

移動ロボットのための Sequence to Sequence モデルを用いた歩行者移動軌跡の予測

Predicting a Pedestrian Trajectory Using Sequence to Sequence Learning for Mobile Robot Navigation

坂田 夏樹^{*1} 木下 裕香^{*1} 加藤 由花^{*1}
Natsuki Sakata Yuka Kinoshita Yuka Kato

^{*1}東京女子大学

Tokyo Woman's Christian University

This paper proposes a method to predict the future trajectory of a pedestrian as sequence data by using massive trajectory records collected by various sensor devices. We aim to use the method for safely and efficient path planning of autonomous mobile robots in a human-robot coexisting environment. For the prediction, we use Sequence to Sequence learning, which is frequently used in the field of natural language processing. That enables to treat long-term sequence data. In order to verify the effectiveness of the proposed method, we conduct experiments using a dataset. As a result, we show that we can predict the trajectory sufficiently by converting the trajectory data to adequate sequence data.

1. はじめに

人・ロボット共存社会の実現に向け、ロボットと人との衝突回避が重要な研究課題になっている。我々はこれまで、Social Force Model (SFM) として知られる Helbing の歩行者モデル [Helbing 95] を拡張した人移動モデルを構築し、この問題を解決する手法の研究を進めてきた [Kato 17]。ここでは、空間の形状を基に人移動をシミュレーションするモデルを構築したが、環境に応じた人の意図（移動の目的等）はモデルに組み込まれず、十分な予測精度を得ることは難しいという課題が残った。

本稿では、この課題に対応するために、環境内に設置したセンサにより収集される人の移動履歴を基に、未来の人の移動軌跡を系列データとして予測する手法を提案する。近年、行動把握のためのセンシング技術が進展し、高い精度で、常時人移動データを取得する環境が現実味を帯びてきている。我々自身も、各種センシング技術とロボットサービスの融合に関する研究を行っており [Sakai 16]、本稿でも、これら大量のセンシング結果を、ロボットの経路計画に利用することを考える。特に、近年、自然言語処理やマーケティング分野等で、時系列データ予測技術が大きく進展しており、時空間データから人の行動を理解・予測する時空間行動理解技術の研究が盛んに行われている。本稿では、系列データを予測する機械学習アルゴリズムの1つである Sequence to Sequence モデル (Seq2Seq) [Sutskever 14] を利用し、移動軌跡を適切な系列データに変換することで、未来の移動軌跡の予測を実現する。

2. 関連研究

人の移動軌跡を系列データとして予測するには、大きく分けて2つの方法が考えられる。一つは、歩行者の移動傾向を数理モデルとして定式化し、このモデルに従い人の移動傾向を推定するものである。もう一つは、大量の観測データを利用し、対象環境における人の移動傾向を推定するものである。

人移動モデルを用いる手法として最もよく知られているのは、前述した Helbing の SFM [Helbing 95] である。これは、

人と人、人と壁の間に、粒子間力のような心理的斥力と物理的斥力を定義し、歩行者の進行方向を決定するモデルである。その他、避難シミュレーションのためにセル・オートマトンモデルを用いて人の移動傾向を予測する手法 [Boukas 15] の提案や、ロボットと共存する環境での人移動モデルの定式化 [Kato 17] なども行われている。これらの手法は、適切なモデルを定義することができれば、事前に大量のデータ取得を行う必要がなく、適用が容易であるという大きな利点がある。一方、モデル化において考慮する必要があるパラメータ数が多く、精度の高いモデル構築が難しいという課題がある。

観測データを用いる方法は、環境ごとに大量の移動軌跡データを収集し、その結果から人の移動傾向を推定する方法である。人移動計測データから移動モデルを生成し、ロボットの経路計画に利用する方法 [Noguchi 12] などがある。近年では、機械学習アルゴリズム、特に自然言語処理技術の進展により、移動データを系列データとして学習させる研究が盛んに行われている。RNN (Recurrent Neural Network) を用いて移動軌跡の系列から目的地を推定する手法 [遠藤 16] や、系列データ間の相互作用を Social LSTM (Long short-term memory) としてモデル化する手法 [Alahi 16] などが提案されている。これらの手法は、事前に大量のデータ収集が必要になるというデメリットはあるが、環境に適したモデルの構築が可能であり、建物の幾何形状や環境要素もモデルに取り込むことができる。本稿でも、観測データを用いる手法を採用し、特に、系列から系列を予測する手法を用いる。

3. 提案手法

本稿では、人の移動軌跡を系列データとして扱い、観測した系列データから未来の系列データを予測する。予測には、Seq2Seq モデルを用いる。これは、LSTM encoder, LSTM decoder の2つの LSTM を組合わせて構成され、 m フレームの系列データを入力すると、予測結果として n フレームの系列データが出力されるものである。モデルのイメージを図 1 に示す。

LSTM encoder は、入力された系列データ $\mathbf{x}_{t-m+1:t}$ に対し \mathbf{c} を出力し、これが LSTM decoder の入力になる。また、LSTM encoder の隠れ変数 \mathbf{h}_t は、LSTM decoder の隠れ変数の初期値となる。LSTM decoder は、入力された \mathbf{c} に対

連絡先: 加藤由花, 東京女子大学数理科学科,
yuka@lab.twcu.ac.jp

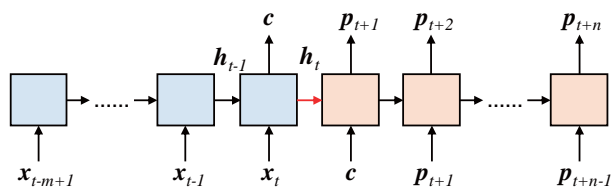


図 1: Seq2Seq モデルのイメージ. 水色の部分が LSTM encoder, ピンクの部分が LSTM decoder に相当する.

し, x_{t+1} の予測に相当する p_{t+1} を出力する. 同様に, p_{t+1} が LSTM decoder に入力され, その結果 p_{t+2} が出力される. この処理を繰り返し, 最終的に系列データの予測値として $p_{t+1:t+n}$ が得られる.

ここで, Seq2Seq モデルを用いて系列データを予測するためには, 対象となる人の移動軌跡を系列データ $x_{t-m+1:t+n}$ に変換する必要がある. 本稿では, 人の位置座標を 1 次元のグリッド空間上にマッピングし, 人の移動をグリッド間の遷移とみなす. そして, 人の移動軌跡を, グリッド ID の系列データとして構成することにした.

4. 実験

4.1 実験の内容

手法の有効性を検証するために, 商業施設内を一般の歩行者が徒歩で移動するデータである ATC pedestrian tracking dataset [Brscic 13] を用いて, 提案手法の評価を行った. これは, 大阪の商業施設 ATC ショッピングモールにおけるのべ 3,758,348 軌跡 (1 日平均 40,851 本) の人移動軌跡と, ユーザ ID などのメタデータを含んだ, 約 1 年間 (2012/10–2013/11 までの毎週水曜・日曜, 9:40–20:20), 通算 92 日分のデータである. 複数の 3 次元距離画像センサーにより, 人の位置が 10 ~ 40 Hz (周期は人の密集度により異なる) で継続的に計測されている. データとしては, ユーザ ID, 位置座標 [mm], 移動速度等が含まれる. 本稿ではこのうち, ショッピングモールの一部のエリア (10 m × 10 m の大きさ) に含まれる 1635 本の軌跡を取り出し, 軌跡ごとに $N = 160$ フレーム分の位置座標 $x_{1:N}$ を抽出して利用した. パラメータとしては, $m = n = 15$ とし, LSTM の隠れ層の次元は 128 次元とした. なお, 提案手法では, 2 次元座標をグリッドに変換するため, 入力データ x_t は 1 次元のデータになる.

入力データと教師データは, $x_{1:N}$ から t を 1 フレームづつずらしながら, 入力データとして m フレーム ($x_{t:t+m-1}$), 教師データとして n フレーム ($x_{t+m:t+m+n-1}$) を続けて切り出すことで取得した. これにより, $N - m - n = 130$ 個の入力データと教師データの組が得られる. つまり, 全体で 1635 本 × 130 個 = 212,550 組のデータを学習データとして利用した (ここでは, 学習データを 9 : 1 に分割し, 前者を訓練データ, 後者をテストデータとした). なお, 重み更新のためのバッチサイズは 200 とし, 100 イテレーション学習させた.

表 1: グリッド幅と Accuracy の関係

グリッド幅	Accuracy
0.3 m	0.02
0.5 m	0.61
1.0 m	0.66

評価指標としては, テストデータに対する Accuracy (教師データと予測データが系列として一致した割合) を用い, 人の位置座標を 1 次元グリッド空間にマッピングする際の各グリッドの幅を, 0.3 m, 0.5 m, 1.0 m と変化させたときの Accuracy を算出した.

4.2 実験の結果

各グリッド幅における Accuracy を表 1 に示す. また, 系列予測結果の一例を以下に示す.

```

-----
グリッド幅 0.3 (不一致)
入力: [271 271 271 302 302 302 302 302 302 301 301 301 301 301]
出力: [300 300 300 300 300 300 300 331 331 331 331 331 331 331]
予測: [304 304 304 304 304 304 304 304 304 304 304 304 304 304]
-----
グリッド幅 0.5 (不一致)
入力: [ 73  73  73  73  73  73  73  73  73  73  73  92  92  92]
出力: [ 92  92  92  91  91  91  91  91  91  91  110 110 110 110]
予測: [ 92  92  92  92  91  91  91  91  91  91  91  91  91  91]
-----
グリッド幅 1.0 (一致)
入力: [ 34  34  44  44  44  44  34  44  44  34  34  44  44  44]
出力: [ 44  44  44  44  43  43  43  43  43  43  43  43  43  43]
予測: [ 44  44  44  44  43  43  43  43  43  43  43  43  43  43]
-----

```

これらの結果から, グリッドの幅, つまり人の移動軌跡の表現方法により, 予測精度は大きく異なることがわかる. グリッドの幅を小さく取る方が細かい動きの表現が可能だが, データを系列として扱う場合には次元が爆発し, 学習がうまく進まない (0.3 m の場合). 一方, グリッドの幅を大きく取ると予測精度は上がる. しかし, これはフレームごとの移動がほとんど発生しないためであり, 人の移動の粒度が適切に表現されているとは言い難い (1.0 m の場合). 適切なグリッド幅は, 人体の大きさ, 人の移動速度, フレーム周期に依存し, 今回の実験条件では, 0.5 m 程度が適切なグリッド幅であると考えられる.

なお, 本稿では評価指標に Accuracy を用いたが, ロボットの経路計画での利用を考えると, 系列が完全に一致する必要はない. 例えば, 上記の 0.5 m の例では, 4 フレームで値の不一致が発生しているが, 軌跡としての予測は十分な精度で行えていると考えることもできる. 今後, 系列データ間の距離等, より適切な評価指標の利用を検討していきたい.

5. おわりに

本稿では, Sequence to Sequence モデルを用いて, 未来の人の移動軌跡を系列データとして予測する手法を提案した. データセットを用いた評価実験により, 移動軌跡を適切な系列データに変更することで, 十分な精度で軌跡の予測が可能であることを確認した.

今回の実験は, 限定的なデータでの検証実験であったが, 今後, グリッド幅に合わせた計測周期の調整, 適切な系列データの表現方法, より長い系列の予測, 評価指標の見直し, Seq2Seq2 以外のモデルの適用等を行っていく予定である. また, 系列間の相互作用を考慮した手法を検討し, 予測結果をロボットの経路計画に利用できるよう検討を進めていく.

謝辞

評価実験にあたっては, 東京女子大学数理科学科の柴田ひかりさん, 西村早也香さん, 吉村美穂さんにご協力頂きました. ここに感謝の意を表します. また, 本研究の一部は, JSPS 科研費 15K00137, 17K00366 の助成を受けたものです.

参考文献

- [Alahi 16] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S.: Social LSTM: Human Trajectory Prediction in Crowded Spaces, in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2016)*, pp. 961–971 (2016)
- [Boukas 15] Boukas, E., Kostavelis, I., Gasteratos, A., and Sirakoulis, G.: Robot Guided Crowd Evacuation, *IEEE Trans. on Automation Science and Engineering*, Vol. 12, No. 2, pp. 739–751 (2015)
- [Brsic 13] Brsic, D., Kanda, T., Ikeda, T., and Miyashita, T.: Person Position and Body Direction Tracking in Large Public Spaces Using 3D Range Sensors, *IEEE Trans. on Human-Machine Systems*, Vol. 43, No. 6, pp. 522–534 (2013)
- [Helbing 95] Helbing, D. and Molnar, P.: Social Force Model for Pedestrian Dynamics, *Physical Review E*, Vol. 51, No. 5, p. 4282 (1995)
- [Kato 17] Kato, Y., Nagano, Y., and Yokoyama, H.: A Pedestrian Model in Human-Robot Coexisting Environment for Mobile Robot Navigation, in *Proc. IEEE/SICE Intl. Symposium on System Integration (SII 2017)*, pp. 992–997 (2017)
- [Noguchi 12] Noguchi, H., Yamada, T., Mori, T., and Sato, T.: Mobile Robot Path Planning Using Human Prediction Model Based on Massive Trajectories, in *Proc. IEEE Intl. Conf. on Networked Sensing Systems (INSS 2012)*, pp. 1–7 (2012)
- [Sakai 16] Sakai, S., Kimura, S., Nomiya, D., Ikeda, T., Matsuhira, N., and Kato, Y.: Classification of Age Groups using Walking Data Obtained from a Laser Range Scanner, in *Proc. the 42nd Annual Conference of the IEEE Industrial Electronics Society (IECON 2016)*, pp. 5862–5867 (2016)
- [Sutskever 14] Sutskever, I., Vinyals, O., and Le, Q. V.: Sequence to Sequence Learning with Neural Networks, in *Advanced in Neural Information Processing Systems 27 (NIPS 2014)*, pp. 3104–3112 (2014)
- [遠藤 16] 遠藤 結城, 西田 京介, 戸田 浩之, 澤田 宏: 長期的な状態依存性を考慮した移動軌跡からの目的地予測, 情報処理学会マルチメディア・分散協調とモバイルシンポジウム 2016 論文集, pp. 1524–1536 (2016)