

A Unified Approach to Handle Human Motion Sequences: From Reconstruction to Enhancement and Analysis

何, 天嘉

<https://hdl.handle.net/2324/7363811>

出版情報 : Kyushu University, 2024, 博士 (工学) , 課程博士
バージョン :
権利関係 :



氏 名 : 何 天嘉

論 文 名 : A Unified Approach to Handle Human Motion Sequences:
From Reconstruction to Enhancement and Analysis
(ヒューマンモーションシーケンスの統一的処理手法:
再構成から強化、解析まで)

区 分 : 甲

論 文 内 容 の 要 旨

近年、ウェアラブル型あるいは環境設置型のセンサを用いて安価に人間の動作を記録する技術が急速に発展・普及しつつある。センサを用いて記録した大量のヒューマンモーションシーケンスを、スマートホームやヘルスケア、スポーツ分析、姿勢矯正、ゲーム、仮想現実、映像監視システムなどの様々な応用分野において有効活用するためには、ヒューマンモーションシーケンスを効果的に処理することのできる統一的な手法を開発することが非常に重要である。本研究では、まずヒューマンモーションシーケンスから人間の動作を高い精度を持って再構成するために、時空間アテンション機構とグラフ畳み込みネットワークを統合した手法 **Spatio-Temporal ATtention based Graph Convolutional Networks (ST-ATGCN)** を提案した。次に、使用するヒューマンモーションシーケンスの品質が低くても人間の動作を復元・改善することができるように、**ST-ATGCN** に改良を施した。さらに、ヒューマンモーションシーケンスの欠損部分を補完できるようにするために **ST-ATGCN** を拡張した **Attention-based GCN with Spatio-Temporal Disentangling mechanism (ASTD-GCN)** を開発した。

まず、ヒューマンモーションシーケンスから人間の動作を高い精度を持って再構成するために、時空間アテンション機構とグラフ畳み込みネットワークを統合した手法 **Spatio-Temporal ATtention based Graph Convolutional Networks (ST-ATGCN)** を提案した。この手法は、グラフ畳み込みネットワークの伝搬プロセスと時空間アテンション機構を密に結合することの特徴としている。評価実験においては、人間行動解析のための大規模データセット **NTU RGB-D 60** を用いて、対象交差型のベンチマーク (**C-Subject**) およびカメラビュー交差型のベンチマーク (**C-View**) によって従来手法との比較を行った。その結果、提案手法は X 軸 Y 軸方向の再構成処理の精度が従来手法と比較して最も高く、特に Y 軸方向において **C-Subject**、**C-View** それぞれのベンチマークにおいて 98.8%、97.4% の精度を達成できることを示した。

次に、使用するヒューマンモーションシーケンスの品質が低くても人間の動作を復元・改善することができるように、**ST-ATGCN** に改良を施した。この改良によって、安価な市販のモーションセンサを用いて取得したデータの、各種応用分野における活用の可能性を広げることができる。評価実験においては、人間行動解析のための大規模データセット **NTU RGB-D 60** および **NTU RGB-D 120** に加えて、これらのデータセットを人為的に改変して品質を低下させたデータセット

NTU-RGB-ER、安価な市販のモーションセンサ **mocopi™**を用いて独自に構築したデータセット **MCP-ER** を利用し、**C-Subject** と **C-View** の2種類のベンチマークによって従来手法との比較を行った。その結果、低品質のヒューマンモーションシーケンスデータセットを用いた場合に提案手法が最も優れた精度改善効果を示すこと、また最大で26%以上の精度改善を達成できることが確認された。

最後に、ヒューマンモーションシーケンスの欠損部分を補完できるようにするために **ST-ATGCN** を拡張した **Attention-based GCN with Spatio-Temporal Disentangling mechanism (ASTD-GCN)** を開発した。この手法は、**CLUB (Contrastive Log-ratio Upper Bound)**損失関数を用いて、時間情報と空間情報の冗長性が低くなるように潜在空間の分割を行い、それらを再統合して効果的な処理を実現することを特徴としている。評価実験においては、人間行動解析のための大規模データセット **NTU RGB-D 120**に加えて、同データセットから2名の人間が同時に動作を行った部分を抽出したデータセット **NTU-RGB-D MP** を利用し、**C-Subject** と **C-View** の2種類のベンチマークを用いて従来手法との比較を行った。その結果、キーフレーム間隔が7以上の場合に、提案手法は出力された動作位置データと正解データの平均 **L2 距離**に基づく **L2P 値**が最小で0.077, 最大でも0.110であり、補完性能が最も高いこと、また複数人の動作を対象とする補完性能についても高い性能を維持できることを示した。

以上のように本研究では、ヒューマンモーションシーケンスを統一的に処理することのできる手法 **Spatio-Temporal Attention based Graph Convolutional Networks (ST-ATGCN)**を提案し、使用するヒューマンモーションシーケンスの品質が低くても人間の動作を復元・改善することができるように改良を施した。さらに、ヒューマンモーションシーケンスの欠損部分を補完できるようにするために **ST-ATGCN** を拡張し、**Attention-based GCN with Spatio-Temporal Disentangling mechanism (ASTD-GCN)**を開発し、スマートホームやヘルスケア、スポーツ分析などの様々な応用分野における大規模ヒューマンモーションシーケンスの有効活用につながる有用な結果が得られた。