

3次元点群に対する機械学習を用いた形状変形判定法

武田隆雅¹ 小林洸陽¹ 宍戸英彦² 亀田能成² 北原格²

概要：本研究では、建築物の損壊判定を目的として、3次元形状に基づきその形状変化を推定する手法を提案する。推定対象となる建築物の多視点映像から復元した3次元点群の情報に機械学習を適用することにより、建築物の形状変化を推定する。倒壊家屋の形状変形データは入手が困難であるため、プリミティブ形状の3Dオブジェクトを多様に変形させることにより学習データを獲得する。プリミティブ形状を用いて学習した結果、どの程度の形状識別が可能となるかを検証した。

キーワード：3次元点群，深層学習，形状変化，生成型学習，建物損壊判定

1. はじめに

大規模自然災害発生時において、被害状況の把握は迅速な救助や二次災害の防止にとって重要である。例えば、建築物の被害状況調査では、建物応急危険度判定士が直接現地に赴き、大規模災害発生直後から三日間の災害初期に被災した建築物を調査し、その破損状況などから応急危険度を判定している。平成23年に発生した東日本大震災で大きな被災を受けた宮城県では、5万棟を超える建築物に対して応急危険度判定を実施することで二次災害の防止に努めたが、広大な地域を迅速に判定するためには、膨大な数の判定士が必要であるなどの課題が浮き彫りになった[1]。

一方、ドローンやカメラ技術の発達により、高解像度空撮画像を容易に撮影可能になりつつある。利根川ら[2]は、被災地域を見下ろすように撮影した空撮画像に対し、畳み込みニューラルネットワークによって学習した識別器を用いた倒壊判定処理を行っている。空撮画像から建築物の倒壊判定の自動処理を実現しているが、単眼の空撮画像での観測であるため、壁の破損や建物の歪みといった、倒壊の3次元的な判定が困難であるという問題が残っている。小林ら[3]は、多視点空撮画像に基づいた被災状況把握に関する研究を実施し、多視点空撮画像から撮影空間の3次元形状を復元することで被災状況を3次元的に把握することに成功している。3次元情報を用いることで、建築物を3次元的に観察することが可能となるが、倒壊判定自体は人間の判断に委ねられるため、大量の建築物を判定するためには多数のユーザーを確保する必要、いわゆるリソース問題が存在する。

上述したリソース問題の解消を目的として、計算機による3次元点群に基づく建物倒壊の自動判定法が提案されている[4]。被災前後の建築物のズレに注目した手法と内部空間の形状の変化に注目した手法によって3段階の倒壊レベル分類を実施し、T検定によってその有効性が検証されて

いる。これらの手法では、被災前後の形状比較に基づいて倒壊度を判定するため事前撮影が必要となるが、何時何処で発生するかわからない自然災害を対象とした場合、解決すべき課題が残る。我々は、事前撮影されていない地域で自然災害が発生した場合に、建築物の倒壊状況推定を実現する手法に関する研究に取り組んでいる。具体的には、建物の形状変形度を判定する汎用性の高い識別器を事前に用意しておくことにより、多様な被災地の建物損壊判定を目指す。形状変形を判定する識別器の学習には多様な形状の3次元建物モデルが必要となるが、実際に形状が変化した建築物の3次元形状データを大量に入手することは現実的ではない。そこで我々は、3DCGソフトウェアを使用して立方体のようなプリミティブ形状を変形させることにより、変形度が既知な3次元形状データを大量に生成し、それらを用いて深層学習を行う。

2. 関連研究

2.1 2次元画像による建築物の倒壊判定

利根川ら[2]は、図1に示す被災地を撮影した空撮画像と建造物の位置や大きさのわかる地図から、建築物の倒壊判定を行うことで、発生するがれき量の推定に取り組んでいる。倒壊状況をラベリングした画像から画像上での見え方と倒壊状況の関係性を学習した畳み込みニューラルネットワークを用いて建築物の倒壊判定を行う。倒壊判定に用いる空撮画像は建築物を上空から撮影した画像を用いるため、壁面の損害や建物の歪みといった形状の把握は難しく、判定可能な倒壊状態は限定的である。

1 筑波大学大学院システム情報工学研究科
University of Tsukuba
2 筑波大学 計算科学研究センター



図1 テストデータで使用された
空撮画像の例[2]

2.2 3次元形状を用いた被災状況把握

小林ら[3]は、3次元画像処理とクラウドソーシングを統合した被災状況マップ生成プラットフォームを提案している。被災地の多視点画像からその3次元形状の復元を行い、復元した3次元モデルに対して、クラウドソーシングのユーザーが建築物の倒壊判定を実施する。図2に示すように、ユーザーは建築物の3次元形状を様々な角度から観察し倒壊判定を行う。判定結果を随時マップに反映させることで被災状況マップを生成する。

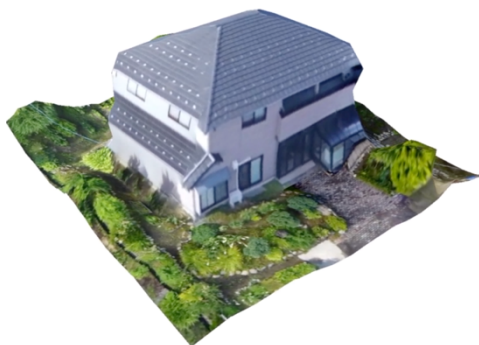


図2 倒壊判定に用いる3次元モデル[3]

2.3 3次元点群を入力情報とする深層学習法

Charlesら[5]は、物体の形状が把握可能な3次元点群を入力として、複数のクラスに分類する深層学習モデルであるPointNetを提案している。3次元点群は近くの点との隣接関係を定義することが困難であるため、隣接関係を利用する畳み込みを適用し難い。また、点の座標はソートされて記述されないために記述順序に依存しない計算を行う必要がある。PointNetでは、すべての点を対称関数で平等に扱うことにより、記述順序に依らない特徴量抽出を実現し、3次元的な形状を情報源として物体を識別している。40種類の物体の3次元モデルのModelnet40[6]を用いた検証実験の結果、89.6%の識別精度を実現している。

2.4 被災前後の情報を利用した建築物の倒壊判定法

武田ら[4]は、被災前後の情報を比較することで建築物の倒壊についての判定を行なっている。3段階の倒壊レベルを定義し、二通りの判定法を提案した。一つは建築物の位置のずれに注目した方法で、もう一つは空間の形状変化に注目した手法である。被災前後の形状情報の比較によって倒壊度を判定するため、事前撮影が必要である。本研究では、汎用的な識別器を学習することで事前撮影の必要がない建築物の形状変形度推定を試みる。

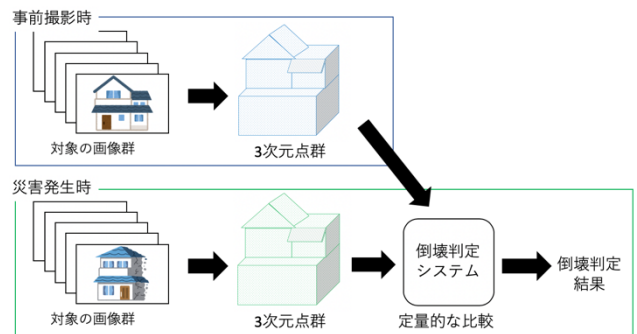


図3 被災前後の比較による倒壊判定法[4]

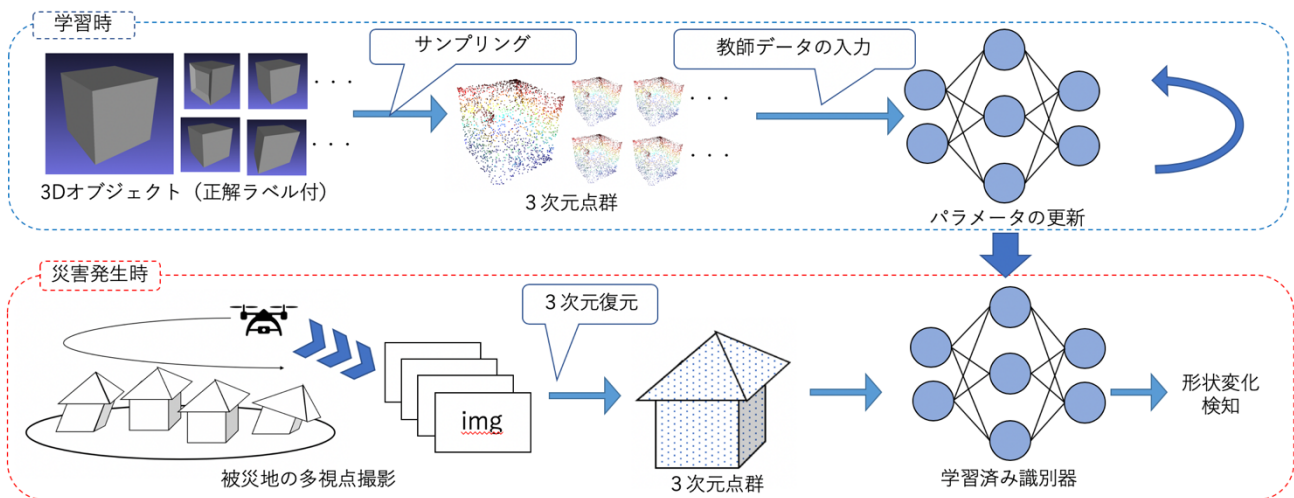


図4 機械学習を使用した形状変化判定

3. 機械学習を使用した形状変化判定

提案する建築物の倒壊判定法の概要を図4に示す。本研究の目的は、被災後の建築物について多視点画像から建築物の形状変化を検知することである。建築物の形状変化は多様であるため、被災後の建築物の3次元形状情報のみから幾何的に変化した部分を検出することは困難である。そこで、事前に物体の形状変化について学習した識別器による形状検知法を提案する。

識別器の学習には大量の学習データが必要となるが、実際に倒壊した建築物の3次元形状情報を得ることは困難であるため、3DCGソフトウェアで生成した多様な形状の3Dオブジェクトとその変形パラメータを学習データセットとする。発災時に計測される3次元形状は3次元点群として取得されることが多いため、生成した3Dオブジェクトの表面を3次元的にサンプリングし、教師データの入力とすることで深層学習を実施する。

実際の発災時には、ドローン等で被災家屋の多視点画像を撮影する。複数の画像間の視差を利用して各画像のカメラの位置・姿勢と対象物体の3次元形状の復元を行う Structure from Motion (SfM) [8]と SfM で生成した点群をパッチの多眼ステレオ処理によって、誤差を軽減して点群を密にする Patch-based Multi-view Stereo Software (PMVS) [9]によって、被災建築物の3次元点群を取得する。各建築物の3次元点群をあらかじめ学習した識別器に入力することで、建築物の形状変化を推定する。

4. 3次元点群を入力とする深層学習

本研究では、形状判定の識別器として、3次元点群を入力としてクラス判別やパーツセグメンテーション、セ

マンティックセグメンテーションを実現する PointNet を導入する。従来の3次元点群を用いた深層学習では、形状の表現法の問題から、ボクセルグリッド等のデータを利用してはいたが、凹凸の情報等が近似によって喪失してしまうなどの問題が存在した。PointNet では、3次元点群を深層学習に直接入力とすることが可能であるため、情報損失の少ない識別器の構築が可能である。

3次元点群の表現法には、深層学習の入力時にいくつかの問題点が存在する。その一つが頂点の順不同性である。画像等では、情報の記述には規則的な並びがあるが、3次元点群では各点の座標値や色情報、法線情報等が羅列した記述方式であるため、図5のように同じ点群でも点の記述順序が異なる場合がある。PointNet では、全ての計算を対称関数で行うことにより、この問題を解決している。対称関数とは、加算や乗算といった変数の入力順序に結果が依存しない関数である。PointNet のネットワーク構造は、図6に示すようになっており、各点ごとに学習や変換を行なった後に、対称関数である Maxpooling 関数によって特徴量を抽出する。

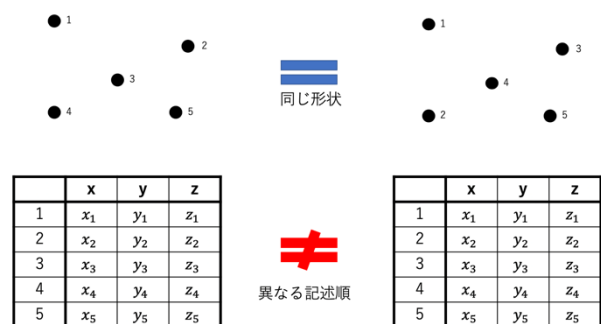


図5 順不同な点群

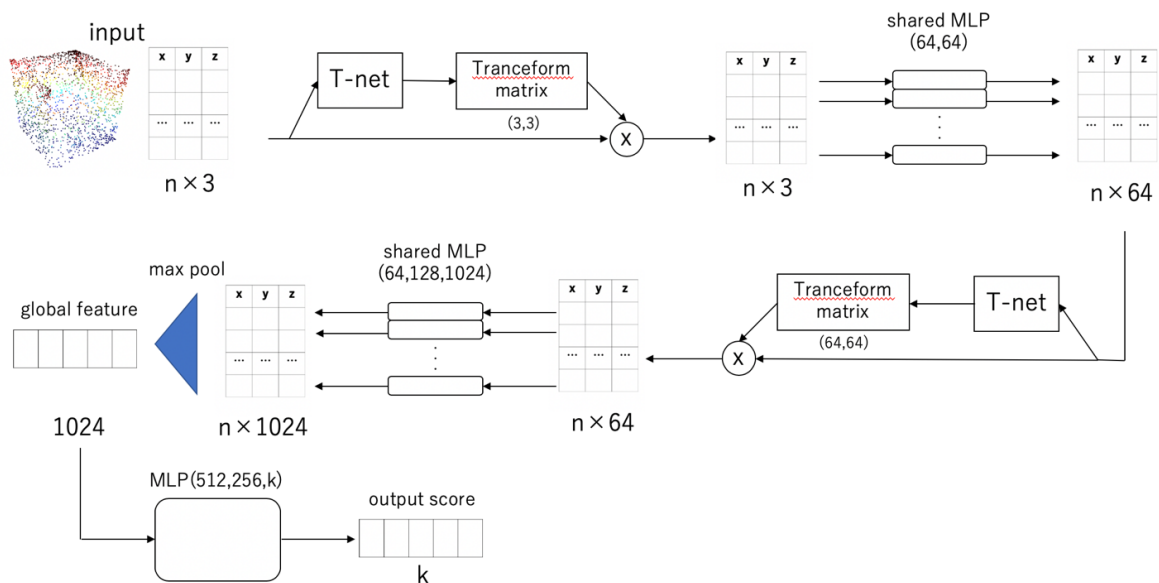


図6 ネットワーク構造

また、同じ形状の点群でも3次元空間上でXYZ軸に対して回転や姿勢が異なるといった、回転や姿勢の普遍性の問題も存在する。PointNetではT-Net(Spiral Transfer Network)[10]を取り入れることでこの問題を解決する。T-Netは、事前に学習したアフィン変換行列によって3次元点群の回転・姿勢を推定し、その結果を用いて点群を変換することにより回転・姿勢の影響を緩和する。図6に示すように、ネットワーク中には点群を入力した直後と、多層パーセプトロンの中の2箇所に取り入れられている。

本研究では、3次元点群とその形状変化ラベルからなるデータセットを入力としてPointNetによる学習を行い、任意の3次元点群の形状変形のクラス分類を行う。

5. 学習データ生成

5.1 学習データ生成の流れ

物体の形状変形を十分に学習できるだけの分量の倒壊建築物の3次元情報を入手することは困難であるため、本研究では、3DCGソフトウェアを利用してプリミティブ形状を多様に変形させた3Dオブジェクトを大量に生成し、その変形パラメータと併せて学習データとする。3Dオブジェクトは、乱数パラメータで指定される形状に変化された後、頂点と面情報から成る3次元メッシュモデルとして出力される。メッシュ表面から3次元点群をランダムにサンプリングすることで学習用点群を取得する。サンプリングする点数はPointNetの推奨に従い、最大で2048点とする。

5.2 推定形状変形の種類

図7に本手法で判別する3種類の形状カテゴリを示す。本研究では建築物の形状変形として、建築物が傾いた形状(skew)と、建築物が凹んだ形状(push)の二つの形状を推対象とする。形状変化が生じていない状態(cube)は立方体を使用する。

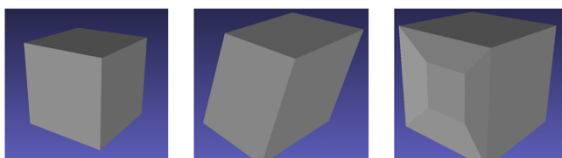


図7 3種類の形状 (左: cube 中央: skew 右: push)

5.3 3Dオブジェクトの生成

skewは建築物の柱や梁の破損によって建築物の壁面が傾いた状況を想定している。skewについては、2軸(X,Y)の対の壁面について図8のように乱数 θ 傾けることによって、生成を行う。

pushは底面を除く5面のうちランダムに選択された1面について“凹みの深さ(depth)”と“凹む面の辺の長さ(thickness)”を乱数で指定することで生成される。ここでは、凹む面は正方形で面の中心に配置する。

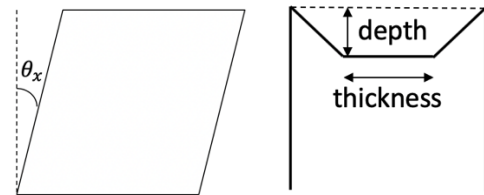


図8 形状変化の様子

5.4 一様3次元点群のサンプリング

5.2節で生成した3Dオブジェクトは、頂点と面情報から構築されるメッシュ情報として保存される。メッシュ形状は、図9に示す三角メッシュを採用する。全ての面に対して均一に3次元点群をサンプリングする。はじめに、サンプリングする三角メッシュを選択する。メッシュを張る全ての三角形の面積を計算し、各三角形の中から面積の比率を確率として一つ選択する。図9に示すように、選択した三角形の内部から重心座標系(Barycentric coordinate system)と一様乱数を使用してサンプリングを行う。三角形の頂点をそれぞれ V_0, V_1, V_2 としたときに図9中の式のように媒介変数 u, v を用いて三角形内の座標を定式化できるので、二つの媒介変数を0から1までの一様乱数でサンプリングすることが可能である。以上の過程をサンプリング数の2048回実行することで図10に示す一様に分布する3次元点群が生成される。

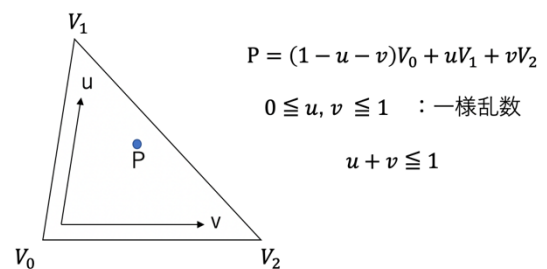


図9 重心座標系による表現

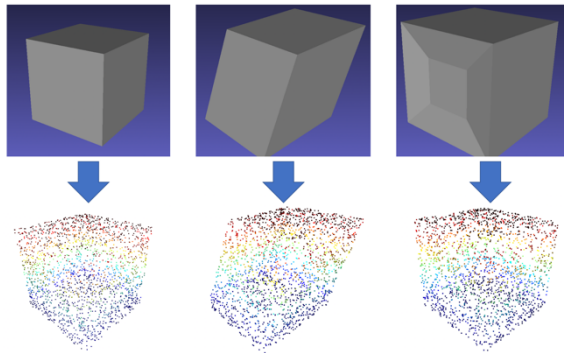


図10 均一な3次元点群のサンプリング

5.5 SfMで生成される3次元点群の分布を考慮したサンプリング

5.4節では、3Dオブジェクトから均一に3次元点群をサンプリングしたが、発災時に撮影した多視点画像にSfMを適用して生成される3次元点群の分布は、図11に示すように必ずしも均一ではない。SfMによって生成される建築物の3次元点群が一様サンプリングと異なる理由は以下三つが考えられる。

一つ目は建築物の底面が復元されないことである。SfMは複数の画像間の特徴量のマッチングを行うことでカメラ位置・姿勢を求めて、マッチングした特徴点が3次元点群として復元されるため、画像上で観測されない建築物の底面は復元されない。本方式では、底面からは3次元点群をサンプリングしないことで対応する。

二つ目は、復元される点群の密度が一様でないことである。SfMでは多視点画像間の対応点情報から、その3次元点群が復元されるが、対応点は濃淡変化の大きい背景と壁面の境界付近に集中する傾向があり、結果としてエッジ付近の点群密度が高くなる。本方式では、サンプリング時にエッジ部分に近いほどサンプリング確率を高く設定することで対応する。一様なサンプリングでは二つの媒介変数 u, v を一様乱数としていたが、一様乱数 x を N 乗することで0付近の値をとる確率を増加させる。実際のSfMで復元した3次元点群と比較して、本稿では $N=5$ とした。

三つ目は、点群の3次元座標値推定時に発生する推定誤差である。建築物を撮影するカメラの空間解像度による撮影限界や特徴点の誤対応によって、復元した建築物の壁面状には一定の誤差が含まれている。本方式では、サンプリングした3次元点群の座標値に一定誤差を付加することで対応する。本稿では、立方体の大きさの1%の誤差を付加している。

以上の3項目を考慮した3次元点群のサンプリングで生成された点群を図12に示す。点数は図10と同じ2048点である。

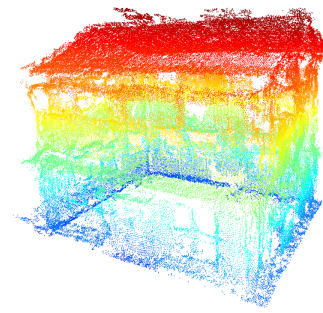


図11 SfMで復元した3次元点群の例

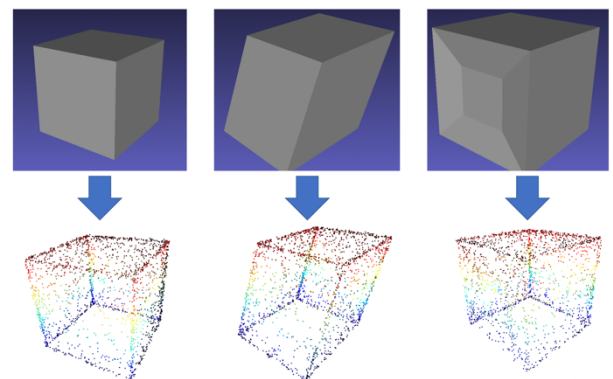


図12 SfMを考慮したサンプリング

6. 実証実験：PointNetの形状変形識別能力の検証

本研究で使用するPointNetの精度検証はModelnet40を使用したものが知られているが、これは飛行機や椅子といった40カテゴリの物体についての認識を目的としていた。本研究で目指すのは、同一物体の形状変化の把握であるため、把握が可能な形状変形について検証実験を通じて調査する。

6.1 一様サンプリングによる検証

プリミティブな形状を対象とした場合の、形状変形判別能力の検証を行う。3Dオブジェクトからの3次元点群のサンプリングは5.4節で述べた一様サンプリングを用いる。3Dオブジェクトは5.2節で紹介した3種類とし、大きさは最大と最小のサイズが2倍の差になる範囲でランダムに設定する。また、回転・姿勢についてもランダムに設定し、skewで変形させる角度は0~20度、pushで凹ませる領域はdepthが最大で立方体のサイズの40%、thicknessが立方体のサイズの80%となるように設定した。学習データ数は3種類の3次元点群を2048個、学習epoch数は150で、学習率は0.001、損失関数はクラス分類で使用される交差エントロピー誤差、パラメータの更新はadam法[10]で行なっている。学習後のテストは、教師データの生成法と同

様に3種類の3次元点群を128個ずつ生成して検出精度を算出する。

教師データによる学習の様子について、損失関数の値の変遷は図13のように、クラス分類の精度は図14のようになった。学習はepoch数10前後から収束を始めて50まで進むと値の変動が小さくなっていることがわかる。

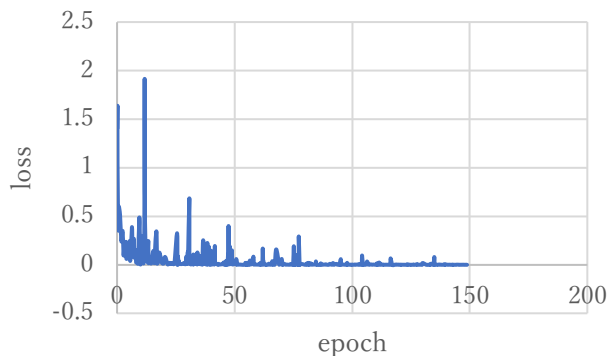


図13 一様サンプリングでの損失関数の様子

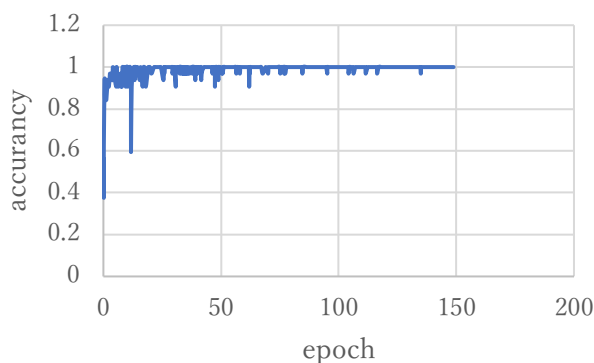


図14 一様サンプリングでの精度の様子

識別処理の結果を表1に示す。全テストデータの内、一つを除いて正しいクラスに分類された。PointNetの検出精度で使用されたModelnet40よりも高い検出率となっている。その要因としては、クラス数がModelnet40では40なのに対して今回の実験では3クラスでの分類であったため、誤検出が少なかったことが考えられる。

表1 一様サンプリングの識別結果

cube	1.000
skew	0.992
push	1.000
All(識別率)	0.997

6.2 SfMを考慮したサンプリングによる検証

一様な3次元点群のサンプリングでは、精度の高い形状変形識別が可能であったため、発被時に復元される3

次元点群に近い状況を想定した5.5節で述べたサンプリングによって得られる3次元点群を対象とした検証実験を行う。オブジェクトの大きさ等、サンプリング法以外の条件は前節と同様である。

教師データによる学習の様子について、損失関数の値の変遷は図15のように、クラス分類の精度は図16のようになった。学習はepoch数10前後まで収束する様子を見せたが、50以降は6.1節の結果よりも大きな振動をしている。検出精度はepoch数10以降から9割を超える精度で収束している。

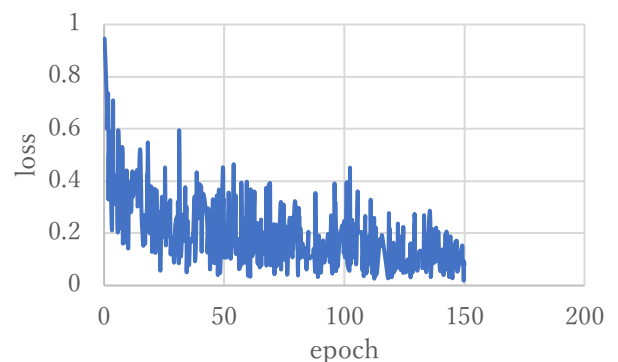


図15 SfMを考慮したサンプリング法での損失関数の様子

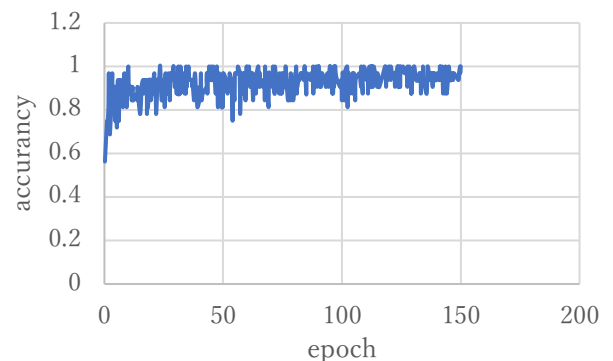


図16 SfMを考慮したサンプリング法での精度の様子

識別処理の結果を表2に示す。全体の検出率は教師データの正解率に比べて63%ほど低い検出率となった。特にcubeの検出については約20%と低い結果となった。その要因としては、pushの凹みが小さいものとノイズをcubeについてほとんど差がなくなり、過学習を起こしたことが考えられる。

表2 SfMを考慮したサンプリングの識別結果

cube	0.195
skew	0.758
push	0.945
All(識別率)	0.632

7. まとめ

本研究では、PointNet の物体の形状変化の認識能力に関する検証実験を実施した。実証実験の結果、一様にサンプリングした3次元点群を対象とした判別実験では高い性能を示したが、SfM を考慮したサンプリングした3次元点群では性能が低下することが確認された。性能が低下した原因として過学習が考えられる。実証実験の結果から、凹みを含むような形状変化の検出は困難であったが、地震災害での家屋損壊は、壁が傾く形状変化が主であるため、その変形の識別にフォーカスした形状変化検知について、検証を進める予定である。

今後は、実際に倒壊した建築物の多視点画像を取得し、そこから復元した3次元点群に基づいた変形識別処理を行うことにより、発災時を想定した実証実験を実施し、識別精度の向上を目指す。

本研究は JSPS 科研費 17H01772 と JST CREST Grant Number JPMJCR16E3, Japan の助成を受けたものである。

参考文献

- [1] 宮城県 “平成23年東北地方太平洋沖地震における応急危険度判定について”
- [2] 利根川 凜 “被災がれき量推定に向けた畳み込みニューラルネットワークを用いた建築物の倒壊建造物の自動抽出” 情報処理学会論文誌 Vol.57 No.6 1565-1575 June 2016
- [3] 小林洸陽, 宍戸英彦, 亀田能成, 北原格, “3次元画像処理とクラウドソーシングによる被災状況マップ生成法” MVE2018-60 pp.37-42
- [4] 武田隆雅, 小林洸陽, 宍戸英彦, 亀田能成, 北原格, “多視点画像から復元した3次元形状に基づく建築物の倒壊判定法” FIT2019
- [5] Charles R. Qi, Hao Su, Kaichun Mo, Leonidas J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation” Computer Vision and Pattern Recognition 2017
- [6] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang and J. Xiao “3D ShapeNets : A Representation of Volumetric Shapes” Proceedings for 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR2015)
- [7] Blender
<https://www.blender.org/>
- [8] C. Wu, “Towards Linear-Time Incremental Structure from Motion”, Proceedings of the 2013 International Conference on 3D Vision pp.127-134 2013
- [9] Y. Furukawa, J. Ponce, “Accurate, Dense, and Robust Multi-View Stereopsis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.
- [10] Max Jaderberg, Karen Simonyan, Andrew Zisserman, Koray Kavukcuoglu “Spatial transformer networks” NIPS (Advances in neural information processing systems) 2015
- [11] Diederik P. Kingma, Jimmy Lei Ba, “ADAM: METHOD FOR STOCHASTIC OPTIMIZATION” arXiv:1412.6980v9. 2017