Accelerated Convergent Motion Compensated Image Reconstruction

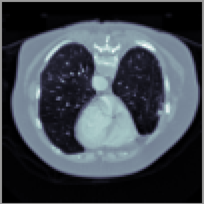
Claire Delplancke

University of Bath

IEEE Medical Imaging Conference 2021

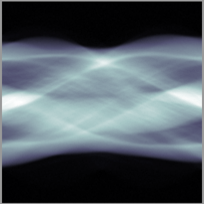
Joint work with Kris Thielemans (University College London)
and Matthias J. Ehrhardt (University of Bath)

Motivation



(a) First motion state

(b) Last motion state



(c) Data

(d) Reconstruction

Figure: Motion in the subject introduces artefacts in the reconstruction

Goal

Proof of concept for a randomized algorithm performing Motion Compensated Image Reconstruction (MCIR) which...

- is faster than the non-randomized counterpart,
- is provenly convergent.

1. MCIR: framework
2. Proposed algorithm and theoretical results
3. Numerical experiments

1. MCIR: framework
2. Proposed algorithm and theoretical results
3. Numerical experiments

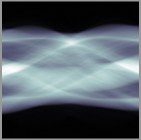
Framework

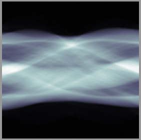
Data divided over N gates:

$$d_i \approx \underbrace{A}_{\text{forward op.}} \underbrace{D_i}_{\text{displacement op.}} x, \quad 1 \leq i \leq N.$$

$\xrightarrow{D_i}$

$\downarrow A$ $\quad \searrow AD_i$ $\quad \downarrow A$





Variational model

With Gaussian noise and l_2^2 regularizer:

$$\begin{aligned}
\min_x \quad O(x) \quad &:= \quad \frac{1}{N}\sum_{i=1}^{N} \|AD_i x - d_i\|^2 \quad + \quad \alpha\|x\|^2 \\
&:= \quad \underbrace{F(Lx)}_{\text{data fit}} \quad + \quad \underbrace{G(x)}_{\text{regularizer}} \ .
\end{aligned}$$

Convex setting.

Primal-Dual Hybrid Gradient (PDHG) algorithm

(also known as Chambolle-Pock algorithm)

Iterate

- $x^{k+1} = \text{prox}_{\tau G}(x^k - \tau L^* \bar{y}^k)$
- $y^{k+1} = \text{prox}_{\sigma F^*}(y^k + \sigma L x^k)$
- $\bar{y}^{k+1} = y^{k+1} + \theta_k (y^{k+1} - y^k)$

$\rightarrow$ convergent algorithm

$\rightarrow$ each iteration requires the evaluation of $L = (AD_1, \dots AD_N)$ and L^*: the computational cost scales linearly with the number of gates.

1 MCIR: framework

2 Proposed algorithm and theoretical results

3 Numerical experiments

Randomized algorithm

- Idea use only one gate, picked at random, for each iteration.
- How? use Stochastic Primal-Dual Hybrid Gradient (SPDHG) algorithm [Chambolle, Ehrhardt, Richtárik, Schönlieb, 2018]

$$\begin{aligned} \min_x \quad O(x) \quad &:= \quad \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \|AD_i x - d_i\|^2 \quad + \quad \alpha\|x\|^2 \\ &= \quad \textstyle\sum_{i=1}^{N} F_i(L_i x) \quad + \quad G(x). \end{aligned}$$

Randomized algorithm

- Idea use only one gate, picked at random, for each iteration.
- How? use Stochastic Primal-Dual Hybrid Gradient (SPDHG) algorithm [Chambolle, Ehrhardt, Richtárik, Schönlieb, 2018]

$$\begin{array}{rcccc} \min_x \quad O(x) & := & \frac{1}{N}\sum_{i=1}^N \|AD_i x - d_i\|^2 & + & \alpha\|x\|^2 \\ & = & \sum_{i=1}^N F_i(L_i x) & + & G(x). \end{array}$$

Convex setting

SPDHG

Iterate

- $x^{k+1} = \text{prox}_{\tau G}(x - \bar{z}^k)$
- Pick a gate i with probability p_i
- $y_i^{k+1} = \text{prox}_{\sigma_i F_i^*}(y_i^k + \sigma_i L_i x^k)$ and $y_j^{k+1} = y_j^k$ for $j \neq i$
- $\delta^k = L_i^*(y_i^{k+1} - y_i^k)$
- $\bar{z}^{k+1} = \bar{z}^k + (1 + \theta_k p_i^{-1})\delta^k$

SPDHG

Iterate

- $x^{k+1} = \text{prox}_{\tau G}(x - \bar{z}^k)$
- Pick a gate i with probability p_i
- $y_i^{k+1} = \text{prox}_{\sigma_i F_i^*}(y_i^k + \sigma_i L_i x^k)$ and $y_j^{k+1} = y_j^k$ for $j \neq i$
- $\delta^k = L_i^*(y_i^{k+1} - y_i^k)$
- $\bar{z}^{k+1} = \bar{z}^k + (1 + \theta_k p_i^{-1})\delta^k$

$\rightarrow$ convergent algorithm

$\rightarrow$ each iteration requires the evaluation of only one $L_i = AD_i$ and L_i^*: the computational cost scales constantly with the number of gates.

Theoretical rates of convergence

In the strongly convex - strongly smooth setting, PDHG and SPDHG converge linearly with known optimal per epoch rates [Chambolle et al., 2011], [Chambolle et al., 2018]:

$$\|x^K_{\text{PDHG}} - x^*\|^2 \leq C(r_N^{\text{PDHG}})^K$$
$$\mathbb{E}\left[\|x^K_{\text{SPDHG}} - x^*\|^2\right] \leq \tilde{C}(r_N^{\text{SPDHG}})^K.$$

Theoretical rates of convergence

In the strongly convex - strongly smooth setting, PDHG and SPDHG converge linearly with known optimal per epoch rates [Chambolle et al., 2011], [Chambolle et al., 2018]:

$$\|x^K_{\text{PDHG}} - x^*\|^2 \leq C(r_N^{\text{PDHG}})^K$$
$$\mathbb{E}\left[\|x^K_{\text{SPDHG}} - x^*\|^2\right] \leq \tilde{C}(r_N^{\text{SPDHG}})^K.$$

Theorem

For N gates and well-chosen step-sizes, it stands that:

$$r_N^{PDHG} = 1 - \frac{2}{1 + \sqrt{1 + \frac{1}{\alpha N}\|(L_1, \dots, L_N)\|^2}},$$

$$r_N^{SPDHG} = \left(1 - \frac{2}{N\left(1 + \sqrt{1 + \frac{1}{\alpha N}\max_i \|L_i\|^2}\right)}\right)^N.$$

Theoretical rates of convergence

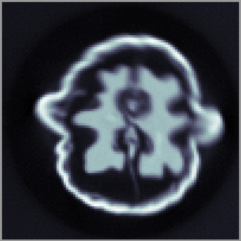
For a moderately conditioned problem such that $\kappa = \frac{\|A\|^2}{\alpha} \geq 16$,

$$r_N^{\mathrm{SPDHG}} \approx \left(1 - \frac{2}{N\left(1 + \sqrt{1 + \frac{\kappa}{N}}\right)}\right)^N \quad < \quad 1 - \frac{2}{1 + \sqrt{1 + \kappa}} \approx r_N^{\mathrm{PDHG}}.$$

Numerical application: rigid motion ($N = 20$ gates)

(a) First motion state

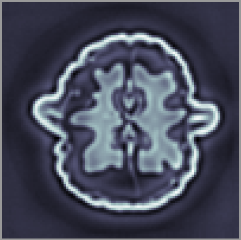
(b) Last motion state



(c) Converged no-MC

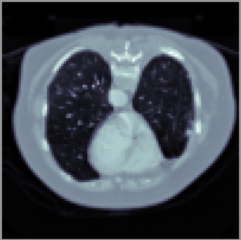
(d) Converged MC

(e) MC SPDHG after 30 epochs



(f) MC PDHG after 30 epochs

Numerical application: non-rigid motion ($N = 10$ gates)


(a) First motion state


(b) Last motion state

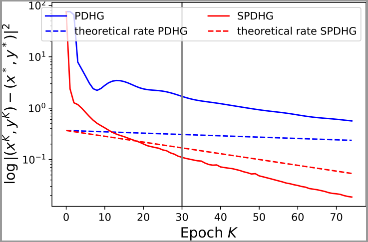

(c) Converged no-MC


(d) Converged MC

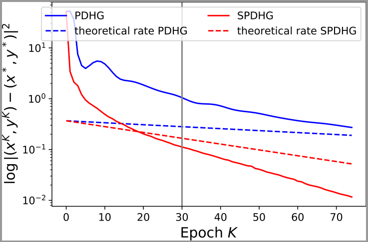

(e) MC SPDHG after 30 epochs


(f) MC PDHG after 30 epochs

Numerical application: convergence rates



(a) Rigid motion



(b) Non-rigid motion

Figure: SPDHG's linear convergence is faster than PDHG's

Contributions

We proposed a randomized algorithm for Motion Compensated Image Reconstruction with the following characteristics . . .

- $\rightarrow$ is provenly convergent,
- $\rightarrow$ requires the same computational effort than the non-motion compensated reconstruction per iteration,
- $\rightarrow$ [in proof-of-concept setting]
  - $\rightarrow$ a theoretical speed-up is proved on linear rates
  - $\rightarrow$ a practical speed-up is observed on synthetic experiments.