

## 学位論文審査結果の要旨

専攻名	システム工学 専攻	氏名	Yunus Andi Prademon
学位論文題目	Study on Human Motion Forecasting using Self-Attention based Approach		
主査・副査	主査	若林 哲史	㊞
	副査	高瀬 治彦	㊞
	副査	高木 一義	㊞
	副査		㊞
<p>審査結果の要旨</p> <p>本学位論文は、映像中の人物の動きを学習して数秒後の位置・姿勢を推定する人体運動予測のための新たなニューラルネットワークモデルを提案し、従来手法との比較を行った研究である。</p> <p>人体運動予測は、自動運転車や自動操縦による物流配送などの自律システムの安全確保や、医療分野での歩行解析などの応用が期待できるため、2次元データ、3次元データを対象にして様々な研究が実施されている。</p> <p>本論文では、まず広く普及している RGB カメラの映像を対象とする2次元平面での人体運動予測を試みた。動画の各フレーム画像から、姿勢推定器である OpenPose を用いて関節などキーポイントの座標群を求め、時系列データを作成する。そして、カルマンフィルターや時系列予測を行うリカレントニューラルネットワークモデルの RNN-LSTM を用いて、一定フレーム数の入力時系列データから、将来の位置・姿勢を推定した。その結果、RGB カメラを適用した人物動作予測の有用性が確認された。</p> <p>次に、ChatGPT に代表される自然言語処理技術のベースである Transformer ネットワークを手がかりにして、Time Series Self-Attention という新たなニューラルネットワークモデルを提案し、短期 400 ミリ秒、長期 1000 ミリ秒の人体運動予測を行った。ただし、前出の OpenPose で求めたキーポイント座標にはノイズが多く、予測性能の評価が難しいため、モーションキャプチャ装置によって観測されたキーポイントデータを含む Human3.6M データセットを用い、2次元キーポイント座標データを対象に RNN ベースの手法と比較する実験を行った。その結果、提案モデルは RNN ベースの手法を上回る性能を示し、キーポイントの観測座標と予測座標のずれは、短期予測で平均 23.5 ピクセル、長期予測で平均 10.3 ピクセルというわずかなものであった。また、RNN ベースの手法がフレーム毎のデータを逐次処理するのに対し、提案モデルでは一括処理が行われるため、予測処理は RNN ベースと比較して 2～3 倍高速であった。さらに、空間方向の特徴を集約する機能を持つ改良手法 Temporal-Spatial Time Series Self-Attention では、3次元キーポイント座標データを対象とする人体運動予測も可能となった。</p> <p>以上の研究成果は、新規性、有用性、信頼性を兼ね備えており、工学的に有意義である。よって、博士（工学）の学位論文として、本学位論文を合格と判定する。</p>			