

## 論文の内容の要旨

論文題目     **Fast Low-Order Finite Element Earthquake Simulations  
on GPUs with Transprecision Computing**  
(精度変動演算を用いたGPUによる高速な低次有限要素地震  
シミュレーション手法開発)

氏     名     山口 拓真

地震災害においてその被害軽減策を立案する上では、より合理的な地震時被害予測を行うことが有効な手段の一つである。既往の研究では、対象領域の不均質性・地盤の複雑構造がシミュレーション結果に大きな影響をもたらすため、対象領域を高分解能で再現できる手法の導入が望ましいと指摘がなされてきた。一方で、このような手法を地震シミュレーションのような広大な領域に適用すると、自由度の増大から計算コストの増加につながる。また、地震災害の一連の現象をシミュレーションするにあたっては、地震シナリオや物性値・地盤地殻構造に不確実性が含まれると考えられる。不確実性を扱うための手法としてはモンテカルロシミュレーションや逆解析が代表的な例として挙げられるが、一般に多数回の順解析を必要とするため、計算コストの増大が課題となる。すなわち、高分解能モデルによる有限要素解析を多数回行うことが理想的であるが、計算時間が現実的なものではなくなる。従来の数値シミュレーションは簡略化されたモデルを使用するか、不確実性の含まれる要素を決定論的に扱うことによって、解析を行ってきた。

近年では計算機環境の急速な発展を受けて、より大規模あるいは多数回のシミュレーションを行うことが現実的になりつつある。対象とするシミュレーションは、その計算コストの大部分が最終的に帰着される連立一次方程式の求解に含まれるため、より高速なソルバーを開発することが計算コストを削減する上で重要となる。本論文では演算加速器とし

で知られている、GPUに着目する。これは多数のコアを用いて計算を並列に処理する演算器で、一般的な科学技術計算に対しても広く使用されるようになっている。このような新しいアーキテクチャを用いることで計算時間の短縮が期待される。

しかし、近年のハードウェアはその計算特性を活用した手法とともに用いることによって初めてその性能を引き出すことが可能となる。GPUの計算特性として、不規則なメモリアクセスが性能を大きく低下させること、GPU-CPU間やGPU-GPU間の通信が計算能力に対して小さいことが知られている。これらの特性を理解したアルゴリズムを使用しなければ、メモリアクセス・データ転送が性能のボトルネックとなる。これに加えて、GPU計算がAI向けのアプリケーションに高い親和性を有することから、新しいGPUがAIアプリケーションに特化したハードウェアを搭載する傾向にあることも考慮に入れる必要がある。このようなハードウェアは通常の計算よりも数倍以上高い理論性能を有するが、低精度変数による行列行列積のような特定の演算にしか機能しないため、対象とするアプリケーションに該当する計算が含まれていなければ恩恵を受けることはできない。以上の点を考慮すると、対象とするソルバーの高速化を実現するためにはアルゴリズム上、実装上の手法開発が課題になると考えられる。

本論文では、以上のような要求を踏まえて、GPUを用いた非構造有限要素解析高速化のための手法開発を行った。精度変動演算をソルバー部に導入することで低精度変数のみに有効なコアを活用し、高速化を実現した。また独自に定義された低精度変数による演算を導入することで演算コストに対するメモリアクセスコストを相対的に減少させ、計算時間の短縮を実現した。

まず、2章においてはCPU向けに開発されたソルバーをもとに、GPUをベースとする計算機環境への移植を行い、性能の評価を行った。計算量だけでなくデータ転送量、メモリアクセスコストが削減されるようなアルゴリズムを用いることで、妥当な高速化が実現することを確認した。加えて、GPU計算のボトルネックを解消するためのアルゴリズムの検討を行った。複数のベクトルに対して一度に解析を行うことで、疎行列ベクトル積計算部分におけるメモリアクセスの不規則性を改善した。またGPUのローカルメモリをキャッシュとして活用し、下層のメモリへのアクセス頻度を減少させることで、主要なカーネル部分における性能を改善した。開発されたソルバーをもとに、地中海周辺の粘弾性地殻変動を対象とした自由度約24億程度の有限要素モデルによる解析を行った。GPUを搭載した大規模計算機環境であるPiz Daintを利用し、2,000時間ステップ分の求解が80分程度で実行できることを示した。

3章では2章のソルバーアルゴリズムを用いて、3次元地盤構造推定を行った。複雑な構造を含む領域に対して信頼性の高い解析結果を得るためには、使用するモデルの信頼性を高める必要があると考えられる。地盤を対象とした振動解析に関しては、既往の研究から対象領域の複雑な形状が地表面速度分布や地中ひずみ分布に大きな影響を与えることが知られている。その一方で、地中構造の高い精度および分解能での測定は難しいことが問題

とされてきた。そのため、本章では地表面上において観測データが利用できるという仮定の下で、観測データが十分に再現されるような有限要素モデルを選択する手法を開発した。すなわち、異なる有限要素モデルを多数回作成して地盤振動解析を行い、観測データとの誤差を最小化することによる、地盤構造の最適化を行った。この最適化において大きな計算コストを占めるのは有限要素モデルの作成とソルバー部分の計算である。ソルバー部分に関しては GPU 計算を導入することが可能であるが、モデルの生成に関しては逐次的計算が一部に含まれているため、GPU の導入が難しく、相対的に計算コストが大きい操作となる。この計算コストを隠蔽するため、ソルバー部分の GPU 計算と同時にモデルの生成を CPU 上で行うようなフレームワークを開発した。これによって、利用可能な計算資源を効率的に使用することが可能となり、解析全体に必要なとされる計算コストを大幅に削減した。適用例としては、開発された手法を用いて  $300\text{ m} \times 400\text{ m} \times 75\text{ m}$  からなる 2 層の地盤構造の境界面の推定を行った。推定には自由度 300 万程度の有限要素モデルを使用した解析が 1,500 回必要となったが、13 時間程度の計算時間で実用上十分な精度で観測データが再現できるモデルを推定した。CPU のみを使用した既往の研究と比較すると、問題サイズ当たりの計算時間が  $1/160$  に短縮されていることから、GPU 計算の有効性を評価することができた。

続く章ではソルバーのアルゴリズムに修正をそれぞれ加え、低精度変数を導入することで、さらなる高速化を実現した。4 章では、地殻変動計算を対象として、NVIDIA Volta GPU に搭載されている AI アプリケーションに特化したハードウェアである Tensor コアに着目した。このコアは半精度変数の行列行列積のみを、倍精度演算の 16 倍の理論性能で処理することのできるコアとなっている。本章では行列行列積としてソルバー内の行列ベクトル積カーネルが処理でき、かつ演算精度を半精度としてもソルバーの収束性が失われないようなアルゴリズムを設計することによって、Tensor コアの導入を可能とした。Tensor コアは通常の計算と比較して非常に高い演算性能を有するため、カーネル部分で十分な性能を得るためには性能のボトルネックに留意する必要がある。そのため、各種のコストの高いメモリアクセスやデータ転送が省略または隠蔽されるような実装を行った。GPU を搭載した大規模計算機環境である Summit を用いた性能測定では、疎行列ベクトル積カーネルが従来の単精度演算を使用した場合の 4 倍高速に計算できることを確認した。またソルバー全体については従来の手法と比較するとより多くの演算量が必要となる一方で、Tensor コアが非常に高い演算性能を発揮することによって、計算時間としては 1.5 倍程度速く求解できることを示した。補遺ではこの実装方法は他の Tensor コアを用いたアプリケーションにおいてもメモリアクセスコスト削減に寄与することを、Matched Filtering を対象に示している。

5 章では都市の地震シミュレーションを対象とした計算コスト削減手法の開発を行った。具体的には、浮動小数点数のビット数を削減した変数を新たに定義し、メモリアクセス量を削減した。ソルバー内の大部分の計算はメモリバンド幅に性能が律速されるため、メモ

リアクセス量の削減が高速化に直結する。低精度変数を使用する場合には、ソルバーの収束性への影響を検討する必要がある。そこで削減するビット数とソルバーの収束性に関する測定を行い、単精度の 32 ビットから下位の 11 ビットを削減した場合に、収束性に大きな影響を与えることなく計算時間が削減されることを確認した。本論文ではこのデータ型を FP21 と定義し、ソルバーの前処理部分においてこの型を活用することで、計算時間の短縮を実現した。GPU 計算機環境である ABCI を用いた性能計測においては、ソルバー部分の求解が CPU 使用時の計算時間の 14 分の 1 で完了した。この性能は CPU と GPU の理論メモリバンド幅比と同等の比率であるから、妥当な高速化が実現したものと評価した。

以上のように、本論文では精度変動演算を有限要素ソルバーに導入することで、半精度演算のみに特化したコアを活用することによる高速化、またメモリアクセス量を削減することによる高速化が実現することを示した。計算機の開発としてはコアが特定の演算に特化することや、理論演算性能に対してメモリバンド幅が低下することは近年よく見られる傾向であるため、本論文で検討した低精度変数によるアプローチは今後より重要性を増すものと考えられる。今後の課題としては、低精度変数を使用した際のソルバーの収束性に関する定量的評価が挙げられる。開発されたソルバーを実際の観測データを用いた解析に適用することが、より合理性の高い地震シミュレーションの実現につながると期待される。