

RGB カメラと MediaPipe を用いた伝統工芸に関わる後継者育成のための 手先動作の 3D モデル化および多視点再生システム

小高 大和[†]
東京電機大学[†]

岩井 将行[‡]
東京電機大学[‡]

1 はじめに

伝統工芸の技術は、文化と芸術の重要な一翼を担っており、この技術を後継者に継承することは貴重な遺産を守る上で不可欠である。しかし、現代の日本において、伝統工芸の技術継承における後継者不足が深刻な問題となっている。そこで本研究では、MediaPipe[1] と安価な RGB カメラを組み合わせ、伝統工芸作家の手先の動きをリアルタイムに捉え、自由視点映像として再生可能なアニメーションを作成するシステムを提案する。このシステムを活用することで、後継者は伝統工芸作家の手の動きを異なる角度から観察することができ、視覚的な学習が促進され、深い理解を得ることが可能となると同時に、デジタルアーカイブ化によって、保存された伝統工芸の技法を未来の世代と共有可能とし、デジタル形式で永続的にアクセスできる環境を提供する。また、本システムは PC と RGB カメラのみの使用となるため、特別な機材の準備を必要とせず、安価で実装することができる。

2 関連研究

近年、骨格認識技術が発展しており、簡単に 3 次元の動きを認識することが可能となっている。稲垣らの研究では、Kinect を用いて、三次元骨格映像のフィードバックを与えることによる運動支援アプリケーションを提案している [3]。学習者には自身の

3D modeling and multi-view playback system of hand movements for training successors in traditional crafts using an RGB camera and MediaPipe.

[†] Yamato Kodaka, Tokyo Denki University

[‡] Masayuki Iwai, Tokyo Denki University

動きを客観的に認識することを可能とし、指導者には、指導のポイントを明確に認識することや学習者の改善点を認識することを可能とした。田中らの研究では、モーションキャプチャを用いてダンスの上達を支援するシステムを提案した [2]。取得したストリートダンスの動きを解析し、ダンス上級者と比較することで自らのダンスを客観的に見ることが可能となった。

3 システム構成

本稿が提案するシステムは、RGB カメラと画像認識によって検出された手のランドマークデータを csv ファイルとして出力し、csv ファイルを元に Unity 上で自由視点で再生可能な手先動作の 3D アニメーションを出力する。

3.1 手検出

手指の検出には Google 社が提供するオープンソースの機械学習ライブラリである MediaPipe を用いる。MediaPipe には、細かな顔の各パーツの動作検出や全身の動作検出といったソリューションがあり、今回は図 1 のように 21 個の各関節の 3D 座標を推測することのできる MediaPipe Hands を使用し、手のランドマーク 21 個、両手で計 42 個の 3D 座標の時系列データをもつ csv ファイルを出力する。



図 1 MediaPipe Hands の landmark の番号

3.2 3D アニメーションの作成

手検出によって出力された両手のランドマーク42個の3D座標のcsvファイルをUnityに渡すことによって、図2のようにランドマークの座標を元に手先の動きの3Dアニメーションが作成される。作成されたアニメーションは、自由視点での再生を可能とし、アニメーションの速度の調節、コマ送り、コマ戻しにも対応している。また、手検出に利用したRGBカメラ映像をmp4ファイルをとって出力しており、図3のように手先動作と実際の映像の両方の確認を可能とする。

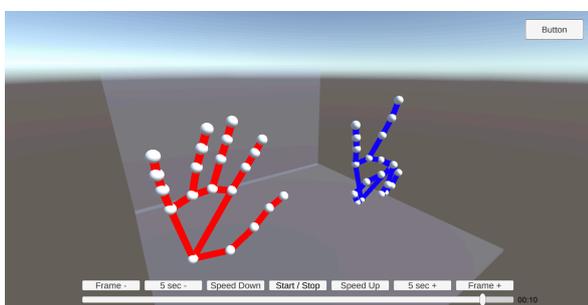


図2 実際の3Dアニメーション

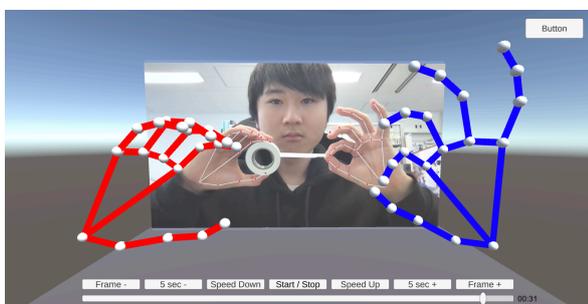


図3 実際の3Dアニメーション (mp4付き)

4 評価

指導者と学習者に分かれ、学習者は用意された指導者の手先動作から作成した3Dアニメーションを動画のみ、3Dアニメーションのみ、その両方の3パターンで確認した後、手先動作の再現のために3Dアニメーションを作成し、指導者と学習者の手のランドマークの三次元座標から平均二乗誤差を求めた結果、動画のみが0.00478、3Dアニメーションのみが0.00591、両方が0.00438という結果となった。

3Dアニメーションのみが最も誤差が大きくなっており、手先だけでは周りの状況を理解できず、再現が困難となる場合もあるが多視点で撮影することやモデルを調節することで解決することが可能である。

5 まとめ

本研究では、伝統工芸作家の手先動作をリアルタイムに捉え、3Dアニメーションとして自由視点で再生可能とするシステムを提案した。MediaPipeとRGBカメラを組み合わせることで、手先動作を3Dアニメーションとして再現し、自由視点により異なる角度から手先の動作を学習することを可能にしたが、問題点として、RGBカメラとMediaPipeでは、正確な手の深度を取得することが難しく、実際の奥行きを完全に再現することが困難となる可能性がある。今後の展望として、取得した深度情報を正確に対処するための追加の手法や改善策の検討や、システムのユーザビリティを向上させ、後継者の効果的な教育を支援するため、直感的な操作や学習プロセスの最適化を目指す。

謝辞

本研究は、科研費20K20414の助成をうけ「東秩父村竹縄技術保存会」の皆様にも多大なご協力をいただきました。深く感謝いたします。

参考文献

- [1] Mediapipe. <https://developers.google.com/mediapipe>. 参照 2024年1月11日.
- [2] 田中佑典, 齊藤剛, et al. モーションキャプチャを用いたダンス上達支援システムの開発. 第75回全国大会講演論文集, 2013(1):225-226, 2013.
- [3] 稲垣潤, 中島寿宏, 春名弘一, 昆恵介, 佐藤洋一郎, 鈴木昭弘, 本郷節之, et al. Kinectを用いた運動学習支援システムの作成と実地における試用. 第82回全国大会講演論文集, 2020(1):271-272, 2020.