

jh210030-DAH

大規模分散医用画像処理に向けた医用画像処理アプリケーションの最適化

大島 聡史（名古屋大学）

概要

発展の著しい医用画像処理技術は今日の医療現場を支える重要な技術である。例えば CT や MRI は既に広く普及し、必要不可欠なものとなっている。しかしこれらにより直接的に得ることができるのは画像や映像であり、具体的な症状の判断や特定などは人間（医師など）が行う必要がある。そのためソフトウェアによる判断支援などの需要は大きく、急速に研究が進んでいる。

医用画像処理では画像などのデータを多く扱っており、並列計算技術の活用が有効である。さらに機械学習技術の活用も始まっている。しかしながら医用画像処理分野の研究者・開発者が並列計算技術に精通しているとは限らない。そこで本研究では、高性能計算分野の研究者と医用画像処理分野の研究者が協力し、スーパーコンピュータを用いた大規模分散医用画像処理アプリケーションの開発に向けた共同研究を行っている。

本年度は二つの対象アプリケーションに対して性能の分析や高速化・大規模化を行った。その結果、いずれも大幅な性能向上や大規模データへの対応が可能となった。

1 共同研究に関する情報

1.1 共同研究を実施した拠点名

東京大学 名古屋大学

1.2 共同研究分野

■超大规模データ処理系応用分野

1.3 参加研究者の役割分担

- 大島 聡史 [代表] (名大)：スパコン向けの実装と評価を行うとともに、代表として全体を取りまとめる。
- 片桐 孝洋 (名大)：大島とともに、スパコン上でのプログラム最適化を担当。
- 森下 誠、杉浦 拓未、栢木 慎也 (名大)：大学院生。大島・片桐の指導の下で実装や評価を実施。

- 本谷 秀堅 [副代表]、クグレ マウリシオ、横田 達也 (名工大)：医用画像処理分野の研究者として対象プログラムやデータの提供と結果の検証、医学的考察を行う。
- 埜 敏博 (東大)：分散機械学習についての知見を提供。

2 研究の目的と意義

医用画像処理技術の発達により生体の内部を視覚的に理解するための様々な方法が開発され、すでに医療現場で広く活用されている。例えば CT や MRI は既に広く普及しており、怪我や病気の特定などにおいて必要不可欠なものとなっている。しかし、それらによって直接的に得ることができるのは画像や映像であり、

具体的な症状の判断や特定などは人間（医師など）が行う必要がある。そのためソフトウェアによる判断支援などの需要は大きく、急速に研究が進んでいる。これらの処理には画像の表示・分析・検索・比較といった処理が多用されるため、最新のマルチコア CPU や GPU による高速化への期待は大きい。さらに近年では機械学習技術の活用も始まっている。一方、医用画像処理分野の研究者・開発者はスーパーコンピュータや PC クラスタなどの活用に必要な知識や経験が乏しいことも多く、高性能な計算資源の活用が十分に行えているとは限らないのが現状である。

そこで本研究では、高性能計算分野の研究者と医用画像処理分野の研究者が協力し、スーパーコンピュータを用いた大規模分散医用画像処理アプリケーションの開発に向けて共同研究を実施している。

本研究は前年度からの継続課題であり、前年度とは研究体制や対象アプリケーションを一部変更して実施している。今年度の主な対象問題（医用画像処理アプリケーション）は以下の2つである。

1. PET 画像の再構成

PET (Positron Emission Tomography) は放射性同位元素で標識した薬剤を患者に投与し、薬剤の密度分布の時間変化を計測（撮影）することで神経受容体機能を推定する技術である。計測したデータを視覚的にわかりやすい画像に変換する際に生じるノイズを削減する際に Deep Learning 技術を用いており、ノイズ軽減処理の高速化や、大規模高解像度画像への対応が課題となっている。高速化が行えれば、診療の際に撮影後にすぐ診断を行うことが可能となり、また再構成がうまく行かなかった場

合にもすぐにやり直して結果を得ることができるようになる。

2. TNR 再構成

TNR (Trajectory-based Non-linear Registration) は、病理顕微鏡画像の三次元再構成を行う技術である。多数の二次元画像から三次元的な位置関係・形状を把握する際に大規模なテンプレートマッチング処理が必要で時間がかかるため、分散並列化による高速化が求められている。我々の最終的なターゲットとしては数 10TB 級の大規模データに対する再構成を予定しており、このような大規模な TNR 再構成の前例はまだない。

大規模な医用画像処理を高速に行えるようになれば、従来は行えなかった規模の医用画像処理が行えるようになったり、処理時間の制約により行えなかった分析などが行えるようになる。その結果、医学的な新しい発見に結びつき学術的に大きく貢献できる可能性がある。もちろん、我々が利用する医療にも大きく貢献できる可能性があり、本研究の学術的な意義や社会的な意義は大きい。さらに、比較的新しい分野である医用画像処理にスーパーコンピュータが有用であることを示すことで、当研究分野におけるスーパーコンピュータの活用のさらなる拡大が期待できる。（センタースパコン活用の幅が広がり、さらなる利用者数の増加などにも貢献できる可能性がある。）

3 当拠点公募型研究として実施した意義

本共同研究では大規模な計算資源を用いて大規模かつ高速な医用画像処理を行いたいという強い要求があり、一方で我々の研究グループはそれを実現するだけの計算資源を有していな

かった。そこで拠点公募型共同研究として提案し、計算資源を確保した。十分な計算資源を利用可能になったことで、対象プログラムの品質や速度の向上、大規模実行への対応が行えるようになったことには大きな意義がある。

また、我々の対象アプリケーションは参加研究者らが独自に有する計算機環境で開発してきたものであり、継続元の研究提案を行うまでは共有のスーパーコンピュータ環境で実行した実績がなかった。今後さらに対象アプリケーションを発展させていくうえで、共有のスーパーコンピュータ環境における利用実績を積むことにも意義がある。さらに、医用画像処理分野におけるスーパーコンピュータのさらなる活用促進という点でも本共同研究の実施には意義があると考えらる。

以上のように、本共同研究は拠点公募型研究として実施することで高性能計算分野の研究者と医用画像処理分野の研究者が十分な計算資源を得て共同研究を行うことが可能となり、対象アプリケーションの高性能化も進んでおり、実施した意義は大きい。

4 前年度までに得られた研究成果の概要

本共同研究は継続課題であり、当年度は2年目にあたる。

前年度（初年度）は、本研究ではこれまで共有のスーパーコンピュータ環境で実行したことのなかったアプリケーションを対象としていたうえに、利用した計算機環境のメモリ容量が不足気味であったりソフトウェア環境の整備に手間取ったりしてしまい、対象アプリケーションを正しく動作させることと基本的な性能評価をすることに多くの時間が割かれた。そのような中でもプログラムの性能分析と改善を行い、PET 画像の再構成について一定の（元のプロ

グラムに対して「不老」Type II サブシステムの1GPU 環境にて3.76 倍）性能向上を得た。

当年度は対象アプリケーションとしていない別の医用画像処理プログラムについてもスーパーコンピュータ環境での性能評価や最適化を行った。2つの臓器画像間の微分同相写像を求めるプログラム（Large Deformation Diffeomorphic Metric Mapping, LDDMM）の高速化については、マルチ GPU 化により高速化を行うことで当面の目標を十分に達成できたため、対象アプリケーションからは外した。CT ボリュームデータの領域分割・対応付けについても、多くの性能データを取得したりプログラムの問題を発見することができたため、継続対象アプリケーションからは外した。TNR 再構成については年度の途中で新たな対象アプリケーションとして追加し、並列化に向けて検討を開始した。

5 今年度の研究成果の詳細

今年度は2つの対象アプリケーションについて大規模実行に向けた分散並列最適化を進めた。詳細は以下の通りである。

5.1 PET 画像の再構成

PET 画像の再構成は前年度から取り組んでいる課題であり、前年度の時点で1GPU による実行の最適化（高速化）はある程度行っていた。今年度は複数 GPU を用いたさらなる性能の向上と、高精細な（容量の大きい）画像データを元にした PET 画像再構成に向けた実装の改善を行った。なお本研究の成果はまだ論文文化できていないため、概要のみを記す。

本研究で用いている PET 画像の再構成手法は、研究参加者である横田らにより提案された手法に基づいている。この手法は非負行列因子分解と Deep Image Prior を組み合わせる手法であり、ノイズの少ない (Signal/Noise

Rate の高い) 高精度な再構成手法である。実装には Python を用いており、TensorFlow を用いた GPU 向け実装が行われている。(PET 画像の再構成自体は機械学習そのものではないが、TensorFlow の計算グラフによって実装されており、GPU による加速も行われている。)

前年度の時点で 1GPU による実行の最適化は進んでおり、元々のプログラムと比べて大幅な性能向上を得られていたが、今年度はその効果を再評価するとともに、さらなるプログラムの改善により性能向上を達成した。(前年度の成果もあわせて) 具体的に行った最適化の内容は以下の通りである。

- CuPy で行っていたノイズ生成処理を NumPy に変更。(単純なノイズ生成の処理の高速化よりも、GPU に処理をまとめることによる効率化が重要。)
- 全体的な処理を見直し、CPU と GPU を行き来するデータを減らすことで、データ通信時間や GPU カーネル起動オーバーヘッドを削減。
- 内部で行っていた U-Net による重み計算処理を複数 GPU に分散させた。(基本的に基底の数だけ並列に行える。現在の問題設定では基底の数は 3 であり、1 ノードに 4GPU 以上を有する現在の実験環境での実行に適している。)

名古屋大学情報基盤センターの「不老」Type II サブシステムと東京大学情報基盤センターの Wisteria-A (Aquarius) で性能評価を行った。結果の傾向に大きな差はなかったため、ここでは「不老」Type II サブシステムにおける結果のみ記す。

図 1 に実行結果を示す。グラフからわかるように、最適化を適用することで実行時間の大幅な削減ができた。オリジナルの実装に対する実

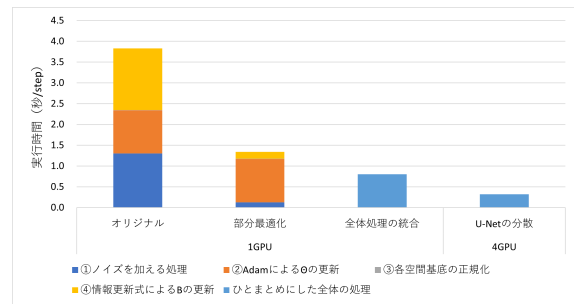


図 1 PET 画像の再構成、実行時間内訳

行時間比は、1GPU による実行 (全体処理の統合) で 35.0% (2.9 倍の性能)、4GPU による実行 (U-Net の分散) で 8.4% (12.0 倍の性能) となった。

さらに、入力画像の解像度を上げた高解像度な PET 画像の再構成にも取り組んだ。従来の 1GPU のみを用いた実装では GPU のメモリが不足して実行できなかったが、Univied Virtual Memory (UVM) の利用、または U-Net の分散により、入力画像の解像度を 2 倍に高めた場合でも正しく計算が行えるようになった。4 倍に高めた場合についても実験してみたが、内部で利用しているライブラリの入力データの制限によって実行できなかった。さらにその制限を回避するような修正も行ってみたが、メモリ不足によりやはり実行できなかった。

図 2 に 2 倍超解像度版の実行結果を示す。グラフからわかるように、2 倍超解像度では 4GPU 使用時の性能向上が大きな結果となった。その主な理由としては、1GPU 実行では多くのデータを扱うため UVM による CPU-GPU 間のデータ転送が行われたのに対して、4GPU では UVM を使う必要がなく暗黙的なデータ転送によるオーバーヘッドがなかったためであると考えられる。オリジナルの実装に対する実行時間比は、1GPU による実行 (全体処理の統合) で 59.1% (1.7 倍の性能)、4GPU に

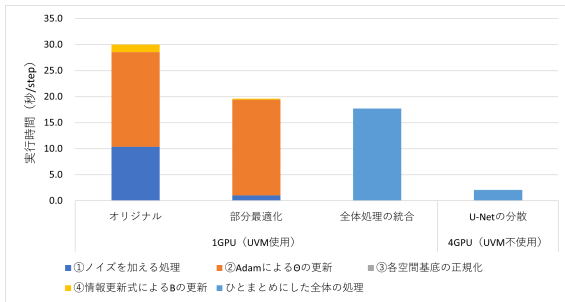


図2 PET 画像の再構成、実行時間内訳 (2倍超解像度版)

よる実行 (U-Net の分散) で 6.9% (14.6 倍の性能) となった。

なお、これら一連の最適化作業の過程において元プログラムの実装上の不具合を発見した。不具合を修正した後に改めて性能評価を行い、論文化を目指す予定である。

5.2 TNR 再構成

TNR 再構成プログラムは C++ で書かれており、元プログラムは並列化されていなかった。昨年度行った検討により基本的な並列化の方針はたててあったため、今年度は実際に並列化を行い小規模なデータで動作と性能の確認を行った。

幸いにも TNR 再構成プログラムの主要な処理部分 (テンプレートマッチング処理) は画像ファイル単位などわかりやすいレベルでの並列性を有していたため、プログラムの構造をある程度変更する必要はあったものの、MPI と OpenMP を用いることで素直に並列化を行うことができた。

性能確認として、表 1 のデータセット小およびデータセット大を用いて TNR 再構成を行った。

実験環境としては、名古屋大学情報基盤センターの「不老」Type I サブシステム (「富岳」型、A64FX 搭載) およびクラウドシステム (Xeon 搭載) を利用した。実験結果は以下の通りで

ある。

- データセット 小
 - Type I 1N*4P*12T 125 秒
 - Type I 20N*4P*12T 23 秒
 - クラウドシステム 1N*4P*20T 19 秒
 - クラウドシステム 4N*4P*20T 8 秒
- データセット 大
 - クラウドシステム 4N*4P*12T 182 分

(aN*bP*cT は、a ノード使用、1 ノードあたり bMPI プロセス使用、1MPI プロセスあたり c スレッド使用、の意味。)

実験の結果、TNR 再構成の実行環境としては主に「不老」Type I サブシステムの利用を考えていたが、クラウドシステムを利用した方が良い性能を得やすいことがわかった。主な理由としては、主要な処理であるテンプレートマッチング部の実装が Xeon CPU に適している可能性が考えられるが、詳細な分析はまだ行っていない。クラウドシステムは Type I サブシステムよりもノードあたりのメモリ容量が多いため、さらに大きなデータセットでも実行しやすくターゲットデータに対する TNR 再構成の実行環境として適している可能性がある。

6 今年度の進捗状況と今後の展望

2つの対象アプリケーションそれぞれについて、高速化や大規模化に向けたプログラムの改良や性能評価を行うことができた。

PET 画像の再構成については、1GPU 実行の性能改善に加えて 4GPU による大規模問題への対応と高速化も達成し、さらに 2 倍超解像についても対応することができた。より大規模な 4 倍超解像については問題点の把握をするに留まったが、当初の目標の多くを達成することができた。論文などの形で对外発表をするまでは至らなかったものの、前章で述べたとおり、

表1 TNR再構成 データセット

	データセット 小 (テストデータ)	データセット 大 (テストデータ)	参考 (ターゲットデータ)
画像ファイル数	5 (4 ペア)	809 (808 ペア)	4000 (3999 ペア)
画像サイズ (ピクセル数)	4k * 4k	15k * 10k	100k * 150k
画像ペアあたりランドマーク数	1400 程度以下	5000 程度以下	未確認

発見した不具合を修正し次第、性能の評価をし直して論文化を目指す予定である。

TNR 再構成については、MPI と OpenMP を用いた並列化に成功し、大規模実行のめどが立った。今後はさらに性能評価を行い問題規模と性能について分析したうえで、TB 級の主ターゲットデータに対する再構成を試みる予定である。この規模のデータに対する TNR 再構成はこれまでに行われておらず、適切に行うことができればそれ自体が大きな成果となる。

7 研究業績一覧 (発表予定も含む)

学術論文 (査読あり)

国際会議プロシーディングス (査読あり)

国際会議発表 (査読なし)

国内会議発表 (査読なし)

公開したライブラリ等

その他 (特許, プレス発表, 著書等)