



第 27 回日本バーチャルリアリティ学会大会論文集 (2022 年 9 月)

ディープラーニングによる画像認識を用いた物体の跳ね返り挙動の再現

Reproducing the bouncing behavior of objects using deep learning image recognition

松川将也¹⁾, 山田孝治¹⁾

Masaya Matsukawa, and Koji Yamada

1) 琉球大学大学院理工学研究科工学専攻 (〒 903-0213 沖縄県中頭郡西原町千原 1)

概要: XR のアプリケーションでは空間認識やトラッキングなど, 現実空間の認識にディープラーニングによる画像認識が用いられる. 本研究は画像認識を用いて平面の跳ね返りの推定を行うことで, 現実空間内の物体の挙動を仮想空間で再現することを目的とする. 80 個のデータを学習させたモデルの精度を検証すると, RSME が 11.00 程度にとどまり, 実測と予測のデータの相関係数が 0.718 で強い相関があることがわかった. また推定された跳ね返りで自由落下のシミュレーションを行うと挙動が実物体の挙動と近似している.

キーワード: 拡張現実, 画像認識, ディープラーニング

1. はじめに

近年, AR(Augmented Reality) や VR(Virtual Reality) といった XR 技術を用いたコンテンツが普及し始めている. VR や AR の体験において, 現実空間の認識技術を用いる場面が多く, 特に AR コンテンツでは仮想物体と現実空間内の物体のインタラクションを実現するための研究やそれを応用したコンテンツなどが数多く存在する.

河出らは AR マーカーを用いて手や現実空間の物体と仮想物体の接触を実現している [1]. Breen らの研究では 3D モデルを実物体の位置に合わせる手法と深度マップを生成する手法で仮想物体の実物体への衝突を実現した [2].

近年では, ディープラーニングによる画像認識を AR に応用するコンテンツや研究も存在する. Meta 社の Meta Quest2 ではハンドトラッキングの際にカメラから隠れたり, 識別できない速さで動いている場合にディープラーニングを用いて手の位置やポーズ推定を行っている [3]. DeTone らによる研究では, XR のアプリケーションにおいて自己位置推定などで用いられる SLAM をディープラーニングを用いた新たなトラッキング方法で発展させた [4]. Innamorati らの研究では単一画像内の実物体に対する仮想物体の衝突の挙動をニューラルネットワークを用いて再現した [5].

これらの研究は物体の位置や形状などの認識を行い仮想物体の実物体に対する衝突判定を与えることで, インタラクションを実現している. しかし, 衝突の判定だけでは現実空間と異なる挙動をとるため, 仮想物体がそこにあるかのような没入感を得るためには力学的な要素を考慮するべきだと考える.

竹内らの研究では環境マップにあらかじめ測定した反発係数を適応させることで, 仮想物体と現実空間の力学的な相

互作用を実現した [6]. この手法では, あらかじめ反発係数を測定する必要があるため, あらかじめ定められた場所以外での活用ができない. そこで, 本研究では不特定の場所で仮想物体と現実空間の力学的な相互作用を実現するために, 反発係数の測定を画像推定モデルを用いて跳ね返りを予測することにより自動化することを試みる.

2. 提案手法

本研究では, 壁や床といった現実空間内の平面の画像から, その平面に対する物体の跳ね返りをディープラーニングによる画像認識を用いて推定することを目的とする画像推定モデルを提案する.

2.1 データセットの作成

物体の跳ね返りを学習させるために独自のデータセットを用意する. データには, 跳ね返りの高さをラベルとした様々な種類の平面の写真を用いる. 跳ね返りの高さの計測には, メジャー等の計測道具を横に配置した状態でボールが一定の高さから落下し, 跳ね返る様子を撮影する. 撮影した動画から跳ね返り後の最高位置のフレームを記録する.

計測には直径 40mm の卓球用のボールを使用した. 安価なこと, 同仕様のボールの入手が容易, 軽く扱いやすいという理由で採用した. また, 計測対象の平面は地面, カーペット, ベッドの上, 机の上等の水平で, 垂直落下で跳ね返りの測定ができる物を対象としている.

2.2 モデルの学習

本研究では跳ね返りの推定モデルを作成するために, Keras の学習済みモデルである Xception[7] に用意したデータセットを転移学習させる. Xception の出力層側にある全結合層を除去し, 新たな全結合層を結合し出力層のユニット数を一

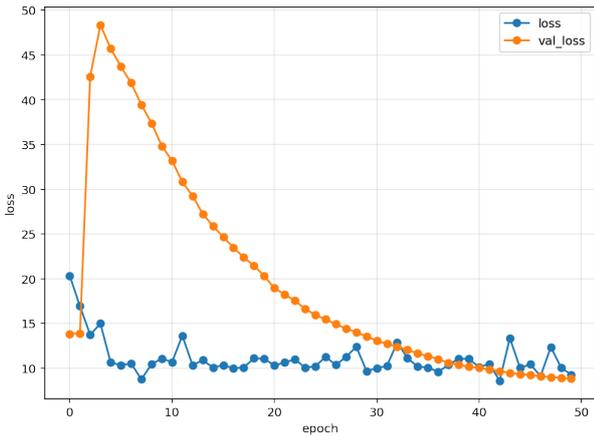


図 1: RMSE の推移

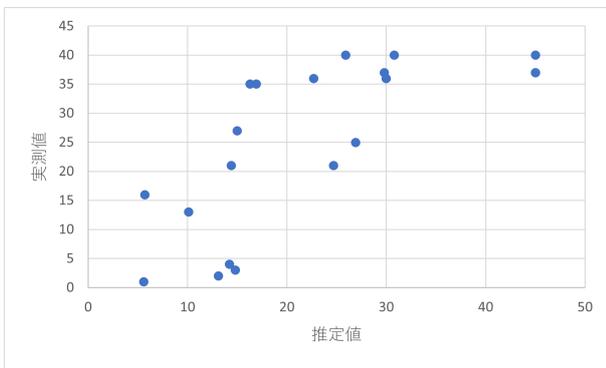


図 2: 散布図

つにした回帰モデルとする。このモデルで推測された跳ね返りの値を仮想空間内のオブジェクトに適応させることで、現実空間の挙動に近づくと考える。

3. 実験 1

3.1 実験概要

本実験では跳ね返りの推定モデルの学習と評価を行う。本実験の学習で用いるデータセットは 25 種類の平面の画像を 100 枚のうち、80 枚をトレーニングデータとし、20 枚をテストデータとして学習を行う。損失関数に RMSE を用いるため、学習による RMSE の推移をグラフに示す。また、テストデータから横軸を予測データ、縦軸を実測データとした散布図を出力し、相関からモデルを評価する。

3.2 実験結果

図 1 は横軸を epoch 数とした RMSE の推移を表したグラフである。トレーニングデータでは epoch5 から変化があまり見られないが、テストデータでは epoch4 以降 val_loss の値が減少し続けている。evaluate_loss は 11.00 となった。また、散布図を図 2 に示す。予測データと実測データの相関係数は 0.718 で、強い相関があることがわかる。



図 3: シミュレータの様子

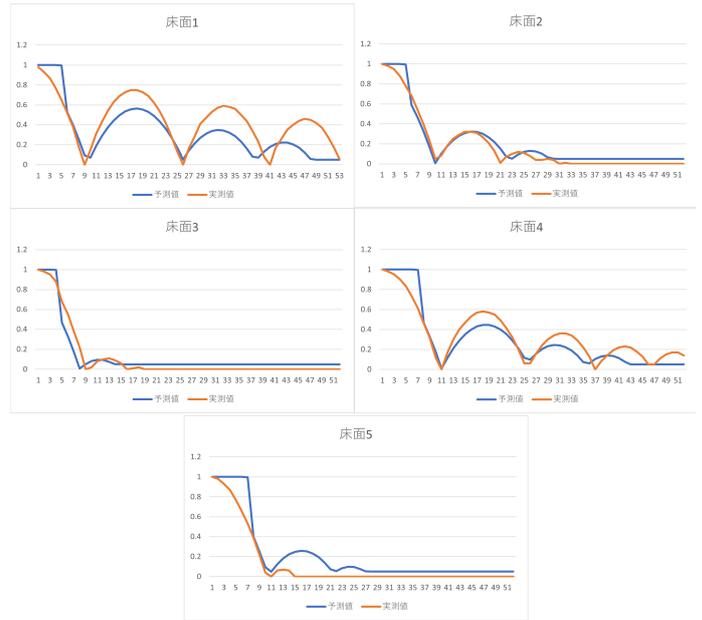


図 4: 自由落下の挙動の比較

4. 実験 2

4.1 実験概要

本実験では、現実空間と仮想空間のそれぞれの自由落下の様子を記録し、比較する。まず、対象となる平面を 5 つ選択し、その平面に対して現実空間での自由落下の様子を 30fps で撮影したものを 1 フレーム単位で高さを記録する。そして、Unity 上で床面とボールを作成し、簡易的な自由落下のシミュレータ (図 3) を用意する。このシミュレータの跳ね返りの値としてモデルが推定した値を設定し、フレーム単位で高さを記録し、記録したフレームごとの高さの推移を、グラフ化することで挙動を比較する。

4.2 実験結果

5 つの平面の自由落下の挙動は図 4 のようになった。このグラフは横軸がフレーム、縦軸がそれぞれのフレームでのボールの高さを示している。平面ごとに予測の挙動に変化があり、それぞれ実測データと近似していることがわかる。

5. 考察

実験1では、トレーニングデータの損失が epoch5以降変化があまりないが、テストデータにおいては減少し続けているため学習自体は進んでおり、誤差は10前後で収束するものと思われる。また、散布図から相関を確認できるため、跳ね返りやすい平面と跳ね返りづらい平面を大まかに推測することができていることがわかる。また、実験2では、5つの平面の現実空間とシミュレータの自由落下を比較し、それぞれの平面で現実空間とシミュレータの挙動が近似していることがわかる。完全に一致してはいないものの、平面の種類に応じて挙動を変化させている。

6. おわりに

本稿では、ディープラーニングによる画像認識を用いて平面の跳ね返りを推定する手法を提案し、検証した。その結果、平面画像から跳ね返りの予測を行うことができ、仮想物体の挙動を実物体の挙動に近づけることができた。しかし、検証のためのサンプルが少ないためより多くの平面に対して検証する必要があると考えられる。また、データセットもデータの数が少ないためより大量のデータを用いて学習をすることでより精度を上げられる可能性もある。また、本推定モデルはARアプリケーションで運用することを目的としているため、今後はより大量のデータで学習させたモデルをARのアプリケーション上で動かし、精度や体験に与える影響を様々な条件で検証していきたい。

参考文献

- [1] 河出 弘志, 萩原 良信, 今村 弘樹, 崔 龍雲: 実物体と仮想物体のインタラクションを考慮したARに基づく情報提示装置の開発, 映像情報メディア学会技術報告, Vol.35, No.9, pp.59-63, 2011
- [2] . E. Breen, E. Rose, and R. T. Whitaker: Interactive occlusion and collision of real and virtual objects in augmented reality, Munich, Germany, European Computer Industry Research Center, 1995.
- [3] Presence Platform's Hand Tracking API Gets an Upgrade, <https://developer.oculus.com/blog/presence-platforms-hand-tracking-api-gets-an-upgrade/>, (アクセス日 2022.7)
- [4] Daniel DeTone, Tomasz Malisiewicz, Andrew Rabinovich: Toward Geometric Deep SLAM, arXiv:1707.07410, 2017
- [5] Carlo Innamorati, Bryan Russell, Danny Kaufman, Niloy Mitra: Neural Re-Simulation for Generating Bounces in Single Images, 2019 IEEE/CVF International Conference on Computer Vision (ICCV) pp8718-8727, 2019
- [6] 竹内 一平: 実物体と仮想物体の自然な力学的相互作用を可能としたARシステム, 法政大学大学院紀要. 情報科学研究科編, Vol.13, pp1-6, 2017
- [7] Francois Chollet: Xception: Deep Learning with Depthwise Separable Convolutions, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp1800-1807, 2017