

数値データの圧縮

～並列計算向けアルゴリズムの開発～

特願2010-081190号

工業用CADやCG、物理現象のシミュレーションなどで用いられるデータを、並列コンピュータを用いて高速に可逆圧縮する技術を開発しました。

工業用データの圧縮

工業用CADや数値シミュレーションなどのデータは、数GB（ギガバイト）に達することも珍しくありません。大容量のデータをそのまま扱おうと、メモリやハードディスクなどの記憶装置の容量が多く必要になったり、インターネットでデータを送受信する時間が長くなったりします。このため、データを圧縮してサイズを小さくします。例えば、JPEG、MP3、MPEGなどの画像や音声・動画フォーマットは、圧縮されたデータを保存します。

画像や音声・動画の圧縮は、多くの場合、人間には認知することが難しいような細かなデータの違いを無視するため、完全に元に戻すことはできません。これに対して、工業用CADなどの用途では、データの精確さが求められます。このため、元のデータを完全に復元できる範囲で圧縮する可逆圧縮方式が用いられますが、例えば3D CADなどに適した可逆圧縮方式は、現在数種しか提案されていません¹⁾。

開発手法

都産技研では、GPU（図1）を用いた並列計算による高速な電磁界シミュレータの開発に取り組んでいます²⁾。その要素技術の一つとして、高速の可逆圧縮方式を考案しました。従来の圧縮手法の多くが並列処理を考慮していませんが、本手法は並列処理のために設計されているため、高速な処理が可能です。

表1は、元のデータサイズを1とした時の、圧縮されたデータの大きさを示したものです。工業用データに対応していない従来方式（ZIP）に比べて圧縮率が向上しています。図2は、表1のデータ1をGPUを用いて圧縮し

た際の所要時間（ミリ秒）を示しています。横軸は並列度を表します。提案手法では、従来手法に比べて高速に圧縮できます。

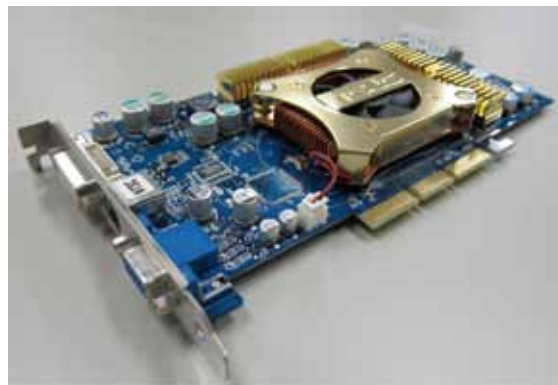


図1 GPUを搭載したビデオカードの例

表1 従来手法と提案手法の圧縮率

圧縮手法	データ 1* ¹	データ 2* ²
GZIP	0.940	0.957
従来手法 ¹⁾	0.884	0.801
提案手法	0.880	0.760

*1 文献1) で扱われているデータ

*2 文献2) の電磁界シミュレーションのデータ

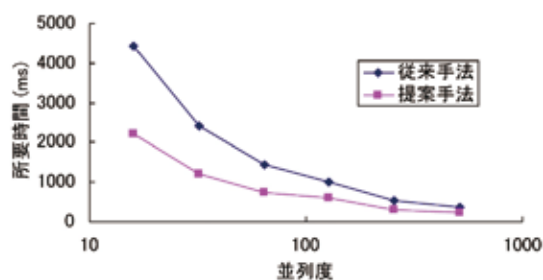


図2 従来手法と提案手法の圧縮率

参考文献

- 1) M. Burtscher et. al., Proc. 2009 Data Compression Conf., pp. 43-52, 2009.
- 2) 山口隆志、TIRI News、2010年5月号、pp. 2-3、2010

経営企画本部 経営企画室<本部>

大原 衛 TEL 03-5530-2426

E-mail:ohara.mamoru@iri-tokyo.jp