Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 18 (2013) 2643 – 2646

Procedia
Computer Science

International Conference on Computational Science, ICCS 2013

A regularized MRI image reconstruction based on Hessian penalty term on CPU/GPU systems

Francesco Piccialli*[a], Salvatore Cuomo[a], Pasquale De Michele[a]

[a] *University of Naples Federico II, Naples, Italy*

Abstract

In this paper we investigate an inverse reconstruction problem of Magnetic Resonance Imaging with few acquired body scanner samples. The missing information in the Fourier domain causes image artefacts, therefore iterative computationally expensive recovery techniques are needed. We propose a regularization approach based on second order derivative of both simulated and real images with highly undersampled data, obtaining a good reconstruction accuracy. Moreover, an accelerated regularization algorithm, by using a projection technique combined with an implementation on Graphics Processing Unit (GPU) computing environment, is presented. The numerical experiments give clinically-feasible reconstruction runtimes with an increase in speed and accuracy of the MRI dataset reconstructions.

Keywords: Compressed Sensing; Numerical Regularization; Graphics Processing Unit; Parallel and Scientific Computing.

1. Introduction

In the diagnostic application of Magnetic Resonance Imaging (MRI), the acquisition speed and the reconstruction quality are very important tasks. A recent methodology in MRI, called Compressed Sensing (CS) [1], is applied to improve the speed of an acquisition. The basic CS idea is the compressibility of medical images that can be given by faster scanning via subsampling. The quality of MRI images can be reached by combing CS with a Parallel Imaging technique [2]. Unfortunately, reconstructing an image, starting from very few acquired samples (k-space points) in the Fourier domain, causes image artifacts; moreover the efficiency of the reconstruction algorithms [1] is a crucial step: improving the quality in an CS-MRI image requires iterative algorithms that are more computationally expensive than traditional inverse Fourier method. To overcame this problem a recent trend in MRI has been to accelerate reconstructions by implementing and optimizing numerical codes on massively parallel processors: the Graphics Processing Units (GPUs) [3] [4]. In this paper we study the reconstruction of a CS k-space both acquired by a real body scanner and simulated by means of a subsampling stochastic algorithm. Here, we propose a regularization technique for the reconstruction taking into account a prior-knowledge of the acquired image. Moreover, we compare this approach with other constrained minimization methods applied to simulated and real k-spaces with highly undersampled data. This work presents the parallelization of an iterative algorithm via GPUs that can gives clinically-feasible reconstruction runtimes with a significant increase in speed. The paper outline is: the Section 2 describes the mathematical background; the Section 3 explains the minimization algorithm and the parallel implementation; the Section 4 presents the experimental results. Conclusions section closes the paper.

*Corresponding author.

E-mail address: salvatore.cuomo@unina.it (Salvatore Cuomo), pasquale.demichele@unina.it (Pasquale De Michele), francesco.piccialli@gmail.com (Francesco Piccialli).

Selection and peer review under responsibility of the organizers of the 2013 International Conference on Computational Science
doi:10.1016/j.procs.2013.06.001

2. Sparse Representation and Mathematical Preliminaries

Let $x \in R^n$ represents a discrete signal and $b \in R^m$ a vector of linear measurements formed by taking inner products of x with a set of linearly independent vectors $a_i \in R^n$, $i = 1, 2, \ldots, m$. In the CS context, only some components of the signal x are acquired, so the signal to be processed is a sparse vector x_{sparse}. We observe that the sparsity is hidden into a signal x, so that it becomes sparse only under a "sparsifying" basis ψ (i.e. a Wavelet representation):

$$x_{sparse} = \psi x$$

By a changing of variable $x_{sparse} = \psi x$, and supposing that ψ is an invertible transformation, the *decoding* becomes:

$$\psi^{-1} x_{sparse} = A^{-1} b.$$

When the number of measurements m is equal to n, decoding simply entails solving a linear system of equations. When $m < n$, the linear systems are typically under-determined, permitting many solutions. Generally speaking, CS refers to the following two-step approach: choosing a measurement matrix $A \in R^{mxn}$ with $m < n$ and taking measurement $b = Ax$ on a sparse signal x, and then reconstructing $x \in R^n$ from $b \in R^m$ by some regularization technique [5]. Let be the image of interest a vector I, ψ denote a sparsifying operator, and F_u be the Fourier transform. The reconstruction problem is:

$$\arg\min_I \quad \|\psi I\|_1 \ + \ \alpha\|TV(I)\|_1 + \ \beta\|H(I)\|_1 + \ \lambda\|F_u I - y\|_2 \tag{1}$$

Here I is the reconstructed image, where y is the measured k-space data from the scanner and $\|F_u I - y\|$ controls the fidelity of the reconstruction to the measured data. It is a well-known approach [1] [6] to include additional terms in the functional that express prior knowledge in the reconstruction. The Total Variation (TV) is generally adopted in many different MRI applications as penalty term [7].
In this work we introduce a regolarization term which is capable to measure, in some sense, image characteristics up to a certain order of differentiation: $H(I)$, that is the Hessian of I. A tradeoff between parameters α and β in (1) cover a key role to weigh the contribution of regularization terms [8]. In this context, we prefer the l_1 norm in which many small coefficients tend to carry a larger penalty than a few large coefficients, therefore small coefficients are suppressed and solutions are often sparse.

3. Minimization algorithm: POCS Method and GPU implementation

A number of inverse image reconstruction problems, can be formulated in terms of finding a vector in the intersection of certain convex sets that serve to constrain the solution [9], as in Projection Onto Convex Sets (POCS) method. Let C_i, $i = 1, \ldots, n$ be closed convex subset of a Hilbert space, with nonempty intersection C, if partial K-space data K(u,v) , (u,v) $\in$ Z are given where Z is a closed region in the Fourier plane, then a convex constraint $C = \{I \in R^{N^2} : F(I)_{u,v}| = K(u, v), (u, v) \in Z\}$ and $C \subset H$ is determined. The objective function of problem (1) can be found by the following projection method:

$$I_{k+1} = P(I_k - (\alpha_k g(I_k) + \beta_k h(I_k))) \text{ with } \alpha_k > 0 \text{ , } \beta_k > 0 \text{ and } I_0 \in C$$

where $g(I)$ and $h(I)$ are respectively the subgradient and the Hessian of I, P is the projection onto C, k stands for k-th iteration. One condition can guarantee convergence is:

$$\textstyle\sum_{k=0}^{+\infty}(\alpha_k + \beta_k) = +\infty \text{ , and } \sum_{k=0}^{+\infty}(\alpha_k^2 + \beta_k^2) < +\infty$$

The minimization schema is reported in the Algorithm 1. The computationally intense operations in MRI reconstructions contain nested data parallelism. In particular, operations such as Fourier transforms are performed over k-dimensional slices through the N-dimensional reconstruction volume, with $k < N$. Parallelizing over independent 2-D reconstructions is very efficient, as the decoupled 2-D reconstructions require no synchronization.
We implement the main tasks of this algorithm using a GPU programming with the CUDA framework [10].

Data Fidelity constraint: This task is based on a matrix-vector multiplication per pixel. We perform this operation by using the CUDA based supported `cublasSgemv()` of CUBLAS library [11].

Algorithm 1 POCS minimization algorithm.

1: Calculate a $mask(i, j)$ starting from the subsampling algorithm
2: Set $M = \{(i, j) : mask(i, j) = 1$ and (i, j) are the coordinates of the $I(i, j)$ grid point$\}$
3: Determinate the contstraint $C = \{I \in R^{N^2} : F(I)_{i,j} = I(i, j), (i, j) \in M\}$
4: Get initial image I_0 and compute $IFFT(I_0)$
5: **while** $k < Iter$ **do**
6: Determine α_k and β_k and calculate $I_t = I_k - (\alpha_k g(I_k) + \beta_k h(I_k))$ and the projection of I_t onto constraint according POCS, assume $F(I_t)$ in the Fourier trasnform of I_t and PF is the frequency after projection:

$$PI_t = \begin{cases} K(i, j), & if(i, j) \in M \\ F(I_t)_{i,j}, & if(i, j) \notin M \end{cases}$$

7: **end while**

Regularization: In order to calculate the Total Variation TV(I) and the Hessian H(I), we perform a pixelwise implementation. Each pixel of these matrices is computed in parallel by using a dedicated thread.
Projection Method: Before starting the iteration process, the algorithm creates a *support mask* of the input image in the Fourier domain; this support mask will be the Convex Set. We project the reconstructed image on the mask, in order to fill the not-acquired values of the input image corresponding to the mask value 0.
In order to improve the parallelism in image reconstrucion if we have a multi-GPU system, we can run our software on it exploiting all the GPUs installed. The k-space is divided among all the GPUs, in order to distribute the data, by launching a CPU-thread for each GPU.

4. Numerical Experiments

In this section, we report in Table 1 few k-spaces reconstruction and a comparison between the values of the Peak Signal-to-Noise Ratio (PSNR) obtained by the reconstruction with H-penalty, IRGNTV and NUFFT [12] [13].
We simulate a sparse acquisition technique by using a subsampling procedure: 1) $S = \{k = (k_x, k_y, k_z)\} \mid \|k\| < \rho\}$ is the preserverd k-space; 2) Fixed a $\bar{x}$ as a reading direction and a powerlaw number $\alpha > 0$, the probability distribution for which the couples $(k_y, k_z) = 0$ is:

$$\frac{1}{\phi^{-\alpha}} \qquad with \quad \phi = \sqrt{k_y^2 + k_z^2}$$

3) The ratio between η_{tot} the total k-space lines and η_p the preserved ones is: $\frac{\eta_p}{\eta_{tot}} = \frac{1}{\gamma}$ where γ is a compression factor. These stochastic rules should reflect fairly faithfully the criteria actually used to really subsample (i.e. directly in acquisition phase) the acquisition of k-space in the many Compressed Sensing application.
In order to test and compare these reconstruction functionals we use two k-spaces. The table 1 underline that

	Type	Size	Compr.	Powelaw	H-penalty	IRGNTV	NUFFT
$K1$	RK	192x220x40	3	1	30.23	28.22	25.53
$K2$	RK	192x220x40	3	2	29.08	27.11	22.28
$K3$	RK	192x220x40	2	0	32.11	33.11	28.44
$K4$	RK	192x220x40	2	1	32.22	33.01	29.68
$K5$	Phantom	256x256x256	3	1	30.64	28.12	26.11
$K6$	Phantom	256x256x256	3	2	27.93	26.77	24.12
$K7$	Phantom	256x256x256	2	1	34.33	32.10	28.45

Table 1: H-penalty vs other reconstruction methods. RK is a full 3D real k-space, Phantom is a 3D Shepp-Logan phantom.

the H-penalty is able to reach better values of PSNR. Clearly, it is due to the contribution of the H-penalty term when the compression factor is high. This result makes it possible to take into strongly account this reconstruction functional in order to obtain acceptable image quality in less execution time.
Moreover, we analyze the CUDA implementation performance in terms of execution time, *GFLOPS* and Speed-up (ratio between the GFLOPs of GPU and CPU). The hardware configuration we used is: *CPU: Intel core i7 950 3.06 GHZ*, *GPU: 3 x TESLA C1060 with (each) core clock: 602 MHZ, 240 Thread Processors*. Adding a regularization term, like H-penalty, increases the amount of computational demand. In table 2 we report only

the execution times of a large 3D dataset in order to highlight that using a multi-thread or massively multi-thread (GPU) architectures leads to overcame the algorithm computational requirement.

	CPU	CPU - 8 Threads	Single GPU	Multi GPU
448x448x448	2513.22	410.32	270.35	107.21

Table 2: Overall application execution times with 50 POCS iterations.

We stress the application in a simulated parallel MRI scenario where a reduced dataset in the phase encoding direction(s) of k-space is acquired to shorten acquisition time, combining the signal of several coil arrays.

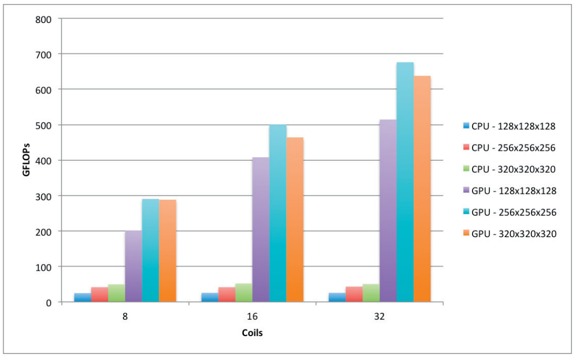

Fig. 1: GFLOPs of a multi coil approach.

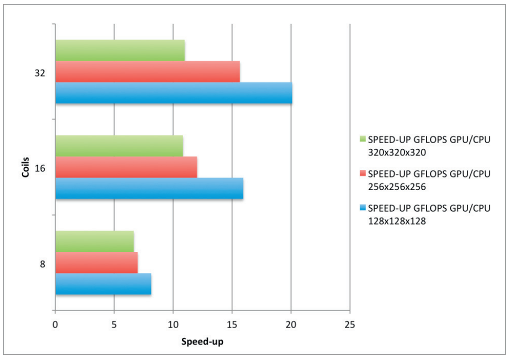

Fig. 2: Speed-up of a multi coil apporach.

In figure 1 and 2 we show how in a multi coil approach, GPU-GFLOPs well scale in growing up the coil number and leading to an increase of the performance in terms of GFLOPs and reaching values over 660.

5. Conclusions

To improve the reconstruction quality of highly undersampled MRI images is necessary to adopt prior-knoledge regularization terms. In this work we propose a penalty term based on the Hessian of the acquired sparse image. Although the computational complexity of the algorithm increases, the improvement in image quality encourages the use of this approach. For making reconstruction times closer to the real time we develop a GPU based application. The performance results highlight the powerful using a parallel approach for MRI reconstruction. These accelerated hardware allow us to take into account algorithms computationally onerous which executed on a serial architecture lead to computing time not reasonable.

References

[1] J. P. M. Lustig, D. Donoho, Sparse mri: The application of compressed sensing for rapid mr imaging, Magn. Res. in Medicine (2007).
[2] L. A. Ricardo Otazo, Daniel Kim, Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion mri, Magn. Reson. Medicine (2010).
[3] S. Cuomo, P. De Michele, R. Farina, A CUBLAS-CUDA Implementation of PCG Method of an Ocean Circulation Model, AIP Conference Proceedings (2011).
[4] S. Cuomo, P. De Michele, F. Piccialli, A performance evaluation of a parallel biological network microcircuit in NEURON, International Journal of Distributed and Parallel System (2013).
[5] Y. Zhang, On theory of compressive sensing via 1-minimization: Simple derivations and extensions (2008).
[6] S. Zaroubi et al, Complex denoising of MR data via wavelet analysis: Application for functional MRI, Magn. Reson. Imaging (2000).
[7] B. W. P. Rodriguez, Efficient minimization method for a generalized total variation functional, Image Proc., IEEE Transactions (2009).
[8] P. Hansen, D. OLeary, The use of the l-curve in the regularization of discrete ill-posed problems, SIAM J. on Scientific Comp. (1993).
[9] C. Byrne, Iterative projection onto convex sets using multiple bregman distance, INVERSE PROBL (1999).
[10] Nvidia cuda programming guide, `http://developer.download.nvidia.com`, nvidia Technical Report.
[11] Cublas library, `http://developer.download.nvidia.com`, nvidia Technical Report.
[12] J. Song, Y. Liu, S. Gewalt, G. Cofer, G. Johnson, Q. Liu, Least-square nufft methods applied to 2D and 3D radially encoded MR Image reconstruction, Biomedical Engineering, IEEE Transactions (2009).
[13] M. Freiberger, F. Knoll, K. Bredies, H. Scharfetter, R. Stollberger, The agile library for image reconstruction in biomedical sciences using graphics card hardware acceleration (2012).