



# 融合身体がフィードフォワード制御に基づく動作の学習に与える影響の予備検討

福田 晴紀<sup>1)</sup>, 橋浦 健太<sup>1)</sup>, 橋本 健<sup>2)</sup>, 峯 大典<sup>1)</sup>, 谷川 智洋<sup>1)</sup>, 鳴海 拓志<sup>1)</sup>

1) 東京大学 (〒 113-8656 東京都文京区本郷 7-3-1, {fukuda-haruki, k.hashiura, mine, tanikawa, narumi}@cyber.t.u-tokyo.ac.jp)

2) ソニーコンピュータサイエンス研究所 (〒 141-0022 東京都品川区東五反田 3 丁目 14-13, hashimoto@csl.sony.co.jp)

**概要:** 融合身体は、複数人の動作を統合して一つの身体として提示する手法で、これまで主にフィードバック制御に基づく運動課題での効果が示されてきた。本研究では、フィードフォワード制御が主となるテニスのスイング学習において、融合身体の使用が教師動作への適応、内部モデルの取得に与える影響を明らかにするため、条件比較による予備的な検討を行った。右利きのテニス未経験者を三条件（統制、JFM、融合）に割り当て、素振りの練習前後で姿勢および重心の改善量を比較した。結果、姿勢は融合条件で良好な改善傾向がみられた一方、重心移動は融合条件の伸びが最も小さかった。

**キーワード:** 融合身体, 運動学習, 教育・訓練

## 1. はじめに

スポーツやリハビリテーション、舞台芸術など多様な実践領域では、正しいフォームを獲得し安定させることが重要である。近年、バーチャルリアリティ(VR)技術の発展により、身体動作を精緻に計測・提示できる学習環境が整い、対面指導を補完する運動学習手法が注目されている。その一つである Just Follow Me (以下 JFM) は、教師アバタを学習者アバタに重ねて提示し、「見たままを追従する」ことで模倣精度を高める方法であり、VR 空間で教師動作を一人称視点で観察できる利点が報告されている [3]。しかし、JFM では模倣依存が強まり、学習後のスキル定着が十分でないという指摘もある [2]。

そこで注目されているのが、教師と学習者の動きをリアルタイムに融合し、同一アバタの共同操作を可能にする融合身体である [1]。児玉らは、左右の手で別々の動きを行うデュアルタスクに融合身体を適用し、学習者が高い行為主体感をもって教師の動きを体験することで、身体図式が効率的に更新され、練習中のパフォーマンスだけでなく単独実施時の定着率も向上することを示した [2]。この成果は、融合身体が学習効率と定着率を高める可能性を示すものであるが、検証は運動実行中にリアルタイムで誤差修正が可能な比較的低速な課題に限られている。テニスのスイングのように、運動開始後のリアルタイム修正が困難な高速なフィードフォワード制御課題でも同様の効果が得られるかは未解明である。そこで本研究では、テニスのフォアハンドスイングを模した素振り課題を用いて、融合身体がフィードフォワード制御に基づく動作の学習に与える影響を、統制条件および JFM 条件との比較によって検証した。

## 2. 実験

### 2.1 研究仮説

本研究では、融合身体がフィードフォワード制御課題における技能学習を促進するかを検証するため、姿勢と重心の一致度に着目して学習前のテスト (pre テスト) と学習後のテスト (post テスト) の成績を比較した。姿勢と重心の一致度の定量化には Dynamic Time Warping (DTW) の累積コスト (以下、DTW コスト) を用いた。統制条件および JFM 条件と比較して、融合条件における学習効果が高いとする以下の仮説を設定した。

**H1: 姿勢の一致度向上** 融合条件では、統制条件および JFM 条件と比較して、姿勢 DTW コスト ( $\Delta DTW_{\text{Posture}}$ ) の改善量が大きくなる。

**H2: 重心の一致度向上** 融合条件では、統制条件および JFM 条件と比較して、重心 DTW コスト ( $\Delta DTW_{\text{COM}}$ ) の改善量が大きくなる。

### 2.2 参加者

本実験には右利きでテニス経験のない 15 名 (男性 9 名, 女性 6 名) (年齢:  $32.9 \pm 11.8$ ) が参加した。運動制限のない健康成人であり、1 時間程度の運動を行っても支障がないことを事前に確認した。参加者は 3 条件に均等に割り振られ、それぞれ 5 名 (男性 3 名, 女性 2 名) とした。実験は所属機関の倫理規定に則って実施した。

### 2.3 装置・計測環境

本実験では、ヘッドマウントディスプレイ (HMD) として VIVE Pro 2 を使用し、参加者の全身動作は VIVE トラッカー (3.0) を計 10 台用いて計測した。トラッカーは、腰、胸、両肘、両手、両膝、両足首の計 10 か所に装着し、頭の位置は HMD によって取得した。アバタの姿勢推定に

は、Unity 2022.3 環境上で動作する VRIK<sup>1</sup>を使用した。また、参加者アバタの位置および関節角度は、50 Hz のサンプリングレートで記録された。本研究で解析対象とした関節は、骨盤、脊椎、両肩、両肘、両手、両股関節、両膝の計 12 関節である。

融合アバタは、PoseSynth<sup>2</sup>[4] を用い、取得した指導者と参加者の各関節角度を 50% : 50% で加重平均することで生成した。頭部位置についても同様に両者の座標を加重平均し、参加者の視点を融合アバタの頭部に追従させた。

視覚提示については、胸部や頭部の 3D モデルが視界を遮らないよう、レンダリング距離を 0.20 m 以上に設定した。また、参加者の前方には鏡を配置し、自身の動作や融合アバタ、教師アバタの動作を鏡越しに視認できるようにした。参加者は右手にコントローラを 1 つ保持し、試行の開始操作などに使用した。

## 2.4 実験条件

本実験は 1 要因 3 水準の参加者間実験で実施した。

**条件 1 : 統制条件** 参加者アバタと、その右隣に教師アバタが存在する。

**条件 2 : JFM 条件** 参加者アバタと、それに重なるような位置に半透明の教師アバタが存在する。

**条件 3 : 融合条件** 参加者と指導者の動作が加重平均された融合アバタだけが存在する。

## 2.5 実験タスク

本実験では 3 回の素振り（低軌道→中軌道→高軌道）を 1 試行とするテニススイングタスクを実施した。各スイングはフォアハンドストロークで構成され、低軌道と中軌道はクローズドスタンス、高軌道はオープンスタンスで実行された。動作は、1 回目のスイング前のスプリットステップ直前から開始し、3 回目のスイング終了後のレディポジション復帰までとした。1 試行の継続時間は 8.43 秒で、参加者はコントローラのボタンを押すことで任意のタイミングで試行を開始できた。

タスクに用いた教師データには、テニス経験年数 23 年のコーチ（172 cm）のスイングデータ<sup>3</sup>を使用した。このデータは 6 回連続のスイングで構成されており、本実験では難易度の観点から前半の 3 回を採用した。

練習中、各試行後には、教師と参加者の間の姿勢誤差（ワールド座標を原点とした 12 関節の回転量の差の平均として算出）、上下半身別の姿勢誤差、および重心位置差について、1 試行内の平均値が VR 環境内にフィードバックとして表示された。参加者はそれらの値を小さくすることを学習の目標とした。

## 2.6 実験手順

参加者は実験の説明を受けた後、HMD やトラッカーを着用して実験を開始した。実験は pre テスト、練習、post テストの 3 フェーズで構成された。pre テストでは、参加者は

条件に関係なく単独で 5 試行を実施した。融合条件の参加者には、練習開始前に融合アバタの仕組みについて説明を行った。続く練習では、参加者は各条件に応じてアバタを直接または鏡越しに視認しながら 120 試行の学習を行った。練習は 40 試行ごとに 3 分間の休憩を設けた。練習終了後、post テストとして再び単独で 5 試行を実施した。実験全体の所要時間は約 75~90 分であった。

## 2.7 アンケート調査

実験前後に、主に参加者の主観的体験を補足的に把握する目的でアンケートを実施した。事前には参加者情報および酔いの症状に関する項目を、事後には酔いの症状、身体感覚、タスクの負荷、および自由記述形式の感想を含む項目を尋ねた。

## 2.8 データ解析

参加者の動作と教師の模範動作との一致度を姿勢と重心の 2 指標で評価した。いずれも時系列類似度指標として DTW コストを用い、この値が小さいほど両者の動作が近いと解釈した。

一致度の変化量（改善度）の指標には、以下のように pre テストと post テストの差分  $\Delta DTW$  を定義して用いた：

$$\Delta DTW = DTW^{\text{post}} - DTW^{\text{pre}}.$$

$\Delta DTW < 0$  (DTW の減少) は一致度の向上を表す。

### 2.8.1 姿勢

解析対象である 12 関節の局所オイラー角  $(x, y, z)$  を、ラップアラウンド問題の回避のため、 $\sin$  と  $\cos$  に展開した。この処理により、1 フレームあたり  $12 \times 3 \times 2 = 72$  次元の姿勢ベクトルが得られる。こうして得られた生徒と教師のベクトル系列に対して DTW を適用した。

### 2.8.2 重心

重心位置の一致度評価には、生徒および教師の腰の位置（三次元座標）を用いた。各ファイルの初期フレームにおける  $x$ - $z$  平面上の差分  $(\Delta x, \Delta z)$  を算出し、これに基づいて生徒側の系列を平行移動させ、位置合わせを行った。その後、両者の三次元重心ベクトル系列間に DTW を適用した。

## 2.9 統計解析

姿勢および重心の一致度の改善量  $\Delta DTW$  について、条件 (3 水準) を要因とする一元配置分散分析 (ANOVA) を行った。有意水準は 5% とした。

## 3. 結果

### 3.1 外れ値除去と指標算出の手順

各参加者について、pre テスト 5 試行、post テスト 5 試行のデータをそれぞれ独立に扱い、各集合内で四分位範囲に基づく外れ値除去を行った。その後、外れ値除去後に残った試行から  $\Delta DTW$  を算出し、その参加者毎の平均を解析指標とした。

### 3.2 姿勢の一致度 ( $DTW_{\text{posture}}$ )

全身姿勢の変化量 ( $\Delta DTW_{\text{posture}}$ ) について Shapiro-Wilk 検定を行った結果、統制条件で正規性が棄却された ( $W =$

<sup>1</sup><https://root-motion.com/>

<sup>2</sup><https://github.com/Cyber-Interface-Lab/PoseSynth>

<sup>3</sup><https://val.cs.tut.ac.jp/project/tennis>

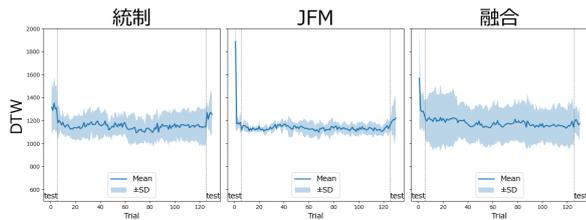


図 1: 各条件における姿勢 DTW コストの時系列推移 (外れ値除去前の生データ) (実線: 平均, 陰影:  $\pm$ SD,  $n = 5$ ). 縦破線は pre テスト (左端 5 試行) と post テスト (右端 5 試行) の区切りを示す. 練習は計 120 試行.

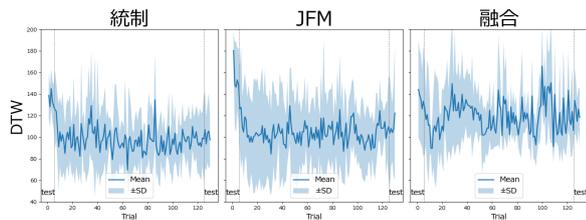


図 2: 各条件における重心 DTW コストの時系列推移 (外れ値除去前の生データ) 表記は図 1 と同じ.

0.719,  $p = 0.015$ ) 一方, JFM 条件 ( $W = 0.901$ ,  $p = 0.413$ ) および融合条件 ( $W = 0.937$ ,  $p = 0.642$ ) では棄却されなかった. このため, 要因「条件」に対してノンパラメトリックの Kruskal-Wallis 検定を適用したところ, 主効果は有意ではなかった ( $H(2) = 1.46$ ,  $p = .482$ ). ペアワイズ比較の効果量 (Cohen's  $d$ ) は統制-JFM:  $d = -0.68$ , 統制-融合:  $d = +0.39$ , JFM-融合:  $d = +0.98$  であった (図 3).

### 3.3 重心の一致度 ( $DTW_{COM}$ )

重心の変化量 ( $\Delta DTW_{COM}$ ) について Shapiro-Wilk 検定を行った結果, いずれの条件でも正規性は棄却されなかった (統制:  $W = 0.921$ ,  $p = 0.536$ ; JFM:  $W = 0.819$ ,  $p = 0.116$ ; 融合:  $W = 0.938$ ,  $p = 0.650$ ). このため, 要因

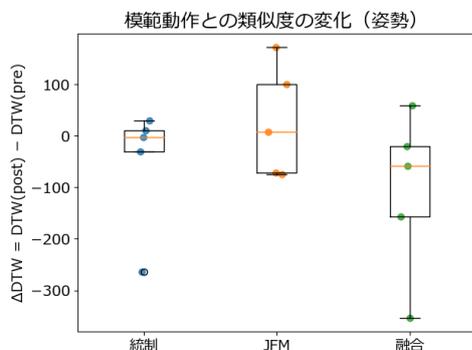


図 3: 姿勢  $\Delta$ DTW コストの条件別分布 (箱ひげ: 四分位範囲, 実線: 中央値, 点: 個人データ). 値が低いほど改善度合いが大きいことを表す.

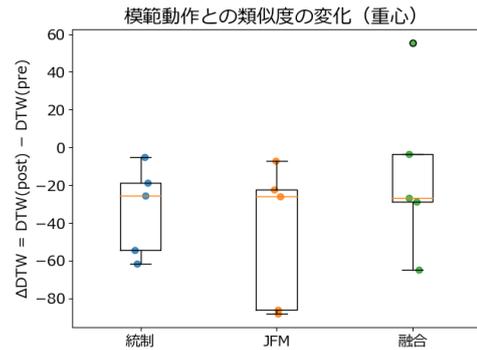


図 4: 重心  $\Delta$ DTW コストの条件別分布. 表記は図 3 と同じ. 値が低いほど改善度合いが大きいことを表す.

「条件」に対して一元配置 ANOVA を適用したところ, 主効果は有意ではなかった ( $F(2) = 0.98$ ,  $p = .403$ ). ペアワイズ比較の効果量 (Cohen's  $d$ ) は統制-JFM:  $d = +0.40$ , 統制-融合:  $d = -0.54$ , JFM-融合:  $d = -0.78$  であった (図 4).

### 3.4 主観指標

主観的評価項目 (Ownership, Agency, Change) について, 各条件 (統制, JFM, 融合) 間での比較を行うために, 一元配置分散分析 (ANOVA) を実施した.

まず, Ownership スコアについては条件間に有意な差は認められなかった ( $F(2, 12) = 0.28$ ,  $p = .76$ ,  $\eta^2 = .044$ ). 次に, Agency スコアについても統計的に有意な差は確認されなかったものの, 効果量は大きかった ( $F(2, 12) = 2.51$ ,  $p = .12$ ,  $\eta^2 = .295$ ). さらに, Change スコアについても有意差は認められなかった ( $F(2, 12) = 0.41$ ,  $p = .68$ ,  $\eta^2 = .063$ ). 記述統計量を表 1 に示す.

表 1: 主観指標における条件ごとの平均値および標準偏差

条件	Ownership	Agency	Change
統制	5.15 $\pm$ 1.56	6.20 $\pm$ 1.02	2.75 $\pm$ 1.46
JFM	4.55 $\pm$ 1.89	4.15 $\pm$ 2.27	3.20 $\pm$ 1.44
融合	4.55 $\pm$ 0.69	4.35 $\pm$ 1.19	3.50 $\pm$ 1.03

## 4. 考察

### 4.1 姿勢の一致度の変化

仮説 H1 「融合条件では, 統制条件および JFM 条件と比較して, 姿勢 DTW コストの改善量が大きくなる」は統計的には支持されなかった. しかし, JFM 条件と融合条件の比較において中程度以上の効果量 ( $d = 0.98$ ) が得られており, 融合条件のほうが改善傾向を示していた.

ただし, 今回の pre テスト以前には模範動作の提示がなく, 多くの参加者が腕だけを動かす「手打ち」にとどまっていたにもかかわらず, pre テストのスコアが post テストと同等となることがあった. 局所関節角度の差を指標に用い

た場合、全身を使ったスイングよりも手打ちの方が見かけ上コストが低くなるという評価上の偏りが生じ、初期スコアのばらつきや指標の歪みに影響した可能性がある。

今後は、動作観察を行う導入フェーズの設定や、練習中の中間時点での一人試行を挟んだ評価を取り入れることで、学習過程をより精密に捉える実験設計が求められる。

#### 4.2 重心の一致度の変化

仮説 H2「融合条件では、統制条件および JFM 条件と比較して、重心 DTW コストの改善量が大きくなる」は支持されなかった。むしろ、効果量の観点では融合条件が他条件に比べて最も小さく、重心の一致度改善において不利に働いた可能性が示唆された。この要因としては、融合条件において、教師の重心移動に関する視覚的な手がかりが希薄であったことが挙げられる。

他の条件では、教師アバタの重心移動が 100% 提示され、その動きを模倣することが可能であったのに対し、融合条件では教師の動きが 50% しか反映されない。さらに、融合アバタには「教師と自分との差異」が直接的に提示されないため、自身がどの程度教師とずれているかを直感的に把握することが難しい。このような提示仕様の違いが、重心一致度の改善を妨げた一因であると考えられる。

#### 4.3 JFM 条件における視認性と主観的体験

JFM 条件の参加者 5 名のうち 1 名からは、「教師の動きが自分の体と重なり、少し見にくかった」との自由記述が得られた。また、別の 3 名は練習の途中から後方に移動し、教師アバタを真後ろから観察する形式に切り替えていた。これらの行動は、JFM による重畳提示が高速動作時に視覚的混乱を生じさせ、動作識別や模倣を妨げていた可能性を示唆している。

さらに、主観評価では JFM 条件の Ownership および Agency スコアは、融合条件と同等かやや低めであった。JFM は一人称視点での追従を促す設計だが、教師アバタと自己アバタとの視覚的干渉により、自己帰属感や行為主体感の形成が阻害されたと考えられる。

このように、JFM によるアバタの重畳提示は、高速な動作を伴うタスクには不向きである可能性があり、今後はユーザアバタの非表示化や視点のオフセット化など、視認性向上の工夫が求められる。

#### 4.4 融合身体の課題と改善可能性

融合条件では、学習中に誤った内部モデルが構築され、修正されない例が見られた。たとえば、3 回目のスイングがオープンスタンスであるにもかかわらず、終始クローズドスタンスで実行した例や、高軌道スイングをスマッシュと誤認したまま学習が進んだケースがあった。これは、融合アバタが教師と学習者の動作を 50:50 で統合しているため、教師の正しい動作を十分に提示できず、誤認が強化されやすいためと考えられる。

この課題への対応としては、融合身体の提示を段階的かつ柔軟に設計することが有効であると考えられる。たとえば、学習初期には教師の動作を 100% 提示する観察フェー

ズを設ければ、学習者に正しい動作イメージの内在化を促せる。さらに、一致度に応じて関節ごとに融合比率を動的に調整すれば、模倣が難しい部位に対しては教師の影響を強めるといった適応的な支援が可能となる。

このように、融合身体の提示方法に工夫を加えることで、誤った内部モデルの固定化を防ぎ、より効果的な学習支援が実現できると考えられる。

## 5. おわりに

本研究は、高速なフィードフォワード制御を要するテニススイング課題を対象に、VR 環境下で教師と学習者の動作をリアルタイムに融合する「融合身体」が技能学習に与える影響を検討した。姿勢 DTW コストの変化については統計的有意差は得られなかったが、融合条件では中程度以上の効果量が確認され、他条件より一貫して高い改善傾向を示した。一方、重心 DTW コストの変化は仮説と逆方向に大きな効果量が観察され、融合身体のみ改善が限定的であった。これは、融合提示における重心移動の視覚的手がかりの不足が、学習を妨げた可能性を示している。これらの結果は、リアルタイム修正が難しい高速運動においても融合身体が姿勢学習や内部モデルの獲得に有効である一方、視覚情報の提示設計が不適切であれば重心移動の改善には効果が及ばないことを示している。今後は、視野提示の最適化や融合身体の制御率の動的調整など、提示手法を改善し、融合身体の有効性をより高精度に評価する必要がある。また、参加者数の拡充や他の動作課題への応用を通じ、長期的な学習定着や技能転移への影響も含めた包括的な検証を進める。

**謝辞** 本研究の一部は、JST ムーンショット型研究開発事業 (JPMJMS2013) および基盤研究 (A) (24H00706) の支援を受けて実施された。

## 参考文献

- [1] Rebecca Fribourg, Nami Ogawa, Ludovic Hoyet, Ferran Argelaguet, Takuji Narumi, Michitaka Hirose, and Anatole Lécuyer. Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals. 2020.
- [2] Daiki Kodama, Takato Mizuho, Yuji Hatada, Takuji Narumi, and Michitaka Hirose. Effects of collaborative training using virtual co-embodiment on motor skill learning. 2023.
- [3] Ungeyeon Yang and Gerard Jounghyun Kim. Implementation and evaluation of “just follow me”: An immersive, vr-based, motion-training system. 2002.
- [4] 橋浦 健太 鳴海 拓志 山口 周, 畑田 裕二. アバタの構造・動作の実装を効率化するツールキット「posesynth」の提案. 2024.