

Co-Tune: 融合身体によるダンス学習のための適応的テンポ調整システム

橋浦 健太* 時田 聡実* 鳴海 拓志* 石黒 祥生*

概要. ダンス初心者には、複雑な全身動作を習得する際、本来よりも遅いテンポで段階的に学習することが一般的である。しかし、本来の速度での習得には長期的な練習が必要であり、学習者のモチベーション維持が課題となる。本研究では、教師と学習者が1つのバーチャルアバターを共同で操作する融合身体技術に、テンポを適応的に調整する機能を組み込んだシステム Co-Tune を構築した。Co-Tune では、元の動作に対する再生速度の割合を段階的に調整することで、学習者が教師と協調しやすい環境を提供し、成功体験を積み重ねながら本来の速度での習得を支援する。また、教師と学習者の動作同期度に基づく自動調整機能と学習者による手動調整機能を組み合わせることで、個人に適したテンポでの学習を実現した。

1 はじめに

ダンスなどの複数の身体部位を協調させた全身動作の習得は、初学者にとって大きな認知的負荷を伴う。認知負荷を軽減するため、振付を小さな単位に分割する手法 [5] や、動作の要点となるキーポーズのみを抽出して提示する手法 [6] が提案されている。しかし、これらの手法はキーポーズ間の滑らかな移行や動きの質を学習者に伝えることが困難であり、断片的な動作習得にとどまるという限界がある。

近年、バーチャルリアリティ (VR) 技術を用いた運動学習支援が注目されており、特に融合身体は運動技能の定着性が高いアプローチとして研究が進められている。融合身体とは、教師と学習者が1つのバーチャルアバターを共同で操作する技術であり、学習者は自身の運動意図を保ちながら教師の正確な動作による身体的な誘導を受けることができる。そのため、学習者は単なる模倣ではなく、主体性を持って運動しながら正しい動作パターンを体験できる。先行研究では、融合身体を用いた学習が、教師の一人称視点を共有する方法や単独での学習と比較して、運動技能の学習効率を向上させることが示されている [3]。

しかし、融合身体をダンスのような複雑な全身動作の学習に適用する際には、新たな課題が生じる。融合身体では教師と学習者が1つのアバターを共同制御するため、両者の運動意図が一致していない状態では、適切な制御が困難となる。特に素早い運動では、学習者が教師の運動意図を予測する時間的余裕がないため、適切な協調制御が困難となる [1]。この問題に対処するため、我々は事前研究としてダンス動作を本来よりも遅いテンポで実行する融合身体学習の効果を検証した。遅い（本来の 1/3）テンポ

と速い（本来の 2/3）テンポでの学習効果を比較した結果、個人によって動作精度やタイミングの習得に最適なテンポが異なることが明らかになった [2]。この知見は、全学習者に対して同一のテンポで練習させるのではなく、個人に適したテンポへの調整が必要であることを示唆している。

そこで本研究では、融合身体技術にダンスのテンポを個人の習熟度に応じて適応的に調整する機能を組み込んだシステム「Co-Tune」を開発した (図1)。Co-Tune は、教師と学習者の動作同期度を Dynamic Time Warping (DTW) によりリアルタイムで評価し、学習者が協調しやすい最適なテンポを段階的に調整することで、成功体験を積み重ねながら本来の速度での動作習得を支援する。具体的には、動作同期度に基づく自動調整機能と学習者による手動調整機能を実装し、個人に適したテンポでの学習を実現する。本論文では、Co-Tune のシステム設計と実装について述べ、WISS2025 でのデモンストレーション内容を紹介する。

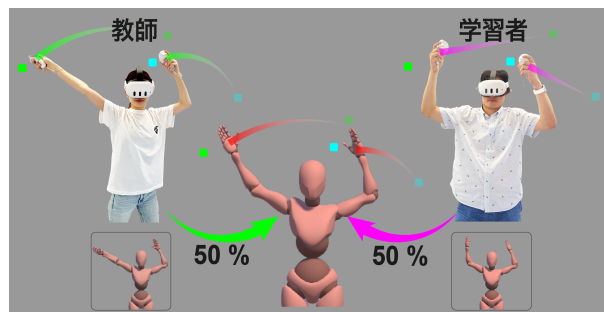


図 1. Co-Tune システムの概要。教師と学習者が1体のバーチャルアバターを 50% ずつの寄与率で共同制御する融合身体システム。色分けされた立方体は目標手首位置（緑：左手首、青：右手首）を示す。

2 システム設計と実装

2.1 設計指針

我々の先行研究 [2] では、融合身体を用いたダンス学習において、学習者個人によって最適なテンポが異なることが明らかになった。この知見から、全ての参加者に固定のテンポで練習させるのではなく、個人に適したテンポ設定を可能にする必要性を示している。また、適切なテンポ設定は教師にとっても重要である。学習者の能力に合わないテンポで融合身体を実施すると、教師は融合アバターを適切な動作に近づけるために元の振り付けから逸脱した動作を行わざるを得なくなる。その結果、融合身体における誘導の効果が低下する。以上の知見に基づき、Co-Tune は以下の 3 つの設計方針を採用する。

(1) 個人に適した初期テンポの決定：学習開始時に、学習者が取り組みやすいダンステンポを決定する。まず 2 つの異なるテンポ（元動作の 30% と 60%）において、融合身体での学習を 10 回実施し、その後 1 人で試行した結果から、それぞれの条件における元動作との運動同期度を DTW により測定する。この測定値から、学習しやすいテンポを初期テンポとして決定する。

(2) 段階的なテンポ上昇による習熟支援：学習者は成功体験を積み重ねながら本来の速度（100%）での動作習得を目指す。特定のテンポにおける習熟度が一定のレベルに到達した時点、すなわち元動作との DTW 距離が設定された閾値を下回った時点で、テンポを 10% ずつ段階的に上昇させる。

(3) 手動調整機能の統合：習熟度評価に基づく自動調整だけでなく、学習者が手動でテンポを変更できる機能も実装する。これにより、学習者が快適に感じるテンポで練習することが可能となり、システムの適応性と学習者の主体性の両立を図る。

2.2 システム構成

Co-Tune は、我々が開発した融合身体実装ツールキット PoseSynth[4] をベースに実装されている。PoseSynth は、複数のユーザが 1 体のアバターを共同操作する融合身体を容易に実装できる Unity 向けパッケージである。アバターの関節をグループ化し、グループごとに定義した重み係数を用いて複数のユーザの動作を球面線形補間により重ね合わせる。システムは 2 台の VR 対応 PC と 2 台の Meta Quest 3 で構成され、開発環境は Unity 2022.3.6f1 を使用した。各 Quest 3 は 3 点トラッキング（頭部および両手の位置）を使用し、その他の関節位置は Final IK によって推定される。本研究では、上半身の動作学習のみに注目しており、視覚的な情報過多を避けるため、融合アバターは両腕部分のみを表示している。

2.3 システム機能

2.3.1 動作可視化

学習者の認知的負荷を軽減するため、学習対象のダンス動画から音楽のビートに同期したキーポーズを抽出し、各キーポーズにおける両手首の位置を色分けされた立方体で可視化する（左手首は青、右手首は緑）。さらに、動作のタイミング学習を支援するため、ビートに同期した効果音も組み込んでいる。

2.3.2 動作同期性の評価

Co-Tune は、教師と学習者の動作同期度をリアルタイムで評価し、テンポ調整の判断材料とする。理想軌道と学習者の実際の軌道との形状的類似度を評価するために、Dynamic Time Warping (DTW) を採用した。DTW は時間軸の非線形な伸縮を許容するため、個人差による実行速度の違いの影響を受けずに、動作の空間的な正確性を評価できる。計算効率を考慮し、ウィンドウサイズを制限した FastDTW アルゴリズムを採用している。

2.3.3 テンポ調整機能

Co-Tune の中核となる適応的テンポ調整機能は、自動調整モードと手動調整モードの 2 つを提供する。自動調整モードでは、学習者の動作同期度が設定された閾値を上回った場合、システムが自動的にテンポを段階的に上昇させる。DTW 距離が閾値を下回ったことを検出すると、テンポを 10% 上昇させ、本来の速度（100%）に到達するまでこのプロセスを繰り返す。手動調整モードでは、学習者は自身の感覚に基づいてテンポを任意のタイミングで変更できる。VR コントローラのボタン操作により、10% 刻みでテンポを上下に調整可能である。この機能により、学習者は自身の学習ペースをコントロールする主体性を保持できる。

3 WISS2025 でのデモンストレーション

WISS2025 では、Co-Tune の機能を体験できるデモンストレーションを実施する。参加者は学習者として、経験豊富な教師役（展示担当者）と融合身体を通じてダンス動作を学ぶことができる。体験の開始時には、参加者自身が快適だと感じるテンポを選び、そのテンポで融合状態のダンス動作を行う。その後、システムが参加者の動作と理想軌道との同期度を解析し、その結果に基づいて自動的にテンポを調整する過程を体験できる。学習動作として AIST Dance Video Database[7] の Waack ジャンルの Basic Dance を使用する。

謝辞

本研究は JST 次世代研究者挑戦的研究プログラム (JPMJSP2108)、JST ムーンショット型研究開発

事業（JPMJMS2013）および科学研究費助成事業
基盤研究 A（24H00706）の支援を受けて行われた。

参考文献

- [1] K. Hashiura, K. Takada, T. Hagiwara, S. Wakisaka, and K. Minamizawa. “When to Take the Lead?” Mediation of Motion Intention in Collaborative Avatar Manipulation. In *Proceedings of the Augmented Humans International Conference 2025*, AHs '25, pp. 139–151, New York, NY, USA, Oct. 2025. Association for Computing Machinery.
- [2] K.Hashiura, S.Tokida, T.Narumi, and Y.Ishiguro. It's Not Just About Slow: How Tempo Control in Co-embodied Dance Training Affects Movement Accuracy and Timing. *IEEE International Symposium on Mixed and Augmented Reality Workshops*, 2025.
- [3] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose. Effects of Collaborative Training Using Virtual Co-embodiment on Motor Skill Learning. *IEEE Trans. Vis. Comput. Graph.*, pp. 1–11, 2023.
- [4] PoseSynth. <https://github.com/Cyber-Interface-Lab/PoseSynth>.
- [5] J.-P. Rivière, S. F. Alaoui, B. Caramiaux, and W. E. Mackay. Capturing Movement Decomposition to Support Learning and Teaching in Contemporary Dance. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), nov 2019.
- [6] S. Tokida and Y. Ishiguro. Understanding the Effects of Rhythmic Frame-by-Frame Presentation for Improving Dance Learning. In *Proceedings of the Augmented Humans International Conference 2025*, AHs '25, p. 7–17, New York, NY, USA, 2025. Association for Computing Machinery.
- [7] S. Tsuchida, S. Fukayama, M. Hamasaki, and M. Goto. AIST Dance Video Database: Multi-genre, Multi-dancer, and Multi-camera Database for Dance Information Processing. In *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019*, Delft, Netherlands, Nov. 2019.