

動画中の物理現象を再現するためのシミュレーションパラメータ推定

多加谷 海斗† 岡部 誠†

静岡大学†

1 はじめに

今日、CG 技術の進歩により編集ソフトを用いれば物体の破壊や爆発といった豪華なシーンの撮影が可能である。しかし、近年の映画にもリアリティを追求するために、実物を破壊や爆破して撮影されたものも少なくない。実物を用いる撮影方法は非常に費用が掛かるため、シミュレーションで再現することについて研究する。例えば、ボールの落下という物理現象をシミュレーションするには、ボールの形状や大きさ、初速度、質量、弾性率などのパラメータを設定する必要がある。このとき、パラメータを設定してシミュレーション映像を確認し、再現したい映像と違うのであれば、パラメータを設定し直すという作業を繰り返さなければならない。この作業はパラメータ数が応じて作業の難易度や修正に必要な時間が変化し、複雑な現象を再現するのであれば、適切なパラメータ設定は難しく、多くの時間が必要となる。そこで、本研究ではボールが落下するなどの物理現象が映っている動画からパラメータを推定するシステムの開発を目標としている。

2 既存研究

2.1 シミュレーション手法

物理シミュレーション分野において、計算処理の効率化やシミュレーション手法の改良など様々な研究が盛んにおこなわれている。我々は、MPM[1]に着目した。MPM は、大規模な変形・破壊シミュレーションが得意であり、弾性、粘性、塑性などの多様な材質を表現可能な最先端な手法である。本研究では、ゴムのような弾性体や雪のような塑性体など様々な物体のパラメータ推定を目標としているため、MPM をシミュレーション手法として採用する。

2.2 水面シミュレーションのパラメータ推定

本研究では、動画から対象物のパラメータを推定するシステムを提案している。システムから推定されたパラメータをシミュレータに与えることで動画中の物

理現象を再現する。また、動画ではなく一枚の写真から水面のシミュレーションをする研究[2]が発表されている。Sugimoto ら[2]が提案する手法のフローチャートを図 1 に示す。入力された画像の水面とそうでない領域に対してセグメンテーションを行い、水面領域の反射テクスチャを生成する。反射テクスチャとは、水面から波によるぼやけや水の色相を除去した反射している風景のみを予測した画像である。水面のレンダーリングには、風、波、カメラ、光源に関する 21 個のパラメータを設定する。カックウ探索を用いてパラメータ推定し、入力画像のシミュレーションを得る。カックウ探索とは、ランダムにパラメータを設定し、そのシミュレーションを評価することの繰り返しによって解を探索する最適化手法の一種である。カックウ探索は何度もシミュレーションを行う必要があるが、水面のシミュレーションは非常に高速であるため、採用している。本研究で用いる MPM はあまり高速でないため、カックウ探索を用いると非常に多くの時間がかかってしまう。そのため、後述する別の手法によってパラメータを推定する。

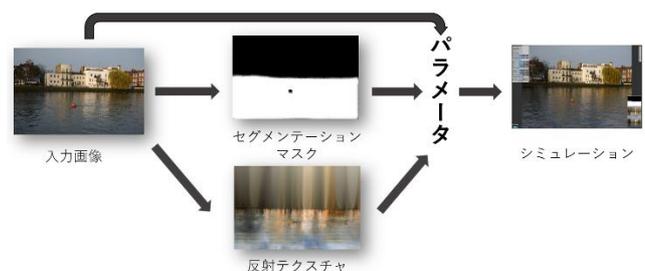


図 1 : Sugimoto らの提案手法のフローチャート

3 提案手法

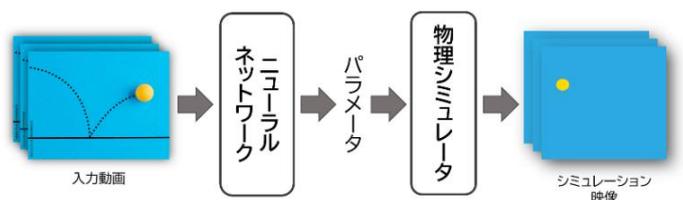


図 3 : 提案手法のフローチャート

Estimation of simulation parameters to reproduce physical phenomena in video

† Kaito Takaya, Makoto Okabe, Graduate School of Engineering, Shizuoka University

図 3 に本研究の提案手法のフローチャートを示す。ボールが地面に当たって跳ねるなどの物理現象の動画をニューラルネットワークに入力すると、適切なパラメータを推定される。推定されたパラメータを物理シミュレータに設定することで入力を再現するシミュレーション映像が得られる。このようにして、動画からシミュレーション映像を作成する。提案手法のニューラルネットワークは動画から特徴量を抽出してパラメータを決定する。動画は時系列データであるので、画像処理において広く用いられている 2 次元の畳み込み演算を用いると時間的情報が得られない。そのため、本研究では空間的特徴と時間的特徴を抽出することができる 3 次元に拡張した畳み込み演算(3D Convolution) [3]を使用している。3D 畳み込み演算の処理を図 4 に示す。3D 畳み込み演算では、幅 W 、高さ H 、奥行き L の 3 次元データを時間方向にも演算しているため、3 次元データが出力されている。

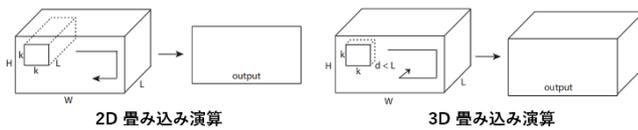
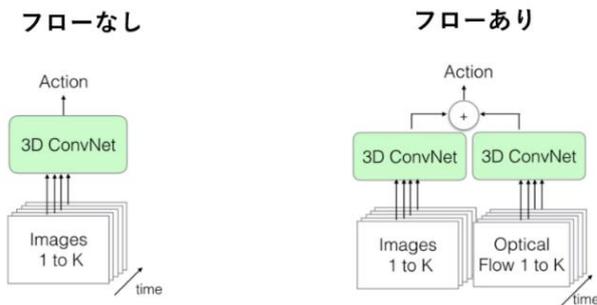


図 4: 2 次元, 3 次元畳み込み演算の比較

また、入力に関しては Tran らが提案する C3D モデル[3]や C3D を改良した I3D モデル[4]において、動画とオプティカルフローを併せて入力することで分類精度が向上すると報告されている。



データセット	UCF-101	HMDB-51	Kinetics
フローなし	83.6	43.2	52.4
フローあり	91.2	58.3	65.6

図 5: オプティカルフローなしモデルとありモデルの比較

オプティカルフローとは、動画フレーム間での移動量を計算したものである。図 5 にオプティカルフローが分類精度に与える影響を示す。オプティカルフローなしのモデルでは動画の特徴量から分類をしているが、

オプティカルフローありのモデルでは動画の特徴量とオプティカルフローの特徴量を合わせたものから分類をしている。また、図 5 の UCF-101, HMDB-51, Kinetics とは、ラベル付けされた日常的な行動やスポーツをしているシーンの動画データセットであり、いずれに対しても、オプティカルフローを入力に追加することで大幅に精度が向上していることがわかる。よって、動画分類においてオプティカルフローは効果的な特徴量と考えられる。

4 今後の方針

既存手法[2]では、シミュレーションの対象を水面に限定しているため、ボールが跳ねるといった物体の運動などの他の物理現象については対応していない。そのため、本研究では、物体の運動や変形、破壊などの物理現象に対しても再現シミュレーションを作成することを目標としている。現在は、ボールが跳ねる映像のパラメータ推定を課題に設定しているが、今後はさらに複雑な課題に取り組むことを考えている。本手法を用いることで数多くの物理現象に対しても適切にパラメータを推定することができることを示したい。また、C3D[3]や I3D[4]では、人間の行動を分類するタスクの精度向上に取り組んでおり、人やモノなどが含まれる大規模なデータセットを用いて事前学習を行うことで精度向上が可能である。しかし、本手法では物体の運動や性質に関する独自のデータセットが必要であるため、学習用データセットを大量に作成することが課題である。また、時間的情報を考慮するモデルとして LSTM が広く用いられており、本研究のニューラルネットワークに組み込むことで精度向上に役立つのかどうかを調査する。

参考文献

- [1] C. Jiang, C. Schroeder, J. Teran, A. Stomakhin and A. Selle, "The Material Point Method for Simulating Continuum Materials", SIGGRAPH (2016)
- [2] R. Sugimoto, M. He, J. Liao and P. V. Sander, "Water Simulation and Rendering from a Still Photograph", SIGGRAPH Asia (2022)
- [3] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks", ICCV (2015)
- [4] J. Carreira and A. Zisserman, "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset", CVPR (2017)