

# 3次元点群を用いた 時空補間的アプローチに基づく人物軌跡構成法の提案

大野 真和<sup>1</sup> 右京 莉規<sup>1</sup> 天野 辰哉<sup>1</sup> Hamada Rizk<sup>1</sup> 山口 弘純<sup>1</sup>

**概要:** 近年, 公共施設や大型商業施設など様々な人々が行きかう空間における歩行者の人流解析の重要性が高まっている. 我々はこれまで3次元測域センサ (LiDAR) を用いた歩行者のトラッキングシステム「ひとなび」を開発してきた. ひとなびは固定設置した複数 LiDAR からの3次元点群からの人物検出およびトラッキングを行うことで, LiDAR の可視範囲内での歩行者移動軌跡のリアルタイム出力を可能にする. しかし, センサコストや建造物の構造上の制約により, 欠損なく関心領域全てをカバーするような LiDAR の設置は困難であることが多い. そこで本研究では, 点群から得られた歩行者軌跡に対する人物再識別手法を提案する. 提案手法では, (1) 点群から得られる人物の特徴の類似性, と (2) 建物レイアウトや過去の移動軌跡パターンにも基づく移動軌跡の整合性の2つの観点から, 全軌跡ペアに対してそれが同一人物によるものである確率を求め, 離れた計測領域からの軌跡のマッチングを実現する. 人物特徴の類似性計算方法として, Fisher Vector を用いた点群からの特徴抽出とそれに基づく距離学習を提案する. 大阪大学箕面キャンパスにおいて5日間にわたって取得した LiDAR データによる, 延べ人数 15,101 人の軌跡データを用いて提案手法の評価を行った. その結果, 2つの軌跡が同一人物によるものであるかどうかの二値分類において, F 値 0.80 という高い精度が得られた.

## 1. はじめに

スマートシティやスマートキャンパス実現に向け, 公共空間における人流や滞留状況, 空間の使われ方を計測する需要が高まっている. 人流計測とその解析により, 屋外では広告や店舗の配置を効率的に行うことが可能となる. 屋内では大型商業施設での売上が高くなるような店舗の配置や, 空港や病院での効率的な施設の配置や動線の確保に役立つ. また, 常に施設の利用者の状況を把握することで突発的な災害に対しても効率的な避難誘導が可能となり, 近年流行している新型コロナウイルス感染症対策として密になっている場所の発見にも役立つ. 在室状況や設備利用状態はスマートビルディングにおいて空調サービス最適化やスマートキャンパスにおいて授業の出席確認にも活用できるなど, 利点は大きい.

深層学習をはじめとする画像処理技術の発展により, RGB カメラからの映像を利用した人物検出やトラッキングを行うシステムが数多く提案されているが, それらは顔などの個人情報を含む情報を直接取得するため, プライバシー保護の面で課題がある. これに対して, 物体への距離情報のみを取得する3次元測域センサ (LiDAR) を用いた公共

空間における人流計測が注目を集めている. LiDAR の計測データは計測対象の3次元点群のみであり, 色情報を含まないため, 歩行者のプライバシーを侵害する恐れが小さい. 我々の研究グループではこの LiDAR を複数用いた公共空間での人物検出・トラッキングシステム「ひとなび」[1, 2] の開発を進めてきた. ひとなびでは, 複数の LiDAR から得られた点群を統合し, オクルージョンや点群ノイズなどを考慮した堅牢な歩行者の検出と追跡を実現している.

しかし, 設置した LiDAR の観測範囲から通行者が退出するとトラッキングが中断され, 同一観測範囲への再進入や, 別の観測範囲への進入の際に同一人物として認識されないという問題がある. 我々が対象とする大型商業施設等の建物全体を網羅したトラッキングを行うには多くの LiDAR が必要となるが, 網羅可能な台数を設置するために必要となるコストが大きく現実的でなく, また建物構造上センサ設置が困難である箇所も存在する. これらのことから計測不可能な領域が生じることは避けられない.

これに対して本研究では, そのような計測不可能な領域によって断片化した人物軌跡の再接続を目的とし, ひとなびにより得られる人物点群および人物軌跡に対し, それらが同一人物によるものかを判別する Person Re-Id のアルゴリズムを提案する. 提案手法では, 3次元の人物点群の特徴に基づく類似度と軌跡の遷移頻度および遷移時間と

<sup>1</sup> 大阪大学 Osaka University

いった遷移パターンの2つを主に組み合わせて、軌跡の再接続を行う。人物点群からの類似性計算のため、点群からの Fisher Vector 特徴量の抽出とその距離学習により、2つの人物点群の非類似度を表現する距離関数を獲得する。同時に、建物構造や確度の高い軌跡の観測から推定できる観測領域の接続関係とその遷移時間の履歴に基づいて、離れた2つの軌跡が同一人物の移動により生成された確率を求め、人物点群からの類似度と合わせることで、最終的に軌跡が同一人物によるものであるかを判定する。

提案手法の有用性を示すために、まず人流シミュレーションにより評価を行った。その結果、シミュレーション上では高い精度で同一人物の再識別を行えることを確認した。さらに実環境での評価のため大阪大学箕面キャンパスの廊下に設置した7台のLiDARから得られる大規模3次元点群に対して、人物追跡を実施し、5日間でのべ15,101人の軌跡データを取得し評価を行いF値0.80の精度を達成した。また、Fisher Vectorと距離学習に基づく人物再識別において、同一人物かどうかの2値分類においてF値0.680を達成した。

## 2. 関連研究

### 2.1 Multi-Object Tracking (MOT)

単一カメラ内で人物や車両などを追跡する手法は Multi-Object Tracking (MOT) として以前より多く研究されている [3,4]。文献 [5,6] では、Deep Neural Network により、単眼のRGB画像からオブジェクトとカメラの距離を計算し3D MOTを実現する。この手法では奥行き方向のトラッキングが可能となる精度が向上する。しかし、Deep Neural Networkを必要とするために計算量が多いという短所もある。奥行きの情報をLiDARなどの距離センサから取得しトラッキングする研究も行われている。RGB画像と3次元点群を両方利用し、RGB画像のみで行うトラッキングに3次元点群の奥行き情報を追加し歩行者検出やトラッキングを行う研究 [7-11] も存在している。これらの手法では、奥行き情報の取得に対する計算量が小さい。また、[12-16] では、RGB画像を利用せず3次元点群のみから歩行者を追跡する手法を提案している。3次元点群を利用した手法では、夜や逆光などRGB画像では明るさの関係で人物検出が困難な場合も、検出漏れを抑制することが可能である。ただし、これらの3次元点群を利用したMOTでは、人物の外見の特徴を利用していない。[17,18] では、自身で設計した特徴量を利用して3DMOTを行なっているが、点群全体の特徴を利用することは難しい。ベンチマークやデータセットとして、MOT Challenge [19] やKITTIデータセット [20] などが公開されている。MOT Challengeでは任意のオブジェクトのトラッキングや、混雑時におけるトラッキングなど様々なシーンにおけるトラッキング手法の比較が可能となっている。KITTIデータセットは通行する車

からLiDARセンサを用いて取得した3次元点群のデータであり、一般の通行者を多く含むデータである。MOTの代表的な手法としてDeepSORT [21] がある。DeepSortでは、物体検出手法Yoloを用いて画像内の対象オブジェクトを検出し、2次元画像フレーム内での移動を想定したカルマンフィルタを用いて連続フレーム間で対象物体の追跡を実現している。

### 2.2 人物再識別 (Person Re-ID)

異なる地点に設置された複数のRGB画像などで検出された各人物が同じ人物であるかどうかの判定を行う手法が研究されている [22]。疎な配置のセンサーでの歩行者トラッキングおよびマッチングを行う研究には、RGBカメラを用いた手法が一般的である。池亀ら [23] は、RGBカメラで得られた歩行者の服の色とカメラから見えなくなってから再び出てきた時間およびカメラの可視領域間の距離を用いて、歩行者同士の類似度を計算しハンガリアン法 [24,25] からマッチングを行なっている。Mogelmooseら [26] は、RGB-Dカメラと温度センサーを用いて人物の再識別を行なっている。また、Wangら [27] はRGBカメラと赤外線カメラを用いて人物の再識別を行なっている。これらの研究ではマップ情報を利用せずに単にカメラから撮影された人物同士の類似度を考えている。文献 [26] では深度情報を用いて同一人物の判定を行なっているが、これは点群から歩行者の骨格情報を取り出してその類似度を計算している。点群から骨格情報を取り出す手法はXuら [28] も提案しているが、個人の特定には多くの点群が必要となり、多数の歩行者が存在する環境においては、多数の点群の処理にかかる計算負荷は無視できない。ARデバイスなどの動くカメラで撮影された映像での人物再識別手法 [29] ではArcFace [30] を用いた人物再識別モデルを構築し、シーンの切り替えが頻繁に発生するラグビーの試合映像に対し評価を行なっている。

RGBカメラからの人物再識別のためのデータセットとして、Market-1501 [31] やMotion Analysis and Re-identification Set(MARS) [32] が知られている。Market-1501はDeformable Part Model(DPM)による人物検出器を用いた3万2千以上のバウンディングボックスがアノテーションされており、50万以上の不正解データも含まれている。またMARSは動画データを対象としており、Market-1501同様、DPMを用いて、1,261の人物による約2万のTracklet特徴量(追跡点の軌跡)を提供している。

### 2.3 軌跡パターン抽出

移動軌跡からパターンを抽出する研究として、軌跡圧縮 [33] や軌跡予測 [34] [35] が存在する。Chenら [33] は、日々増え続ける軌跡情報を保管するために、複雑な軌跡をシンプルな碁盤の目の地図上にマッピングして、方位とそ

の変化を使って圧縮する手法を提案している．これによって圧縮率と計算時間のトレードオフにおいて優れた結果が得られている．また，Qiao ら [34] は過去の移動パターンに基づいて，交通ネットワークにおける軌跡予測モデルを提案している．軌跡をクラスタに分割し頻出する軌跡パターンを発見することで，将来の移動軌跡を高い精度で予測する．同じく軌跡予測の研究を行なった Habibi ら [35] はスパースコーディング [36] を用い，それぞれの軌跡を，あらかじめ定義しておいた軌跡パターンの組み合わせで表現することで，新たな軌跡に対しても続く軌跡を予測可能にしている．また，それぞれが学習したデータをエージェント同士で通信し知識を更新する類似性ベースのモデル融合アルゴリズムを提案している．これらの研究は軌跡のルートから将来を推測して危険予知などを行っている．本研究でも LiDAR で得られた軌跡を基にルートに関する確率を計算して利用することで同一人物かどうかを判定する．

## 2.4 提案手法の位置付け

前述した人物再識別の手法は RGB カメラを用いており，歩行者の顔がそのままデータとして使われるといったプライバシー等の問題がある．これに対して提案手法では LiDAR を使うことによりプライバシーの心配を軽減している．RGB カメラを使わないため歩行者の顔や衣服等の色を把握することができないが，LiDAR を利用するため，より正確に歩行者の位置を捉えることが可能となり，歩行者の身長や速度を把握できるため，これらから同一人物の判定をすることが可能である．また，点群自体の類似度を使って同一人物の判定を行う．

## 3. 提案手法の概要と前処理

### 3.1 手法概要

提案手法は「ひとたび」が一定期間に出力する歩行者軌跡群を入力として受け取り，すべての軌跡ペアに対して同一人物によるものかどうか推定を行い，その結果を出力する．「ひとたび」による LiDAR からの歩行者軌跡の推定手法については 3.2 節で簡単に説明する．歩行者の軌跡は人物 ID，フレーム数，2 次元平面上での座標，そして歩行者の人物点群からなる．人物 ID は歩行者それぞれに与えられている ID で，一度軌跡が途切れて再び軌跡が検出された場合異なる人物 ID が与えられる．2 次元座標と高さは検出された人物のバウンディングボックスから算出する．2 次元座標はバウンディングボックスの 2 次元平面の中心の点として，高さは高さ方向のバウンディングボックスの長さを使う．同一の人物の連続するフレームでの 2 次元座標を繋げたものを人物の軌跡とする．出力はどの軌跡の人物 ID とどの軌跡の人物 ID が同一人物であるかのマッチングリストとする．

提案手法の流れを図 1 に示す．提案手法では，LiDAR

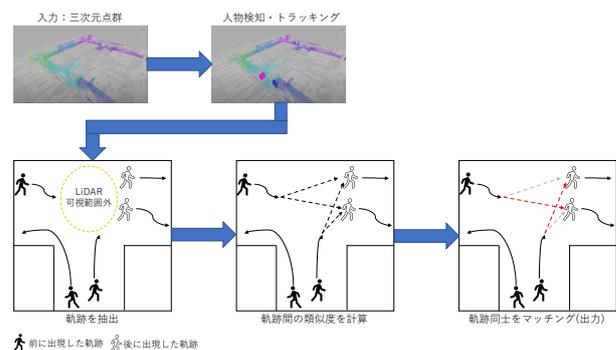


図 1 提案手法の流れ

にて取得した 3 次元点群にて歩行者検出・トラッキングを行い，生成された軌跡のうち同一人物の軌跡のマッチングを行う．入力された軌跡群に対して，2 つの全ペアに対して同一人物確率を計算する．軌跡間の同一人物確率の計算は，身長と歩行速度という人物の特徴から得られる人物類似度，人物点群の Fisher Vector 特徴量に対する距離学習に基づく類似度，軌跡パターンから得られる遷移経路と遷移時間による人物類似度をそれぞれ求め，それらの相乗平均を取ることで軌跡同士の同一人物確率を求める．この軌跡同士の同一人物確率から，それぞれの軌跡の終端および始端の接続可能かのマッチングを行い，出力はどの軌跡の人物 ID とどの軌跡の人物 ID が接続されるかのマッチングリストとする．

計算した同一人物確率  $P$  をもとに，どの二つの軌跡が同一人物であるかのマッチングを行う．それぞれの確率の総和が最大となるようなマッチングを行うため，ハンガリアンアルゴリズム [24] を利用する．ハンガリアンアルゴリズムは最小コスト問題を解くアルゴリズムであるので，それぞれの確率を 1 から減算したものを使って最大コスト問題から最小コスト問題に置き換える．これにより軌跡ペアのマッチングが定まるが，その中でも低い同一人物確率の軌跡ペアは新しく出現したり後半が途切れたりした軌跡のマッチングなどで正しくない可能性が高い．そのため，ある閾値以下の確率を持つマッチングはハンガリアンアルゴリズムによるマッチング結果から除外する．

### 3.2 LiDAR による歩行者の検出とトラッキング

本節では我々がすでに提案している複数 LiDAR を用いた歩行者トラッキング手法 [1] について簡単に述べる．本トラッキング手法では，まず 3 次元点群のデータを各フレームごとに，前景点群の抽出とクラスタリングによって歩行者の検出を行い，検出された歩行者に対してトラッキングを行う．移動物体を認識するために前景点群の抽出を行う．そのためにあらかじめ移動物体が存在していない時の背景点群を用意しておき，各フレームごとに取得された点群と背景点群との差分を取ることで移動物体の点群（前

景点群)を取得する。次に複数のLiDARから得られた点群を一つの全体の前景点群にまとめる必要がある。統合された点群において複数センサの視野が重複する領域では点群の密度が高くなってしまふ。これは後述する密度ベースのクラスタリングに影響を与える。そのためにVoxel Grid Filter [37]を使用し、得られた統合前景点群全体に対して各ボクセルグリッド内で重複する複数の点を一つの点にまとめるダウンサンプリングを適用する。統合された前景点群に対してクラスタリングを適用し、異なる複数の移動物体のセグメントに分割する。クラスタリング時の処理時間を圧縮するために、鉛直方向の軸を取り除き3次元点群を2次元点群に圧縮する。公共空間では鉛直方向の複数人の重複を考慮する必要はないためである。2次元に変換した点群に対し密度ベースのクラスタリング手法であるDBSCANを適用し、歩行者セグメントを得る。検出された歩行者セグメントの状態をカルマンフィルタにより歩行者セグメントの合体・分裂を予測しながらそれぞれの歩行者の位置と速度を予測し、観測された歩行者との割り当てを行うことでトラッキングを行う。カルマンフィルタを用いることでオクルージョンやノイズによる点群の欠損及び複数人物の接近による点群の結合などに対応し、堅牢なトラッキングを実現している。

### 3.3 身長と歩行速度の計算

歩行者トラッキングにより出力された軌跡から、4.1節での同一人物確率計算時に必要となる、歩行者の身長と歩行速度を計算する。LiDARの精度は距離に近いほど高くなるので、ある人物の軌跡の中でLiDARからの距離が最短の時の身長をその人物の身長とする。また歩行速度については、フレームごとの2次元座標間の絶対距離から歩行速度を得ることができ、それぞれ座っている状態や止まったり走ったりしている状態が考えられるので、閾値内の速さのデータのみを有効な観測として保存する。奥行き情報は高さに比べて正確に測定されるので有効な観測すべての中央値をそれぞれの人物の歩行速度とする。

## 4. 点群を用いた人物類似度計算

人物点群自体から得られる特徴から2つの軌跡が同一人物によるものである確率を求める。ここでは身長・歩行速度のユニークさを考慮した類似度計算と、人物点群から得られるFisher Vectorを特徴量とする距離学習に基づく類似度計算方法について述べる。

### 4.1 人物の特徴から得られる人物類似度

身長や歩行速度など観測で得られる情報から2つの軌跡が同一人物である確率を計算する。ここでの同一人物である確率とは、二つの観測が発生した時にそれらが同一人物である条件付き確率である。つまり二つの観測情報が似て

いれば確率は高くなり、さらにそれらの観測が珍しいものであればあるほどその確率は高くなる。あらかじめ分かっている観測の存在分布を $D_1$ とし、それぞれの観測が起こす測定誤差の分布 $D_2$ 、観測された二つの軌跡 $T_1, T_2$ に対して、それぞれの観測の値を $O_{T_1}, O_{T_2}$ として以下の計算を行う。

二つの観測 $O_{T_1}, O_{T_2}$ が発生し、 $T_1, T_2$ が同一人物であるとマッチする( $m(T_1, T_2)$ と書く)確率

$P(O_{T_1}, O_{T_2}, m(T_1, T_2))$ は、観測された一人の実際の観測情報の値 $x$ が発生する確率 $D_1(x)$ と二つの観測誤差( $O_{T_1} - x$ ), ( $O_{T_2} - x$ )が発生する確率 $D_2(O_{T_1} - x), D_2(O_{T_2} - x)$ を掛け合わせたものを、全ての観測情報 $x$ に対して計算して足し合わせたものとなる。これを式で表すと式(1)となる。

$$P(O_{T_1}, O_{T_2}, m(T_1, T_2)) = \int_0^{\infty} D_1(x) \times D_2(O_{T_1} - x) \times D_2(O_{T_2} - x) dx \quad (1)$$

また、二つの観測が発生しそれらが同一人物でない確率 $P(O_{T_1}, O_{T_2}, \neg m(T_1, T_2))$ は、軌跡 $T_1$ の実際の観測情報の値 $x$ が存在して測定誤差( $O_{T_1} - x$ )が発生する確率を掛け合わせたものを全ての観測情報 $x$ に対して計算して足し合わせたものと、軌跡 $T_2$ についても同様の計算を行なったものの積とする。式で表すと式(2)となる。

$$P(O_{T_1}, O_{T_2}, \neg m(T_1, T_2)) = \int_0^{\infty} D_1(x) \times D_2(O_{T_1} - x) dx \times \int_0^{\infty} D_1(x) \times D_2(O_{T_2} - x) dx \quad (2)$$

よって二つの観測情報 $O_{T_1}, O_{T_2}$ が観測された時に、それらが同一人物である条件付き確率

$P(m(T_1, T_2)|O_{T_1}, O_{T_2})$ を求めると式(3)となる。

$$P(m(T_1, T_2)|O_{T_1}, O_{T_2}) = \frac{P(O_{T_1}, O_{T_2}, m(T_1, T_2))}{P(O_{T_1}, O_{T_2})} = \frac{P(O_{T_1}, O_{T_2}, m(T_1, T_2))}{P(O_{T_1}, O_{T_2}, m(T_1, T_2)) + P(O_{T_1}, O_{T_2}, \neg m(T_1, T_2))} \quad (3)$$

観測情報として、身長と歩行速度について同一人物確率について計算する。身長の存在分布 $D_1$ は20歳日本人の男女の身長分布を組み合わせて混合ガウス分布として表現し、測定誤差分布 $D_2$ は平均を0とした正規分布で表現する。測定誤差の標準偏差は観測環境によってパラメータを決定する。歩行速度については存在分布と測定誤差分布の両方を正規分布で表現する。

この計算式で全ての身長間についての確率を実際に計算した例を図2に示す。この例ではx,y軸は観測された身長 $O_{T_1}, O_{T_2}$ を表しており、120cmから200cmまでの身長について計算している。z軸は身長に関する測定 $O_{T_1}, O_{T_2}$ が

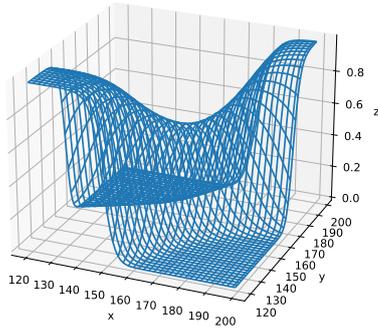


図 2  $P_1$  の例

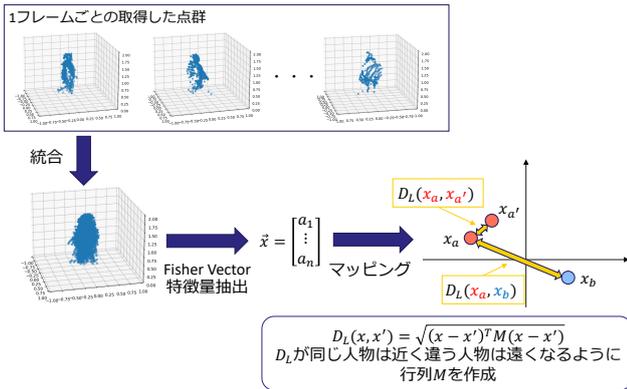
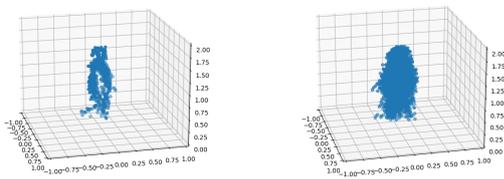


図 3 人物点群の Fisher Vector 特徴量を用いた距離学習



(a) 1 フレーム分の人物点群 (b) 10 フレーム統合後の人物点群

図 4 距離学習における入力人物点群の例

観測された時に軌跡  $T_1, T_2$  が同一人物である条件付き確率  $P(m(T_1, T_2) | O_{T_1}, O_{T_2})$  で、測定誤差分布の標準偏差は 5cm としている。このように、極端に低い値と高い値の場合は、もう片方の身長が多少ずれていても高い確率が与えられ、160cm あたりの出現頻度の高い身長については、もう片方の身長が全く同じであっても、およそ 0.6 の確率しか与えられない。

#### 4.2 Fisher Vector を用いた距離学習に基づく人物類似度計算

点群から特徴量を抽出し、距離学習を用いてそれらがどのくらい類似しているかという値から人物間の類似度を求める。手法の概要を図 3 に示す。

取得した点群に対して以下の操作を行う。1 フレーム分

の点群を可視化すると図 4(a) のようになる。これだけでは点群が疎なので複数フレームを統合する。統合はそれぞれの点群を上から見た XY 平面の中心の点が 0 座標に来るように平行移動し、それらの点群を重ね合わせた。重ね合わせた点群を図 4(b) に示す。また、統合した点群同士を比較する際に同じ行動同士を比較する方が良いので直進状態のみ使用する。そのために統合する前景点群の最初と中央と最後の XY 平面の中心座標  $P_1, P_2, P_3$  を用いて、 $\overrightarrow{P_1 P_2}$  と  $\overrightarrow{P_2 P_3}$  のコサイン類似度が閾値以下の統合点群はデータから除去する。最後に統合した点群同士の向きを合わせるために、 $\overrightarrow{P_1 P_3}$  のなす角を -1 倍した分だけ原点中心に回転させる。

次に点群から特徴量を抽出するために Fisher Vector [38] を用いる。点群を深層学習に用いるのは点群の不規則な形式から困難である。点群をボクセルグリッドに変換する方法も存在するがメモリサイズが大きくなってしまいう問題がある。これに対し本手法では、点群の特徴量を混合ガウスモデル (GMM) に最尤フィットさせるための対数尤度の勾配である Fisher Vector で表現し、小さいメモリサイズで点群を表現する。

得られた特徴量に対して距離学習 [39] を使って同一人物の点群が近づくような行列を学習する。距離学習では 2 つの入力  $x, x'$  に対して、それらのマハラノビス距離  $D_L$  を  $D_L(x, x') = \sqrt{(x - x')^T M (x - x')}$  として同一のクラス同士が近くなり異なるクラス同士が離れるように行列  $M$  を決定する。この決定するためのアルゴリズムとして LMNN (Large Margin Nearest Neighbor) [40] を用いる。LMNN ではヒンジ損失に基づく凸最適化によってクラス分類を行う。

学習した行列  $M$  を用いて 2 つの点群間のマハラノビス距離を求める。求めたマハラノビス距離が短いほど高い同一人物確率を与える。

## 5. 遷移パターンに基づく類似度計算

### 5.1 接続関係

計測開始当初は、LiDAR の可視範囲間の移動がどの程度の頻度で発生し、どの程度の時間を要するかのデータが十分に集積しておらず、遷移経路や移動時間に関する確率を得ることができない。したがって、この計算のために、軌跡の接続関係を調べる。軌跡の接続関係とは、マッチングによって接続された二つの軌跡について、前に出現した軌跡の終点と後に出現した軌跡の始点の座標、ならびにそれらの間の移動にかかった時間のことを指すものとする。この接続関係によって LiDAR の可視範囲の境界間の遷移確率と移動時間を求める。接続関係の初期値として、どのエリアからどのエリアへどのくらい時間がかかるのかという情報を事前に与えておく。この初期値として入力した接続関係だけではデータ数が少なく正確に表すことができ

いない。マップ上での移動経路の発生しやすさと移動時間分布を正確なものにするために接続関係の更新が必要になる。しかし移動した人物の同一人物判定はLiDAR スキャナから得た情報だけでは難しいので、マッチングした結果の中でも確度が高く信頼できるデータを探して接続関係の更新に用いる。軌跡同士の同一人物確率が高く、両方の軌跡に対して他に確率の高い候補が存在しない場合に、再接続されるこの軌跡ペアを信頼できるデータと判定する。この接続関係の更新は一定期間ごとに行う。

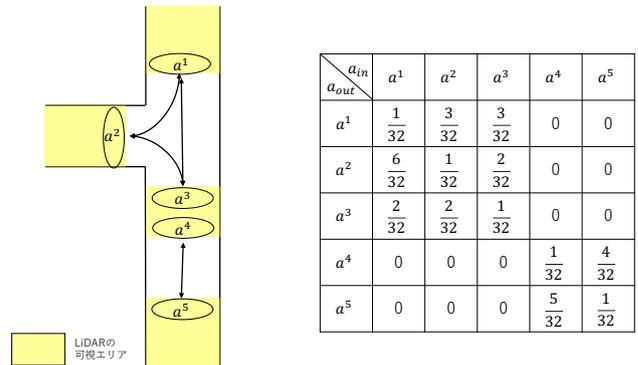


図 5 遷移経路による同一人物確率の例

## 5.2 時空間的整合性から得られる同一人物確率

信頼できるデータから得られた接続関係 (LiDAR の可視範囲外での移動経路や移動所要時間) を用いて、2つの軌跡が同一人物由来のものであるかの確率を求める。そのために遷移経路と遷移時間の二つの観点から求める確率を定義する。

まず、経路の中で通りやすい経路の確率を高くして通りにくい経路の確率を低くするために、遷移しやすさを表す確率を遷移経路による同一人物確率  $P_3$  とする。軌跡の遷移経路を定めるために、軌跡の終点と始点をそれぞれクラスタリングする。終点をクラスタリングしたエリアを  $a_{out}^1, a_{out}^2, \dots$  とし、始点をクラスタリングしたエリアを  $a_{in}^1, a_{in}^2, \dots$  とする。クラスタリングにはクラスター数が未定であることからウォード法 [41] を用いる。信頼できるデータを使ったクラスタリングの結果から、どのエリアからどのエリアへ何人が遷移したかが分かるので、それぞれの終点からそれぞれの始点への遷移確率を要素とするエリア間の遷移行列  $A$  が作成できる。ここで求めたい確率は全体に対する遷移の頻度ではなく、ある終点があった時にそれぞれの始点への遷移のしやすさなので、終点を条件とした時の条件付き確率を求めたい。終点のエリアが  $a_{out}^i$  である軌跡  $T_1$  と始点のエリアが  $a_{in}^j$  である軌跡  $T_2$  の条件付き確率  $P_3(T_1, T_2)$  を求めると式 (4) となる。

$$P_3(T_1, T_2) = \frac{A(a_{out}^i, a_{in}^j)}{\sum_k A(a_{out}^i, a_{in}^k)} \quad (4)$$

例えば図 5 のようにマップが存在し、 $a^1$  から  $a^5$  までを終点・始点をクラスタリングした結果のクラスターとして、簡単のために終点と始点のクラスター番号を同じものとして考える。遷移行列は総和が全て 1 になるように生成される。この中で終点を条件とした時の条件付き確率  $P_4$  は対象の要素をその行の総和で割ったものとなるので  $a_{out}$  に  $a^1$ 、 $a_{in}$  に  $a^2$  が与えられたとき  $P_3 = \frac{3}{7}$  となる。この遷移行列の初期値は手入力で終点・始点とその間にかかる時間をおおまかに設定し、それぞれの遷移確率は視点から遷移する可能性のある終点について全て等確率で与える。

新たに出現した二つの軌跡  $T_1, T_2$  に対して、先に出現した  $T_1$  の終点と後に出現した  $T_2$  の始点について最も近い

クラスターを探し、そのクラスター間の遷移確率を遷移行列  $A$  から遷移経路による同一人物確率  $P_3(T_1, T_2)$  を計算する。

次に終点と始点が正しいとした時の遷移時間の尤もらしさを遷移時間による同一人物確率  $P_4$  とする。軌跡  $T_1$  から  $T_2$  へ遷移した確率を考える。 $T_1$  の終点を観測した時間を  $t_{out}^{T_1}$  として  $T_2$  の始点を観測した時間を  $t_{in}^{T_2}$  とすると、遷移にかかった時間  $t_{diff}$  は  $t_{diff} = t_{in}^{T_2} - t_{out}^{T_1}$  と表すことができる。正解データの中で終点が  $T_1$  の終点と同じクラスターに属するもので始点が  $T_2$  の始点と同じクラスターに属するものを取り出す。それらのデータを使って実際にかかった遷移時間の頻度から遷移時間分布の確率密度関数  $f(x)$  を推定する。この関数に先に求めた  $t_{diff}$  を代入することで遷移時間による同一人物確率は、 $P_4(m(T_1, T_2) | P_3(T_1, T_2)) = f(t_{diff})$  と表すことができる。

## 5.3 信頼できるデータの判定と接続関係の更新

LiDAR スキャナのみでは正解データの判定ができず、どのデータを学習に使うべきか分からない。よってここで接続関係を調べるために使う信頼できるデータの判定方法と、そのデータを用いた接続関係の更新方法について述べる。

信頼できるデータを判定するために、二つの閾値パラメータ  $\alpha, \beta (0 < \alpha < \beta < 1)$  を用意する。 $\alpha$  は低い確率のパラメータでこれより低い確率を持つ軌跡の組み合わせは確実に誤ったマッチングであると、 $\beta$  は高い確率のパラメータでこれより高い確率を持つ軌跡の組み合わせは信頼できるデータとして使える可能性があるマッチングとする。既にマッチングされた二つの軌跡  $T_1, T_2$  が存在して、それらの同一人物確率  $P(T_1, T_2)$  が  $\beta$  より大きく、 $T_1, T_2$  を除く全ての軌跡  $T_k$  に対してそれらの同一人物確率  $P(T_1, T_k)$  と  $P(T_k, T_2)$  が  $\alpha$  より小さい場合にそのマッチングを信頼できるデータと判定する。つまり同一人物確率が高いマッチングのペアで、その軌跡と他の軌跡の同一人物確率は全て低いマッチングのみを信頼できるデータとする。

信頼できるデータには、マッチングされた前の軌跡の終点と後の軌跡の始点の座標、軌跡間の移動にかかった時間

表 1 使用した 3 次元 LiDAR の仕様および設定

	Livox Avia	Hokuyo YVT-35LX
最大計測点群数	240,000 点/フレーム	2,664 点 / フレーム
フレームレート	10 フレーム/秒	10 フレーム / 秒
最大計測可能距離	460 m	35 m
走査視野角 (水平)	70.4 度	210 度
走査視野角 (垂直)	77.2 度	40 度
距離精度	± 2.0 cm	± 0.1 cm
角度精度	± 0.05 度	± 2.0 度

が含まれており、この信頼できるデータを接続関係として 1 日ごとに加算され蓄積されていく。蓄積された接続関係に対して以下の操作を行う。まず終点の座標と始点の座標をそれぞれクラスタリングする。クラスタリングで得られたクラスターごとに遷移確率行列と全てのクラスター間の遷移時間分布を求める。求めた遷移確率行列と遷移時間分布を遷移経路と遷移時間による同一人物確率の計算に利用する。

## 6. 性能評価

大阪大学箕面新キャンパスの 1 フロアで収集したデータを用いて点群の特徴から得られる同一人物確率の精度の評価を行った。また、提案手法を評価するために Rhino 7 および Ped Sim を用いた人流シミュレーションと、大阪大学箕面新キャンパスの別のフロアで収集したデータで実験を行った。

### 6.1 点群特徴の距離学習に基づく類似性

まず 4 章で述べた Fisher Vector による点群抽出と距離学習に基づく人物点群の類似度計算の精度について評価を行った。2 台の LiDAR がを用いて 32 人の歩行者が移動する様子の人物点群データを取得した。3 次元 LiDAR として LIVOX 社の Avia を使用した。LiDAR の性能を表 1 に示す。被験者一人につき 100 秒ほど計 1000 フレームの点群を取得した。

取得した人物点群に対して 10 フレーム単位で点群を統合し、それを 1 サンプルとして Fisher Vector 特徴量を抽出し学習を行った。Fisher Vector 抽出のための混合ガウス分布の個数は 27 個とした。得られた統合した点群のうち 75% をトレーニングデータとして、残りの 25% をテストデータとして使用する。16 人のトレーニングデータを使って距離学習でモデルを作成 LNMN を用いて距離関数を得た。学習した距離関数適用後のトレーニングデータを t-SNE [42] により二次元空間で表現すると図 6 になる。距離関数適用後のマハラノビス距離  $D_L$  を使って、それぞれのテストデータ  $x_1$  と全てのトレーニングデータ  $x_2$  との距離の平均  $\bar{D}_L(x_1, x_2)$  を計算した。距離 (非類似度) の行列を図 7 に示す。

この非類似度行列から正しい組み合わせとそれ以外の組

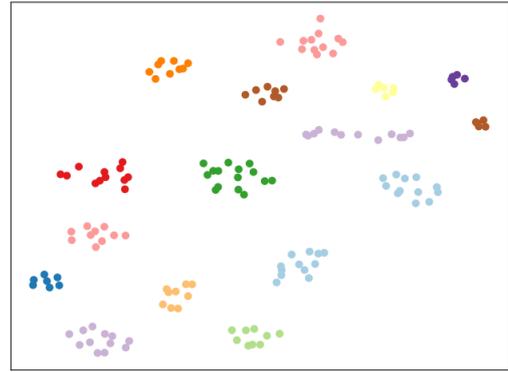


図 6 提案手法による 16 人の人物点群の非類似度行列を t-SNE により 2 次元可視化した図。1 点が 1 つの人物点群を示し、同一色は同一人物による点群であることを示す。

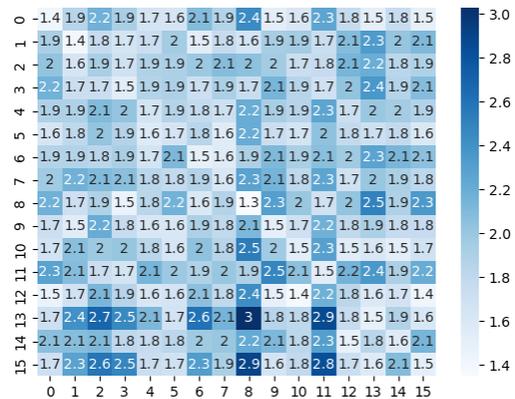


図 7 学習により得られた距離関数による人物の類似性 (学習データに含まれる 16 人)

表 2 距離の平均と標準偏差 (学習データに含まれる 16 人)

	平均	標準偏差
同一人物による点群のペア	1.56	0.144
異なる人物による点群のペア	1.93	0.279

表 3 距離の平均と標準偏差 (学習データに含まれない 16 人)

	平均	標準偏差
同一人物による点群のペア	1.99	0.307
異なる人物による点群のペア	2.39	0.432

み合わせの距離  $\bar{D}_L$  の平均と標準偏差を計算し表 2 に示す。またに学習データに含まれない 16 人のデータで非類似を計算した。結果を図 8 に、得られた距離の平均と標準偏差を表 3 に示す。この結果の 2 つの正規分布から、両方の平均が出現した時にどちらのペアであるかを計算する。その結果距離が 1.99 の時に同一人物である確率は 68%、距離が 2.39 の時に同一人物である確率は 38%であった。

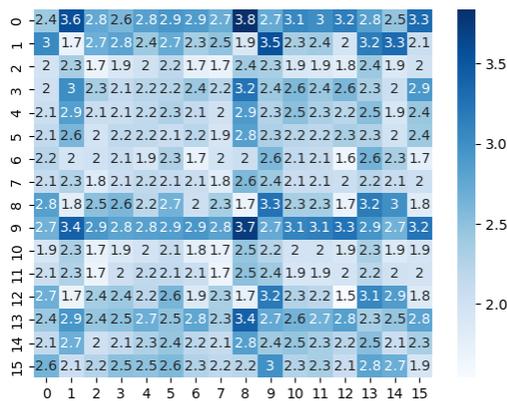


図 8 学習により得られた距離関数による人物の類似性（学習データに含まれない 16 人）

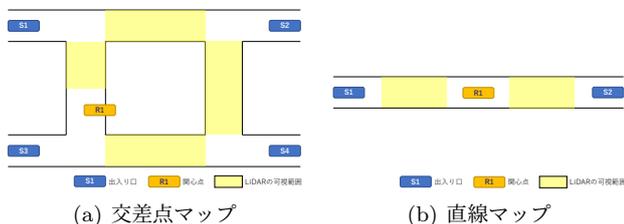


図 9 シミュレーションで用いたマップ

## 6.2 人流シミュレーション上での実験・評価

シミュレーションでは、十分に正解データが存在した時に提案手法によって正しくマッチングできるかを評価する。人物の特徴から得られる同一人物確率を使った場合と点群の特徴から得られる同一人物確率を使った場合に別けて評価した。

人流シミュレーションで用いたマップを図 9 に示す。マップの端にある青の点はそれぞれ人の出入り口を表し、人が生成される。これらの点からランダムな二点を選択し、始点もしくは終点とすることで人流を再現する。中央の黄色の点は関心点を表し、歩行者の視界に関心点が入った場合、その関心点に向かって一定時間滞在した後再度目的地向かって移動する。この関心点に向かうかどうかは歩行者の生成時に 50% の確率で決まる。関心点に向かわない 50% の歩行者は、視界に関心点が入っても決められた終点に向かう。したがって図 9(a) で右側の通路 (S2 と S4 の出入り口の間) を行き来する歩行者はたとえ関心点を持っていたとしても視界に入らないので関心点を持っていない人物と同一の動きをする。歩行者は一定時間ごとにランダムな出入り口から生成される。歩行者には一定の大きさが設定されており、対面から別の歩行者が来た場合は互いに避けるように移動する。黄色いエリアはセンサーの可視範囲を表す。センサーの可視範囲内の軌跡同士について同一人物かどうかを判定したいので、可視範囲外の軌跡データは削除しておく。図 9(a) のマップでは四隅の出入り口

から他の出入り口に向かう軌跡を再現し、センサーの無い交差点部分で遷移の分岐が発生するように設計した。センサーがある場所から来た歩行者が交差点に入り、センサーがあるエリア、もしくはセンサーが無いエリアへ遷移する。これによりトラッキングが途切れた人がどのセンサーで検出されるのかそれともセンサー範囲外に出てしまうのかを判断したい。また、図 9(b) のマップでは直線を中心に関心点を置くことにより、ルート別の遷移確率には大きな違いはないが遷移時間が大きく二つに別れるようにマップを設計した。

どちらのマップでもシミュレーションデータを 30 日分ずつ用意して、10 日分をトレーニング用データ、残りの 20 日分をテスト用データとした。ここで、1 日分のデータは 50 人分の軌跡を持つ。トレーニング用データでは同一の歩行者の軌跡が分かっている状況で全ての出入り正解データとして保存した。テスト用データでは一つの ID に対して異なるセンサーの可視範囲ごとに異なる ID を与えた。異なる二つの軌跡が同一人物であるかどうかの二値分類に対して、精度の評価を行なった。

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

与えられた二つの軌跡が同一人物のものであるかどうかの二値分類に対して、TP (True Positive), FP (False Positive), FN (False Negative), TN (True Negative) の数を求め、Precision, Recall, F 値を求める。それぞれを求める方法は式 (5) に示す通りである。シミュレーションにより得られた 30 日分の軌跡群データセットに対して、10 日分のデータで学習した後、残りの 20 日分のデータで提案手法の精度と生成した信頼できるデータの個数を評価した。身長と歩行速度を仮定したデータを使って、交差点マップでの評価結果を表 4 に示す。この結果データ数が増えるたびに Precision の値は上昇したが、Recall の値は減少した。また、F 値は 0.9 を超えていた。次に直線マップでの評価結果を表 5 に示す。交差点マップとは異なり常に高い結果が得られた。こちらでも誤った信頼するデータを生成することはなかった。この理由は二つあり、一つ目は交差点マップで問題であった移動経路確率の低下が解消されたためである。直線マップの境界は分かれ道が存在せず、全てが新たに現れるかそれ以降存在しない軌跡であるか、全てがマッチングする軌跡が見つかる軌跡かのどちらかしか無いからである。二つ目は移動時間による一致確率によって、直接向かった歩行者と関心点を中継した歩行者の移動時間どちらについても考慮できているからである。一方で少ないながらもマッチングに失敗しているものも存

表 4 交差点マップでのシミュレーション評価結果

日付	Precision	Recall	F 値
1~5	0.959	0.891	0.921
6~10	0.952	0.880	0.914
11~15	0.983	0.845	0.908
16~20	0.984	0.866	0.920
1~20	0.970	0.871	0.916

表 5 直線マップでのシミュレーション評価結果

日付	Precision	Recall	F 値
1~5	0.987	0.987	0.987
6~10	1	1	1
11~15	0.987	0.987	0.987
16~20	0.987	0.987	0.987
1~20	0.990	0.990	0.990

表 6 点群特徴を用いた時の交差点マップでのシミュレーション評価結果

日付	Precision	Recall	F 値
1~5	0.931	0.922	0.926
6~10	0.958	0.864	0.960
11~15	0.874	0.917	0.895
16~20	0.954	0.939	0.946
1~20	0.929	0.935	0.932

在する。これは似たような身長、歩行速度、経路、移動時間が重なった時に、総合確率について誤ったマッチングの和が正しいマッチングの和より大きい場合に入れ替わりが起こってしまう。この誤りの対策としては、それぞれの同一人物確率に対して重みを付け、相乗平均を取れば性能が上がる可能性がある。

次に、身長と歩行速度の情報の代わりに点群から得られる特徴量を使った場合のシミュレーションを行った。点群情報をシミュレーションで実現するのは困難なため、6.1節にて示した評価の結果から同一人物の場合は68%、違う人物の場合は38%の確率を与える。他の設定はそのまま実験を行なった。交差点マップでの評価結果を表6に示す。次に、直線マップでの評価結果を表7に示す。この結果からほとんどの場合でF値が向上していることが分かった。特に軌跡が途中で失われない直線マップでは精度が1になり信頼できるパラメーターが複数あればかなり高い精度で予測することができるということが分かった。接続関係について、初期のトレーニング段階での可視範囲の境界における座標のクラスタリングと遷移時間を調べた。シミュレーションでは動き方のパターンがある程度決まっているため、図10で示されるような境界のクラスタリングや図11で示されるような遷移時間分布が綺麗に生成できた。

### 6.3 キャンパスでの実データでの評価

大阪大学箕面キャンパスにおいて5日間にわたって収集したデータを用いて提案手法を精度を評価した。

表 7 点群特徴を用いた時の直線マップでのシミュレーション評価結果

日付	Precision	Recall	F 値
1~5	1	1	1
6~10	1	1	1
11~15	1	1	1
16~20	1	1	1
1~20	1	1	1

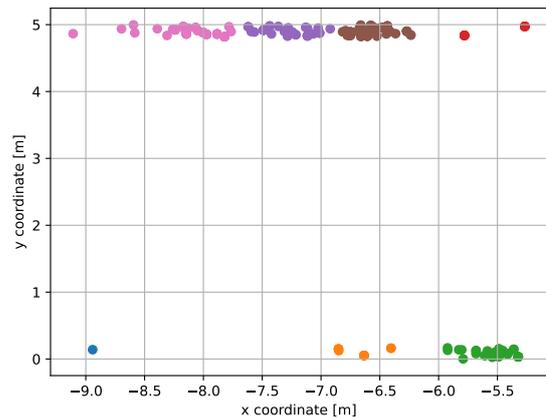


図 10 交差点マップの左側の通路における境界のクラスタリング結果

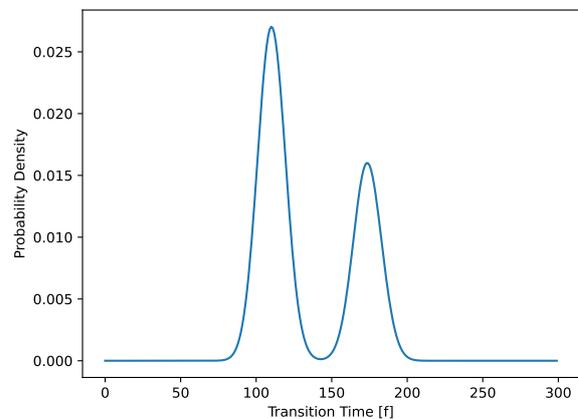


図 11 直線マップにおけるエリア間の遷移時間分布

#### 6.3.1 実験環境および収集データ

3次元LiDARとして北陽電機株式会社のYVT-35LX-F0を使用した。LiDARの性能を表1に示す。大阪大学箕面新キャンパスの5階のフロアに図12のように7台のLiDARを設置した。それぞれのLiDARは図13のように天井に貼り付くように設置されている。また、図12で青枠で囲われている部分の軌跡データを削除しその前後の軌跡が同一人物のものかどうかを判定した。データを削除する部分は図13に示すように、1から3まで三通りの場所を一ヶ所ずつ別々に削除して検証する。1の三叉路ではほとんどの場合三叉路の右と下を行き来する。2のエリアは直線だが

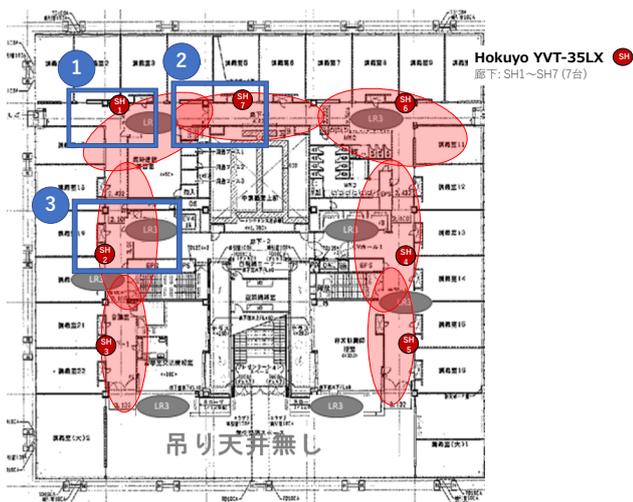


図 12 建物のマップと LiDAR の設置位置および評価に使用する領域 (図の青枠 1~3)



図 13 LiDAR の設置方法

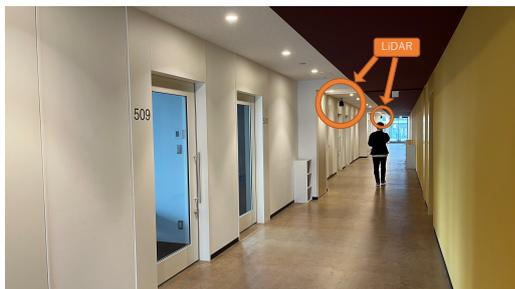


図 14 LiDAR を設置した廊下 (外周)

間の教室に入出入りする場合もある。3の三叉路では1と違い遷移経路の偏りが少なく、上・右・下の全てに同じような確率で遷移する。このフロアは外側に教室が並び、その内側に廊下が一周通っている。図 14 は図 12 において SH3 から SH2 方向に撮影されたもので、外周の廊下には側面にいくつか教室がある。また、図 15 は図 12 において SH2 と SH4 の間の廊下を撮影したもので、中央には自動販売機が設置されている。LiDAR の可視範囲から外周の廊下では軌跡を検出することができるが、中央の廊下とそれぞれの教室内では軌跡を検出することができない。

7 台の LiDAR で観測された点群を統合して得られた大規模点群を図 16 に示す。紫や青などで描かれている点群は背景点群で色によってどの LiDAR で観測されたものかが異なる。廊下に描かれている緑の点群は前景点群を表し、歩行者と識別された前景には薄い色のバウンディングボックスがかけられている。また、バウンディングボックス



図 15 LiDAR を設置した廊下 (中央)

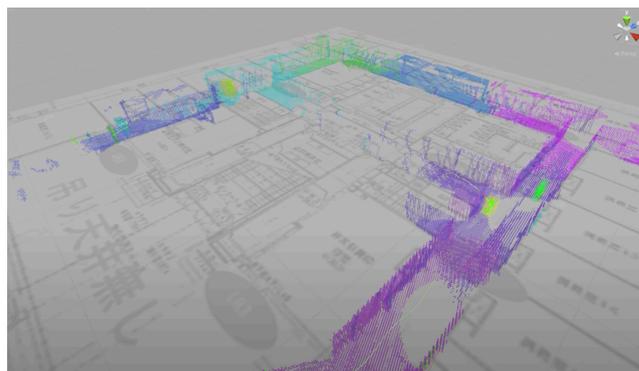


図 16 LiDAR によって得られる点群

表 8 検出された軌跡の数

日付	軌跡の数
1月20日(木)	3691
1月21日(金)	3901
1月22日(土)	437
1月23日(日)	5864
1月24日(月)	1208
合計	15101

スの中心の軌跡が歩行者の軌跡としてバウンディングボックスと同じ色の線分で描かれている。

LiDAR から得られた点群のみを使って正解データのラベリングを行うことは困難なので、あらかじめトラッキングできている外周の廊下で得られたデータを使い提案手法の精度を検証する。データの一部を削除することでその部分は点群が得られなかった事として、そのエリアに入る前後の軌跡を繋げる作業を本手法で行う。データを削除する前にトラッキングしていた時の人物 ID を保持しておくことにより同一人物判定が正しく行える。

2022 年 1 月 20 日 11 時から同年 1 月 24 日 13 時まで合計 15,101 個の軌跡データを収集した。それぞれの日で得られた軌跡の数をカウントしたところ表 8 のようになった。ここで、土曜日は学生が少なく軌跡の数も少なかった。日曜日は工事があり検出された軌跡の数も多く、平日の歩行者の動きとは異なる傾向の軌跡が得られた。

これらのデータを 1 時間ごとのデータに分け、それぞれマッチング・評価・接続関係の更新を順に行った。この時、軌跡が途中から始まっていたり、軌跡がデータの最後

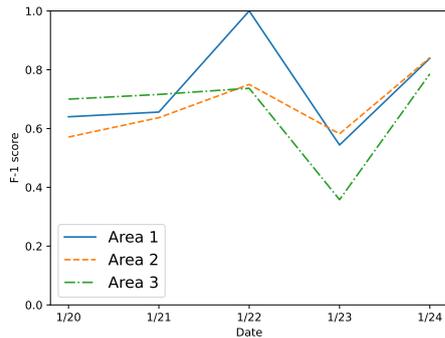


図 17 5日分のデータに対する一日ごとの精度評価

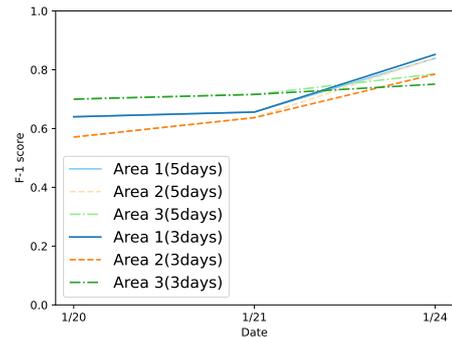


図 18 3日分の平日のデータに対する一日ごとの精度評価

表 9 5日分のデータに対する信頼できると判定したデータの評価

場所	日付	正信頼数	誤信頼数
1	1/20	40	4
	1/21	58	6
	1/22	8	0
	1/23	34	7
	1/24	11	0
2	1/20	67	63
	1/21	55	34
	1/22	6	2
	1/23	65	201
	1/24	21	2
3	1/20	122	72
	1/21	118	27
	1/22	7	3
	1/23	53	227
	1/24	44	4

まで続き途中で区切れた可能性のある軌跡は評価の対象外とした。1日ごとにF値を求め評価する。また、学習に使用する信頼できるデータとして判定されたものの中で、実際に正しいマッチングであった「正信頼数」と、実際は誤りだった「誤信頼数」をカウントする。これを全ての5日分のデータと、土日を除いた3日分のデータについて実行し性能を評価する。

### 6.3.2 5日分のデータに対する実験結果と考察

5日間のデータセットに対して図12のように三ヶ所のボックスで囲まれた部分の軌跡を別々に削除し、削除された前後の軌跡で同一人物かどうかの判定を行った。マッチングの精度についてのグラフを図17に示す。

また、学習した正解データの中で実際に正しかったものと誤りだったものの数をカウントし、その結果を表9に示す。

収集したデータの中で1月22日は歩行者の数が少なくマッチングが単純に行うことができ評価が良く、逆に1月23日は歩行者が多く普段とは異なる軌跡が多く評価が悪かった。しかし、シミュレーションとは異なり、誤ったマッチを信頼できるデータとして判定してしまった。特に

1月23日の結果では、FP正解データ数がTP正解データより多く、誤った経路情報モデルが生成されてしまっている。これは、来校していた工事業者の方が普段の歩行者とは異なる動きをしていたからだと考えられる。

また、Mac Book Pro (Apple M1 Pro) を使用してPythonでプログラムを実行したところ、一箇所につき約17時間の実行時間を要した。

このような結果となった原因として、観測によって得られる身長と歩行速度にばらつきがあり、それらによる同一人物確率に大きな差が出なかったこと、一度誤ったデータを学習してしまうとそれらに従って次の確率も計算されてしまうので、連鎖的に誤った結果を出してしまったことが挙げられる。また、誤ったデータを正解データとして利用してしまったのは、初期値として与えた接続関係が適切でなかったことが考えられる。誤ったデータを学習しても時間経過に伴い精度が上昇したのは、正しい移動時間に似た移動時間を信頼できるデータとして蓄積していくので遷移時間関数のグラフは大きくは変わらないからである。しかし、この手法をより長期間の実験で得られたデータで実行すると、次第に遷移時間関数のグラフが大きく変わってしまい、精度が下がる可能性がある。

### 6.3.3 3日分の平日のデータに対する実験結果と考察

平日と休日では歩行者の動きに違いがあることがわかったので、土日のデータを除いたもので同様の条件で再度実行した。マッチング精度についてのグラフを図18に示す。それぞれ薄い色は5日分のデータに対する評価で、3日分を抽出し重ね合わせた。また、信頼できるデータと判定したものの数の評価も同様に行い表10に示す。

平日のみのデータでは、F値が1日ごとに上昇していったことが分かった。5日分のデータと比較すると、全体的に大きな変化は無かったが、最終日の精度が向上した場合としていない場合があった。これらの違いとして、誤った信頼できるデータを接続関係として蓄積した数が少ないほど、最終日の精度が上昇した。また、最終日に信頼できるデータと判定したものの中で正しいものの割合が変わらないものもあったが、全体的に見ると僅かにその割合は上昇

表 10 3日分の平日のデータに対する信頼できると判定したデータの評価

場所	日付	正しい信頼データ数	誤った信頼データ数
1	1/20	40	4
	1/21	58	6
	1/24	11	0
2	1/20	67	63
	1/21	55	34
	1/24	14	2
3	1/20	122	72
	1/21	118	27
	1/24	40	2

した。これは、休日に多く接続関係として蓄積されていた誤った信頼できるデータの数が減少したからだと考えられる。

## 7. おわりに

本研究では、複数のLiDARによる3次元点群からの歩行者追跡システムにおいて、LiDARの計測不可能範囲により歩行者軌跡が断片化する問題に対して、歩行者軌跡を再接続（再識別）する手法を提案した。提案手法では、点群からの人物特徴抽出と距離学習による類似性計算、建物構造と移動にかかる時間を考慮した移動パターンによる整合性計算に基づいて、2つの軌跡が同一人物によるものであるかを決定する。提案手法の評価のため、大阪大学の箕面キャンパスで5日間にわたって実際に収集した3次元点群および歩行者軌跡に対して軌跡の再接続を適用した。その結果、2つの軌跡が同一人物によるものであるかどうかの二値分類においてF値0.80を達成した。

今後の課題として、今回の提案手法はオフライン処理により一日ごとに得られた軌跡の全ペアに対して類似性計算を実施したが、キャンパス全館のトラッキングを実現するため、逐次処理によってリアルタイムでの再識別を実現する予定である。また今回点群による人物再識別のためFisher Vector特徴と距離学習の一種であるLMNNを用いた結果、学習データに含まれる人物に対しては比較的高い精度が得られたものの、学習データに存在しない人物同士の類似性計算では大きく精度が低下することが分かった。今後TripletLossによる深層距離学習を用いることで精度の向上を目指す。

## 参考文献

[1] Hirozumi Yamaguchi, Akihito Hiromori, and Teruo Higashino. A human tracking and sensing platform for enabling smart city applications. In *Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking (ICDCN 2018)*, 2018.

[2] 右京莉規, 天野辰哉, 廣森聡仁, 山口弘純, and 守屋充雄. 複数LiDARによる大規模三次元点群を用いた歩行者トラッキング手法の実装と評価. Technical Report 28, nov

2021.

[3] Patrick Dendorfer, Aljosa Osep, Anton Milan, Konrad Schindler, Daniel Cremers, Ian Reid, Stefan Roth, and Laura Leal-Taixé. Motchallenge: A benchmark for single-camera multiple target tracking. *International Journal of Computer Vision*, 129(4):845–881, 2021.

[4] Zheng Tang and Jenq-Neng Hwang. Moana: An online learned adaptive appearance model for robust multiple object tracking in 3d. *IEEE Access*, 7:31934–31945, 2019.

[5] H. N. Hu, Q. Z. Cai, D. Wang, J. Lin, M. Sun, P. Krahenbuhl, T. Darrell, and F. Yu. Joint monocular 3d vehicle detection and tracking. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2019)*, 2019.

[6] S. Scheidegger, J. Benjaminsson, E. Rosenberg, A. Krishnan, and K. Granström. Mono-Camera 3D Multi-Object Tracking Using Deep Learning Detections and PMBM Filtering. In *Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV 2018)*, pages 433–440, 2018.

[7] D. Frossard and R. Urtasun. End-to-end Learning of Multi-Sensor 3D Tracking by Detection. In *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA 2018)*, pages 635–642, 2018.

[8] A. Sheno, M. Patel, J. Gwak, P. Goebel, A. Sadeghian, H. Rezafofighi, R. Marten, and S. Savarese. JRMOT: A Real-Time 3D Multi-Object Tracker and a New Large-Scale Dataset. In *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020)*, pages 10335–10342, 2020.

[9] E. Baser, V. Balasubramanian, P. Bhattacharyya, and K. Czarnecki. FANTrack: 3D Multi-Object Tracking with Feature Association Network. In *Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV 2019)*, pages 1426–1433, 2019.

[10] A. Rangesh and M. M. Trivedi. No Blind Spots: Full-Surround Multi-Object Tracking for Autonomous Vehicles Using Cameras and LiDARs. *IEEE Transactions on Intelligent Vehicles*, 4(4):588–599, 2019.

[11] M. Munaro, F. Basso, and E. Menegatti. Tracking People Within Groups with RGB-D Data. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, pages 2101–2107, 2012.

[12] H. Wu, W. Han, C. Wen, X. Li, and C. Wang. 3D Multi-Object Tracking in Point Clouds Based on Prediction Confidence-Guided Data Association. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–10, 2021.

[13] M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz, and H. M. Gross. ComplexYOLO: Real-Time 3D Object Detection and Tracking on Semantic Point Clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019) Workshops*, June 2019.

[14] X. Weng, J. Wang, D. Held, and K. Kitani. 3D Multi-Object Tracking: A Baseline and New Evaluation Metrics. In *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020)*, pages 10359–10366, 2020.

[15] L. Spinello, K. Arras, R. Triebel, and R. Siegwart. A Layered Approach to People Detection in 3D Range Data. *Proceedings of the AAAI Conference on Artificial In-*

- telligence, 24(1):1625–1630, Jul. 2010.
- [16] A. Carballo, A. Ohya, and S. Yuta. People Detection Using Range and Intensity Data from Multi-Layered Laser Range Finders. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, pages 5849–5854, 2010.
- [17] J. Choi, S. Ulbrich, B. Lichte, and M. Maurer. Multi-Target Tracking Using a 3D-Lidar Sensor for Autonomous Vehicles. In *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 881–886, 2013.
- [18] S. Song, Z. Xiang, and J. Liu. Object Tracking with 3D LIDAR via Multi-Task Sparse Learning. In *Proceedings of the 2015 IEEE International Conference on Mechatronics and Automation (ICMA 2015)*, pages 2603–2608, 2015.
- [19] P. Dendorfer, H. Rezatofghi, A. Milan, J. Shi, D. Cremers, I. Reid, S. Roth, K. Schindler, and L. Leal-Taixé. Mot20: A benchmark for multi object tracking in crowded scenes. *arXiv:2003.09003[cs]*, March 2020. arXiv: 2003.09003.
- [20] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Proceedings of the 2012 IEEE conference on computer vision and pattern recognition (CVPR 2012)*, pages 3354–3361. IEEE, 2012.
- [21] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP 2017)*, pages 3645–3649. IEEE, 2017.
- [22] Yixiao Ge, Feng Zhu, Dapeng Chen, Rui Zhao, et al. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. *Advances in Neural Information Processing Systems*, 33:11309–11321, 2020.
- [23] 池亀幸久, 平野真, 玉木徹, 山本正信. 視野重複のないマルチカメラを用いた人物の確率的歩行経路推定. *信学技報 PRMU2004-186*, Feb, 2005.
- [24] James Munkres. Algorithms for the assignment and transportation problems. *Journal of the society for industrial and applied mathematics*, 5(1):32–38, 1957.
- [25] Francois Bourgeois and Jean-Claude Lassalle. An extension of the munkres algorithm for the assignment problem to rectangular matrices. *Communications of the ACM*, 14(12):802–804, 1971.
- [26] Andreas Mogelose, Chris Bahnsen, Thomas Moeslund, Albert Clapés, and Sergio Escalera. Tri-modal person re-identification with rgb, depth and thermal features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR 2013)*, pages 301–307, 2013.
- [27] Guan’an Wang, Tianzhu Zhang, Jian Cheng, Si Liu, Yang Yang, and Zengguang Hou. Rgb-infrared cross-modality person re-identification via joint pixel and feature alignment. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2019)*, pages 3623–3632, 2019.
- [28] Tianxu Xu, Dong An, Yuetong Jia, and Yang Yue. A review: Point cloud-based 3d human joints estimation. *Sensors*, 21(5):1684, 2021.
- [29] 高橋直也, 天野辰哉, 山口弘純, 東野輝夫. 深層距離学習を用いた AR デバイス向けの人物識別手法. *マルチメディア, 分散協調とモバイルシンポジウム 2021 論文集*, 2021(1):388–394, 2021.
- [30] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, pages 4690–4699, 2019.
- [31] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision (ECCV 2015)*, pages 1116–1124, 2015.
- [32] Liang Zheng, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. Mars: A video benchmark for large-scale person re-identification. In *Proceedings of the European Conference on Computer Vision (ECCV 2016)*, pages 868–884. Springer, 2016.
- [33] Chao Chen, Yan Ding, Xuefeng Xie, Shu Zhang, Zhu Wang, and Liang Feng. Trajcompressor: An online map-matching-based trajectory compression framework leveraging vehicle heading direction and change. *IEEE Transactions on Intelligent Transportation Systems*, 21(5):2012–2028, 2019.
- [34] Shaojie Qiao, Nan Han, William Zhu, and Louis Alberto Gutierrez. Traplan: an effective three-in-one trajectory-prediction model in transportation networks. *IEEE Transactions on Intelligent Transportation Systems*, 16(3):1188–1198, 2014.
- [35] Golnaz Habibi and Jonathan P How. Human trajectory prediction using similarity-based multi-model fusion. *IEEE Robotics and Automation Letters*, 6(2):715–722, 2021.
- [36] Yu Fan Chen, Miao Liu, and Jonathan P How. Augmented dictionary learning for motion prediction. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA 2016)*, pages 2527–2534. IEEE, 2016.
- [37] Radu Bogdan Rusu and Steve Cousins. 3D is here: Point Cloud Library (PCL). In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2011)*, Shanghai, China, May 9-13 2011. IEEE.
- [38] Yizhak Ben-Shabat, Michael Lindenbaum, and Anath Fischer. 3d point cloud classification and segmentation using 3d modified fisher vector representation for convolutional neural networks. *arXiv preprint arXiv:1711.08241*, 2017.
- [39] William De Vazelhes, CJ Carey, Yuan Tang, Nathalie Vauquier, and Aurélien Bellet. metric-learn: Metric learning algorithms in python. *J. Mach. Learn. Res.*, 21:138–1, 2020.
- [40] Kilian Q Weinberger and Lawrence K Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of machine learning research*, 10(2), 2009.
- [41] Joe H Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.
- [42] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.