



間接的生体情報を用いた 人間状態推定とその応用

東京科学大学 教授 小池英樹

編纂

人間の姿勢や運動状態の推定は、リハビリテーションや技能習得支援など多くの分野で重要な技術であり、将来的には脳波（EEG）を用いた高次の状態推定が期待されているものの、現段階ではノイズ耐性やリアルタイム性に課題があり、実用化には至っていない。そこで小池研究室では、筋電（EMG）・視線・足圧といった、より安定的かつ取得容易な間接的生体情報に着目し、これらを用いた姿勢推定技術を構築する（図1）。その中でも本研究では足圧センサに焦点を当て、小型の装着型センサから得られる情報のみで高精度な三次元姿勢推定手法を開発することで、高額で場所の制約がある光学ベースのデバイスに依存せず、現場での実用に足るリアルタイムかつ汎用的な推論システムの確立を目指す。

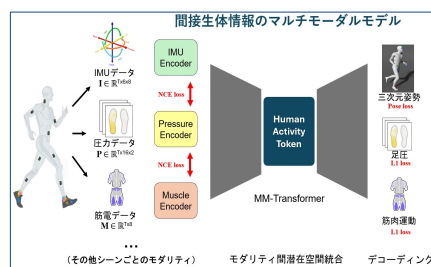
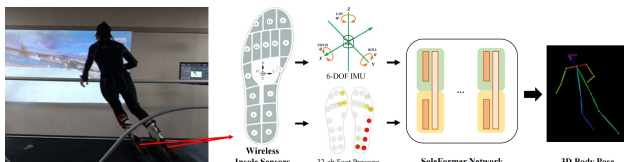


図1 研究全体の概要

SolePoser: 足圧から姿勢へ

本研究では、インソール型足圧センサのみを利用し、軽量かつリアルタイムに全身の3次元姿勢を推定する手法「SolePoser」を提案した。従来のモーションキャプチャは高価なカメラ設備や多数のIMU装着を必要とし、日常利用やスポーツ現場への適用に制約があった。一方で本手法は、インソールセンサから得た足裏の圧力分布の情報を入力とし、Dual Stream Transformer「SoleFormer」によって全身17関節の3次元姿勢を推定する。特に、出力姿勢からIMU値・足圧を再構成するDual Cycle Lossと、IMU・圧力間の関係性を強調するCross Attentionを導入することで、物理的一貫性を保ち高精度な推定を実現している。



さらに、本研究では足圧による姿勢推定の評価のために、新たに大規模なマルチモーダルデータセットを構築した。スポーツに特化したSP-Sデータセットでは、スキー・スノーボード・ゴルフスイング・卓球など高難度の動作を収録し、28名の被験者から計606Kフレームを得た。日常動作に焦点を当てたSP-Eデータセットでは、歩行・スクワットなどを対象に10名（平均年齢25.9歳）から302Kフレームを収集した。両者を合わせると計908Kフレームに及び、既存のTMMデータセットを大きく超える多様性と規模を実現した。実験の結果、スポーツ動作データで平均関節位置誤差が65.3mm、日常動作で51.0mmを達成し、推定角度誤差が29.7°および22.3°であった。推論速度は1フレームあたり11ms（@RTX3090Ti）であり、実時間応用に十分可能である。

Network	Inference Time (ms)	Results w/o Elbow and Wrist (13 Joints)						Results w/ Elbow and Wrist (17 Joints)					
		SP-S	SP-E	TMM100 [40]	MPJPE	MPJAE	MPJAE	SP-S	SP-E	TMM100 [40]	MPJPE	MPJAE	MPJAE
PoseFormer (RGB) [76]	22.5	63.8	29.9	49.2	22.2	42.7	19.0	65.2	31.3	48.0	21.4	46.9	20.4
IMUPoser (2 IMUs) [31]	10.2	85.0	39.1	80.2	37.4	X	X	103.7	47.9	99.9	46.1	X	X
AvatarPoser (3 IMUs) [18]	6.1	74.6	37.6	69.7	31.3	X	X	70.7	34.9	66.7	30.0	X	X
Simple Foot IK [8]	5.0	119.3	49.3	100.7	46.7	X	X	135.3	56.6	118.5	53.3	X	X
Direct Transformer	8.7	72.2	35.3	69.5	31.0	X	X	87.6	40.3	77.4	36.6	X	X
PressNet (Inversed) [40]	38.9	87.3	40.9	82.8	38.2	54.3	25.8	94.4	44.6	89.9	41.9	59.7	27.5
SoleFormer (Pressure Only)	10.2	81.7	39.0	79.9	37.6	40.0	18.3	90.4	42.4	86.3	40.1	48.5	22.0
SoleFormer w/o CA	10.9	65.4	30.2	51.0	22.1	X	X	71.5	35.2	70.2	31.1	X	X
SoleFormer (Ours)	11.0	65.3	29.7	51.0	22.3	X	X	70.9	34.6	67.1	30.3	X	X

比較実験において、RGBカメラベースのPoseFormerや、2〜3個のIMUを用いるIMUPoser・AvatarPoserを勝る性能を示し、足圧のみを利用した既存法PressNetを大幅に上回った。また、アブレーション実験では、足圧入力を加えることでIMUのみの場合に比べて関節の平均位置誤差17mm・角度誤差8°改善された。さらにDual Cycle LossとCross Attentionの導入が精度向上に寄与することも確認された。実環境での評価として、アルペンスキーや自然環境下での歩行における検証を行った。ここでは、17個のIMUを用いたXsensや、胸部・頭部カメラを用いた第三者視点映像を比較対象とした。その結果、SolePoserは6個のIMUを用いるTransPoseやPIPIに次ぐ性能を示し、RGBベース手法PoseFormerよりもスキー動作では優位、歩行動作では関節位置誤差0.4mmの差と同等の精度を達成した。

以上の成果を国際学会のUIST2024にて登壇発表を行った^[1]。

- [1] Erwin Wu, Rawal Khirdkar, Hideki Koike, Kris Kitani, "SolePoser: Full Body Pose Estimation using a Single Pair of Insole Sensor", ACM Symposium on User Interface Software and Technology (UIST) 2024.
[2] Shino Ito, Yichen Peng, Erwin Wu, Hideki Koike, "SoleLoadEvaluator: A Real-Time Feedback System for Walking Posture with Anterior Load Using Insole Sensors", Augmented Humans (AH) 2025.
[3] Toshihiro Hirano, Yuki Tabei, Yichen Peng, Chen-Chieh Liao, Erwin Wu, Hideki Koike, "SkiTechCoach: A Multimodal Alpine Skiing Dataset with 3D Body Pose, Sole Pressure, and Expert Coaching", ACM Multimedia (MM) 2025.

今後の展望

これまで技能獲得のために熟練者の外観、つまり身体姿勢を真似することが中心であったが、今後は、脳波、筋電、視線といった生体情報を知ること重要である。現在は、こうした生体情報の計測のコストが大きいが、これを画像処理やAI技術を用いて簡易に計測、あるいは推定する技術が重要となる。

Peng Yichen
東京科学大学 特任助教Wu Erwin
東京科学大学 特任准教授小池英樹
東京科学大学 教授

東京科学大学（旧東京工業大学）情報理工学院教授。1991年東京大学博士課程修了。工学博士。ヒューマン・コンピュータ・インタラクション(HCI)、特に画像認識を利用したVision-based HCIの研究に従事。最近では、JST CREST, JST ASPIREなどの研究代表者として、コンピュータビジョン、AI、VR/ARを用いたスポーツアシスト、熟練ビュースト、外科医のための技能獲得支援に関する研究を実施。IEEE論文賞、日本ソフトウェア学会基礎研究賞、情報処理学会フロンティア。

Liao Chen-Chieh
東京科学大学 研究員