

3D ポイントクラウドを用いたマイクロ行動認識の検討

Investigation of Micro Activity Recognition Using 3D Point Cloud

三嶋 祐輝[†] 松井 智一[†] 松田 裕貴[‡] 諏訪 博彦[‡] 安本 慶一[‡]
Yuki Mishima Tomokazu Matsui Yuki Matsuda Hirohiko Suwa Keiichi Yasumoto

1. はじめに

近年, IoT 技術の発展による宅内環境のスマート化に伴い, 居住者の行動認識に関する研究・システム開発が盛んに行われている. 家庭内での生活行動を認識することにより, 高齢者の生活の補助 [1] や居住者の生活習慣に応じた家電の自動制御 [2][3] といった生活・行動支援への応用が期待できる.

高齢者の生活の補助では, 高齢者の宅内行動を認識することで, 転倒などの異常な行動を検知し, 非常時の迅速な対応が可能となる. 総務省の調査によると, 2007 年に 65 歳以上の割合が全人口の内 21% を超える超高齢化社会となった. 2022 年現在, 65 歳以上人口は 3,621 万人となり, 総人口に占める割合も 28.9 % となっている [4]. これらを踏まえると今後の日本の高齢化社会において, 高齢者の精神的負担の軽減や生活の質を向上させるという観点から, 高齢者の生活を補助することは, より重要度が増してくると考えられる.

家電の自動制御では, 居住者の行動に合わせて, お風呂の湯を沸かす, エアコンの温度調整を行う, 電気を点灯または消灯するといったことが可能となる. これにより本来であれば, 居住者が自ら行わなければならない操作を自動的に行うことができるようになり, 精神的なゆとりが生まれ, 結果として居住者の生活の質向上が期待される. ここで, 日常生活行動をより正確に認識するためには, より細かい行動であるマイクロ行動を認識することが重要である.

行動認識の手法としてカメラを用いるものがある [5]. この手法では認識のために動画・画像情報を用いるためプライバシーの侵害が懸念される. また, 人物以外の物体によりオクルージョンが発生するという問題がある. ジャイロセンサや加速度センサなどのウェアラブルセンサを用いる行動認識手法 [6] が提案されているが, ウェアラブルセンサを常に装着しなければならないという手間がかかる問題がある. 振動センサを利用した行動認識手法も多数提案されている. San ら [7] は, レーザードップラー振動センサによりセンシングした振動データを学習データとして, 機械学習モデルにより解析を行い, ド

ライヤーやシャワーなどの, 18 種類の行動を認識するシステムを提案している. レーザードップラー振動センサを天井のある一点に設置することで, 微小な振動の測定を行う. Bonde ら [8] は, 卓上の上に振動センサを設置し, 周囲の構造物の振動から, 歩くやタイピング, 話すなどの行動を識別した. 振動センサを用いた手法は, カメラなどのビジョンベースの手法に比べて, オクルージョンが発生しないという利点がある. しかし一方で, 物体と接触しない行動は原理的に区別することが不可能であるという問題がある. また, これらの研究ではマイクロ行動の認識には至っていない.

そこで本研究では, LiDAR センサによってセンシングされる 3D ポイントクラウドを用いてプライバシーに配慮した行動認識を目指す. 3D ポイントクラウド [9] は, 対象物を 3 次元の XYZ 座標上の点の集合で表したものである. 3D ポイントクラウド上では人物と物体との位置関係の情報を扱うことができる. その点, 従来の動画・画像による行動認識の手法に比べて, 物体を考慮した処理を行うことができるため, より精度の高い行動認識を行えることが期待される.

本システムは, 奈良先端科学技術大学院大学にあるスマートホーム (以下, NAIST スマートホーム) 内に構築する. 14 台の LiDAR センサを用いて 3D ポイントクラウドを収集し, 従来認識が難しかった日常的なマイクロ行動認識を行う. マイクロ行動認識により, より細かい粒度での行動を把握することができ, 一連のマイクロ行動からより複雑な行動を認識することを目指す.

2. 関連研究

本章では, 宅内における日常行動認識およびマイクロ行動認識に関する研究について概観する. また, 3D ポイントクラウドおよび 3D ポイントクラウドを用いた行動認識に関する既存研究について述べる.

2.1 宅内における日常行動認識

Balli らは, スマートウォッチに搭載された三軸加速度センサ・ジャイロセンサ・歩数計・心拍センサから得られるデータを, 主成分分析による次元削減を行った上で RandomForest に適用することにより, 「歩く」, 「座る」, 「PC 操作」, 「静止状態」, 「座る」など 8 種の動作の識別において, F 値で 0.985 の精度を示している [10]. Cook ら

[†] 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

[‡] 理化学研究所, RIKEN

が提案した CASAS は、光量センサ、気温センサ、モーションセンサ、ドアセンサなどから取得したデータを用いて、10種の行動を平均で0.581の精度で認識することに成功している [11]. Sasaki らは家電の利用状況から行動認識を行う手法を提案し、9種の行動を平均で0.675の精度で認識している [12]. Mastui らは SALON と呼ばれる PIR センサ、気温・湿度・騒音などが計測可能な環境センサ、ドアセンサ、環境発電によって駆動するアノテーション用ボタン、データ収集・解析用サーバからなるシステムを提案し、「入浴」、「食事」、「調理」、「睡眠」の4種の行動認識において、再現率で0.723、F値で0.407の精度を達成している [13]. これらの研究では一般的な日常行動の認識対象としているが、より適切な生活補助や家電制御のためには、より細かい行動（マイクロ行動）の認識が必要と考える。

2.2 マイクロ行動認識に関する研究

Michibata ら [14] は、CNN を用いて一人称視点での動画情報から調理作業におけるマイクロ行動認識を行なった。調理作業は「調理器具を取り出す」、「食材を取り出す」、「洗う」などの行動を対象としている。Mizumoto ら [15] は、深度カメラ KinectV2 を用いて、調理作業中のマイクロ行動認識を行なっている。深度カメラより得られる深度データをもとに算出した骨格情報から、調理者のキッチンにおける相対位置ならびに胴体と上肢の位置関係と角度を特徴量としている。これらの特徴量をランダムフォレストの学習用データに用いることで、行動認識を行なっている。クロスバリデーションを用いて、「PICKUP foodstuff」、「PICKUP tool」、「PRE-PROCESSING cutting slicing」、「PREPROCESSING poring powdering」、「WASHING」の5種類の調理作業を78%の精度で分類可能とした。しかし、これらのマイクロ行動は調理作業に限定したものであり、その他のマイクロ行動についての検討はされていない。

Rai ら [16] は、階層構造に注目した行動認識の新しい手法および課題に取り組むためのデータセットを提案している。データセットには、複数視点から撮影されたマルチモーダルな情報と、階層的な行動ラベル・マイクロ行動ラベルが含まれる。行動の階層構造を学習することにより、行動認識の精度向上を示している。表1にマイクロ行動の定義に利用した接頭語一覧を示す。接頭語と Object 名を合わせることでマイクロ行動の定義をしている。本稿におけるマイクロ行動の認識には、Rai らの提案したマイクロ行動の定義に基づいて、いくつかを抜粋して扱う。

表 1: マイクロ行動の定義に利用する接頭語一覧

In front of	Wiping	Wearing
Looking at	Beneath	Eating
Holding	On the side of	Lying on
Not contacting	Sitting on	Above
Touching	Leaning on	Behind
Not looking at	Carrying	In drinking from
Twisting	Standing on	Unsure
Covered by	Writing on	Have it on the back

2.3 3D ポイントクラウドを用いた行動認識

LiDAR センサの小型化、低価格化に伴い、3D ポイントクラウドの研究が盛んになっている。Charles ら [17] は 3D ポイントクラウドに深層学習を用いてオブジェクト推定する手法を提案している。Lio ら [18] は、3D ポイントクラウドにおけるノイズ除去のアルゴリズムを提案している。

3D ポイントクラウドを用いた行動認識に関する研究も行われている。Liu ら [19] は、動的な 3D ポイントクラウドシーケンスを学習するための新しいニューラルネットワークアーキテクチャである MeteorNet を提案している。MeteorNet を用いて行動認識、セマンティックセグメンテーション、シーンフロー推定など、様々な 3D 認識タスクでベンチマークを行なっている。行動認識タスクでは、深度マップで構成されている MSRAction3D データセット [20] から 3D ポイントクラウドを構築し、動的な 3D ポイントクラウドシーケンスを MeteorNet に学習させている。その結果、最大で 88.5% の認識精度を得ている。このことから MeteorNet は、動的な 3D ポイントクラウドシーケンスを用いた行動認識タスクにおいて有用であると考えられる。しかし、MSRAction3D データセットの内容が、「jogging」、「tennis swing」、「tennis serve」、「golf swing」、「pickup&throw」等から構成されていることを踏まえると、マイクロ行動に対する行動認識は行われていない。そこで本研究では、Liu らの提案した MeteorNet を用いて、動的な 3D ポイントクラウドからマイクロ行動の認識を行うことを検討する。

3. 提案システム

本章では、3D ポイントクラウドを用いたマイクロ行動認識システムについて述べる。LiDAR センサからセンシングされる人物の 3D ポイントクラウドを分析し、マイクロ行動認識を行うことを目的とする。

3.1 提案システムの流れ

NAIST スマートホーム内に設置されている LiDAR センサから、人物のマイクロ行動を取得するシステムの流

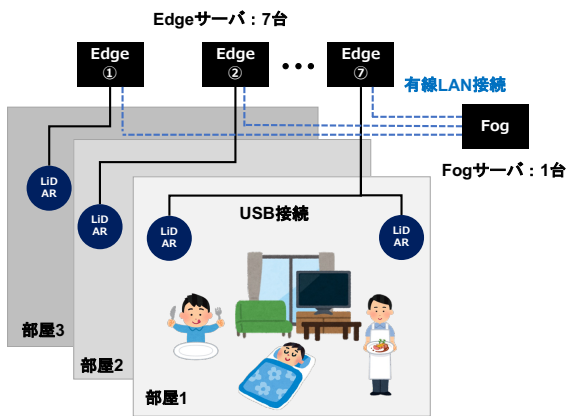


図 1: システム構成図

れを以下に示す。Fog サーバ, Edge サーバ, LiDAR センサから構成されるシステムを図 1 に示す。

1. Fog サーバから各 Edge サーバを起動し, Edge サーバに接続されている LiDAR センサより, 各視点の 3D ポイントクラウドを取得する。
2. LiDAR センサの位置を考慮した上で, 各視点の 3D ポイントクラウドを Robot Operating System (ROS) 上で, 一つのデータに統合する。図 3 に統合されたデータを示す。
3. 3D ポイントクラウドをひとたびシステムの技術を用いて背景データと前景データに分けることにより, 前景データを人物の 3D ポイントクラウドとする。
4. 正解ラベルを付与した前景データを学習用データとし, MeteorNet に学習させる。
5. 学習済みモデルに前景データを入力することで, マイクロ行動の推定を行う。

本システムによって認識されるマイクロ行動から, 行動認識をより細かい粒度で扱うことができる。そのため, 同じ行動が認識された場合でも, 構成されるマイクロ行動によって異なる家電の制御やサービス提供が可能となり, 高齢者の見守りや家電制御への応用が期待される。

3.2 マイクロ行動の定義

Rai ら [16] は, マイクロ行動を動詞とオブジェクトの関係から定義した。例えば, 動詞として「holding」, 「touching」, 「wearing」, オブジェクトとして「pen」, 「detergent」, 「apple」があるとする。これらの関係から合計 9 種類のマイクロ行動が定義される。本研究では, Rai らの手法に基づいてマイクロ行動を定義する。

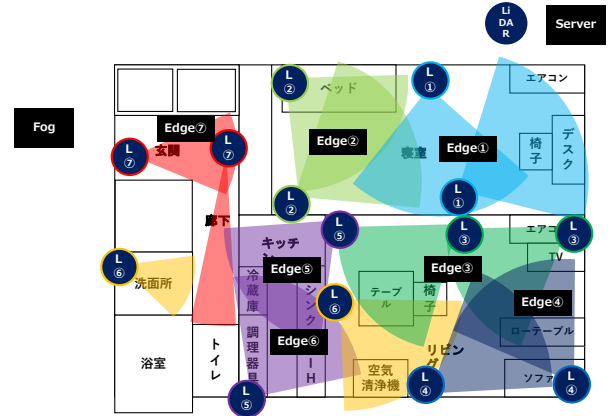


図 2: NAIST スマートホームの見取り図および LiDAR センサの設置箇所

3.3 3D ポイントクラウド収集システム

3D ポイントクラウド収集システムは, 奈良先端科学技術大学院大学内に設置されたスマートホームに構築している。NAIST スマートホーム内に見取り図および LiDAR センサの設置箇所を図 2 に示す。NAIST スマートホーム内には, 玄関に 2 台, 洗面所に 1 台, キッチンに 2 台, リビングに 5 台, 寝室に 4 台の計 14 台の LiDAR センサが設置されている。そのため, 室内をあらゆる角度から網羅的にセンシングをすることで, さまざまな場面のマイクロ行動を取得することができる。本研究では LiDAR センサとして図 5 に示す RealSense L515 を使用する。

3.4 マイクロ行動認識モデルの構築

LiDAR センサによって得られるデータは, 背景差分により背景データと前景データに分けられる。前景データには, 人物の 3D ポイントクラウドの情報のみが含まれる。データは LiDAR センサによって逐次的にセンシングされるため, 人物の行動は前景データ上で動的な 3D ポイントクラウドシーケンスとして表現される。この動的な 3D ポイントクラウドシーケンスを MeteorNet に学習させることで, マイクロ行動認識モデルの構築を行う。

4. 予備実験

4.1 実験概要

本研究では, 3D ポイントクラウドを用いることによりマイクロ行動の認識を目指す。そのために, まず RealSense よりセンシングされる人間の関節の位置からマイクロ行動の認識を行う。センシングされる関節の位置を学習データとし, ランダムフォレストを用いて行動の分類および評価を行う。関節の位置の取得には, RealSense L515 とスケルトントラッキングミドルウェアで



図 3: NAIST スマートホームの 3D ポイントクラウド

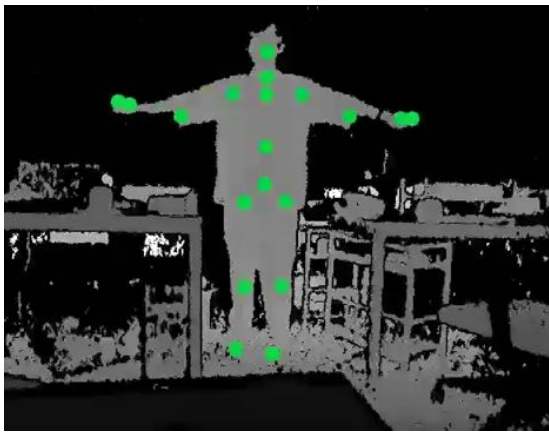


図 4: NuiTrack SDK による関節位置の認識の例

ある NUI TRACK SDK を用いる。NuiTrack SDK を用いて、関節の位置の認識結果を図 4 に示す。NUI TRACK SDK は、1 秒間に関節位置を 30 フレームで記録ができる。NUI TRACK SDK によって得られる関節を表 2 に示す。

本実験で対象とする行動は、移動とスマホ作業である。移動は、「walk」、「stand」、「sit」、スマホ作業は、「take out smartphone/take out」、「use smartphone/use」というマイクロ行動から構成されているものとする。「walk」は、カメラを横切るように歩く動作である。「stand」は椅子から立ち上がる動作であり、「sit」は椅子に座る動作である。「take out」はズボンのポケットからスマホを取り出す動作であり、「use」はスマホをスワイプする動作である。

被験者 1 名 (20 代男性) が、4 種類のマイクロ行動 walk, stand&sit, take out smartphone/take out, use smartphone/use のそれぞれを 8 分~10 分間繰り返し行い RealSense で撮影し、合計 4 セッション (各マイクロ行動に対し 1 セッション) のデータを作成した。なお、1 回のマイクロ行動の長さは 0.6 秒から 2 秒であった。撮



図 5: Intel RealSense LiDAR Camera L515

表 2: NUI TRACK SDK 上で扱われる関節の一覧

head	right_shoulder
neck	right_elbow
torso	right_wrist
waist	right_hand
left_collar	left_hip
left_shoulder	left_knee
left_elbow	left_ankle
left_wrist	right_hip
left_hand	right_knee
right_collar	right_ankle

影したデータから、NUI TRACK SDK を用いて関節の位置を取得した。

このとき、行動の開始と同時に動画撮影を開始し、タイムスタンプを用いて各行動のアノテーションを行った。認識に使用するデータサンプルは 1 セッションの各行動の中から 1 秒粒度で取り出したものである。特徴量は、1 秒間のマイクロ行動における関節の位置の XYZ 座標の平均、および 1 フレーム後における XYZ 座標の変化量の絶対値の総和とした。これらの特徴量をランダムフォレストに学習させ、マイクロ行動の分類を行った。この時、各関節の位置の XYZ 座標は、torso の座標を基点とする相対座標とした。各データサンプルの詳細としては、「walk」が 135 点、「stand」が 106 点、「sit」が 107 点、「take out」が 92 点、「use」が 135 点である。

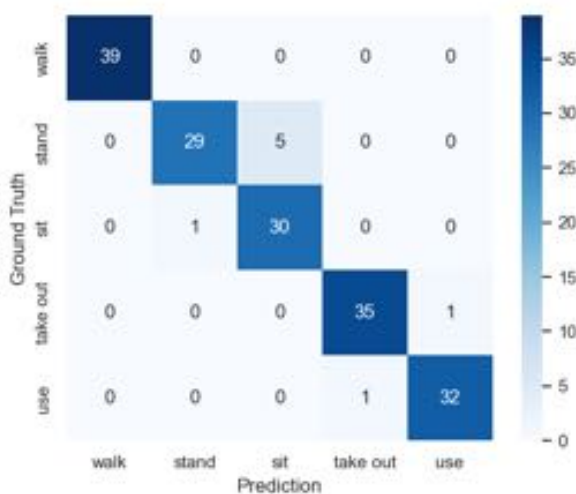


図 6: ランダムフォレストによる各行動における混同行列

表 3: ランダムフォレストによる各行動の分類精度

	walk	stand	sit	take out	use	平均
F 値	1.00	0.90	0.91	0.98	0.99	0.956

4.2 実験結果

各マイクロ行動データに対して、ホールドアウト法により検証を行った。ホールドアウト法とはデータ全体を、モデルを作る学習データと、モデルを評価するテストデータに分割して評価を行う方法である。今回は、各セッションを前半7割と後半3割に分割することにより、それぞれを学習データとテストデータとした。ランダムフォレストによる各マイクロ行動認識の混同行列を図6に、F値を表3に示す。実験の結果、Accuracyは96%であり、最も精度の高い行動は「walk」、最も精度が低い行動は「stand」であった。

4.3 考察

本実験では、ホールドアウト法による検証を行った。その結果、平均F値は95.6%、Accuracyは96%となった。各行動のデータセットは、被験者が一人かつ、一つのセッション内から取り出したものである。そのため、さまざまな場面を想定したデータセットによる実験を行うことができていない。これを踏まえると、環境や人物の違いによっては精度がある程度低下すると思われる。

5. まとめ

本稿では、3Dポイントクラウドを用いたマイクロ行動について検討した。予備実験により、マイクロ行動の認識を確認するために、まずは人物の関節の位置をランダムフォレストに学習させることで、行動の分類を行った。

中でも、スマホ作業と移動の行動におけるマイクロ行動の分類を行った。「walk」、「stand」、「sit」、「take out」、「use」のいずれの状態であるかを平均F値95.6%で推定可能であることを示した。今後は、実験のデータセットの汎用性を高めるために被験者とセッション数を増やす、また、3Dポイントクラウドによるマイクロ行動の認識システムの構築を行う予定である。

謝辞

本研究の成果はNICT課題222「ウイルス等感染症対策に資する情報通信技術の研究開発」の支援によるものです。

参考文献

- [1] 玉水一柔, 榎原誠司, 佐伯幸郎, 中村匡秀, 安田清. 宅内の環境変化と声掛けに基づく在宅高齢者の日常生活行動センシングシステムの検討. IEICE technical report 信学技報 116 (404), 2017-01. pp.7-12.
- [2] John Krumm Brian Meyers Michael Hazas Stephen Hodges James Scott, A.J. Bernheim Brush and Nicolas Villar. *Controlling Home Heating Using Occupancy Prediction*. Proc. UbiComp 2011 (2011).
- [3] David Irwin Prashant Shenoy Sean Barker, Aditya Mishra and Jeannie Albrecht. *Smart-Cap: Flattening peak electricity demand in smart homes*. Proc. Percom 2012. pp.67-75.
- [4] 総務省. 令和4年版高齢社会白書. 2022.
- [5] Oscar T.-C. Chen, Ching-Han Tsai, Hung Ha Manh, and Wei-Chih Lai. *Activity recognition using a panoramic camera for homecare*. 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017. 1-6.
- [6] Timo Sztyler, Heiner Stuckenschmidt, and Wolfgang Petrich. *Position-aware activity recognition with wearable devices*. Pervasive and Mobile Computing 38, 2 (2017). pp.281-295.
- [7] Wei Sun, Tuochao Chen, Jiayi Zheng, Zhenyu Lei, Lucy Wang, Benjamin Steeper, Peng He, Matthew Dressa, Feng Tian, and Cheng Zhang. *VibroSense: Recognizing Home Activities by Deep Learning Subtle Vibrations on an Interior Surface of a House from a Single Point Using Laser Doppler Vibrometry*, Vol. No.21. September 2020.

- [8] Amelie Bonde, Shijia Pan, Mostafa Mirshekari, Carlos Ruiz, Hae Young Noh, and Pei Zhang. *OAC: Overlapping Office Activity Classification through IoT-Sensed Structural Vibration*. 2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI), 2020. pp.216-222.
- [9] S. Cousins R.B. Rusu. *3D is here: Point Cloud Library (PCL)*. in: IEEE International Conference on Robotics and Automation, 2011. pp. 1–4.
- [10] Serkan Balli, Ensar Arif Sağbaşı, and Musa Peker. Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. *Measurement and Control*, Vol. 52, No. 1-2, pp. 37–45, 2019.
- [11] Diane J. Cook, Aaron S. Crandall, Brian L. Thomas, and Narayanan C. Krishnan. CASAS: A Smart Home in a Box. *Computer*, Vol. 46, No. 7, pp. 62–69, 2013.
- [12] Wataru Sasaki, Masashi Fujiwara, Hirohiko Suwa, Manato Fujimoto, Yutaka Arakawa, Aki Kimura, Tomoko Miki, and Keiichi Yasumoto. Daily Living Activity Recognition through Time Series Analysis of Data Obtained with ECHONET Lite Appliances and Motion Sensors. *JSAI Technical Report, Type 2 SIG*, Vol. 2018, No. SAI-031, p. 05, 03 2018.
- [13] Tomokazu Matsui, Kosei Onishi, Shinya Misaki, Manato Fujimoto, Hirohiko Suwa, and Keiichi Yasumoto. SALON: Simplified Sensing System for Activity of Daily Living in Ordinary Home. *Sensors*, Vol. 20, No. 17, 2020.
- [14] Shinya Michibata, Katsufumi Inoue, Michifumi Yoshioka, and Atsushi Hashimoto. *Cooking Activity Recognition in Egocentric Videos with a Hand Mask Image Branch in the Multi-Stream CNN*. CEA '20. Association for Computing Machinery, 2020. pp.1–6.
- [15] Teruhiro Mizumoto, Alberto Fornaser, Hirohiko Suwa, Keiichi Yasumoto, and Mariolino De Cecco. *Kinect-Based Micro-Behavior Sensing System for Learning the Smart Assistance with Human Subjects Inside Their Homes*. 2018 Workshop on Metrology for Industry 4.0 and IoT, 2018. pp.1-6.
- [16] Jingwei Ji Rishi Desai K. Kozuka Shun Ishizaka Nishant Rai, Haofeng Chen and Juan Carlos Niebles E. Adeli. *Home Action Genome: Cooperative Compositional Action Understanding*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11 May 2021.
- [17] Charles Ruizhongtai Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. *PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation*. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [18] Shitong Luo and Wei Hu. *Score-Based Point Cloud Denoising*. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021. 4583-4592.
- [19] Xingyu Liu, Mengyuan Yan, and Jeannette Bohg. *MeteorNet: Deep Learning on Dynamic 3D Point Cloud Sequences*. IEEE/CVF International Conference on Computer Vision (ICCV), 2019. 9245-9254.
- [20] Zhengyou Zhang Wanqing Li and Zicheng Liu. *Action recognition based on a bag of 3d points*. In CVPR Workshops, 2010.