

Strip-Based Unfoldingに基づく3次元データ圧縮

川 尻 崇 博[†] 宮 崎 大 輝^{††}
高 丸 尚 教[†] 横 井 太 一[†]

近年の自然科学シミュレーションでは、ある時刻における一つの状態変数の空間構造データ量が1TB以上になることが珍しくない。しかし、単純に可逆圧縮法を用いても、遠隔地点間のデータ転送には適さないほど大きなデータにしか変換できない。そこで我々は、圧力等の3次元スカラー型データに対して自然科学者が注目する対象データの分離点分布に基づいてオブジェクトを特定化した後、Striped-Based Unfolding法を用い複数の2次元データ群に切り出すことでデータ縮約を行う手法を開発した、また、この手法を並列化実装し、その性能評価も行った。

Compression For 3D-data Using Strip-based Unfolding

TAKAHIRO KAWAJIRI,[†] DAIKI MIYAZAKI,^{††} HISANORI TAKAMARU[†]
and YOKOI TAICHI[†]

In recent years, when We simulate in the field of natural science, amount of spacial structure data for an condition variable is generally more than one-terabyte at a given time. But, compressed data by lossless compression method are too big to be suitable for data transfer between distant point. Therefore, We developed system for this problem that identify objects based on split off point distribution of intended data which natural scientist focus on for 3D scalar-type data (pressure et al), and then compress the 3D model data by clipping it as 2D data-constellation using the Striped-Based Unfolding method. In addition, We implemented parallelization system and evaluated its performance.

1. はじめに

スーパーコンピュータを利用したシミュレーション研究において、生み出されるデータはある一時刻における一つの変数データでさえも数十TBという大きさを持っている。このようなデータをスーパーコンピュータが設置されている研究設備以外で直接解析することはできない。しかし、大学に所属する研究者は、長期間にわたってそのような研究施設に行くわけにはいかない。この状況は25年以上も前、N1ネットワークに接続し、拠点研究施設に附置されているスーパーコンピュータに接続しながら遠隔地である大学からアクセスしながら研究をしていたころと、本質的に変わらない。大学の研究室で持てるリソースには限りがありTBを超える主記憶を持つコンピューティング設備を持つことは決して容易ではない。従って、データの可逆圧縮にこだわらず、自然科学研究者として着目し

ている情報量を抽出しやすいようにデータマイニングし、重要度の高い部分については解像度を維持し、代わりに重要度の低い部分については不可逆圧縮し、転送データ容量の大幅な削減を行う必要がある。転送後、適切に復元する技術を加味し、元々数十TBのデータを数GBの主記憶容量しかない小規模計算機で解析を可能にする手法を提案する必要がある。特に、国策として開発運用する理化学研究所の「K Computer」においても、全国からの利用を促すには本研究が提案する簡易データ解析手法を利用することが必要であると考えられる。

簡易データ解析手法は次の2つである。第1は図1に示すように圧力分布における特異点を検出し、特異点に近い領域を高解像度で、遠い領域を低解像度で圧縮する技術¹⁾であり、本研究室で開発してきたものである。しかし、自然科学シミュレーション実験の初期段階ではもっと大雑把なクイックルックができる程度のシミュレーションデータの観測で十分であり、この第1の手法はクイックルックのためには無駄な情報量を含んでしまう。そこで、クイックルックのために第2の手法を提案する。この手法実現のためにStrip-Based

[†] 中部大学工学研究科情報工学専攻
Chubu University

^{††} 中菱エンジニアリング株式会社
Churyo Engineering Co.,Ltd.

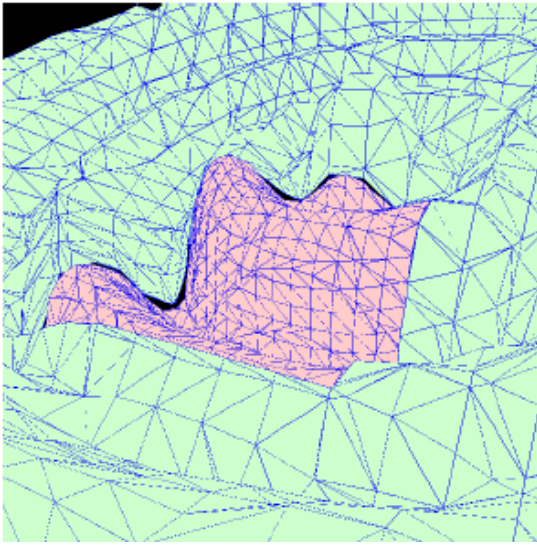


図 1 注視領域の高解像度化

Unfolding (以下 SBU) 手法²⁾を用いた。

自然科学シミュレーションには主にスカラー型データ構造のものとベクトル型データ構造を扱うものがある。本研究では圧力などのスカラー型データの分離点分布に基づくデータ圧縮手法を提案する。

2. 提案手法

2.1 手法 1

スカラー型データである圧力分布などを扱うシミュレーション実験の場合、自然科学者が観測したい領域は圧力分布の谷にあたる分離点付近である。このことを考慮した方法が手法 1 であり、自然科学者が注視する領域を検出し、そこから近いところを高解像度化、遠いところを低解像度化するシステムである。このシステムの流れを次に示す。

- (i) シミュレーションデータに $\frac{\sin(x)}{x}$ 型フィルタを適応し、低解像度化する。これを再生環境の処理能力とメモリ量にて再生が可能になるまで繰り返す。
- (ii) 自然科学者の注視領域を取得する。
- (iii) 注視領域近辺を 3 次補間法である tricube 法を用いて高解像度化する。高解像度化の範囲は任意にできる。

2.2 手法 2

自然科学シミュレーション実験の初期段階では手法 1 適用の圧縮データより、もっと大雑把なデータの観測 (クイックルック) で十分である。手法 2 はシミュレーションデータを圧力の強さに応じて段階的に通信するクイックルックのためのプログラムであり、これを実現するために SBU 手法を使用する。この手法は

ペーパークラフト作成のために 3D モデルデータを平面空間上に自動展開するというツールのアルゴリズムを応用させたものである。ここで、SBU 手法の特徴と適用理由を説明する。3D モデルデータを複数パーツに分割できるため、各パーツごとの圧力の強さを計算し、遠隔利用者の要求に応じてスパコンは各パーツを送信できる。また、この手法は 3D データを 2D データに変換するため、3D モデルデータの圧縮技術にも応用できる。さらに、SBU 手法は 3D モデルデータの多くの特徴を保持するため、特徴情報が必須の自然科学シミュレーションに最適である。

3. 実装

今回、システム検証用モデルデータにはシミュレーションデータではなく、仮想データを使用した。このデータは Marching Cube 法を用いて圧力分布などのデータを等値面表現したものである。利用者側のデータ表示プログラムには OpenGL を利用した。提案手法 1 によって圧縮したデータのサイズは大きくなる傾向があるため、利用者のコンピュータで表示するためにデータ表示プログラムの CPU 並列化記述を行った。並列化率は 98% である。

4. まとめと展望

自然科学シミュレーション実験の中期段階では提案手法 1 を、初期段階では提案手法 2 を使い分けてデータ圧縮することによって、スパコン遠隔利用シミュレーションを支援するシステムを開発した。今後はさらに自動視点位置決定機能を搭載することによってスパコン遠隔利用システムとしての利便性を向上させる。

参考文献

- 1) 宮崎大輝, 高丸尚教, 大石智彦: Mass Data Reconstruction on parallel system, 情報処理学会第 70 回全国大会 (2008)
- 2) MITANI J., SUZUKI H.: *Making Papercraft Toys from Meshes using Strip-based Approximate Unfolding*, ACM Transactions on Graphics, 23, 3(2004), 259-263.
- 3) LEVY, B., PETITJEAN, S., RAY, N., AND MAILLOT, J.: *Least squares conformal maps for automatic texture atlas generation*. ACM Transactions on Graphics, 21, 3(2002), 362-371.