

京都大学	博士 (情報学)	氏名	西村 真衣
論文題目	View Birdification: On-Ground Pedestrian Movement Estimation and Prediction from Ego-centric In-Crowd Views (混雑環境下における自己位置及び周辺歩行者の軌跡復元・予測)		
(論文内容の要旨)			
<p>自動運転や街中での配送ロボットなどの研究開発の進展とともに、人が介在する環境における自律移動技術の重要性はますます高まりつつある。本博士論文は、この社会的背景を受け、実世界における自律移動のための基盤技術として、混雑環境における自己位置推定及び周辺環境の変化予測を実現する空間認識手法の確立を主題としている。</p> <p>従来の自己位置推定技術では、静止した特徴点 (キーポイント) を複数視点から観測し、幾何拘束を用いてこれらにより記述される環境地図と観測カメラの位置 (自己位置) を同時復元する方策 (Simultaneous Localization and Mapping) が広く用いられてきた。このような従来手法は、混雑環境においては静的な特徴点を安定して観測・追跡し続けることが困難であるため、機能しない、あるいは精度が落ちることが知られていた。これに対し本研究では、静的な特徴点の発見および抽出を一切必要とせず、観測者の周囲環境における動的な歩行者自身を直接特徴点として用いる方策を示し、その基礎理論の導出、評価基盤の構築、および他手法との実験的比較を行った。</p> <p>本論文では、上述の方策に基づき、一人称視点で観測された歩行者の動きから俯瞰視点における観測カメラ及び周辺歩行者の移動軌跡を復元する View Birdification と名付ける新たな問題設定を示し、その解法及び予測モデルへの拡張についてまとめたものであり、7章から構成される。</p> <p>第1章では、混雑環境における空間認識における自己位置推定、環境マップの構築及び動的な環境マップの変化予測の重要性を説き、従来の静的特徴点を用いた幾何学的再構成の限界について議論するとともに、動的な特徴点のみを用いた俯瞰視点における周囲環境内の動的物体 (主に歩行者) と観測カメラ位置の復元をおこなう View Birdification と呼ぶ新たな課題の定義を行っている。また、観測から抽出された動的な特徴点を雑踏環境の抽象表現として捉えることにより、世界モデルと呼ばれる、観測者の動きに条件付けられた動的周囲環境の予測モデルとして View Birdification を拡張しうることを示している。さらにこれらの研究課題について、視点変換を含め、歩行者間の相互作用を深層系列処理の代表的手法である Transformer を構成する自己注意機構を用いることにより効果的に表現できることを示している。これら研究課題の解決が混雑環境下における空間認識の基盤の確立につながることを議論し、研究主題をまとめている。</p> <p>第2章では、動環境における幾何学的再構成について俯瞰するとともに、従来手法に対する本研究の位置付けを行っている。本研究は既存の Visual SLAM、三次元物体検出・追跡では不可能であった、密に混雑した環境における自己位置と周辺歩行者位置の復元を、歩行者を動的な特徴点として追跡することのみから実現するものであり、他に類を見ないものであることを確認している。雑踏環境における歩行者の動き予測として捉える歩行者世界モデルを物体中心の世界モデルの一種として位置づけ、視点変換と予測モデルの構築を同時に行う点、観測主体であるカメラが動的である点などが歩行者世界モデルに固有の問題設定であることを論じている。</p> <p>第3章では、一人称視点で観測される歩行者の位置及び速度から、俯瞰視点における観測カメラ及び周辺歩行者位置への二次元から二次元への変換問題として View Birdification を定義し、その問題設定と主な仮定について解析するとともに、透視投影及び円筒投影などの適用可能な線形カメラモデルについて述べている。</p>			

第4章では、歩行者の身長分布と線形カメラモデルに基づく観測モデルを尤度分布、俯瞰視点上での歩行者の移動モデルを事前分布としたベイズ推定の枠組みによってView Birdificationを定式化し、事後分布最大化により観測カメラ及び歩行者位置を求める最適化アルゴリズムを導出している。俯瞰視点における歩行者及びカメラの移動軌跡とその一人称視点で観測される動きの対によって構成されたデータセットでの評価を通じ、この2つの拘束条件から観測カメラ及び歩行者位置が安定して復元でき、歩行者の数が増えるほどより正確にカメラ位置を推定できることを明らかにしている。

第5章では、前章に示したView Birdificationの基本解法の課題として、歩行者の動きモデルが既知と仮定されており、また反復最適化による計算コストが非常に大きい点に言及し、深層学習に基づいたデータ駆動の解法を導出することによりこれらの限界を克服することを提案している。Transformerにおける自己注意機構を用いて歩行者間の相互作用を表現することで、歩行者の移動モデルを観測データから同時に学習可能にするとともに、観測カメラ及び歩行者位置を1回の順方向演算によって高精度に推定できることを示し、大幅な高速化と実時間処理を実現している。

第6章では、歩行者の位置及び速度情報を混雑環境における周囲状況の抽象表現として捉えることで、観測カメラの動きに条件付けられた歩行者の将来位置の予測を物体中心の世界モデルとして記述できることを示している。一人称視点観測から俯瞰視点における歩行者位置の確率マップへの変換を深層生成モデルとして定式化し、俯瞰視点における各歩行者の将来位置を確率マップとして出力することで視点変換及び軌跡予測の不確かさを考慮した周囲環境予測モデルの構築を実現している。

第7章では、本論文の目的と提案手法のまとめを行うとともに、今後の課題と応用について議論している。

(続紙 2)

(論文審査の結果の要旨)

本研究は混雑環境下における自己位置推定及び周辺環境における人物の動き予測を実現する空間認識手法の創出を主題とし、静的な特徴点に一切依存せず、観測主体の周辺に存在する動的な歩行者を主たる特徴点として用いる方法論を新たに着想し、その基礎理論の導出及び評価基盤の構築を行ったものであり、得られた成果は以下の通りである。

(1) 一人称視点で観測された歩行者の動きのみから俯瞰視点における観測カメラ及び周辺歩行者の移動軌跡を復元する View Birdification と名付ける問題設定を示し、俯瞰視点における歩行者及びカメラの移動軌跡とその一人称視点で観測される動きの対によって構成されるデータセット及び評価基盤を確立した。

(2) 歩行者の身長分布と線形カメラモデルに基づく観測モデルを尤度分布、俯瞰視点上での歩行者の移動モデルを事前分布としたベイズ推定の枠組みによって View Birdification を定式化し、制約を満たすカメラ位置及び周辺歩行者位置を事後分布最大化によって求める最適化アルゴリズムを導出した。

(3) Transformerを構成する交差注意機構 (Cross-Attention) 及び自己注意機構 (Self-Attention) が View Birdification における観測カメラ運動の推定を伴う視点変換を効果的に表現可能であることを示すとともに、歩行者の移動モデルを注意機構によって観測データから同時に学習可能とする深層学習モデルを提案した。観測カメラ及び歩行者位置を1回の順方向演算によって高精度に推定することで、解析的解法と比較して大幅な高速化を達成し、実時間処理を実現した。

(4) 観測者の動きに条件付けられた動的周囲環境の予測モデルとして View Birdification を拡張するとともに、一人称視点観測を入力とした俯瞰視点における歩行者位置への変換を Conditional VAE (CVAE) に基づく深層生成モデルとして定式化し、歩行者位置推定の不確かさをガウス分布として表現することで俯瞰視点における確率的な将来位置マップの予測を実現した。

以上本論文は、従来手法で困難であった密に混雑した環境における自己位置推定及びその周辺歩行者位置の変化予測を初めて可能とし、実際の群衆軌跡データを利用して構築されたデータセットにおいてその有効性を実証したもので、学術上、實際上寄与するところが少なくない。よって、本論文は博士 (情報学) の学位論文として価値あるものと認める。また、令和5年2月21日論文内容とそれに関連した口頭試問を行った結果合格と認めた。また、本論文のインターネットでの全文公表についても支障がないことを確認した。