

人物追従ロボットのための歩行者経路予測における機械学習用データ構成法

赤羽根 里奈¹ 加藤 由花²

概要: サービスロボットの普及に伴い, 家庭や職場, 公共空間などで人と共存するロボットに対する期待が高まっている. 我々はこれまで, そのようなサービスロボットの一種である人物自動追従ロボットを対象に, ターゲットとなる人の未来の経路を予測することで, 人物追従の精度を向上させる手法に関する研究を進めてきた. ここでは, 人の移動軌跡 (位置座標の時系列) を系列データとみなし, 機械学習モデルにより, 入力系列から出力系列を予測する手法を提案してきた. しかし, 個別の予測結果を見たときに, 十分な精度が得られないという問題が残されていた. 本稿では, この手法を拡張し, これまで人の位置座標から構成していた系列データを, 速度ベクトルにより構成する系列データへと変更する. この変更により, 追従ターゲットの移動傾向を予測モデルに取り組みることが可能になり, 特定の追跡対象に対する経路予測の精度向上が期待できる. 本稿では, 人移動軌跡データセットを利用し, 提案手法と位置座標を用いた手法の経路予測精度を比較する. その結果から, 平均的な予測精度に大きな違いは無いものの, 速度ベクトル系列を用いた手法の方が, 特定ターゲットに対して真の軌跡と近い経路を予測できることを示す. また, 位置座標と速度ベクトルの両方を考慮することで, 群衆内における特定ターゲットの経路予測の精度が向上する可能性についても議論する.

1. はじめに

近年, サービスロボットの普及に伴い, 家庭や職場, 公共施設などで人と共存するロボットに対する期待が高まっている. これらサービスロボットの一種に, 人物自動追従ロボットがある. これは, ターゲットとなる人を追従し続けるロボットであり, 重い荷物の運搬, 屋内外の施設案内, 医療現場や工場での作業補助等, 様々な場面での利用が期待されている. 人物追従には通常, LiDAR 等によるセンシングデータが用いられるが, 障害物による遮蔽等が発生した場合に, ロボットが人を見失ってしまうケースが多発している [1]. この問題を解決するための一手段として, 我々はこれまで, センサーで計測された過去の人の移動軌跡を基に, 未来の移動経路を予測することで, その結果をロボットの追従制御に利用する手法の研究を進めてきた [2]. 手法のイメージを 図 1 に示す. ここでは, 人の移動軌跡を系列データとみなし, 機械学習モデルにより, 入力系列から出力系列を予測する問題として定式化を行った.

機械学習モデルとしては, RNN (Recurrent Neural Network) ベースの手法である Social LSTM [3] を採用し, 特徴量は, 人の位置座標 (ターゲットとなる人の初期位置を

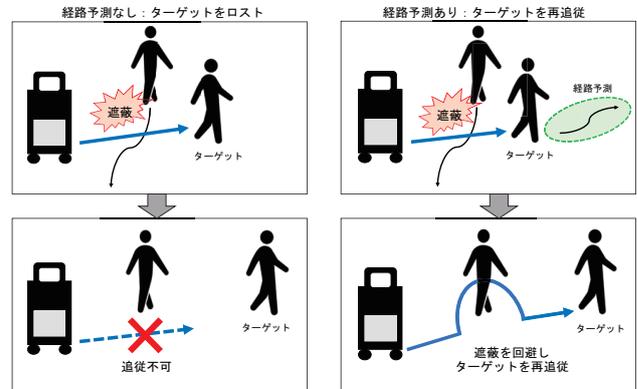


図 1 提案手法の利用イメージ. ターゲットとなる人の過去の移動軌跡データから, 未来の移動経路を予測し, その結果を人物自動追従ロボットの追従制御に用いる.

基準とした相対座標) の時間変化を系列データとして構成した. Social LSTM は, 同じ空間に存在する人と人の相互作用をモデルに内包することができる点に特徴があり, 位置座標から人と人の相対距離を把握してモデル構築に利用している. ただし, 群衆内での平均的な経路予測結果の向上を目的としているため, 個別の予測結果を見たときには十分な精度が得られない事例が見受けられた. 一方, 予測結果を追従ロボットで利用することを考えた場合には, 追従対象は固定されるため, ターゲットの移動傾向やその他の特徴を何らかの形でモデルに取り込むことで, 推定精度

¹ 東京女子大学 大学院理学研究科
Suginami, Tokyo 167-8585, Japan
² 東京女子大学 現代教養学部 数理科学科

の向上が期待できる。

このような仮説の下、本稿では、ターゲットの特徴の一つとして歩行速度に着目し、歩行速度ベクトルの系列を入力とする経路予測モデルを提案する。考慮すべき特徴には、ターゲット自体のその他の特徴（歩容、年齢など）、環境側の特徴（空間形状など）も考えられるが、本稿ではまず、歩行速度についての評価を行う。

本稿の貢献は以下の2点である。

- 速度ベクトルの系列を入力とする LSTM ベースの経路予測モデルを提案する。
- 既存のデータセットを利用し、位置座標系列と速度系列を用いたモデルの性能を比較する。

2. 関連研究

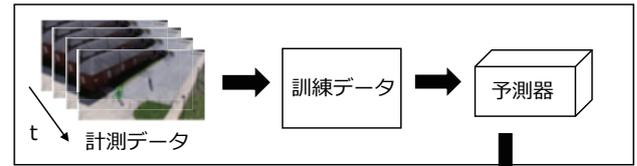
歩行者経路予測（歩行者の過去の移動軌跡から未来の移動経路を予測する手法）は、近年、特にビジョン研究の分野で活発に研究が行われている課題である [4]。本章では、それら経路予測に関する研究、およびロボット人追従手法に関する研究について概観する。

2.1 歩行者経路予測に関する研究

歩行者経路予測については、まず、古くから研究が行われてきたベイズフィルターに基づく一連の研究が存在する。これらの手法では、ターゲットのトラッキングを目的に、短期的な経路予測が行われる。例えば、観測された歩行速度を利用し、近接時の人と人の相互作用を移動モデルに組み入れることで拡張カルマンフィルターを構成する手法 [5]、複数の手がかり（Appearance, Motion, Interaction）に基づき、LSTM (Long Short-Term Memory) [6] によりそれぞれ独立に構築された予測器を組み合わせ、次周期の複数ターゲットの位置を予測する手法 [7]、マルコフ決定過程を用いてターゲットの状態（Active, Inactive, Tracked, Lost）を推定し、状態ごとに処理内容を定義する手法 [8]、CNN (Convolutional Neural Network) に基づく歩行者検出と組み合わせることで、リアルタイムで追跡候補を予測する手法 [9] などがある。これらの手法は、ターゲット再検出のための様々な工夫がなされているが、いずれも、ある程度以上の長期間の予測は困難である。

そのため、近年では、RNN ベースの手法を用いて予測器を構築し、系列データとして経路予測を行う手法も提案されている。各自の移動経路（系列データ）を LSTM でモデル化し、隠れ層の状態を共有することで人と人との相互作用を表現する Social LSTM [3] や、これを GAN (Generative Adversarial Networks) を用いた手法に拡張し、複数人の経路を同時に予測（生成）可能にした Social GAN [10] などがある。さらに、経路予測チャレンジとして、Trajectory

●学習フェーズ



●予測フェーズ



図 2 経路予測の流れ。予測器は、訓練データを用いて事前に構築しておく。この予測器と、追従ロボットが計測したセンシングデータを用いて、歩行者の未来の経路を予測する。

Forecasting Challenge*¹なども行われている。

本稿では、これらの系列データ予測手法を利用し、未来の移動経路を予測する。

2.2 ロボット人追従手法に関する研究

ロボットによる人追従手法としては、LiDAR によるセンシング結果を利用する手法がよく用いられる。精度が高く、日照など環境の影響を受けにくいこと、単体でターゲットの大きさ、位置、移動方向の推定が可能であることから、様々な環境で利用されている。カメラを用いる手法と異なり、視覚情報の取得（一般物体認識等）はできないが、様々な工夫により、ターゲットを検出した上で追従を実現している。例えば、検出対象（人の足の見え方）のテンプレートを用意しておき、テンプレートマッチングにより位置を推定する手法などが提案されている [11]。追従にはベイズフィルター（Particle Filter）が用いられている。

本稿でも、ロボットの人物追従にはこれらの手法を用いるが、前述の系列データ予測手法と組み合わせることにより、追従の精度向上を目指す。

3. 予測モデル

3.1 モデルの概要

本稿では、博物館等の公共空間に設置され、人に追従しながら案内を行うロボットを対象に、追従精度を向上させるための歩行者経路予測モデルを提案する。ここでは、あらかじめ対象環境で計測しておいたデータを用いて予測器を構築しておき、追従ロボットに搭載された LiDAR によるセンシング結果を用いて、歩行者の未来の経路を予測する。経路予測の流れを 図 2 に示す。

予測手法は、予測器を構築する学習フェーズと、経路を予測する予測フェーズの2つのフェーズに分かれており、以下の手順に従って予測を行う。

*1 <http://trajnet.stanford.edu>

- 学習フェーズ
 - (1) 計測データから訓練データセットを生成する。
 - (2) 訓練データを用いて予測器を構築する。
- 予測フェーズ
 - (1) 移動軌跡を計測する。
 - (2) 計測データから予測器への入力データを生成する。
 - (3) 移動経路を予測する。

3.2 学習フェーズ

3.2.1 訓練データセットの生成

提案手法における訓練データセットは、以下の情報を1レコードとし、制御周期ごとに観測されたユーザID分のデータとして生成する。

- フレーム番号 (タイムスタンプに相当)
- ユーザID (1トラッキングごとに割り当てられるID)
- 測定時刻における移動速度 v [m/s]

ここでは、歩行者軌跡を系列データとして扱うため、時間軸方向に離散化した一単位としてフレームを用いる。つまり、1フレームを1単位時間とみなす。

通常、計測により人の位置追跡を行うことで、トラッキング対象の位置座標 $\mathbf{x} \in \mathbb{R}^2$ が得られるので (本稿ではこれを2次元座標とする)、この値を用いて上記の v を算出する。具体的には、ある時刻 m (この場合はフレーム番号) における、ある歩行者 i の移動速度 $\mathbf{v}_m^i = (vx_m^i, vy_m^i)$ を、位置座標 $\mathbf{x}_m^i = (x_m^i, y_m^i)$ を用いて、

$$vx_m^i = (x_{m+1}^i - x_m^i) / \Delta t \quad (1)$$

$$vy_m^i = (y_{m+1}^i - y_m^i) / \Delta t \quad (2)$$

と算出する。ここで Δt は1フレームに対応する単位時間であり、座標系としては、対象環境に固定されたワールド座標系を用いる。

訓練データは、後述する予測フェーズの節で説明するデータ計測手法を用いて、対象環境において取得することを想定する。ただし、何らかの手段で位置座標の時系列データが得られているのであれば、ドメイン適応により、他の環境で計測・観測されたデータや、既存のデータセットに含まれるデータを用いることも可能である。

3.2.2 予測器の構築

予測器の構築には、前述した Social LSTM を用いる。Social LSTM では、定められたある領域内に存在する全ての歩行者が、隠れ層の状態を互いに共有することにより、混雑状態での人と人との相互作用 (お互いに衝突を避けようとする行動等) を予測モデルに組み込んでいる。Social LSTM のイメージを 図 3 に示す。

ここでは、歩行者ごとに1つの LSTM モデルを定義し、近隣の歩行者同士で隠れ層 (Social Pooling 層) を共有することで、お互いの空間情報の共有を実現している。その結果、互いが干渉すると考えられる領域内での相互作用

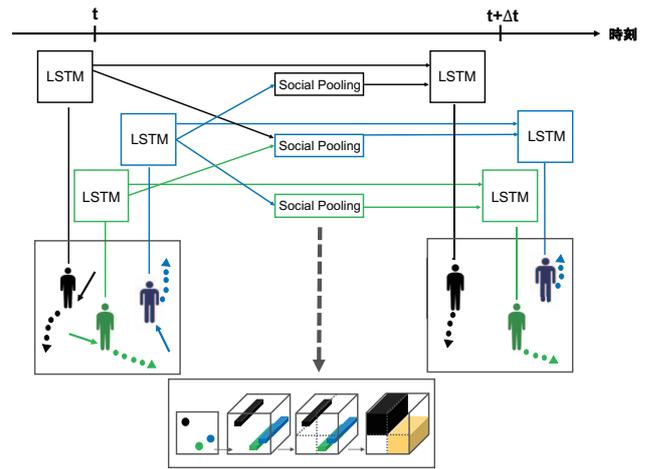


図 3 Social-LSTM のイメージ。歩行者ごとに1つの LSTM モデルを定義し、近隣の歩行者同士で隠れ層を共有する。その結果が次周期の予測結果に反映される。文献 [10] に基づき作成。

用がモデルに組み込まれ、混雑状態であっても比較的精度高く未来の経路を予測可能なモデルを実現している。オリジナルなモデルでは、ターゲットとなる歩行者の初期位置 $\mathbf{x}_j^{(\text{target})}$ を原点とした相対座標として周辺の歩行者の初期位置を補正した上で、周辺の歩行者全ての位置座標の系列データを入力と出力に分割し、予測器に入力することで学習を行っている。具体的には、歩行者 i における訓練データは、 k フレームからなる入力データ列 $\{\mathbf{x}_j^i, \mathbf{x}_{j+1}^i, \dots, \mathbf{x}_{j+k-1}^i\}$ と、 l フレームからなる出力データ列 (正解データ、予測フェーズではこの系列を予測することになる) $\{\mathbf{x}_{j+k}^i, \mathbf{x}_{j+k+1}^i, \dots, \mathbf{x}_{j+k+l-1}^i\}$ となる。

本稿では、前項で述べたように、この系列データを速度ベクトルに置き換える。つまり、入力データ列として、式 (1), (2) から算出される v を用いて $\{v_j^i, \dots, v_{j+k-1}^i\}$ を、出力データ列として $\{v_{j+k}^i, \dots, v_{j+k+l-1}^i\}$ を用いる。なお、Social Pooling 層における空間情報の共有は、位置座標に基づき行う。

3.3 予測フェーズ

3.3.1 データの計測

予測フェーズでは、LiDAR の計測範囲内に入った人の位置追跡を行い、LiDAR から人までの距離と角度を計測する。人位置追跡には、文献 [12] に記載されたものと同様の方法を用いる (ただし、文献 [12] では人の足首を追跡対象としているが、本稿では胴体を追跡対象とする)。ここでは、追跡対象の位置を確率分布として扱い、追跡対象に応じた遷移モデル (人の移動) と、センサデータに基づく観測モデル (LiDAR により測定される距離) から、再帰的に追跡対象の状態量に関する確率密度関数を推定する。人位置 (追跡器の位置) の確率分布には、Particle Filter を用いる。Particle Filter では、確率分布は標本 (Particle) の集合として定義され、Particle を重みに従ってリサンプ

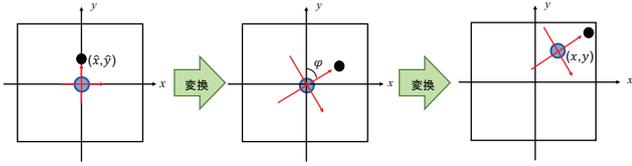


図 4 ローカル座標系からワールド座標系への変換のイメージ。黒丸がターゲットとなる人の位置座標である。ワールド座標系に対するロボットの姿勢 φ により座標軸を回転させ、ロボットの自己位置推定結果により原点を移動する。

リングすることで、事後確率（同じく Particle の集合として定義される）が生成される。

歩行者の位置は、LiDAR により計測される距離と角度を、フレーム周期ごとに極座標形式から直交座標形式に変換することで求める。この結果から、フレームごとに、計測範囲内に存在する歩行者ごとの位置座標 $(\hat{x}_m^i, \hat{y}_m^i)$ が算出される。なお、ここでの座標系は、ロボットに固定されたローカル座標系である。

3.3.2 入力データの生成

前項の手順で得られた計測結果から、予測器への入力データを生成する。予測器には、対象環境に固定されたワールド座標系での移動速度の系列データを入力する必要があるため、この変換処理を行う。具体的には、前項で取得される $(\hat{x}_m^i, \hat{y}_m^i)$ を、 k フレームごとに観測された歩行者の人数分 (i の最大値分) 収集し、ローカル座標からワールド座標に変換した上で、1 歩行者、 k フレームを 1 レコードとして、移動速度のデータ列を生成する。

まず、座標系の変換を行う。ワールド座標系における歩行者 i の m フレーム目の位置座標 (x_m^i, y_m^i) は、ワールド座標系での追従ロボットの位置座標 (X_m^R, Y_m^R) (車軸の中心座標) および姿勢 φ_m (ロボットの正面が向いている方向) を用いて、以下のように計算される。なお、ロボットの自己位置については、他の手法 (環境地図を用いた自己位置推定手法など) を用いて別途算出しておく。

$$\begin{pmatrix} x_m^i \\ y_m^i \end{pmatrix} = \begin{pmatrix} \cos \varphi_m & -\sin \varphi_m \\ \sin \varphi_m & \cos \varphi_m \end{pmatrix} \begin{pmatrix} \hat{x}_m^i \\ \hat{y}_m^i \end{pmatrix} + \begin{pmatrix} X_m^R \\ Y_m^R \end{pmatrix} \quad (3)$$

座標変換のイメージを図 4 に示す。

次に、算出された k フレーム分の (x_m^i, y_m^i) を用いて、歩行速度の系列データ、つまり k フレーム分の (vx_m^i, vy_m^i) を生成する。

3.3.3 移動経路の予測

前項で示した手法により、観測された歩行者分の $\{v_j^i, \dots, v_{j+k-1}^i\}$ を取得し、これを前節で構築した予測器に入力する。その結果、未来の移動速度系列として、 $\{v_{j+k}^i, \dots, v_{j+k+l-1}^i\}$ を得ることができる。ターゲットとなる歩行者の、予測経路の初期位置座標は $x_{j+k}^{(target)}$ であるので、これと以下の関係を用いて、未来の移動経路が系列データとして予測される。



図 5 UCY データセットのシーン例 [13]。市街地の歩行者が鳥瞰視点で撮影されている。比較的シンプルな環境における人移動軌跡データで構成される。

$$x_{m+1}^{(target)} = x_m^{(target)} + v_m^{(target)} \cdot \Delta t. \quad (4)$$

4. 評価実験

本稿では、追従ターゲットの特徴の一つとして歩行速度に着目し、歩行速度ベクトルの系列を入力とする経路予測モデルを提案した。本章では、提案手法の有効性を検証するために、既存のデータセットを利用し、位置座標系列と速度系列を用いた場合のモデルの性能を比較する。

4.1 実験の方法

4.1.1 データセット

実験には、公開されている人移動軌跡データセット (human-trajectory datasets) として、UCY Dataset [13] を用いる。これは、市街地の歩行者を鳥瞰視点で撮影したシーンからなるデータセットであり、比較的シンプルな環境における人移動軌跡データから構成されている。本稿における実験では、映像データから各歩行者の軌跡を抽出し、位置座標系列の形に変換した後のデータを利用する。UCY データセットのシーン例を図 5 に示す。なお、選択したデータセットの影響度合いを評価するために、比較対象として ETH Dataset [5] (UCY と同様、市街地の歩行者を鳥瞰視点で撮影したデータセット) を合わせて利用する。

4.1.2 実験の内容

データセットから、3 章で述べた手順に従って訓練データを生成する。ここでは、入力フレーム数 $k = 8$ 、出力フレーム数 $l = 12$ とした。対象データセットにおけるフレーム間隔 (1 フレームに対応する単位時間) は $\Delta t = 0.4$ 秒である。これより、 v の系列としての入出力の組を、219,320 フレームのデータのうち 213,960 フレームを訓練データ、5,360 フレームをテストデータとして生成し、評価を行う。比較対象としては、同数の x の系列としての入出力の組を生成し、同様に評価する。

なお、速度ベクトルの系列を用いた場合であっても、本来は、Social LSTM における空間情報の共有は位置座標に

基づき行う必要がある。しかし、本稿では実装の都合上、以下の3パターンの比較を行うこととした。

- 実験1: 予測モデルとして Social LSTM を利用し、速度ベクトルの系列を入力した場合と、位置座標の系列を入力した場合の結果を比較する。
- 実験2: 速度ベクトルの系列を入力とし、Social LSTM と通常の LSTM (隠れ層の共有を行わないもの) の結果を比較する。
- 実験3: 予測モデルとして Social LSTM, 入力系列として速度ベクトルを利用し、データセットとして UCY を利用した場合と、ETH を利用した場合を比較する。

4.1.3 評価の指標

評価の指標としては、予測経路と真の経路 (Ground Truth) とのフレームごとの誤差 (ユークリッド距離) の平均値 (ADE: Average Displacement Error), および最終的な到達地点の予測値と真値の差 (FDE: Final Displacement Error) を用いる。 \mathbf{x}_m を m フレーム目の予測座標, $\hat{\mathbf{x}}_m$ を m フレーム目の真値とし、予測系列長を M としたとき、それぞれ以下のように算出される。

$$\text{ADE} = \frac{1}{M} \sum_{m=1}^M \|\hat{\mathbf{x}}_m - \mathbf{x}_m\|_2 \quad (5)$$

$$\text{FDE} = \|\hat{\mathbf{x}}_M - \mathbf{x}_M\|_2 \quad (6)$$

また、代表的な予測系列の視覚化を行うことで、定性的な比較を合わせて行う。

4.2 実験の結果

各実験の結果を表1, 表2, 表3に示す。単位はメートルである。予測の誤差は 50 cm 程度であり、人追従ロボットでの利用を想定した場合には、いずれも十分な精度

表1 実験1: 入力系列の違いによる精度の比較。
(予測モデルは Social LSTM を利用)

入力系列	ADE	FDE
速度ベクトル	0.55	0.56
位置座標	0.47	0.67

表2 実験2: 予測モデルの違いによる精度の比較。
(入力系列は速度ベクトルを利用)

予測モデル	ADE	FDE
Social LSTM	0.55	0.56
LSTM	0.47	0.49

表3 実験3: データセットの違いによる精度の比較。
(予測モデルは Social LSTM, 入力系列は速度ベクトルを利用)

データセット	ADE	FDE
UCY	0.55	0.56
ETH	0.55	0.55

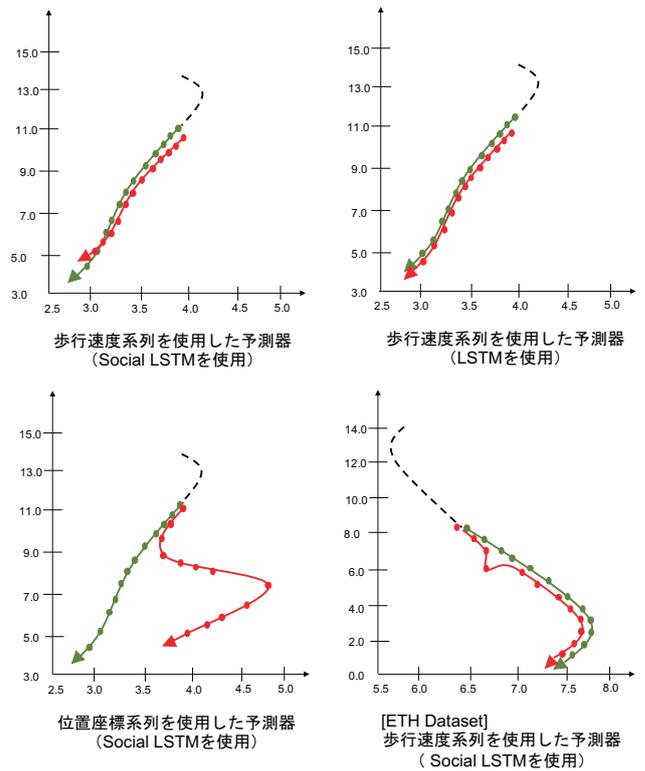


図6 ある系列データに対して予測経路を可視化した結果。黒点線で表現した観測データ (入力データ) に対し、赤線で表現した予測データが出力された。緑線が真値である。

で予測ができていると結論づけることができるであろう。一方、1章で述べたとおり、位置座標を用いた手法は、群衆内での平均的な予測精度が高い一方、個別の予測結果においては十分な精度が得られていない可能性がある。そこで、いくつかの系列データを抽出し、予測経路の可視化を行った。可視化の結果を図6に示す。これらの結果からは、位置座標を用いた手法は経路の振れ幅が大きいことがわかる。一方、速度ベクトルを用いた手法は、系列データをよく予測できている。以上より、ターゲットの特徴として移動速度を用いることで、位置座標をそのまま用いる場合より、推定精度の向上を見込めることがわかった。

以下、各実験結果の考察を行う。実験1については、入力系列の違いが予測結果に大きな影響を与えているようには見えないが、実験内容の部分で述べたとおり、今回の実験では速度ベクトルを用いた場合に、位置座標に基づく空間情報の共有が行われていない。この部分が不利に働いていると考えると、速度ベクトルを用いる方法は、今後、実装方法の変更による精度の向上が見込めるであろう。実験2からも同様の結論が導ける。予測モデルとして、複数の LSTM 間の相互作用を考慮しない、通常の LSTM を用いた予測結果は、実験1で位置座標を用いた結果よりも良くなっている。実験2において Social LSTM を用いたモデルが、不適切な相互作用によるマイナスの影響を受けているだろうことを考えると、実装方法の変更による今後の精

度向上が期待できる。

なお、実験3ではデータセットの違いが精度に与える影響を評価したが、今回利用した2つのデータセットに関しては、この影響はほぼないように思われる。いずれのデータセットに対しても、十分な精度が得られているが、今後はデータセットの特徴に対して、より詳細な検討を行っていく必要があるだろう。

5. おわりに

本稿では、人物追従ロボットのための歩行者経路予測手法として、速度ベクトルにより構成される系列データを利用し、機械学習モデルにより未来の経路を予測する手法を提案した。これまで、人の移動軌跡（位置座標の時系列）を系列データとみなし、入力系列から出力系列を予測する手法を提案してきたが、利用する系列データを速度ベクトルに変更することで、追従ターゲットの移動傾向を予測モデルに組み込むことを可能にした。本稿ではさらに、人移動軌跡データセットを利用し、提案手法と位置座標を用いた手法の経路予測精度を比較した。その結果、速度ベクトルを用いた手法の方が、特定ターゲットに対して真の軌跡に近い経路を予測できることがわかった。

本稿の結論は以下の2点である。

- 追従ターゲットの特徴の一つとして歩行速度に着目し、この系列データを機械学習モデル（Social LSTM）の入力とする予測手法を提案した。また、速度ベクトルとともに、空間情報共有のために位置座標を合わせて利用することにより、さらなる精度向上が期待できる可能性を示唆した。
- 既存のデータセットを利用し、位置座標系列と速度系列を用いたモデルの性能を比較した。その結果、いずれの場合も予測誤差は50 cm程度であり、人物追従には十分な精度であることがわかった。ただし、予測系列の可視化の結果からは、位置座標を用いた手法の経路は振れ幅が大きく、速度系列を用いた手法が優位であることもわかった。

4章で述べたように、今回、速度ベクトルを用いた手法において、Social LSTMにおける位置座標に基づく空間情報の共有は行われていない。今後、この機構を実験プログラムに組み込み、評価実験を行う予定である。また、追従ターゲットや環境側の特徴については、移動速度以外にも様々な特徴が考えられるため、それらのモデルへの組み込みも検討していく。さらに、実環境において、追従ロボットが取得したデータを用いた評価実験も行っていく予定である。

謝辞 本研究はJSPS 科研費20K11776の助成を受けたものである。

参考文献

- [1] 加藤宏一朗, 浅田郁弥, 松日楽信人, 赤羽根里奈, 加藤由花: 歩行者路特徴抽出 RTC を用いた人物追従ロボットの追従精度の向上, 計測自動制御学会 SI 部門講演会 (SI2019) (2019).
- [2] 赤羽根里奈, 加藤由花: 人物自動追従ロボットのための機械学習による歩行者経路予測手法, 情報処理学会マルチメディア通信と分散処理ワークショップ, pp. 62–67 (2019).
- [3] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L. and Savarese, S.: Social LSTM: Human Trajectory Prediction in Crowded Spaces, *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, pp. 961–971 (2016).
- [4] 平川 翼, 山下隆義, 玉木 徹, 藤吉弘亘: 動画を用了経路予測手法の分類, 電子情報通信学会論文誌, Vol. J102-D, No. 2, pp. 53–67 (2019).
- [5] Pellegrini, S., Ess, A., Schindler, K. and van Gool, L.: You'll Never Walk Alone: Modeling Social Behavior for Multi-target Tracking, *Proc. IEEE International Conference on Computer Vision (ICCV 2009)*, pp. 261–268 (2009).
- [6] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780 (1997).
- [7] Sadeghian, A., Alahi, A. and Savarese, S.: Tracking the Untrackable: Learning to Track Multiple Cues with Long-Term Dependencies, *Proc. IEEE International Conference on Computer Vision (ICCV 2017)*, pp. 300–311 (2017).
- [8] Xiang, Y., Alahi, A. and Savarese, S.: Learning to Track: Online Multi-Object Tracking by Decision Making, *Proc. IEEE International Conference on Computer Vision (ICCV 2015)*, pp. 4705–4713 (2015).
- [9] Chen, L., Ai, H., Zhuang, Z. and Shang, C.: Real-Time Multiple People Tracking with Deeply Learned Candidate Selection and Person Re-Identification, *Proc. IEEE International Conference on Multimedia and Expo (ICME 2018)*, pp. 1–6 (2018).
- [10] Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S. and Alahi, A.: Social GAN: Socially Acceptable Trajectories With Generative Adversarial Networks, *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, pp. 2255–2264 (2018).
- [11] 奥迫伸一, 坂根茂幸: レーザレンジファインダを用いた移動ロボットによる人の追跡, 日本ロボット学会誌, Vol. 24, No. 5, pp. 605–613 (2006).
- [12] 坂井 栞, 木村純麗, 池田貴政, 野見山大基, 松日楽信人, 加藤由花: 測域センサにより取得される歩行パターンを利用した高齢者/若年者弁別手法, 情報処理学会論文誌, Vol. 58, No. 2, pp. 375–383 (2017).
- [13] Lerner, A., Chrysanthou, Y. and Lischinski, D.: Crowds by Example, *Computer Graphics Forum*, Vol. 26, No. 3, pp. 655–664 (2007).