

Performance Evaluation of MAP Algorithms with Different Penalties, Object Geometries and Noise Levels

Yu-Jung Tsai, *Student Member, IEEE*, Alexandre Bousse, Matthias J. Ehrhardt, Brian F. Hutton, *Senior Member, IEEE*, Simon Arridge and Kris Thielemans *Senior Member, IEEE*,

Abstract—A new algorithm (LBFGS-B-PC) which combines ideas of two existing convergent reconstruction algorithms, relaxed separable paraboloidal surrogate (SPS) and limited-memory Broyden-Fletcher-Goldfarb-Shanno with boundary constraints (LBFGS-B), is proposed. Its performance is evaluated in terms of log-posterior value and regional recovery ratio. The results demonstrate the superior convergence speed of the proposed algorithm to relaxed SPS and LBFGS-B, regardless of the noise level, activity distribution, object geometry, and penalties.

I. INTRODUCTION

IN emission tomography, *maximum a posteriori* (MAP) reconstruction is one of the strategies to obtain reconstructed images with desired properties, such as accurate reconstruction with low statistical noise [1]. MAP reconstruction maximises the posterior likelihood. Various optimization algorithms can be used to obtain this maximum. Since each algorithm approaches the MAP solution in a different way, the convergence rate varies between algorithms. For achieving quantitative accuracy in a clinically feasible reconstruction time, algorithms that converge in a relatively low number of iterations are preferable. The aim of this study is to propose a new algorithm LBFGS-B-PC which combines preconditioning (PC) with the use of a quasi-Newton optimization algorithm (LBFGS-B [4]). Its performance is evaluated and compared with relaxed SPS [3] and LBFGS-B.

II. METHOD

A. Objective Function and Preconditioning

MAP reconstruction is performed by maximizing a function Φ consisting of 2 parts, the likelihood L and the penalty

Manuscript received November 20, 2015. This work was supported by GE Healthcare, and by the National Institute for Health Research, University College London Hospitals Biomedical Research Centre.

Yu-Jung Tsai is with the Division of Medicine, University College London, London, UK (e-mail: yu-jung.tsai.14@ucl.ac.uk).

Alexandre Bousse is with the Institute of Nuclear Medicine, University College London Hospital, London, UK (e-mail: a.bousse@ucl.ac.uk).

Matthias J. Ehrhardt is with the Department of Computer Science, University College London, London, UK (e-mail: matthias.ehrhardt.11@ucl.ac.uk).

Brian F. Hutton is with the Institute of Nuclear Medicine, University College London Hospital, London, UK. He is also with the Centre for Medical Radiation Physics, University of Wollongong, Wollongong, Australia (e-mail: b.hutton@ucl.ac.uk).

Simon Arridge is with the Department of Computer Science, University College London, London, UK (e-mail: S.Arridge@cs.ucl.ac.uk).

Kris Thielemans is with the Institute of Nuclear Medicine, University College London Hospital, London, UK (e-mail: k.thielemans@ucl.ac.uk).

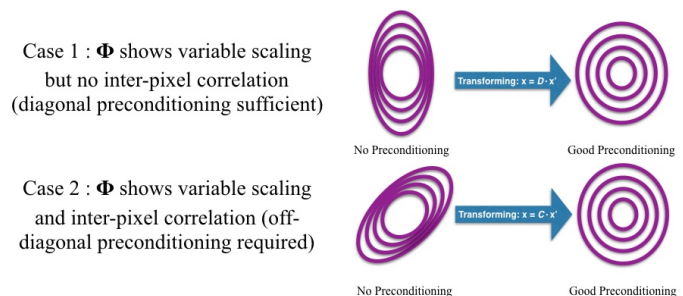


Fig. 1. Illustration of contour-plots of the objective function before and after preconditioning for different cases of problems.

function R with a parameter β controls the strength of the penalty.

$$\Phi = L - \beta R \quad (1)$$

In the optimization literature, preconditioning is used to achieve faster convergence rate. Fig. 1 illustrates different types of problems that need different preconditioners.

B. Optimization Algorithms

1) *Relaxed SPS*: Instead of considering the true objective function, the optimization can be transferred to separable paraboloidal surrogates (SPS). Ahn [2] introduced a modification of the original SPS algorithm [3] using relaxation to be able to cope with subsets. This “relaxed SPS” version uses a precomputed denominator related to the Hessian. Given the current image estimate \mathbf{x}^{old} and the measured data g , the image is updated by:

$$\mathbf{x}^{new} = \mathbf{x}^{old} + \zeta D^2 \nabla \Phi, \quad \zeta = \frac{\alpha}{1+\gamma} \quad \text{and} \quad (2)$$

$$D^2 = \text{diag}\{P^t \text{diag}\{\frac{1}{g+1}\} P \cdot \mathbf{1}\}$$

where D represents the preconditioner, ζ is the relaxation parameter, P is the system matrix and $\nabla \Phi$ is the gradient of the objective function.

The algorithm can be interpreted as preconditioned gradient descent using a diagonal approximation of the Hessian as preconditioner. The algorithm is independent of global scaling¹ and the preconditioner can be precomputed. However, as only a diagonal preconditioner is used (as in Fig. 1 Case 1), potential slow convergence is expected.

¹The global scaling is in the sense that $\tilde{g} = P\tilde{x} = \lambda P\tilde{x}$ leads to the same results for all $\lambda > 0$.

2) *LFBGS-B*: This is a quasi-Newton algorithm that estimates non-diagonal approximations of the Hessian by using previous iterations. The optimization is achieved by minimizing the objective function Φ along a search line p which is constructed by the LFBGS-B algorithm [4]:

$$\mathbf{x}^{new} = \mathbf{x}^{old} + \delta^* \mathbf{p}, \quad \delta^* = \arg \min_{\delta \geq 0} \Phi(\mathbf{x}^{old} + \delta \mathbf{p}), \quad (3)$$

where where $\mathbf{p} = \tilde{H}^{-1} \nabla \Phi(\mathbf{x}^{old})$ with \tilde{H} the approximation of the Hessian at the current iteration. Although the algorithm can cope with the Case 2 in Fig. 1 in a low-dimensional problem, for image reconstruction the number of variables to estimate is very large and the approximated Hessian can potentially lead to slow convergence. Besides, the line search part of the algorithm makes it sensitive to global scaling.

3) *LFBGS-B-PC*: We propose to use analytical information about the objective function to circumvent the above mentioned deficiencies of LFBGS-B. In particular, LFBGS-B-PC uses LFBGS-B in a transformed coordinate system that corresponds to the relaxed SPS preconditioner D :

$$\Phi'(\mathbf{x}') = \Phi(\mathbf{x}), \quad \mathbf{x}' = D^{-1} \mathbf{x} \quad (4)$$

where \mathbf{x}' and Φ' are the image and objective function in the transformed coordinate, respectively. By combining ideas of the two algorithms, we intend to achieve faster convergence and being independent of scale.

C. Penalty Functions

We use Gibbs-type penalties which penalize the difference between pixels within a given neighborhood N :

$$R(x) = \frac{1}{2} \sum_{k \in N} \omega_{jk} \psi(x_j - x_k) \quad (5)$$

where ω_{jk} indicates the weight between pixel j and its neighboring pixel, k . We compare results for 2 potential functions ψ : the Quadratic penalty (QP) and the Rescaled logcosh penalty (LP)

$$\psi(x) = \frac{1}{\rho^2} \log(\cosh(\rho x)) \quad (6)$$

The factor $\frac{1}{\rho^2}$ was introduced in order to have the same behaviour as the QP around 0. The prior is more edge-preserving for larger ρ . $\rho = 1.8$ in the experiments.

D. Data

The projection data were generated to simulate the GE Discovery STE in 2D. To study the effect caused by different phantom geometries, both cylindrical and XCAT phantom were used (Fig. 2). The corresponding attenuation maps can be found in Fig. 3.

1) *Cylindrical phantom*: The phantom contained 4 regions of interest (ROIs), 2 hot spots and 2 cold spots. The activity ratios of the cold and hot regions to the background were 0.5 and 2, respectively. For each group of activity levels, different attenuation materials simulating the effects of bone and soft tissue were applied to each spot. To assess possible noise effects, we generated three sets of data with different level of Poisson noise (total counts were 297 K, 594 K and 1.18 M respectively).

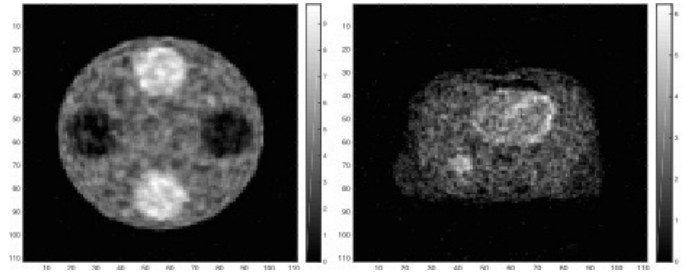


Fig. 2. Reconstructed images from relaxed SPS at 1000th iteration for cylindrical phantom with total counts 594 K (left) and XCAT phantom (right). Quadratic prior with $\beta = 0.1$ was applied for both images.

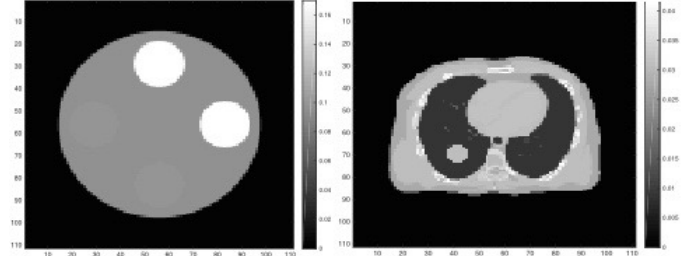


Fig. 3. The corresponding attenuation maps.

2) *XCAT phantom*: A slice from the XCAT torso phantom [5] was used for evaluating all the algorithms under a more realistic condition. We inserted a hot lesion into the right side of the lung region. The activity ratio of the lesion to the normal lung region was 2. Similarly with the cylindrical phantom, both non-uniform attenuation effect and Poisson noise were considered. However, only one noise level with total counts of 838 K was generated.

E. Reconstruction

All the data were reconstructed by relaxed SPS, LFBGS-B and the proposed LFBGS-B-PC with penalty factors of 0.1 and 0.3. The first algorithm had been implemented in STIR [6]. For the last two, we used the Fortran implementation from [4] but using STIR for calculating the objective function and its gradient. The initial image for all algorithms was the reconstructed image from maximum-likelihood expectation-maximization (MLEM) [7] at 1st iteration.

F. Analysis

The convergence rate of each algorithm was evaluated in terms of log-posterior value and normalized regional recovery ratio (RR). We define convergence of ROI values if the total RR reaches the value obtained at 1000 iterations of SPS with error less than 1%. The RR and total RR were calculated using the following equations:

$$RR = \frac{RBR(\mathbf{x})}{RBR(\mathbf{x}^{true})}, \quad RR_{total} = \sqrt{\sum_i^n RR_i^2} \quad (7)$$

where RBR is the region of interest (ROI) to background ratio and n is the number of ROIs. The size of the analyzed ROI were 11 by 11 pixels and 7 by 7 pixels squares for cylindrical

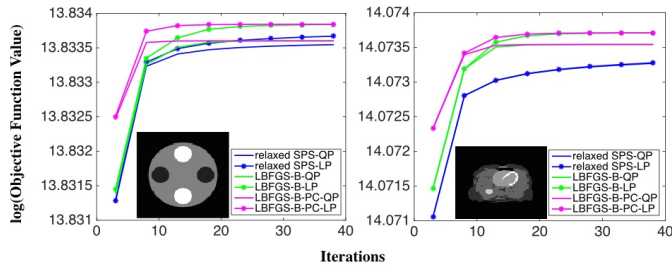


Fig. 4. The logarithm of objective function value (Φ) plotted against iteration numbers for cylindrical (left) and XCAT (right) phantom with $\beta = 0.1$. The results from relaxed SPS with QP and LP are almost overlapped with each other for XCAT phantom.

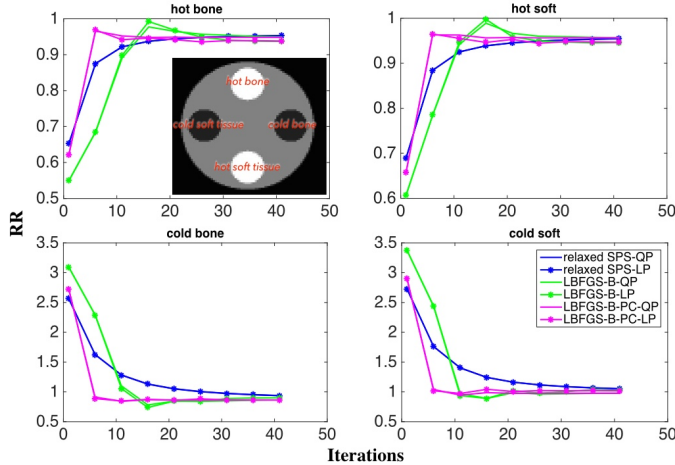


Fig. 5. The RR values plotted against iteration numbers for $\beta = 0.1$ (cylindrical phantom).

and XCAT phantom, respectively. The ROIs were drawn in the center of the hot or cold spot. For all the conditions with cylindrical phantom, we tabled the required number of iterations for achieving convergence of ROI to have a better comparison.

III. RESULTS

The objective function values plotted against the iteration numbers for both phantoms with $\beta = 0.1$ are shown in Fig. 4. Fig. 5 shows the RR values plotted against iteration numbers for cylindrical phantom with $\beta = 0.1$. Independent of the penalty type, object geometry and activity distribution, LFBFGS-B-PC converged faster than other methods. From table I, it can be observed that the convergence rate was independent of data noise level for all algorithms. We did not show results for larger penalty strength ($\beta = 0.3$) and the RR values and comparison table for XCAT phantom with $\beta = 0.1$, since they were consistent with those from cylindrical phantom with $\beta = 0.1$.

IV. DISCUSSIONS

For MAP reconstruction, ideally an algorithm is used that converges to the MAP solution, such that the resulting image is independent of the chosen algorithm, as long as sufficient number of iterations were used. While this was indeed the

TABLE I

THE REQUIRED NUMBER OF ITERATIONS FOR ACHIEVING CONVERGENCE OF ROI VALUES FOR ALL CONDITIONS (CYLINDRICAL PHANTOM). THE VALUES FROM DIFFERENT NOISE LEVELS ARE LISTED FROM HIGH TO LOW NOISE LEVEL (FROM LEFT TO RIGHT) AND SEPARATED BY A SLASH. FOR LFBFGS-B AND LFBFGS-B-PC, WE LIST IN BRACKETS THE NUMBER OF CALLS TO THE OBJECTIVE FUNCTION (*i.e.* A FORWARD- AND BACK-PROJECTION) AS THIS CAN BE A FAIR COMPARISON OF RUN-TIME TO RELAXED SPS.

	relaxed SPS	LFBFGS-B	LFBFGS-B-PC
QP, $\beta = 0.1$	91/121/125	11/13/19 (28/32/44)	9/9/9 (22/22/22)
QP, $\beta = 0.3$	338/342/342	23/16/15 (52/38/36)	7/7/7 (18/18/18)
LP, $\beta = 0.1$	75/91/102	9/18/18 (24/42/42)	6/12/12 (16/28/28)
LP, $\beta = 0.3$	309/326/322	15/12/16 (36/30/38)	12/7/7 (28/18/18)

case for the algorithms used in this study, the required number of iterations depended on the algorithm, but also on the object and other factors. The convergence rate of RR for all 3 algorithms was largely insensitive to the penalty type (Fig. 5) and noise level (table I). Relaxed SPS was slower for higher penalty strength and its convergence rate depends on activity distribution (Fig. 4). Both LFBFGS-B and LFBFGS-B-PC were insensitive to all factors we have studied. However, the convergence rate of LFBFGS-B-PC is significantly superior to the others. Our initial experiments indicate that LFBFGS-B-PC is a promising algorithm for MAP image reconstruction. In future work, we will investigate its performance on more realistic data.

REFERENCES

- [1] A. Alessio and P. Kinahan, *PET Image Reconstruction*, 2nd ed. Amsterdam: Elsevier, 2006.
- [2] S. Ahn and J. A. Fessler, *Globally Convergent Image Reconstruction for Emission Tomography Using Relaxed Ordered Subsets Algorithms*, *IEEE TMI*, vol. 22, no. 5, 2003.
- [3] H. Erdogan and J. A. Fessler, *Ordered Subsets Algorithms for Transmission Tomography*, *Phys. Med. Biol.*, vol. 44, 1999.
- [4] C. Zhu, R. H. Byrd, P. Lu and J. Nocedal, *Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-scale Bound-constrained Optimization*, *ACM Trans. Math. Softw.*, vol. 23, no. 4, 1997.
- [5] W. P. Segars, G. Sturgeon, S. Mendonca, Jason Grimes and B. M. W. Tsui, *4D XCAT Phantom for Multimodality Imaging Research*, *Med. Phys.*, vol. 37, no. 9, 2010.
- [6] K. Thielemans, C. Tsoumpas, S. Mustafaovic, T. Beisel, P. Aguiar, N. Dikaos and M. W. Jacobson, *STIR: software for tomographic image reconstruction release 2*, *Phys. Med. Biol.*, vol. 57, 2012.
- [7] L. A. Shepp, Y. Vardi, *Maximum Likelihood Reconstruction for Emission Tomography*, *IEEE TMI*, vol. 1, no. 2, 1982.