

# 3次元LiDARを用いたキャンパスビル全域での人流と行動のセンシングプラットフォーム構築と運用

天野 辰哉<sup>1</sup> 右京 莉規<sup>1</sup> 山口 弘純<sup>1</sup> 守屋 充雄<sup>1,2</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りつつある。人流データは、人々の接触・混雑状況モニタリング、施設内の動線・広告配置の最適化、災害時の避難計画の設計などに利活用される。我々の研究グループでは、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」を開発してきており、2021年度に大阪大学箕面新キャンパスへ68台のLiDARを設置し、現在箕面キャンパス全域での人流計測を実現している。本稿ではキャンパス環境へのシステムの構築及び運用について説明するとともに、キャンパス実環境におけるトラッキング性能を評価する。またビル全体での人流計測システムおよび取得される大規模3次元点群や人流データの活用・諸課題について述べる。

## 1. はじめに

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。COVID-19感染拡大防止のため、ビルや施設の事業者・管理者は人流計測による対資材者数の把握や密検知を行う必要性に迫られている。感染防止のみならず、滞在者数を常時把握しておくことは、地下街などの構内において突発的な火災やゲリラ豪雨等による水害の危険性に対してでも老率的な避難誘導につなげることができる。また建築設備設計やスマートビルディングにおける空調サービス最適化などにも活用できるため利点は大きい。屋内・屋外に設置する様々なセンサから人流検知や行動センシングを行う手法が数多く提案されているが、ビルの内部や教室などあらゆる場所において高精度に行動に認識するセンシング技術とそれらとAIやビッグデータ技術の併用を可能にするプラットフォームが求められている。

これに対して我々の研究グループでは3次元測域センサ（LiDAR）を用いた人流解析システム「ひとなび」[1]を開発してきている。赤外線を用いた測距センサであるLiDARを用いることで、RGBカメラによるモニタリングとは異なり、プライバシー配慮型でのセンシングを実現している。「ひとなび」は各LiDARから得られる距離情報を3次元点群に変換し、複数LiDARの点群を合成するとともに、

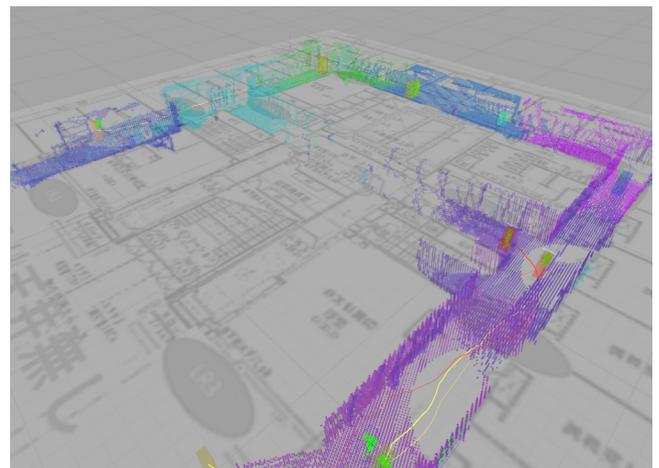


図1 大阪大学箕面新キャンパス5F廊下を捉えた様子。明るい緑の点群が動体、バウンディングボックスが検出された人物、色付きのラインが過去10秒間の人物軌跡を示す。その他の色の点群は各LiDARが捉える非動体の点群であり、LiDAR個体ごとに異なる色で示されている。

人に相当する点群を発見し、それらを核とした高速クラスタリングを行うことでリアルタイムに検出・セグメント化する。得られた点群セグメントをもとに、オクルージョンや近接による点群セグメントの合体を考慮した、カルマンフィルタとセグメント信頼度を組み合わせたトラッキングアルゴリズムを用いて人の点群を軌跡化する。

本稿では、この「ひとなび」を用いたキャンパスビル全域のセンシングプラットフォームの運用とその活用における課題について述べる。我々は学内の人流解析および学

<sup>1</sup> 大阪大学大学院情報科学研究科

<sup>2</sup> 株式会社 HULIX



(a) Hokuyo YVT-35LX (b) Livox Avia

図 2 LiDAR の天井吊り固定

生の行動や学生間交流の計測などを目的とし、大阪大学箕面新キャンパスに複数の LiDAR を設置している。箕面新キャンパスは 10 階建てのビルであり、6 階までが学生の利用するキャンパスとして利用されており、この 1～6 階までに現在までに 68 台設置し、3 次元点群による人流と行動データをキャンパス内の広範囲で取得可能にしている。本プラットフォームでは水平視野角 210 度、垂直視野角 40 度程度の広角 LiDAR と水平・垂直視野角がともに 70 度程度で最大検知距離約 500m の狭角長距離 LiDAR を複数組み合わせ配置し、それぞれの LiDAR からの点群ストリームを 3 台のサーバで処理することで、キャンパス内の学生が利用するほぼすべての廊下および複数の教室内でのセンシングを実現する。例えば図 1 に示すように、教室から教室への移動パターンや学生ラウンジの利用状況、エレベーター前の待ち時間や食堂の混雑具合などをリアルタイムにプライバシーの大きな侵害なく取得可能である。

設置 LiDAR 全体から毎秒約 700MB の頻度で生成される点群ストリームを処理し、リアルタイムに人物軌跡を出力するシステムを設計・実装したとともに現在までに約 3 か月間運用した。また実際の環境にけるトラッキングの精度を評価した結果、全体としてのトラッキングの性能を示す MT 指標は 67.3%、ML 指標は 24.0%であった

以降の章構成は以下の通りである。2 章では歩行者トラッキング・センシングに関連する研究について述べる。3 章では本プラットフォームシステムの設計および実装について、4 章では Multi-Object Tracking アルゴリズムについて説明する。5 章ではシステム評価およびトラッキング性能の評価を、6 章ではプラットフォームの課題について述べる。

## 2. 関連研究

複数の歩行者トラッキングを行っている研究には、3 次元点群ベースの手法 [2] や Deep SORT[3] などがあげられる。文献 [2] は、3 次元点群を入力とし、SECOND[4] とよばれる深層ニューラルネットワークを使用して人物検出を行うとともに、検出した人物セグメントに対して 3D カルマンフィルタとハンガリアンアルゴリズムを使ってトラッキングを行っている。同手法では KITTI データセットに対して評価実験を行っており、オクルージョンした物体

が再出現した時にトラッキングを継続できることを示している。しかし同手法では、観測点群が不完全である場合は考慮されていないため、本稿の手法が想定するような単一のセンサによるセンシングに適用することは難しい。我々の研究グループでも、複数の LiDAR を連結して用いることで、広範囲において人の軌跡を高精度で検出する“ひとなび”システムを開発しており、大型ショッピングモールなどへの設置実験などを実現している [5] が、2 次元データを対象としており、また不完全な点群に対しては堅牢でない。画像に対する最近のトラッキング手法としては Deep SORT[3] がよく知られている。Deep SORT はその前身である SORT[6] の欠点である、オクルージョン後の再出現時に元の ID でトラッキングを行えないといった問題を深層学習を利用して解決している。この手法では、RGB 画像から YOLOv3 を用いて人物検出を行ったものを入力としており、カルマンフィルタを使って前後のフレームで大きさと動きの近いバウンディングボックスを対応させてトラッキングし、KITTI データセットに対して高精度なトラッキングを実現している。

個々の歩行者の観測からの移動推定は Multiple Object Tracking (MOT) とよばれており、カメラ映像を用いたトラッキング技術チャレンジやデータセット共有なども盛んである [7] が、視野が限られる環境における 3 次元点群データを対象とした試みはほとんどみられていない。なお、既存のトラッキング用のデータセットには MOTChallenge[7] や KITTI[8] が存在する。MOTChallenge[7] は歩行者検出用のデータセットであるが、入力が RGB 画像である。そのため、3 次元点群でのトラッキングに利用することはできない。KITTI[8] 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。しかし、通常の道路や大学構内などで取得されたデータにより構成されており、スーツケースやベビーカーを持つ歩行者のデータは含まれない。また背景が一定でないため、固定センサでのデータ処理や評価には向いていない。

## 3. システム設計と実装

### 3.1 システムの物理的構成

1～6 階キャンパス全域からの 68 台の LiDAR 点群データをリアルタイムに処理するため、1・2・5 階、3 階、4・5 階のエリアに分けてそれぞれ独立して 1 台のサーバで人流解析処理を行う。キャンパスビル 1 階・2 階・5 階のセンサデータを収集する部分の物理的なシステムの構成を図 3 に示す。使用した LiDAR は Livox 社の Avia および北陽電機社の YVT-35LX である (図 2)。両 LiDAR ともに動作音を抑える防止ケース内に固定し、天井吊りによりビルに設置している。

各 LiDAR の仕様を表 1 に示す。Livox Avia は狭角長

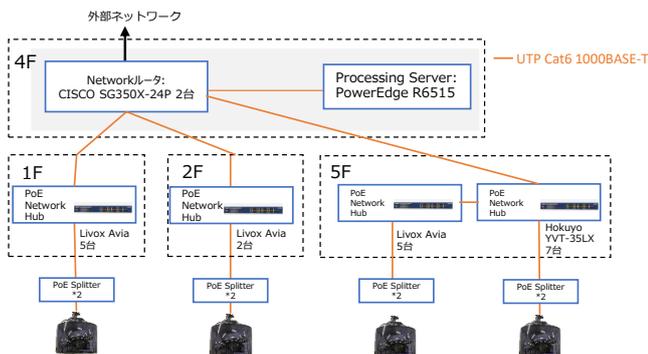


図3 ハードウェア構成 (1階・2階ホールおよび5階)

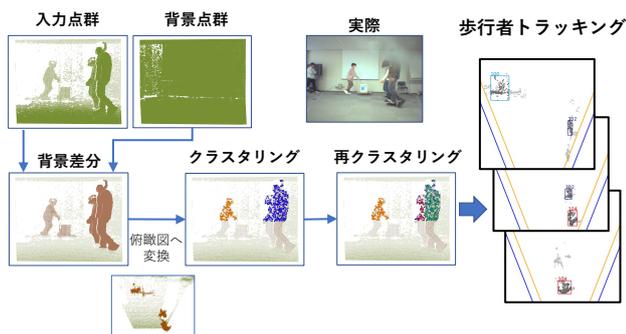


図4 歩行者トラッキングのための点群処理フロー

距離を計測可能であり高密度な点群を取得可能であり、Hokuyo YVT-35LX は点群密度は疎であり、また最大検知距離は、210 度という非常に広い水平視野角を持つ。この特性を考慮し、歩行者の移動を途切れなく検出する必要がある比較的細長い空間 (主に廊下) には YVT-35LX を、空間内での行動検出を目的とする空間 (教室や交流スペースなど) には Avia を設置している。各 LiDAR への電力は PoE により供給する。Livox Avia は最大で 30W の電力を消費するため、これとネットワーク容量を考慮し複数台の PoE ハブを設置している。

システムの中核ウェアとして ROS (Robot Operating System) を採用し、以降のすべてのデータ処理は ROS ノードとして実装し、点群データフォーマットとしては ROS の PointCloud2 形式を利用している。

### 3.2 点群データ処理フロー

各 LiDAR から得られる点群データの処理フローを図に示す。提案手法では、LiDAR から得られる測位距離データを入力とし、データフレーム毎に、3次元空間の座標系上の点群に変換、事前に取得した背景点群との差分をとることで、移動物体群に相当する3次元点群データを抽出する。抽出した3次元点群データから俯瞰視点の2次元点群データに変換しクラスタリングを適用することで、ノイズ除去と歩行者のセグメント化を行う。このフレーム単位で実行する背景差分および移動体抽出については 3.2.1 章

で、クラスタリングについては 3.2.3 章で詳細を述べる。

各フレームごとに得られる歩行者セグメントに対して、以前のフレームで検出された歩行者セグメントと同一人物であるかどうかの対応付けを、セグメント同士の距離や大きさに基づいて決定する。またこの連続フレーム間での対応情報をカルマンフィルタに入力することで、次フレームにおける各歩行者セグメントの位置を予測する。この繰り返しにより歩行者セグメントのトラッキングを実現する。あるセグメントの予測値に対して、割り当てられるセグメントの観測値が存在しない場合、以前のフレームからその歩行者の移動方向および移動速度を予測し推測を行うことで補う。

#### 3.2.1 背景差分による前景点群の抽出

システムではまず背景差分法により移動物体による点群を含んだ前景点群を抽出する。この処理は各 LiDAR 単位で実施する。そのための背景点群として、事前に各センサごとに捕捉可能な領域に移動物体が存在しない時の点群を取得しておいたものを利用する。移動物体が LiDAR により捕捉可能な領域に進入すると、取得された点群と背景点群の間に差分が生じる。その差分の点群を抽出することにより、移動物体の点群 (前景点群) を取得する。

#### 3.2.2 複数センサからの前景点群の統合

複数センサを用いる場合、全センサからの前景点群を統合し、一つの全体の前景点群を生成する。計測する空間の座標系 (絶対座標系) 内での各センサの3次元位置と向きを基にそれぞれの前景点群の座標を絶対座標系に変換し統合する。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、同一物体の同一領域に対して重複する点群データが生じるため、センサ視野が重複しない領域と比較し、得られる前景点群の密度が高くなる。これは後述する密度ベースの点群クラスタリングに影響を与える。そのため Voxel Grid Filter [9] 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。ボクセルグリッドのサイズは使用するセンサの計測距離精度から決定する。

#### 3.2.3 クラスタリングによる歩行者セグメントの抽出

統合された前景点群に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短縮するために、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。公共空間内に存在する移動物体は2次元平面を移動する人およびその人が所持、利用しているものであるため、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点群密度に基づくクラスタリングである DBSCAN を適用し、移動物体セグメントを得る。本システムではこのようにして得られた移動物体のセグメントを歩行者セグメントであるとする。各セグメン

表 1 3次元LiDAR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

トは空間座標の軸に沿ったバウンディングボックス (セグメント内の各点の座標の各軸ごとの最小値および最大値の組み合わせ) で表現する。

#### 4. トラッキングアルゴリズム

歩行者の3次元点群セグメントに基づいて複数歩行者のトラッキングを行う手法について説明する。検出された歩行者セグメントを観測とし、観測されトラッキングの対象となった歩行者を歩行者オブジェクトと呼ぶ。提案手法では歩行者オブジェクトの状態をカルマンフィルタにより補正・更新・予測しトラッキングする。各オブジェクトはカルマンフィルタの状態として、オブジェクトの位置座標の  $xy$  座標成分、およびオブジェクトの速度ベクトルのうち  $xy$  軸と並行なベクトル成分からなる4項組を用いる。

ここでは  $k$  番目のフレームで検出された観測 (歩行者) セグメントの集合を  $O_k = \{o_{k,j}\}$  とする。  $j$  は観測されたセグメントの番号を示す。また  $k$  番目のフレームの処理終了後のトラッキング中のオブジェクトの集合を  $T_k = \{t_{k,i}\}$  とする。  $i$  はトラッキング中のオブジェクトの番号を示す。  $T_k$  中のオブジェクト  $t_{k,i}$  はカルマンフィルタにより補正されたものである。以降の説明において、  $k$  番目のフレームの処理中に  $T_k$  に新たに追加されたオブジェクトの速度は  $xy$  方向のどちらも 1.0 とする。  $T_k$  の各オブジェクト  $t_{k,i}$  から予測した  $k+1$  番目のフレームでのそのオブジェクトの予測値を  $p_{k+1,i}$  とし、その集合を  $P_{k+1}$  とする。  $T_0$  と  $P_1$ , フレーム処理開始時の  $P_k$ ,  $T_k$  はすべて空集合とし、  $k \geq 1$  であるとする。

##### 4.1 同一人物判定

$k$  番目のフレームが得られた際の処理について述べる。フレーム処理開始時にあらかじめ  $T_{k-1}$  をもとに現在のフレームでのオブジェクトの予測値  $P_k$  を求めておく。  $P_k$  が空の場合、  $T_k \leftarrow S_k$ , つまり現フレームで観測されたセグメントすべてをトラッキングの対象としてこのフレームの処理を終了する。

$T_{k-1}$  が空でない場合、以下の3つの処理を行う。

- (1) 現フレームのセグメント集合と前フレームでのオブジェクト集合の対応付けを、距離と体積に基づくコストを最小化するように決定する。

- (2) (1) で対応付けが与えられなかったオブジェクトおよびセグメントに対し、その原因を推定し、それに応じた対応付けを決定する。

- (3) (1) と (2) で決定された対応付けの情報を用い、カルマンゲインと状態更新 (カルマンフィルタ処理) を行う。

(1) の対応付けについては、与えられたセグメントとオブジェクトの組に対し、セグメントに含まれる点をすべて内包し、各辺が  $xyz$  軸のいずれかと平行な直方体 (バウンディングボックスとよぶ) の体積、およびオブジェクトとセグメントの組から計算される「コスト」の和が最小になるように求める。ここで、現フレーム  $k$  のセグメント数 ( $|O^k|$  とする) とフレーム  $k-1$  のオブジェクト数 ( $|T^{k-1}|$  とする) に対し、以下の条件を満足する対応付けを発見する。

- どのセグメントも高々1つのオブジェクトに対応付けられる
- どのオブジェクトも高々1つのセグメントに対応付けられる
- $\min\{|O^k|, |T^{k-1}|\}$  だけの対応が存在する
- 対応付けのコスト総和が最小である

これにより、オブジェクトあるいはセグメントの数の少ないほうがすべて対応付けを持つようにする。

(2) では、(3) でのカルマンフィルタの更新方法を決定するために、対応付けの存在しないオブジェクトあるいはセグメントが生じた原因を推定する。オブジェクトやセグメントはオクルージョン (遮蔽) や領域外への移動による消失や、領域内への移動による出現が発生するが、それらの発生原因の領域内の位置に応じた推定を行うことで (3) のカルマンフィルタの精度向上を図ることを目的とする。

最後に (3) では、得られた対応付けに基づき、カルマンフィルタを用いてフレーム  $k$  でのオブジェクトを予測する。カルマンフィルタの状態として、オブジェクトの位置座標の  $xy$  座標成分、およびオブジェクトの速度ベクトルのうち  $xy$  軸と並行なベクトル成分からなる4項組を用いる。また観測は、現フレームのセグメントを内包するバウンディングボックスの中心座標成分とセグメントの3辺の大きさからなる6項組で表す。これらを用い、例えばあるオブジェクトに対し対応するセグメントが見つからない場合でも、それがオクルージョンによるセグメント消失と想定される場合にはトラッキングを継続する。(2) での対応づけが存在しないオブジェクトについては、観測値なしでカルマンフィルタを更新する。  $T_k$  での速度でセグメントが移動したと仮定しカルマンフィルタを更新する。対応付けが存在するオブジェクトについてはカルマンフィルタを用いて現在の位置を更新する。

- (1), (2) の処理について、以下の4.2節から4.3節で詳細を述べる。

## 4.2 オブジェクトとセグメントの対応付け

前節で述べたように、前フレームでのオブジェクトと、現フレームでのセグメントの各組に対し、それらの距離をその組の「コスト」とし、前節で述べたような条件を満たしながらコストを最小とする対応付けを発見する。オブジェクト  $i$  とセグメント  $j$  に対するコスト  $c_{i,j}$  は以下で定義される..

$$c_{i,j} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (1)$$

$x_j, y_j$  はそれぞれ観測セグメント  $o_{k,j}$  の  $xy$  座標を表し、 $x_i, y_i$  はそれぞれ歩行者セグメントの予測値  $p_{k,i}$  の  $xy$  座標を表す。このコスト  $c_{i,j}$  が最小となる組み合わせから順に割り当てを決定する。なお、決定された組み合わせのコストが閾値（経験的に 1.0 を利用）を超える場合、その割り当てを棄却する。

割り当てられた、観測と  $P_k$  中の歩行者の組については、その観測セグメントとカルマンフィルタを用いて、対応する  $T_{k-1}$  中の歩行者の現在位置を更新し、 $T_k$  に追加する。

## 4.3 対応付けされなかったオブジェクトあるいはセグメントの処理

前節の方針に基づいて決定された対応付けに含まれないオブジェクトあるいはセグメントが存在する場合、本節の追加対応付け手法を適用し、それらの対応付けを発見する。同手法は、対応付けがなされなかった原因を推定し、その理由に応じ、最も適切と思われる対応付けを提供する。

あるセグメントが対応するオブジェクトを持たなかった原因として、本研究では、以下の3つであると考える。

- (S1) 新しい歩行者が領域に進入したためにセグメントを得るが、その前フレームに（それと対応付け可能な）オブジェクトが存在しない。
- (S2) オクルージョンの影響で、1人の歩行者に対し  $l$  個に分裂したセグメントを得るが、その前フレームに、それらに対応するオブジェクトは1つしか存在しない（すなわち残りの  $l-1$  個のセグメントに対する対応付けが得られない）。
- (S3) オクルージョンの影響で観測されていなかった歩行者が再度観測されたためにセグメントを得るが、その前フレームに（それと対応付け可能な）オブジェクトが存在しない。

また、あるオブジェクトが対応するセグメントを持たなかった原因として、本研究では以下の2つであると考える。

- (O1) 前フレームでオブジェクトとして認識されていた歩行者が領域外に退出し、対応するセグメントを失った結果、対応付けが得られない。
- (O2) 複数人が接近して通行したことにより、2人以上の歩行者に対し1つのセグメントしか得られない。

これに対し、(S1) および (S3) の場合は、新たに歩行者が

出現したとみなしてよいいため、次フレーム以降でカルマンフィルタによる位置予測を開始すればよい。したがって、このセグメントに対するカルマンフィルタは適用せず、これを現フレームでのカルマンの状態（オブジェクト）として追加する。(S2) の場合は、前フレームで1つのオブジェクト  $i$  として認識されていたものが、現フレームではオクルージョンにより  $l$  個のセグメントとして観測された場合（観測が「分裂」した場合）である。これが理由と想定される場合、このオブジェクト  $i$  のトラッキングをこの分裂発生時で終了し、以降は分裂により発生したセグメントをそれぞれトラッキングする。

(O1) の場合は、領域外へ歩行者が退出したとみなしてよいいため、このセグメントに対するカルマンフィルタは適用しない。また、 $F_{thr}$  フレーム以上連続してセグメントとの対応付けがないオブジェクトはトラッキングを終了する。(O2) の場合は、前フレームで  $n$  個のオブジェクト  $i_l$  ( $l = 1, 2, 3, \dots, n$ ) として認識されていたものが、現フレームではオクルージョンにより1個のセグメントとして観測された場合（観測が「合体」した場合）である。これが理由と想定される場合、各オブジェクト  $i_l$  の予測値  $p_{k,i_l}$  と観測  $o_{k,j}$  の重なり合う領域を仮想セグメントとして生成し、この仮想セグメントを  $i_l$  のセグメントとして対応付ける。

## 5. 評価

2022年1月20日11時から同年1月24日13時までに箕面キャンパス5Fで評価用のデータを収集し、合計15,101個の軌跡データを得た。日別の検出軌跡数を表2に示す。土曜日は学生が少なく軌跡の数も少なかった。1月23日は日曜日はキャンパス内で工事が行われていたために、工事業者の出入りによって多数の軌跡が検出されている。このうち平日である1月20日の10時~11時台のある2分間におけるトラッキングの精度をMulti-Object Trackingで用いられるCLEAR指標により評価した。用いた指標を以下に示す。また結果を表3に示す。全体としてのトラッキングの性能を示すMT指標は67.3%であった。一方MLは24.0%であった。トラッキングの失敗のほとんどはFN指標が捉えるものであり、密に並んで歩く歩行者が、観測されている全フレームに渡って、一人の歩行者であると認識されていたことが原因である。また計測領域から歩行者がでるタイミングで同じ個所から別の歩行者が侵入してきた場合に、出る歩行者が引き返して戻ってきたと追跡が継続する失敗も見られた。

- MT: 全歩行者のうち、80%以上の期間で正しくトラッキングを行えた歩行者の割合を表す。
- ML: 全歩行者のうち、50%以下の期間でしか正しくトラッキングを行えなかった歩行者の割合を表す。
- FP: 全歩行者・全期間を通じた誤ID割当て数を表す。
- FN: 全歩行者・全期間を通じた見逃し数を表す。

表 2 1月20日～24日にかけて5Fで検出された軌跡の数

日付	軌跡の数
1月20日(木)	3,691
1月21日(金)	3,901
1月22日(土)	437
1月23日(日)	5,864
1月24日(月)	1,208
合計	15,101

表 3 CLEAR 指標による MOT 評価結果

MT(%)	ML(%)	FP	FN	IDsw	Frag
67.3	24.0	2	75	6	4

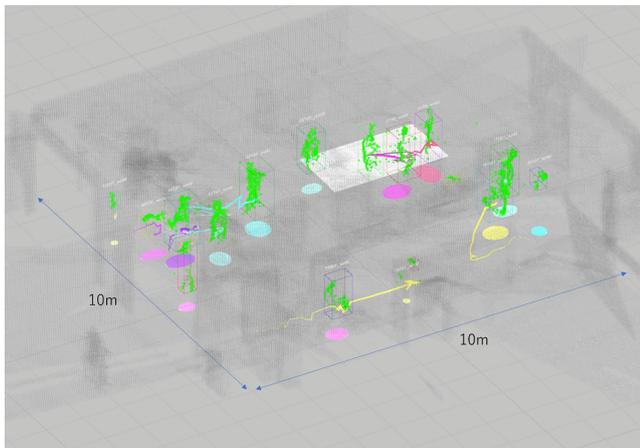


図 5 ポスターセッションを行った際の様子。緑色の点群が動体、灰色の点群が背景点群を示す。

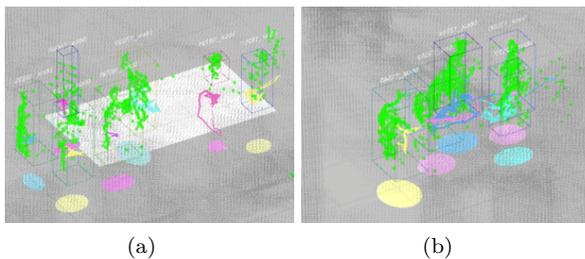


図 6 過密環境下でのトラッキング

- IDsw: 全歩行者・全期間を通じた ID 割当てにおける ID 入替り数を表す。
- Frag: 全歩行者・全期間を通じ、FN に相当する見逃しにより軌跡が分割された数を表す。

図 6 に示すような人が密集する過密環境下においても、提案トラッキングアルゴリズムにより適切に人物を識別し追跡できていることが確認できた。例えば図 6(a) 右側のピンク色で軌跡が表示されている歩行者は、オクルージョンにより下半身の点群が計測フレーム上にはほぼ欠損しているものの適切にマッチングが行われている。

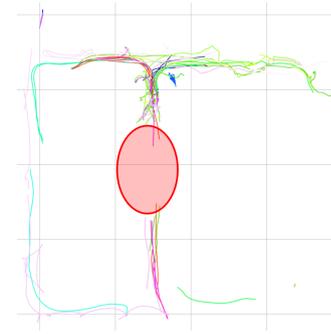


図 7 5階におけるある10分間の軌跡を2次元可視化したもの。同一人物による軌跡は同一色で表示される。中央にセンサの不可視領域が存在し、軌跡が途切れている。

## 6. 課題

### 6.1 超過密環境下におけるトラッキング精度低下

提案アルゴリズムでは複数の歩行者による点群セグメントの合体に対して前後フレームの状態をもとに検出し補正する。そのため、計測範囲に侵入した段階から合体状態の歩行者（例えば密に並列で廊下を歩く学生）に対しては、適切にトラッキングできない場合があった。これに対して、現在歩行者セグメントの点群に対して深層学習を用いて人数を検出する手法を検討しており、これを提案するトラッキング手法に組み込む予定である。

### 6.2 不可視領域の存在

キャンパスビル広域をカバーするように LiDAR を設置したものの、建物の構造および意匠上物理的なセンサ設置が不可能な領域が存在する。例えば建物の防護区画周辺などでは物理的な配線作業が禁止されている場合がある。そのような結果として図 7 に示すようにセンシングデータの不可視領域が発生し、同一人物であっても軌跡が途切れるという問題が生じている。

これに対して、我々がすでに提案している点群特徴量抽出手法 [10] と深層距離学習、空間内の移動軌跡解析を組み合わせた人物 Re-ID 手法を現在設計中であり、断片化された軌跡セグメントからの再構成を行うことで対処する予定である。

### 6.3 点群データの保存・圧縮

行動解析のための学習データとして本システムにより得られる点群データの利用を検討しているものの、出力される点群容量が膨大であり、キャンパスビル全域での点群の保存が困難である。Livox Avia の場合、1台でファイルシステム上の実測で最大 10MB/s のデータが保存される。1・2・5階に限定したとしても、最大 200MB/s 程度のデータストリームが生じる。サーバで使用しているハードディスクドライブの連続書き込み速度の仕様値が 250MB/s であ

り、システムによる読み出し処理も定期的に並列処理されることを踏まえると単体のハードディスクストレージへの常時書き出しは困難であった。また容量の点でも上記速度で生成される点群を保存すると無圧縮ではたった24時間で約20TBに相当する。RAIDなどによりストレージ容量を活かした書き込み速度の向上が現状必須であるものの生成される容量そのものも大きいいため効率的な点群圧縮が必要となる。

## 7. まとめ

本稿では大阪大学箕面キャンパスで現在運用している「ひとなび」を用いたキャンパス全域での人流および行動センシングプラットフォームについて述べた。本プラットフォームでは68台の3次元測域センサ(LiDAR)を用いて大規模な3次元点群を取得するとともに、そこに含まれる人物を検出し、トラッキングを行う。トラッキング手法として、点群の合体や分裂を考慮したカルマンフィルタベースのアルゴリズムを提案した。

今後は本プラットフォームを用いてプライバシー配慮型での人物行動・属性収集手法の開発を進めるとともに、スマートフォンユーザのジェスチャを用いて、3次元点群から検出された人物をスマートフォンで認証・紐づけし、許可を得たユーザからの個人データの活用を進める予定である。

## 参考文献

- [1] R. Ukyoh, T. Amano, A. Hiromori, and H. Yamaguchi. Pedestrian tracking in public passageway by single 3d depth sensor. In *Proceedings of the 2022 IEEE International Workshop on Pervasive Computing for Vehicular Systems Co-located with IEEE PerCom 2022 (PerVehicle 2022)*, 2022.
- [2] B. Sahin, D. Wang, C. Huang, Y. Wang, Y. Deng, and H. Li. A 3D Multiobject Tracking Algorithm of Point Cloud Based on Deep Learning. *Mathematical Problems in Engineering*, Vol. 2020, p. 8895696, 2020.
- [3] N. Wojke, A. Bewley, and D. Paulus. Simple Online and Realtime Tracking with a Deep Association Metric. In *2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3645–3649, 2017.
- [4] Y. Yan, Y. Mao, and B. Li. SECOND: Sparsely Embedded Convolutional Detection. *Sensors*, Vol. 18, No. 10, 2018.
- [5] H. Yamaguchi, A. Hiromori, and T. 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*, Workshops ICDCN '18, pp. 13:1–13:6, New York, NY, USA, 2018. ACM.
- [6] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft. Simple Online and Realtime Tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3464–3468, 2016.
- [7] P. Dendorfer, H. Rezatofighi, 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.
- [8] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous Driving? The KITTI Vision Benchmark Suite. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3354–3361, 2012.
  - [9] Radu Bogdan Rusu and Steve Cousins. 3d is here: Point cloud library (pcl). In *2011 IEEE International Conference on Robotics and Automation*, pp. 1–4, 2011.
  - [10] S. Yamada, R. Hamada, and H. Yamaguchi. An accurate point cloud-based human identification using micro-size lidar. In *Proceedings of the 2022 IEEE International Workshop on Pervasive Information Flow Co-located with IEEE PerCom 2022 (PerCom 2022) Demo*, 2022.