

## E(2)-同変グラフニューラルネットワークによる人流予測

川上 健太<sup>1,2</sup>堀江 正信<sup>3</sup>大西 正輝<sup>2,1</sup>竹内 孝<sup>4</sup><sup>1</sup>筑波大学<sup>2</sup>産業技術総合研究所<sup>3</sup>株式会社 RICOS<sup>4</sup>京都大学

## 1. はじめに

人流予測では Social Force Model (SFM) [1] に代表される仮説駆動型シミュレーション, すなわち人の移動に関する知見に基づいてモデルを定義し予測する手法が主に使用されてきた. 一方, 計測機器や画像処理技術の発展によって人流の計測が容易となり, 大量の人流データからニューラルネットワークのような汎用関数によって人流を予測するデータ駆動型シミュレーションが注目を集めている. この手法は仮説によるモデル定義が必要ないため, 応用と拡張の容易さに利点を持つ.

グラフニューラルネットワーク (GNN) は, ノードとエッジと特徴量から成るグラフ構造のデータを解析する手法として盛んに研究されており, 分子の化学的特性予測, 物理シミュレーションやスポーツ選手の軌跡予測などの分野に応用されている. 近年 GNN を用いた人流予測モデルが提案された [2]. GNN は高い表現力を持つが, その柔軟性から生じる過学習が問題となっている. この過学習の一因は物理現象の備える制約をモデルが正確に満たせないことにあると考えられる.

多くの物理現象は並進・回転などの対称性を備えており, 物理現象の記述はそれらの対称性を満たさなければならぬ. 物理現象の対称性を機械学習で正確に取り扱うためには, 並進・回転などについての同変性があるモデルを使用する必要がある. そのため, 現象の同変性に着目し, GNN に同変性を課したモデル [3] が提案され, 計算速度や予測精度の向上が報告されている.

人流現象は二次元ユークリッド空間での現象として捉えられる. そこで本研究では人流現象は回転, 並進および鏡映 (E(2) 群) に対して対称と仮定し, E(2)-同変性を課した GNN によるデータ駆動型シミュレーション法を新たに提案する. 実データを用いた実験から提案手法による人流予測の高精度化を確認した.

## 2. 関連研究

## 2.1 GNN を用いた人流予測

Zhang らは GNN を用いた人流予測手法として, SFM [1] の関数をニューラルネットワークに置き換えることで学習可能な物理学に基づく人流予測モデルを提案した [2]. 歩行者, 障害物の集合をそれぞれ  $\mathcal{P}$ ,  $\mathcal{O}$  とし, 時刻  $t$  における歩行者  $i$  の特徴量を  $\mathbf{h}_i^t$  とすると, 式 (1) のように次の時刻  $t+1$  における加速度  $\mathbf{a}_i^{t+1}$  が求められる.

なお,  $\varphi_e$  と  $\varphi_n$  は他の歩行者から,  $\varphi'_e$  と  $\varphi'_n$  は障害物から受ける反発力を表す学習可能な関数である.  $\mathbf{f}_{iD}$  は目的地  $D$  への引力を表しており, 歩行者  $i$  の質量  $m_i$ , 理想的な速さ  $v_{id}$ , 目的地への単位方向ベクトル  $\mathbf{n}_{iD}$ , 現在の速度  $\mathbf{v}_i$  と緩和時間  $\tau$  から決定される.

$$\begin{aligned} m_i \mathbf{a}_i^{t+1} &= \mathbf{f}_{iD}^t + \varphi_n \left( \sum_{j \in \mathcal{P} \setminus \{i\}} \varphi_e(\mathbf{h}_i^t, \mathbf{h}_j^t) \right) + \varphi'_n \left( \sum_{o \in \mathcal{O}} \varphi'_e(\mathbf{h}_i^t, \mathbf{h}_o^t) \right), \\ \mathbf{f}_{iD}^t &= m_i \frac{v_{id}^t \mathbf{n}_{iD}^t - \mathbf{v}_i^t}{\tau}. \end{aligned} \quad (1)$$

得られた  $\mathbf{a}_i^{t+1}$  を式 (2) のオイラー法に用いることで, 次の時刻における速度  $\mathbf{v}_i^{t+1}$  と位置座標  $\mathbf{x}_i^{t+1}$  を求める.

$$\mathbf{v}_i^{t+1} = \mathbf{v}_i^t + \Delta t \cdot \mathbf{a}_i^{t+1}, \quad \mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \Delta t \cdot \mathbf{v}_i^{t+1}. \quad (2)$$

この手法は, SFM や既存のデータ駆動型人流シミュレーション手法である Social-LSTM [4], Social-GAN [5] などよりも高い人流予測の精度を達成している.

## 2.2 同変な GNN

Satorras らは E( $n$ )-同変な GNN である EGNN (式 (3)) を提案した [3]. 式 (3) は  $n$  次元ユークリッド空間に埋め込まれたグラフにおいて, ノード  $i$  の特徴量を  $\mathbf{h}_i$ , 座標を  $\mathbf{x}_i$ , ノードのペア  $(i, j)$  を結ぶ辺の特徴量を  $\mathbf{a}_{ij}$  として定義している. また,  $N$  はノードの数で,  $\varphi_e$ ,  $\varphi_x$ ,  $\varphi_h$  は学習可能関数である.

$$\begin{aligned} \mu_{ij} &= \varphi_e(\mathbf{h}_i^t, \mathbf{h}_j^t, \|\mathbf{x}_i^t - \mathbf{x}_j^t\|^2, \mathbf{a}_{ij}), \\ \mathbf{x}_i^{t+1} &= \mathbf{x}_i^t + \frac{1}{N-1} \sum_{j \neq i} (\mathbf{x}_i^t - \mathbf{x}_j^t) \varphi_x(\mu_{ij}), \\ \mu_i &= \sum_{j \neq i} \mu_{ij}, \\ \mathbf{h}_i^{t+1} &= \varphi_h(\mathbf{h}_i^t, \mu_i). \end{aligned} \quad (3)$$

EGNN は多体系の力学シミュレーション, コミュニティグラフの生成, および分子化合物の化学的特性の回帰予測という複数のタスクに対して高い性能を示した.

## 3. 同変性

多くの物理現象は対称性を備えており, それらの現象の記述が正確であるためには, 記述が同変性を備えていることが要求される. ここでは, まず同変性の定義を行う.

$\mathbb{V}$  と  $\mathbb{W}$  を集合とし,  $f: \mathbb{V} \rightarrow \mathbb{W}$  を関数とする. 群  $G$  が  $\mathbb{V}$  と  $\mathbb{W}$  に作用可能で, この作用が関数  $f$  と交換可能であれば, 関数  $f$  は群  $G$  に対して同変であるという.

Towards Human Movement Simulation using Graph Neural Networks with Physical Priors: K. Kawakami, M. Horie, M. Onishi, K. Takeuchi

$$\forall v \in \mathbb{V}, \forall x \in G, f(x \cdot v) = x \cdot f(v). \quad (4)$$

また、特別な場合として  $G$  の  $\mathbb{W}$  への作用が恒等変換である時には、関数  $f$  は  $G$  に対して不変であるという。

機械学習モデルが同変であるためには、モデルを関数  $f$  として、関数  $f$  が回転や並進のような群  $G$  に対して式 (4) を満たす必要がある。例として、コンピュータビジョンなどで用いられる畳み込み演算は並進に対して式 (4) を満たすため、並進-同変である。

#### 4. E(2)-同変 GNN による人流予測手法

人をノード、人同士の関係をエッジと考えると人流はユークリッド空間に埋め込まれた時間発展するグラフ構造のデータと捉えられる。そこで人流現象は回転、並進および鏡映 (E(2) 群) に対して対称と仮定し、人流予測を行う E(2)-同変な GNN モデルを提案する。

提案モデルは式 (1) と同様の変数を用いて式 (5) のように定式化される。式 (1) と同様に、 $\mathbf{a}_i^{t+1}$  から式 (2) のオイラー法を用いて速度、位置座標を求める。

$$\begin{aligned} m_i \mathbf{a}_i^{t+1} &= \mathbf{f}_{iD}^t + \varphi_n(\boldsymbol{\mu}_i) + \varphi'_n(\boldsymbol{\mu}'_i), \\ \mathbf{f}_{iD}^t &= m_i \frac{v_{id}^t \mathbf{n}_{iD}^t - \mathbf{v}_i^t}{\tau}, \\ \boldsymbol{\mu}_i &= \sum_{j \in \mathcal{P} \setminus \{i\}} (\mathbf{x}_i^t - \mathbf{x}_j^t) \varphi_e(\mathbf{h}_i^t, \mathbf{h}_j^t, \|\mathbf{x}_i^t - \mathbf{x}_j^t\|^2), \\ \boldsymbol{\mu}'_i &= \sum_{o \in \mathcal{O}} (\mathbf{x}_i^t - \mathbf{x}_o^t) \varphi'_e(\mathbf{h}_i^t, \mathbf{h}_o^t, \|\mathbf{x}_i^t - \mathbf{x}_o^t\|^2). \end{aligned} \quad (5)$$

式 (2), (5) は  $G = E(2)$  と時に式 (4) を満たすため、提案手法は E(2)-同変であるといえる。

#### 5. 数値実験

本研究では 2.1 節で述べた Zhang らの実験設定に則り、提案手法の有効性を検証する。目的関数、データ、学習に関するハイパラメータは同一の設定を用いる。式 (5) の特徴量  $\mathbf{h}_i^t$  には、 $\mathbf{v}_i^t$  や  $\mathbf{a}_i^t$  を用いた。

数値実験のため、米国のグランド・セントラル駅で観測された歩行者のデータセット (GC)、市街地の歩行者を撮影したシーンからなるデータセット (UCY) を用いた。Zhang らの実験設定に則り、評価指標には Mean Absolute Error (MAE) に加え、歩行者分布としての類似度を評価する Maximum Mean Discrepancy (MMD) と Optimal Transport (OT) を採用した。

人流の予測結果を、Zhang らの手法と比較して、表 1 に示す。また、GC データセットのある場面での提案手法による予測結果を図 1 に示す。表 1 より、提案手法が既存手法よりも高い精度で人流予測を行えることがわかる。

#### 6. 結論

本研究では、GNN を E(2)-同変にすることは高精度な人流予測を行うために有用であることを明らかにした。

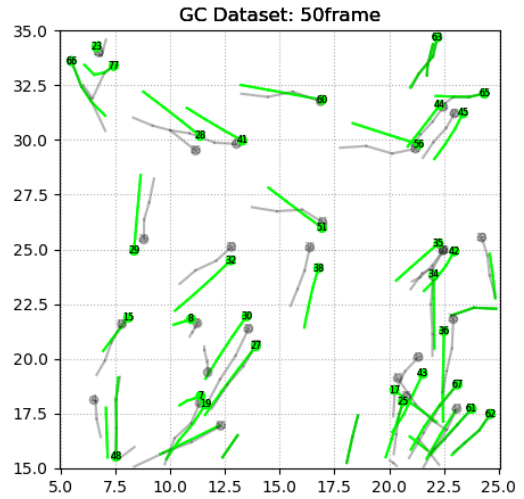


図 1: 提案手法による人流予測。緑、灰色の点は提案手法による予測と正解データである

表 1: Zhang ら [2] と提案手法の人流予測の精度比較

		MAE	MMD	OT
GC	Zhang ら	1.166	0.014	1.896
	提案手法	<b>1.041</b>	<b>0.011</b>	<b>1.593</b>
UCY	Zhang ら	2.372	0.114	6.353
	提案手法	<b>2.327</b>	<b>0.110</b>	<b>6.155</b>

今回、人流現象は E(2)-変換により現象の本質が変化しないと仮定し、E(2)-同変なモデルによる人流予測の手法を提案した。実際、提案手法により人流予測の精度が向上したことから、人流が並進、回転、そして鏡映の変換に対して対称性を有すると考えられる。

今回の提案モデルでは、前の時刻から次の時刻へ伝達される情報は位置座標、速度、加速度に限られていた。前の時刻での潜在層の特徴量を次の時刻での計算に用いることによって、より高精度に人流予測が行える可能性があると考えが今後の課題とする。

#### 参考文献

- [1] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995.
- [2] Guozhen Zhang, Zihan Yu, Depeng Jin, and Yong Li. Physics-infused Machine Learning for Crowd Simulation. In *Proc. of SIGKDD*, 2022.
- [3] Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) Equivariant Graph Neural Networks. In *Proc. of ICML*, 2021.
- [4] Alexandre Alahi, Kratarth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, and Silvio Savarese. Social LSTM: Human Trajectory Prediction in Crowded Spaces. In *Proc. of CVPR*, 2016.
- [5] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. In *Proc. of CVPR*, 2018.