

流体シミュレーション動画のリアリティ 評価手法の一検討

郭富皇^{1,a)} 竹若和之¹ 葛籐トウ² 飯田全広^{3,b)}

概要：現在の画像評価アルゴリズムは画像内のオブジェクトのリアルさを評価することはできない。一方、シミュレーション後に得られる流体の映像のリアルさを評価する手法が求められている。本研究では、人間の感性によるリアリティの評価をラベル値として機械学習させ、流体のリアルさを分類するシステムを提案する。学習用のデータセットは、細分度が異なる3種類の水の流体シミュレーションをblenderを用いて実行し、その動画からラベル付きの画像データを生成する。それを転移学習することで分類器を作成する。未知のシミュレーションデータセットを用いて評価した結果、リアリティを3分類した時の精度は平均81.7%に達成し、実際の水の動画を用いた評価でもリアルと分類できた。

キーワード：機械学習, 画像分類, シミュレーション・応用計算の実装と評価

A study of reality evaluation methods for fluid simulation videos

Fuhuang GUO^{1,a)} Kazuyuki TAKEWAKA¹ Shutong GE² Masahiro IIDA^{3,b)}

Abstract: In this study, a system for classifying the realism of fluid images is proposed. The classifier is trained on a dataset of labeled images generated from water fluid simulations run using Blender. Transfer learning is used to build the classifier. The evaluation achieved an average accuracy of 81.667 for three reality classifications and was able to classify real water videos as realistic. This system provides a new way to evaluate the realism of fluid images in VR applications.

Keywords: Machine Learning, Image Classification, Implementation and Evaluation of Simulation and Applied Computation

1. はじめに

バーチャルリアリティ (VR) 技術の一つとして、流体シミュレーションがある。流体シミュレーションは、水や空気、気体などの流体をデジタルで再現することが可能な技

術である。映画や視覚効果などでは長い間利用されているが、リアルタイム VR 環境での流体シミュレーションは膨大な計算資源が必要なため困難である。しかし、近年のハードウェアとソフトウェアの進歩により、リアルタイムの流体シミュレーションが可能になりつつある。

現在、シミュレーション動画の評価は、画質やノイズに関するアルゴリズムはあるが、人が感じるリアリティを評価する方法はない。そこで本研究では、機械学習を用いて人がラベル付けした判断基準を学習させ、流体シミュレーション動画を自動的に評価するシステムを提案する。このシステムは、人間の主観的評価基準を使って機械的に評価することで、より迅速かつ効率的な評価が可能になる。また、リアルタイムの流体シミュレーションにも応用できる可能性があり、ゲームやバーチャルリアリティなど、流体

1 熊本大学大学院自然科学教育部

Graduate School of Natural Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto, 860-8555 Japan

2 熊本大学工学部情報電気電子工学科

Department of Information and Electrical Engineering, Faculty of Engineering, Kumamoto University, 2-39-1 Kurokami, Kumamoto, 860-8555 Japan

3 熊本大学 大学院先端科学研究部

Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto, 860-8555 Japan

a) guo@st.cs.kumamoto-u.ac.jp

b) iida@cs.kumamoto-u.ac.jp

シミュレーションを含むインタラクティブなアプリケーションの開発が期待される。

2. 流体シミュレーション

本章では、この章では、特に粒子ベースの流体シミュレーション技術を掘り下げ、2つの手法の長所と短所を比較する。さらに、粒子状態から液体状態にレンダリングするプロセスについても解説する。最後に、本研究の目的について説明する。

2.1 流体シミュレーションの方法

流体シミュレーションは、デジタル環境で水やガスなどの流体の動作をシミュレートするために使用される手法である。この技術には、グリッドベースとパーティクルベースの2つの主要な方法がある。

グリッドベースの流体シミュレーションは、流体ドメインをグリッドセルに分割し、各セルが少量の流体を表すように設計されている。次に、流体力学の法則に基づいて、各セルの密度、速度、圧力などのプロパティを更新することにより、流体の動作をシミュレートする。この方法は実装が簡単だが、計算コストが高くなり、複雑な流体の動作を正確に捉えることができない場合がある。

一方、粒子ベースの流体シミュレーションは、相互に作用し、周囲の環境と相互作用する個別の粒子の集まりとして流体をシミュレートする。この方法は計算量が多いが、特に複雑な形状や挙動を持つ流体に対してはより現実的な結果を生成することができる。

粒子ベースの流体シミュレーションの主な利点は、グリッドベースのシミュレーションよりも現実的な結果を生成することができる。流体シミュレーションにおける粒子ベースのシミュレーション手法と計測時間の関係をさらに明らかにするため、2,680、4,012、4,684、5,754、7,284、9,072、12,392個という様々な粒子数の粒子シミュレーションを実施する。シミュレーションで生成された画像を解析し、粒子数の違いによってシミュレーションの見た目や性能がどのように変化するかを評価する。

この評価に用いたコードは、文献[1]のc言語ソースコードを改変したものである。図1に、ソースコードの粒子数と、粒子数12392個を追加したシミュレーション結果を示す。このシミュレーションは、静止した液体の水の塊が重力の影響を受けて崩壊する様子を示しており、粒子数がシミュレーションの見栄えや性能に影響を与えることを説明している。

流体シミュレーションにおけるパーティクル数と実行時間の関係を図1に示す。見ての通り、両者の間には指数関数的な関係があり、粒子数の増加に伴い、粒子数、実行時間ともに急激な増加を示している。

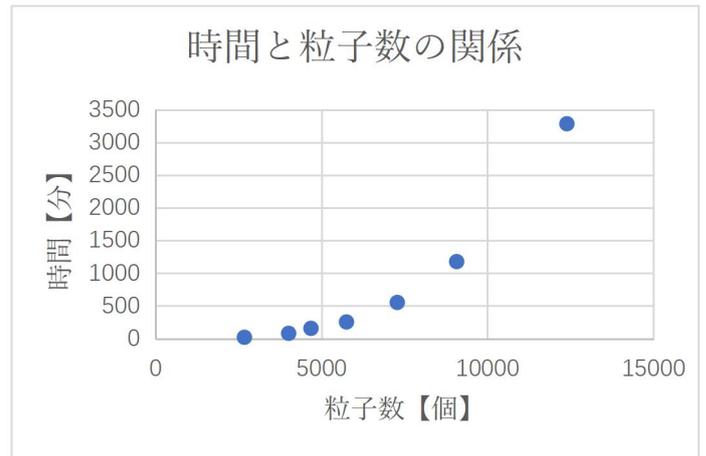


図1 時間と粒子数の関係

Figure 1 Relationship between time and number of particles

また、粒子数がシミュレーション時間に与える影響に加え、シミュレーション性能に影響を及ぼす可能性のある他の要因についても調査する。たとえば、シミュレーション対象の流体挙動の複雑さは、流体を正確にシミュレーションするために必要な計算資源量に影響を与える可能性がある。また、使用するコンピュータのハードウェアの性能も考慮し、従来のCPUではシミュレーションに数時間から数日を要するため、高い計算能力を持つ実験用プロセッサを使用する必要があることもわかる。

流体シミュレーションの研究では、流体の挙動を効率的にシミュレートするアルゴリズムの開発、シミュレーションの精度を向上させるための機械学習技術の使用、流体シミュレーションの統合など、取り組むべき課題が数多くある。

2.2 表面再構築方法

現在、パーティクルを水にレンダリングするには、2つの一般的な方法がある。表面メッシュ再構築と画面空間に基づく再構築である

画面空間に基づく流体レンダリング手法は、スクリーン空間で流体表面を復元することに焦点を当てており、2つのステップで構成されている。最初のステップでは、流体粒子をスクリーン空間にレンダリングし、深度と厚み情報を得て、テクスチャに格納し、後処理演算を行って、流体表面をぼかし、滑らかにする。次のステップでは、深度と厚みのテクスチャを使用して、流体のライティング計算を行い、最終的な流体表面を生成する。

実験で使った方法は表面メッシュ再構築法である。図2のように、粒子を大量のメタボールと見なし、流体メッシュを再構築する。このスキームは効果が高く、一般的な映画・テレビレベルの流体レンダリング効果はこの方法で実

現されており、フレームごとにメッシュを再構築しなければならず、性能コストが高いという欠点がある。

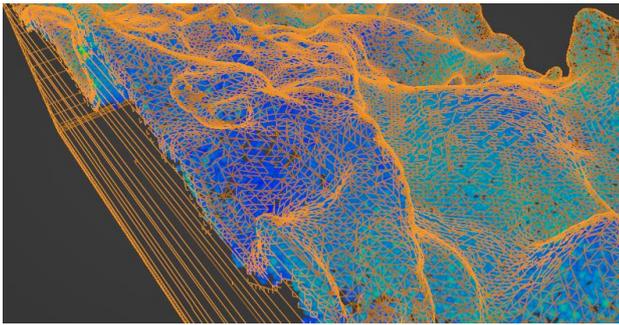


図 2 表面再築法

Figure 2 surface re-building method

2.3 本研究の目的

現在の画像評価アルゴリズムは、画像そのものの品質や元画像との比較による歪みの度合いを評価することはできるが、画像内のオブジェクトのリアルさを評価することはできない。一方、VR への適応をターゲットとしている流体シミュレーションの研究では、シミュレーション後に得られる流体の映像のリアルさを評価する手法が求められている。本研究の目的は、流体シミュレーション映像の主観評価方法を見出すことである。

3. 研究手法

本研究では、人間の感性によるリアリティの評価をラベル値として機械学習させ、流体のリアルさを分類するシステムを提案する。学習用のデータセットは、粒子数などのパラメータが異なる 3 種類の水の流体シミュレーションを blender を用いて実行し、その動画からラベル付きの画像データを生成する。それを転移学習することで分類器を作成する。本章では、本研究で用いた具体的な実験手順や技術など、研究方法を説明する。

3.1 データセット

データセットを作成するため、Blender を使って、照明の位置や水流の初速などを調整し、細分化精度を変えて表面がどのように振る舞うかを広範囲に渡って実験を行う。流体レンダリングは非常に時間がかかるため、最終的な実験データとして、10、32、96 の細分化精度を選ぶ。8 種類の水の動きをシミュレーションし、照明などの環境を細かく調整することで、リアルで視覚的にわかりやすいシミュレーションを実現する。そして、各動画のスクリーンショットを 100 枚ずつ撮影し、学習データセットとする。液面をより適切に分類するために、水の色など他のパラメータも 3 つの分類シミュレーションで可能な限り標準化されている。

blender では、細分化（粒子数）のレベルが異なる 3 つの映像品質をシミュレートすることができ、次にその映像をスクリーンショットして画像のデータセットを作成する。水の動きや粒子の数を多様に表現するため、学習用に 15 種類の異なる流体シミュレーション動画、推論用に 4 種類の別の流体シミュレーション動画を作成している。これにより、様々な条件下で評価システムの性能をより良く評価することができるようになる。

表 1 データセット

Table 1 Dataset

class	trainset	testset	total
悪い類	1437	400	1837
普通類	1473	400	1873
良い類	1259	400	1659

このデータセットは、水の動きや粒子の数を幅広く表現しており、流体シミュレーションの有効性を評価する我々の研究にとって、貴重な資料になる。

本研究で使ったデータセットに関しては、いくつかの分類基準がある。我々は、流体シミュレーション後の表面の粗さによってデータセットを分類する。図 3 に示すように細分度 10 を「悪い」にラベリングし、同様に図 4 の細分度 32 は「普通」に、図 5 の細分度 96 は「良い」にラベリングする。「悪い」類で、流体シミュレーションの最低基準のレンダリングを表す。このクラスは、細分化精度が低いため、流体の表面に特徴があり、はっきりとした三角形の外観を持つようにレンダリングされる。このクラスの流体は、図 3 のように、ポリゴンが流れる溶けるダイヤモンドのようなイメージになるかもしれない。「普通」類で、表面の粗さは他の 2 つの類の中間に位置する。この分類では、流体レンダリングの痕跡がいくつか見られるが、表面は「悪い」類のような三角形にはなっていない。このカテゴリの例は、提供された図 4 に示されている。「良い」類で、表面に三角形がある非常に細かい、ほとんど見えないレンダリングを表している。これは、提供された画像例図 5 で見られるように、表面が滑らかで細かいことが特徴である。



図3 悪い類例
Figure 3 Bad analogy



図4 普通類例
Figure 4 General analogy



図5 良い類例
Figure 5 Good analogy

全体として、この分類システムによって、流体シミュレーション・アルゴリズムの性能を評価し、滑らかで精巧なサーフェスを生成する能力を評価することができる。この

分類基準は、表面の滑らかさに影響を与える要因を特定し、その後の実験でアルゴリズムを改善するために使用することができる。

3.2 学習と転移学習

データセットを作成した後、データセットの規模が小さいため転移学習を採用を用いる。評価に使用する機械学習のモデルは、以下に示す3モデル12種類である。VggNet[2]は画像分類分野で非常に良い結果を出しているため、実験においてVggNet19を採用する。Vgg19は、19層からなるモデルであり、層が多いため精度が高いとされているが、一方で、層数が増えるとモデルが複雑化するため、過学習が起りやすくなるという問題がある。このため、Vgg19bnはバッチ正規化を導入することで、層数を増やしつつも過学習を防ぐことができるので同時にVgg19bnも選択する。ResNet[3]は深いアーキテクチャを効果的に扱い、勾配が消失する問題を軽減できることで知られている。ResNetが深くなるにつれて学習が難しくなる問題があるため、浅い層から始めて徐々に深い層に向かって学習を行うことが一般的である。このように異なるレイヤーのResNetを選択することで、より効率的に学習を進めることができると考えられる。また、転移学習が本当に分類精度を向上させるのかを検証するために、事前学習を行わないresnet34の実験も追加する。ResNetに加えて、トランスフォーマーアーキテクチャ、より具体的にはVision Transformer[4]の実験も行う。ViTを比較対象としてResNetと比較し、その性能を評価する。しかし、ViTの精度が十分でない場合、事前学習したモデルの最後の層を調整するために、転移学習を用いることにする。この方法は、ファインチューニングと呼ばれている。

私たちがこのチューニングを行ったが、最後の層のファインチューニングほど精度は高くなかった。このことは、この特定のタスクに対するニューラルネットワークの精度を向上させるために、事前学習済みモデルの最後の層をファインチューニングすることの重要性を示している。全体として、このアプローチは、流体シミュレーションのタスクに対するニューラルネットワークのパフォーマンスを向上させるのに有効である。

全層を調整する戦略をとったResNetニューラルネットワークとは異なり、VGG19は構造が異なるため、別のアプローチが必要である。VGG19ニューラルネットワークは、畳み込み層とプーリング層を含むVGG19.featuresと、3つの線形層とソフトマックス分類器を含むVGG19.classifierの2つの部分から構成されている。

VGG19ニューラルネットワークを我々のタスクに適合させるため、事前にトレーニングされたVGG19モデルを利用し、VGG19.featuresの畳み込み層とプーリング層を凍結させる。そして、最終的な線形層を我々の3クラスの要

件に合うように修正する。実験では、マルチクラス損失関数として CrossEntropyLoss を用い、学習率 0.0001、運動量 0.9 の SGD オプティマイザを使用する。

4. 評価

本章では、粒子ベースシミュレーションとニューラルネットワークの技術を用いた、流体シミュレーション表面の分類に関する評価方法実験結果を示す。

4.1 評価方法

我々のシステムの性能評価に関しては、画像分類の分野で一般的に使用されているいくつかの指標を使用する。これらは、精度、正確さ、再現性、F1 スコアなどである。我々のシステムの精度は、予測されたラベルとデータセットの画像の真のラベルを比較することによって測定される。精度と再現率は、データセット内の特定のクラスを識別する際の我々のシステムのパフォーマンスを評価するために使用される。F1 スコアは、精度とリコールの組み合わせであり、システムの性能の包括的な評価を提供する。

混同行列を使って、我々の分類システムの性能を評価することができる。混同行列は、分類アルゴリズムの性能を説明するためによく使われる表である。これは、真陽性、真陰性、偽陽性、偽陰性の数を見ることで、アルゴリズムがどの程度の性能を発揮したかを確認することができる。真陽性は、システムが液面を高品質と正しく認識した場合であり、偽陽性は、システムが液面を高品質と誤って認識した場合である。同様に、真陰性とは、システムが液面を低品質または通常品質と正しく認識した場合であり、偽陰性とは、システムが液面を低品質または通常品質と間違っ

て認識した場合である。これらの指標に加え、我々はシステムの性能を評価するために目視検査も行う。これは、我々のシステムが予測したラベルを手作業で分析し、画像の真のラベルと比較するものである。この方法は、定量的な評価基準では見落とされる可能性のあるエラーや誤分類を特定するのに特に有効である。

4.2 評価結果

表 2 8 種類時 Resnet34 評価
Table 2: Resnet34 ratings at 8 types

	precision	recall	f1-score	support	accuracy
悪	0.944444	0.510000	0.662338	100.0	0.510000
普	0.633094	0.880000	0.736402	100.0	0.880000
良	0.906542	0.970000	0.937198	100.0	0.970000
macro avg	0.828027	0.786667	0.778646	300.0	0.786667
weighted avg	0.828027	0.786667	0.778646	300.0	0.786667

表 3 精度対比

Table 3 Accuracy Comparison

ネットワーク	事前学習処理	転移学習処理	精度%
VGG19	ImageNet1000	全結合層を調整	76.75
VGG19_bn	ImageNet1000	全結合層を調整	76.583
Resnet18	事前学習処理なし	処理なし	38.417
Resnet18	ImageNet1000	全結合層を調整	67.77
Resnet18	ImageNet1000	全層を調整	68.417
Resnet34	ImageNet1000	全結合層を調整	68.25
Resnet34	ImageNet1000	全層を調整	81.667
Resnet50	ImageNet1000	全結合層を調整	72.667
Resnet50	ImageNet1000	全層を調整	77.5
Resnet152	ImageNet1000	全結合層を調整	65.917
Resnet152	ImageNet1000	全層を調整	73.167
VisionTransformer	事前学習処理なし	処理なし	75.0

表 2 に示すように、8 種類の動画を切り出した初期データセットを用いて、初期実験を行った。各モデルの精度を測定しているが、事前学習を行わない場合の精度は悪く、転移学習後は大幅に精度が向上していることが分かる。評価には、各カテゴリー 100 枚の画像を収録した動画 1 本を検証用のテストセットとして用い、得られた結果を表 3 に示す。よく分類されたカテゴリーの解像度は高いものの、

平均や悪いカテゴリーの分類能力は低く、この結果はデータサイズが足りない問題なのかと反省している。

そのため、データ量を大きくし、学習用の動画を8本から15本に増やす。同時に、テストセットも元の3倍に拡張する。表3に示すように、ネットワークはResnetの未学習と移行学習の両方を使用することをテストする。一連の比較を行った結果、どちらのネットワークでも、転移学習後に得られる正解率に大きな差はないことがわかった。分析の結果、表面に三角形の構造が見える「低品質」と、表面に非常に小さな三角形が見える「良品質」が分類に影響を与えたと判断し、Transformer実験も追加する。評価結果は表4で、様々な分類をより直感的に理解できるように、分類の精度指標に加えて、表5に示すような混合行も追加する。推論に失敗するケースは、悪い類を普通類と推論するケースがほとんどで、それ以外は少なかった。

表4 Vit 評価

Table 4 Vit Evaluation

	precision	recall	f1-score	support	accuracy
悪	0.951220	0.39	0.553191	100.0	0.39
普	0.425926	0.46	0.442308	100.0	0.46
良	0.649007	0.98	0.780876	100.0	0.98
macro avg	0.675384	0.61	0.592125	300.0	0.61
weighted avg	0.675384	0.61	0.592125	300.0	0.61

表5 Vit 混同行列

Table 5 Vit Mixing Rows

正解	推論		
	悪い類	普通類	良い類
悪い類	28	72	0
普通類	1	98	1
良い類	0	1	99

さらに、実環境における本システムの有効性を検証するため、動画ファイルを読み込んで動画中の水の画質を分類するウィンドウを構築して追加実験を行う。これにより、流体のシミュレーション映像だけでなく、実際に動いている水の映像でシステムの性能を検証することができる。

私たちは、水を注いでいる映像を録画し、それをシステムのテストケースとして使用する。その結果、私たちのシステムは実際の水面を高品質と正確に分類できることがわかり、私たちのシステムが単なるシミュレーション映像にとどまらず、実世界のシーンにも適用できることが証明された。また、各フレームの分類精度は0.8以上であり、我々

の分類システムがロバストであり、実際のアプリケーションで使用できることを強く示唆している。図6にスクリーンショットの一例を示す。

全体として、これらの実験は、我々のシステムが実世界のシナリオに適用可能であり、ビデオ内の水の画質を正確に分類できることを示す強力な証拠であり、映画制作、ゲーム開発など様々な産業で有用なツールとなっている。



図6 リアルタイム予測例

Figure 6 Example of real-time forecast

5. 考察

流体シミュレーションは計算負荷が高く、かなりの時間を要する作業である。私たちのチームの実験によると、たった1万個のパーティクルを使った比較的単純な流体シミュレーションでさえ、特に流体が壁や障害物に衝突する場合には、実行にかなりの時間がかかることが分かっている。そのため、強力なハードウェアを用いても、リアルタイム性を実現することは困難である。そのため、リアルタイムな流体シミュレーションを実現するためには、人工知能や機械学習などの高度な技術の活用が不可欠であると認識している。

流体レンダリングに関しては、リアルタイム性とリアルさを両立させるためには、適切なサンプル数とノイズ低減指数の選択に大きく依存することがわかった。これは複雑で時間のかかる作業であるが、高品質のシミュレーションを作成するためには必要なことと言える。

また、液面シミュレーションの品質分類の精度を向上させるために、転移学習を採用する。実験の結果、Vision transformerは、Resnetよりも高い精度を達成することができたが、転移学習を使用したResnetほど高くはない。さらに、システムは画質の良いものと悪いものを区別することができたが、画質の悪いものと普通のものとの区別は、人間が液面上の多くの三角形を区別することが難しいため、弱いことがわかった。また、異なる画質でも交換可能であることが浮き彫りになった。

このシステムのもう一つの欠点は、現実的な物理学の常識に従わない流体シミュレーションを認識しないことであ

る。現実的な流体では、流れの過程で様々な形態の障害物に遭遇することになる。例えば、水の流れが壁を通過する可能性があるが、これは明らかに物理的にありえないことだが、流体の表面が非常に滑らかであれば、我々のシステムはそれでも非常にリアルなシミュレーションであると判断する。しかし、リアルな商業広告では、夏のビールのCMで水しぶきが浮いていたり、化粧品のCMで水の流れが浮いていたり、レンダリングを高めるために流体シミュレーションは物理的にリアルでないことが一般的である。

このシステムの大きな特長の一つは、流体シミュレーションのリアルタイム解析が可能である。ビデオ検証用の拡張機能を搭載することで、ライブカメラ映像を解析し、行われている流体シミュレーションを素早く予測することができる。これは、ライブストリーミングのように、リアルタイムでの解析が不可欠なアプリケーションで特に有効である。さらに、ローカルカメラの呼び出し機能を搭載することで、他のソフトウェアやハードウェアに簡単に組み込むことができ、幅広い用途に対応できる汎用性の高いツールとなっている。

6. まとめ

本論文では、流体シミュレーションで作成された動画における液面の画質分類システムの開発に焦点を当てた。まず、流体シミュレーションの基本的な手法である格子法、粒子法、および流体シミュレーションで用いられるレンダリング技術について紹介した。また、それぞれの手法の利点と欠点、および実験で使用した手法についても説明した。

試行錯誤の結果、移動学習の手法とソフトウェア Blender を用いてデータセットを開発した。このデータセットを用いて分類用のニューラルネットワークを学習させたところ、最大で 81.67% の精度を達成することができた。さらに、ネットワークの精度を向上させるために、転移学習法を適用した。

この研究で直面した主な課題の 1 つは、大量のアノテーションデータが必要で、その作成に時間がかかることだった。移行学習は、少ないデータセットで高精度のネットワークを学習できるため、この問題に対する効率的な解決策であることが証明された。

また、実際的水流映像を用いて本システムの実用性を検証した結果、本システムが現実的水流の有無を正確に識別できることを確認した。今回の実験結果は、最新技術である Transformer が視覚分類の分野で有効であることを示すとともに、今後の研究の基礎となった。

最後に、常識的な物理学に適合しない流体シミュレーションを認識する上での本システムの限界についても述べ、ターゲット検出や画像分割の技術を追加するなどの解決策の可能性を示唆した。また、流体シミュレーションにおけるリアルタイム性能と計算量のトレードオフの問題についても、予測

のための計算量を少なくすることや、この場合の評価システムを作ることで解決できるかもしれない。

参考文献

- [1] 粒子法入門 流体シミュレーションの基礎から並列計算と可視化まで C/C++ソースコード付 単行本 (ソフトカバー) - 2014/6/28 越塚 誠一 (著), 柴田 和也 (著), 室谷 浩平 (著)
- [2] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [3] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [4] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.