMAN & YANG ALGORITHM: APPLICATION TO MRI RECONSTRUCTION

*Sabine Husse\**, *Yves Goussard\** & *Jérôme Idier†*

\*École Polytechnique  
Institut de génie biomédical  
C.P. 6079, succ. Centre-ville  
Montréal (QC), H3C 3A7, Canada  
firstname.lastname@polymtl.ca

†IRCCyN  
BP 92101  
1 rue de la Noë  
44321 Nantes Cedex 3, France  
jerome.idier@irccyn.ec-nantes.fr

## ABSTRACT

The main contribution of this communication is the derivation of a generalized form of the Geman & Yang construction for minimization of convex, non-quadratic criteria. The generalization provides means of obtaining a normal matrix with predefined structure. We show that this property can be used to improve the numerical efficiency of edge preserving MRI reconstruction. The improvement is assessed experimentally using synthetic data. In order to illustrate the practicality of the method, an example of large size, 3D real data processing is also provided.

## 1. INTRODUCTION AND PROBLEM STATEMENT

The starting point of this work is the problem of image reconstruction in magnetic resonance imaging (MRI). In order to improve the performance of standard image reconstruction techniques, an inverse problem approach is adopted in which the data formation process is modeled by the following relation:

$$\mathbf{y} = \Phi_2 \mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{y}$  and  $\mathbf{x}$  respectively denote the vector of complex data measured in the  $k$ -space, and the set of pixels of the unknown image rearranged in vector form through, e.g., a raster scan.  $\Phi_2$  represents the 2D discrete Fourier transform matrix and  $\mathbf{n}$  is a noise vector that accounts for unmodeled phenomena. The reconstructed image is defined as  $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} J(\mathbf{x})$  with:

$$J(\mathbf{x}) = J_0(\mathbf{x}) + F(\mathbf{x}) \quad (2)$$

where  $J_0(\mathbf{x}) = \|\mathbf{y} - \Phi_2 \mathbf{x}\|^2$  is a least-squares criterion which accounts for the fidelity of the solution to the measured data, and where penalty term  $F(\mathbf{x})$  expresses prior information about the solution.

Here, we assume  $F(\mathbf{x})$  to be an *edge-preserving* function of the form:

$$F(\mathbf{x}) = \sum_{l=1}^2 \lambda_l \sum_{c \in \mathcal{C}_l} \varphi(\mathbf{d}_c^T \mathbf{x}) \quad (3)$$

where  $\cdot^T$  denotes the transconjugation operation, where  $\mathcal{C}_1$  (resp.  $\mathcal{C}_2$ ) represents the set of horizontal (resp. vertical) adjacent pixel pairs called *cliques* and where each vector  $\mathbf{d}_c$  has a support region restricted to clique  $c$ . In order for  $F$  to be edge-preserving while yielding a *convex* criterion  $J$ , function  $\varphi(\cdot)$  is assumed to be strictly convex, coercive with quadratic behavior near 0 and linear behavior toward infinity [1]. Several functions meet these criteria and here  $\varphi(\cdot)$  is defined as:

$$\varphi(u) = \sqrt{\delta^2 + |u|^2} \quad (4)$$

where  $\delta$  is a scaling factor that determines the transition between the quadratic and linear regions.

As reported in [2, 3], criterion  $J(\mathbf{x})$  can be minimized in a very efficient manner using Geman & Yang (GY)-type algorithms. Derivation of these procedures is based upon the introduction of the following pair of auxiliary functions:

$$\begin{aligned} G_{l\alpha}(\mathbf{u}) &= \frac{\mathbf{u}^T \mathbf{u}}{2} - \alpha \sum_{c \in \mathcal{C}_l} \varphi(u_c) \\ G_{l\alpha}^*(\mathbf{b}) &= \frac{\mathbf{b}^T \mathbf{b}}{2} + \psi_{l\alpha}(\mathbf{b}) \end{aligned} \quad (5)$$

Assume that  $\alpha$  is chosen such that  $G_{l\alpha}(\cdot)$  is convex; this condition is met when  $\alpha < \sup_u \{\varphi''(u)\}^{-1}$  [4, 5]. For our choice of  $\varphi$ ,  $G_{l\alpha}$  is convex when  $\alpha \in ]0, \delta[$ . Then  $G_{l\alpha}$  and  $G_{l\alpha}^*$  form a pair of convex conjugate functions [6] provided that  $\psi_{l\alpha}$  be defined as:

$$\psi_{l\alpha}(\mathbf{b}) = \sup_{\mathbf{u}} \left\{ \alpha \sum_{c \in \mathcal{C}_l} \varphi(u_c) - \frac{(\mathbf{u} - \mathbf{b})^T (\mathbf{u} - \mathbf{b})}{2} \right\} \quad (6)$$

It follows from the convex duality relationships that the minimizer of criterion  $J$  can be found by global minimization of *augmented criterion*  $K_\alpha(\mathbf{x}, \mathbf{b})$  with respect to  $\mathbf{x}$  and

© 2004 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

$\mathbf{b}$ ,  $K_\alpha$  being defined as:

$$K_\alpha(\mathbf{x}, \mathbf{b}) = J_0(\mathbf{x}) + \sum_{l=1}^2 \lambda_l \frac{(\mathbf{u}_l - \mathbf{b}_l)^T (\mathbf{u}_l - \mathbf{b}_l)}{2\alpha} + \frac{1}{\alpha} \psi_{l\alpha}(\mathbf{b}_l)$$

with  $\mathbf{u}_l = \mathbf{D}_l \mathbf{x}$ ;  $\mathbf{D}_l = [d_1^T \dots d_{M_l}^T]$  and  $M_l = |C_l|$ ;  $l \in \{1, 2\}$  ( $\mathbf{D}_1$  and  $\mathbf{D}_2$  represent the first difference operators in the horizontal and vertical directions, respectively). In addition, the global minimum of  $K_\alpha$  can be reached by *iterative block-minimization* of  $K_\alpha$  with respect to  $\mathbf{x}$  and  $\mathbf{b}$ , alternatively. This is particularly interesting because each partial minimization operation is simple. Since  $K_\alpha$  is quadratic with respect to  $\mathbf{x}$ , the minimizer can be obtained in closed-form and takes the expression:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1} \left( \Phi_2^{-1} \mathbf{y} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{D}_l^T \mathbf{b}_l \right) \quad (7)$$

$$\text{with} \quad \mathbf{A} = \mathbf{I} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{D}_l^T \mathbf{D}_l \quad (8)$$

where images are assumed to be of size  $N_c \times N_l$ . Minimization of  $K_\alpha$  with respect to  $\mathbf{b}$  follows from the convex duality relationships and the solution can also be expressed in closed-form as:

$$\hat{\mathbf{b}}_l = \mathbf{u}_l - \alpha \frac{\mathbf{u}_l}{\sqrt{\delta^2 + |\mathbf{u}_l|^2}} \quad l \in \{1, 2\} \quad (9)$$

where the ratio in the right hand side of (9) should be understood as a componentwise operation, and where the above expression is valid for real or complex variables  $\mathbf{x}$  and  $\mathbf{b}$ <sup>1</sup>.

One of the main advantages of the GY procedure lies in the fact that normal matrix  $\mathbf{A}$  remains constant along the iterations and therefore needs to be inverted only once. However, for medical images of realistic sizes, the one-time inversion of  $\mathbf{A}$  and the matrix multiplication in (7) at each iteration still represent a significant computational burden. We now turn to the derivation of formulations of the GY construction which provide means for faster implementation of these algorithms.

## 2. NEW FORMS OF THE GY CONSTRUCTION

### 2.1. GY standard formulation and application to MRI

The expression given in (8) shows that normal matrix  $\mathbf{A}$  is Hermitian with size  $(N_c N_l \times N_c N_l)$ . It should be noticed that the terms  $\mathbf{D}_l^T \mathbf{D}_l$  are quasi block-circulant. Therefore, the whole matrix  $\mathbf{A}$  is quasi block-circulant and can be expressed as:

$$\mathbf{A} = \mathbf{C} - \Delta$$

<sup>1</sup>This result can be derived by considering complex variables as 2-component vectors.

where  $\mathbf{C}$  is block-circulant and  $\Delta$  is a low-rank perturbation matrix. Due to the Hermitian symmetry of  $\mathbf{A}$ ,  $\Delta$  can be further factored as:

$$\Delta = \mathbf{U} \mathbf{U}^T$$

where  $\mathbf{U}$  is a rectangular matrix with typical size equal to  $(N_c N_l \times N_c)$  or  $(N_c N_l \times N_l)$ . Inversion of  $\mathbf{A}$  can thus be carried out in an efficient manner using the matrix inversion lemma [7]. Using the fact that  $\mathbf{C}$  can be diagonalized by applying the 2D Fourier operator  $\Phi_2$  which is also the projection matrix of the MRI reconstruction problem, the amount of computation required for evaluating (7) can be reduced by two orders of magnitude [3]. However, for realistic MRI images, the computational burden per iteration remains significant due to the rather large size of  $\mathbf{U}$ . Additional reduction of the amount of computations can be achieved by neglecting perturbation matrix  $\Delta$  and assuming that  $\mathbf{A} \approx \mathbf{C}$  [3], at the expense of undesirable correlations between opposite boundaries of the reconstructed image. Our goal in deriving the next two techniques is to achieve a high numerical efficiency without undesirable boundary effects.

### 2.2. Generalized form of the GY construction

In order to obtain a normal matrix that lends itself to efficient inversion, our approach consists of introducing an additional degree of freedom in the GY construction by replacing the canonical inner product which appears in the convex duality relationships by a more general inner product defined as:

$$\langle \mathbf{u} | \mathbf{v} \rangle = \mathbf{u}^T \mathbf{M} \mathbf{v}$$

where  $\mathbf{M}$  is a symmetric positive definitive matrix. The relationships between convex function  $f$  and its convex conjugate  $f^*$  become<sup>2</sup>:

$$\begin{aligned} f^*(\mathbf{v}) &= \sup_{\mathbf{u}} \{ \mathbf{v}^T \mathbf{M} \mathbf{u} - f(\mathbf{u}) \} \\ f(\mathbf{u}) &= \sup_{\mathbf{v}} \{ \mathbf{u}^T \mathbf{M} \mathbf{v} - f^*(\mathbf{v}) \} \end{aligned}$$

We introduce the following pair of auxiliary functions:

$$\begin{aligned} G_{l\alpha}(\mathbf{u}) &= \frac{\mathbf{u}^T \mathbf{M} \mathbf{u}}{2} - \alpha \sum_{c \in C_l} \varphi(u_c) \\ G_{l\alpha}^*(\mathbf{b}) &= \frac{\mathbf{b}^T \mathbf{M} \mathbf{b}}{2} + \psi_{l\alpha}(\mathbf{b}) \end{aligned} \quad (10)$$

Provided that  $\alpha$  is chosen such that  $G_{l\alpha}(\cdot)$  be convex, which holds for  $\alpha \in ]0, \delta[$ ,  $G_{l\alpha}$  and  $G_{l\alpha}^*$  still form a pair of convex conjugate functions when  $\psi_{l\alpha}$  is defined as:

$$\psi_{l\alpha}(\mathbf{b}) = \sup_{\mathbf{u}} \left\{ \alpha \sum_{c \in C_l} \varphi(u_c) - \frac{(\mathbf{u} - \mathbf{b})^T \mathbf{M} (\mathbf{u} - \mathbf{b})}{2} \right\} \quad (11)$$

<sup>2</sup>These relationships can be derived in an elementary manner by applying standard convex duality formulas [6] to convex function  $g(\mathbf{w}) = f(\mathbf{T}^{-1} \mathbf{w})$ ;  $f$  convex and  $\mathbf{T}$  such as  $\mathbf{T}^T \mathbf{T} = \mathbf{M}$ .

The augmented criterion can then be expressed as:

$$K_\alpha(\mathbf{x}, \mathbf{b}) = J_0(\mathbf{x}) + \sum_{l=1}^2 \lambda_l \frac{(\mathbf{u}_l - \mathbf{b}_l)^T \mathbf{M}(\mathbf{u}_l - \mathbf{b}_l)}{2\alpha} + \frac{1}{\alpha} \psi_{l\alpha}(\mathbf{b}_l)$$

and the update equations of  $\mathbf{x}$  and  $\mathbf{b}$  take the following form:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1} \left( \Phi_2^{-1} \mathbf{y} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{D}_l^T \mathbf{M} \mathbf{b}_l \right) \quad (12)$$

$$\text{with } \mathbf{A} = \mathbf{I} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{D}_l^T \mathbf{M} \mathbf{D}_l \quad (13)$$

$$\text{and } \mathbf{M} \hat{\mathbf{b}}_l = \mathbf{M} \mathbf{u}_l - \alpha \frac{\mathbf{u}_l}{\sqrt{\delta^2 + |\mathbf{u}_l|^2}} \quad (14)$$

As shown by (13), the main interest of this method is to provide a way of imposing a specific structure to normal matrix  $\mathbf{A}$  through an appropriate choice of  $\mathbf{M}$ . As an illustration, we now show how  $\mathbf{M}$  can be chosen so that  $\mathbf{A}$  be block-circulant.

As mentioned previously,  $\mathbf{D}^T \mathbf{D}$  is quasi block-circulant and can be written as  $\mathbf{D}^T \mathbf{D} = \mathbf{C} - \mathbf{U} \mathbf{U}^T$ , where  $\mathbf{C}$  is block-circulant. We look for a matrix  $\mathbf{M}$  such that:

$$\mathbf{D}^T \mathbf{M} \mathbf{D} = \mathbf{C}$$

Let  $\mathbf{S}$  be a matrix that satisfies  $\mathbf{D}^T \mathbf{S} = \mathbf{U}^3$  and define:

$$\mathbf{M} = \mathbf{I} + \mathbf{S} \mathbf{S}^T \quad (15)$$

By construction,  $\mathbf{M}$  is symmetric positive definite and:

$$\mathbf{D}^T \mathbf{M} \mathbf{D} = \mathbf{D}^T \mathbf{D} + \mathbf{U} \mathbf{U}^T = \mathbf{C}$$

Therefore,  $\mathbf{A}$  can be diagonalized by application of Fourier operator  $\Phi_2$  which makes it easy to invert  $\mathbf{A}$  and update  $\mathbf{x}$  according to (12). In addition, update of  $\mathbf{M} \mathbf{b}_l$  is made simple by using the decomposition of  $\mathbf{M}$  given in (15). As shown in the next section, the resulting procedure presents virtually the same numerical cost per iteration as the approximate form of the standard GY procedure, without the disadvantage of undesirable boundary effects.

### 2.3. Vectorial form of the GY construction

This form was proposed in [8, 4] and yields an even simpler form of normal matrix  $\mathbf{A}$  in the case of MRI reconstruction. The auxiliary functions are defined as:

$$\begin{aligned} G_\alpha(\mathbf{x}) &= \frac{\mathbf{x}^T \mathbf{x}}{2} - F_\alpha(\mathbf{x}) \\ G_\alpha^*(\mathbf{b}) &= \frac{\mathbf{b}^T \mathbf{b}}{2} + \psi_\alpha(\mathbf{b}) \end{aligned} \quad (16)$$

$G_\alpha$  is convex when  $\alpha < \sup_{\mathbf{x}} \{\rho(\nabla^2 F(\mathbf{x}))\}^{-1}$ , where  $\rho(\cdot)$  denotes the spectral radius. Here, this condition is met when

<sup>3</sup>Simple solutions to this equation exist when  $\mathbf{D}$  is a first difference matrix.

$\alpha \in ]0, \delta/\rho(\mathbf{D}^t \mathbf{D})[$  and in this case,  $G_\alpha$  and  $G_\alpha^*$  form a pair of convex conjugate functions provided that:

$$\psi_\alpha(\mathbf{b}) = \sup_{\mathbf{x}} \left\{ F_\alpha(\mathbf{x}) - \frac{(\mathbf{x} - \mathbf{b})^T (\mathbf{x} - \mathbf{b})}{2} \right\} \quad (17)$$

The augmented criterion takes the following expression:

$$K_\alpha(\mathbf{x}, \mathbf{b}) = J_0(\mathbf{x}) + \sum_{l=1}^2 \lambda_l \frac{(\mathbf{x} - \mathbf{b}_l)^T (\mathbf{x} - \mathbf{b}_l)}{2\alpha} + \frac{1}{\alpha} \psi_\alpha(\mathbf{b}_l)$$

and the update equations of  $\mathbf{x}$  and  $\mathbf{b}$  take the following form:

$$\hat{\mathbf{x}} = \mathbf{A}^{-1} \left( \Phi_2^{-1} \mathbf{y} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{b}_l \right) \quad (18)$$

$$\text{with } \mathbf{A} = \mathbf{I} + \sum_{l=1}^2 \frac{\lambda_l}{2\alpha N_c N_l} \mathbf{I} \quad (19)$$

$$\text{and } \hat{\mathbf{b}}_l = \mathbf{x} - \alpha \mathbf{D}_l^T \frac{\mathbf{u}_l}{\sqrt{\delta^2 + |\mathbf{u}_l|^2}} \quad (20)$$

The main advantage of the vectorial form is that no matrix inversion is needed since  $\mathbf{A}$  is proportional to identity. This simplicity is obtained at the expense of a convergence speed lower than for the other procedures, as illustrated in the next section.

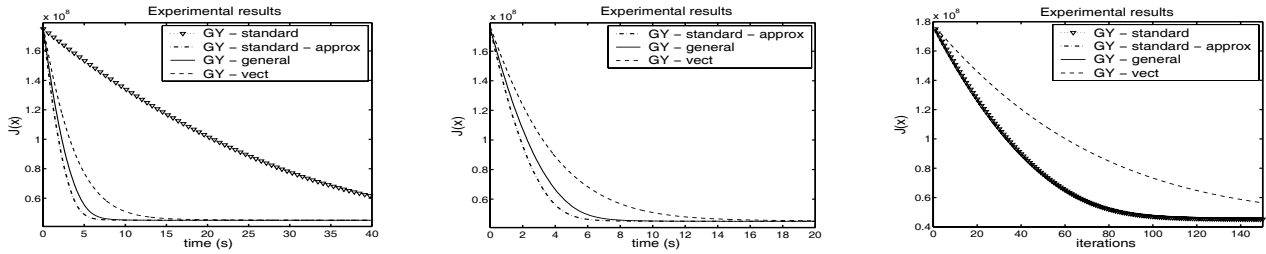
## 3. RESULTS

### 3.1. Experimental comparison on the convergence speed

In order to evaluate the effectiveness of the algorithms, tests were performed on synthetic data with a 2D object of size  $128 \times 128$ . The four methods, i.e., standard exact and approximate GY, generalized GY, and vectorial GY were implemented. Figure 1 shows the evolution of criterion  $J$  with respect to time (left and center) and with respect to the number of iterations (right) for the four methods. It can be observed that in terms of number of iterations, the vectorial form of GY performs poorly whereas the other three techniques exhibit very similar behaviors. However, in terms of time or, equivalently, in terms of amount of computations, the vectorial form of GY can be considered as an interesting option due to its acceptable convergence speed and ease of implementation, thanks to the simple expression of the normal matrix. The exact standard form of GY performs poorly whereas the approximate and generalized forms of GY present the best performance. The generalized form should certainly be preferred due to the absence of undesirable boundary effects.

### 3.2. Reconstruction on real data

In order to demonstrate the ability of these methods to process real world data, reconstruction of a  $110 \times 512 \times 384$  3D



**Fig. 1.** Convergence rate comparison. Left:  $J(x)$  vs. time; center:  $J(x)$  vs. time, zoom without the exact standard GY form; right:  $J(x)$  vs. number of iterations.

image of a knee<sup>4</sup> was performed. The data were processed in less than 100 minutes on a standard PC computer. Figure 2 shows a standard reconstruction (top) and the result provided by the GY approach (bottom). As expected, reconstruction with an edge preserving penalty term provides a significant improvement in noise reduction, especially in the cartilage area, while adequately preserving discontinuities.

#### 4. CONCLUSION

The major contribution of this communication is the derivation of a generalized form of the GY construction which provides means of obtaining a normal matrix predefined structure. This property was used to significantly improve the numerical efficiency of edge preserving MRI reconstruction. These techniques can thus be implemented with moderate computing power, thereby increasing their availability.

#### 5. REFERENCES

[1] Jérôme Idier, “Convex half-quadratic criteria and interacting auxiliary variables for image restoration,” *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1001–1009, July 2001.

[2] R. Deriche, P. Kornprobst, M. Nikolova, and M. Ng, “Half-quadratic regularization for MRI image restoration,” in *Proc. IEEE ICASSP*, 2003, pp. 585–588.

[3] Sabine Husse and Yves Goussard, “Image reconstruction in MRI: Regularized approach by Markov random fields,” in *Proceedings of the International Conference on Image Processing*, Barcelona, Spain, September 2003.

[4] Marc Allain, *Approche pénalisée en tomographie hélicoïdale. Application à la conception d’une prothèse*

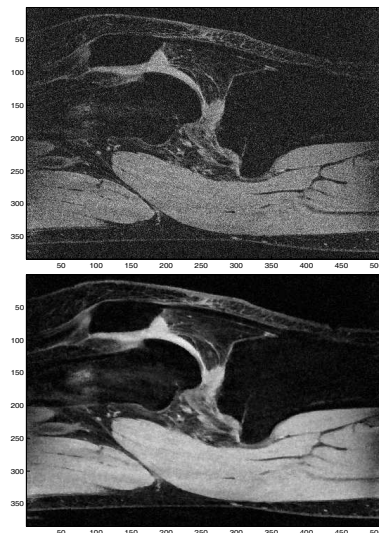
*personnalisée du genou*, Ph.D. thesis, Université de Paris-Sud, Orsay, France, December 2002.

[5] P. Ciuciu and J. Idier, “A half-quadratic block-coordinate descent method for spectral estimation,” *Signal Processing*, vol. 82, no. 7, pp. 941–959, 2002.

[6] R. Tyrell Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[7] A. K. Jain, “Fast inversion of banded Toeplitz matrices by circular decompositions,” *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-26, no. 2, pp. 121–126, April 1978.

[8] Myriam Fiani, Jérôme Idier, and Stéphane Gautier, “Algorithmes ART semi-quadratiques pour la reconstruction à partir de radiographies,” in *Actes 18<sup>e</sup> coll. GRETSI*, Toulouse, sep. 2001, pp. 183–186.



**Fig. 2.** Results with real data: image obtained with classical reconstruction (top) and with the GY approach (bottom).

<sup>4</sup>We thank Dr. Gilles Beaudoin, Centre hospitalier de l’université de Montréal, for providing these data.