

ラフスケッチによる二段階モーション編集技術

Yichen Peng^{1,a)} Chunqi Zhao^{2,b)} 黄 正宇^{1,c)} 福里 司^{2,d)} 謝 浩然^{1,e)} 宮田 一乗^{1,f)}

概要：モーションキャプチャデータからユーザの意図に沿うモーションデータを選出・編集するには、多大な労力が必要で面倒な作業である。従来のキーワードベースのモーションデータ検索システムでは、ラベリングされたキーワードの定義が曖昧で、モーションの詳細部まで網羅することは困難である。この課題を解決するために、本研究は二段階のスケッチ操作を用いてモーションデータを直感的に検索及び編集するためのユーザインタフェースを提案する。提案システムでは、モーションの大域特徴と局所特徴を考慮し、大域的なモーションをデザインする段階と、手足の動き等の局所的なモーションをデザインする段階で構成される。最後に、評価実験により、キャラクタアニメーション制作における提案システム手法（モーション検索と編集）の有効性を検証した。

Two-Stage Motion Editing Interface Using Freehand Sketch

1. はじめに

モーションキャプチャ（Mocap）データとは、現実の人物や物体の動きを記録したものである。Mocap データを 3D キャラクタに適用することで、自然かつ生き生きとしたキャラクタアニメーションを製作できることから、アニメーション業界やゲーム業界で広く利用されている。しかし、目的のモーションデータを製作するためには、Mocap データベースの中からベースとなるモーションデータを選択し、選択されたモーションデータの編集（例：複数のモーションデータをつなぎ合わせる or パーツ単位でモーションデータを組み合わせる）をしなければならず、専門的な知識を持たないユーザ（初心者）にとって非常に難しいタスクである。これは、データベース内のモーションデータの数が増大することに加え、各フレームでの関節角度情報（時系列情報）を有することに起因する。これまでに、キーワードを入力とするモーションデータ検索エンジンが公開されているものの、各モーションに対するキーワード

は主観的に定義されているため、モーションデータの詳細部まで表現することが難しい。更に、これらのエンジンには、モーション自体を編集する機能は実装されていない。

このような背景から、キーワードの代わりにラフなスケッチ操作（例：二次元キャンバス上にキャラクタの姿勢や動きの軌跡を描く）を利用するモーション検索及び編集手法が多数考案されてきた [3], [9]。しかし、スケッチ操作は、専門的な訓練なしに比較的簡単に描くことができる長所を持つものの、（時系列情報を含む）モーションデータの詳細部を考慮した検索や編集を行うことは難しい。そこで本研究では、スケッチ操作を (1) 大域的な動作を指定する段階（グローバル段階）と (2) 局所的な動作を指定する段階（ローカル段階）に分けることで上記の問題を解決し、より高品質なモーション検索と編集を実現する新たなユーザインタフェースを提案する（図 1）。グローバル段階では、ユーザは、キャラクタのルート（中心部）が空間内でどのように動くかを表すストローク線をキャンバス上に描くことで、初期ポーズを編集する。ローカル段階では、手足の動きを表すストローク線をキャンバス上に描くことで、グローバル段階で生成されたモーションデータの手足の動き情報を編集する。更に各段階において、ユーザがストローク線を描くたびに、データベースから関連するモーションデータを検索し、検索されたモーションデータの軌跡をキャンバス上に表示する機能（ガイダンス機能）を実装した [8]。

提案する二段階のスケッチ操作法の有効性を示すため

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology
² 東京大学
The University of Tokyo
a) yichen.peng@jaist.ac.jp
b) shunnki.chou@ui.is.s.u-tokyo.ac.jp
c) huang.zhengyu@jaist.ac.jp
d) tsukasafukusato@is.s.u-tokyo.ac.jp
e) xie@jaist.ac.jp
f) miyata@jaist.ac.jp

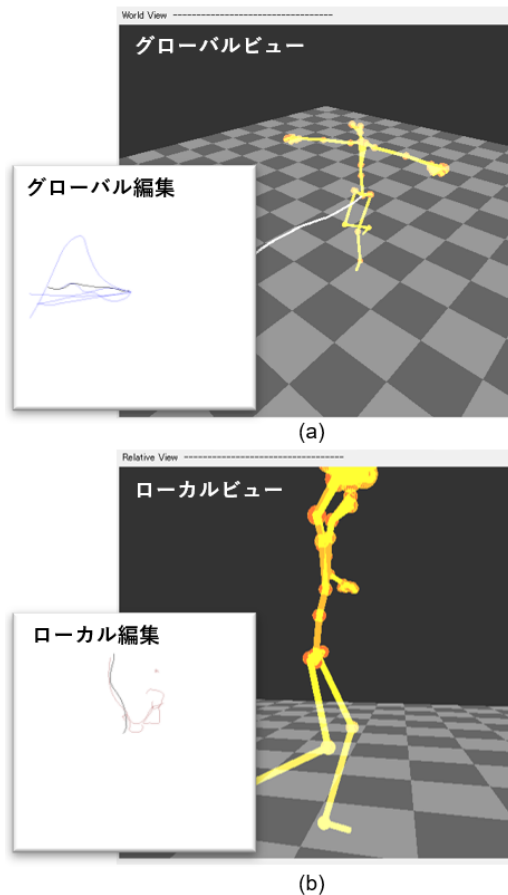


図 1: ラフスケッチによる二段階モーション編集の手法. (a) グローバルの編集段階, (b) ローカルの編集段階.

に、小規模のデータベースを基に目的のモーションデータの生成を行うユーザテストを実施した. その結果から、提案インターフェースが操作しやすく、キャラクターモーションの検索及び編集の手間を削減できることを確認した.

2. 関連研究

2.1 ラフスケッチによるモーション検索と編集

スケッチ操作（キャンバス上で絵を描く操作）の直感的な操作性に着目し、キャラクターモーションの検索や編集を支援するスケッチインターフェースが多数研究されている. 例えば、Choi ら [3] は、二次元の Stick Figure（線分の集合による図）を描くことで、モーションデータを検索する手法を提案している. また、Chao ら [2] は、ユーザが描いた二次元の動きの軌跡（モーションストローク）から 3D モーションデータを検索する際に、階層的な符号化スキームを用いることで検索の速度や精度向上を実現した. ラフスケッチによるモーション生成の研究では、Thorne ら [9] は、ユーザが描いたストロークに対して、事前に定義されたモーションデータ（例：歩く動作やジャンプ動作）を割り当てる手法を提案している. しかし、この手法では、キャラクターの全体的な動きをストローク一本で表すため、詳細なモーション編集を行うことは難しい. そこで本研究は、

大域的な動きと局所的な動きの編集を実現するために、スケッチ操作を二段階に分ける手法を検討する.

2.2 スケッチ操作の支援技術

近年、ユーザによるスケッチ操作自体を支援するためのインターフェースが提案されている. Lee ら [6] は、ユーザが二次元キャンバス上に描いた絵（途中結果）とデータベース中の絵を逐次比較することで、ユーザが描こうとしている絵のガイダンスをリアルタイムに提示する ShadowDraw を提案した. このように絵のガイダンス（以下、「影ガイダンス」とする）を提示するアイディアは、ユーザの画力向上を目的とする研究に広く利用されている. 例えば、He ら [4] は影ガイダンスを用いた書道の練習支援システム、Peng ら [7] は影ガイダンスを用いたモーション検索システムを提案している. 更に、Huang ら [5] は人物顔の大域的な情報（全体的なバランス）と局所的な情報（各部位）に対する影ガイダンスを生成する技術を提案し、より詳細な人物画の作成支援を実現した. そこで我々は、これらの研究を応用し、複数の影ガイダンスを用いたモーション編集システムを検討する.

3. 提案システム

提案システムでは、スケッチ操作をグローバル段階とローカル段階の二段階に分けている（図 2）. グローバル段階では、ユーザはキャラクターの大域的な動きを設計することができる、ローカル段階では、キャラクターの局所的な動き情報を編集し、編集結果をグローバル段階で得られた動きに組み合わせることで、最終的な出力結果を得る.

3.1 モーションデータの前処理

提案システムのプロトタイプを実装するにあたり、CMU Mocap ライブラリ [1] を基に独自のデータベースを構築した. 具体的な手順として、（モーションの編集（及び合成）による不自然な結果を防ぐために）ライブラリ中の Mocap データのうち、背の高さや手足の長さ（比率）が比較的近いものを手作業で収集する. 次に、収集したモーションデータの時間の長さを 100 フレーム分にトリミングし、1 フレーム目のルートノード（股関節）の位置を仮想空間座標系の原点に移動させることで、正規化処理を行う. 正規化されたモーションデータのルート部の動きを「大域的な動き」、ルート部から見たときの四肢の相対的な動きを「局所的な動き」とし、四肢の動きをそれぞれ分割した状態でデータベースに保存した.

3.2 ユーザインターフェース

提案システムは、スケッチキャンバスとグローバルビュー、ローカルビューの 3 つのパネルで構成されている（図 3）. スケッチキャンバスでは、ユーザは本研究で提案する二

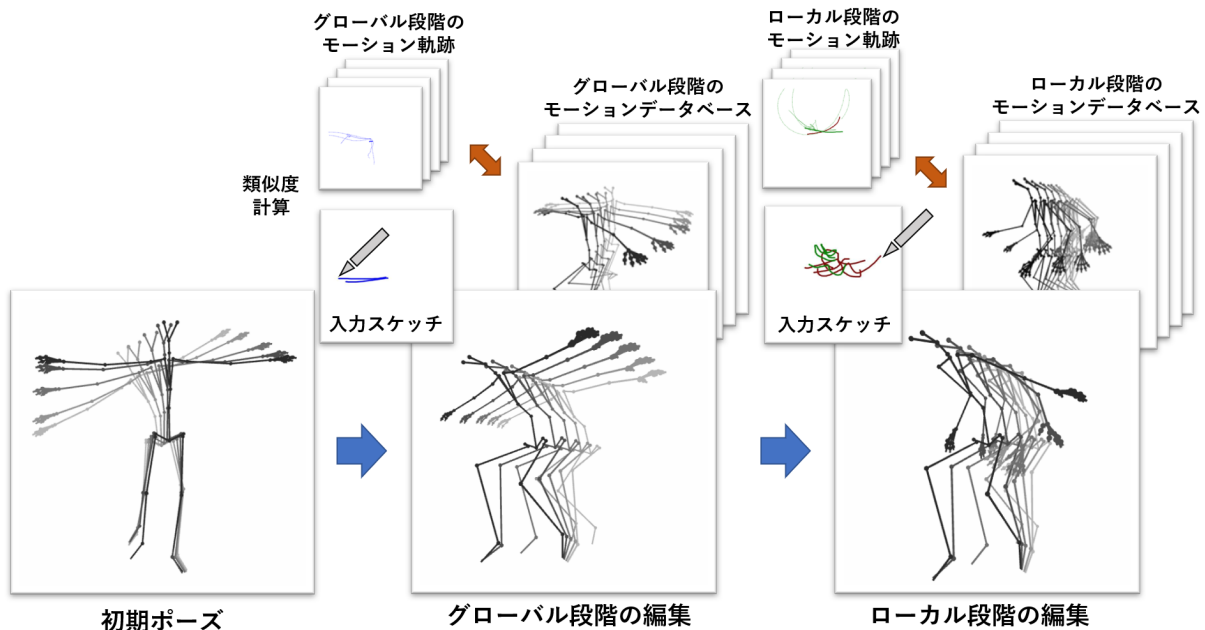


図 2: 本システムの概要. (中央) ユーザはキャラクターの大域的な動き（ルートの動き）を表すストロークを描くことで、初期ポーズのモーションデータを編集する（グローバル段階）. 次に局所的な動き（手足の動き）を表すストロークを描くことで、グローバル段階で得られたモーションの手足の動きを編集し、最終的なモーションデータを生成する（ローカル段階）.

段階のスケッチ操作（グローバル段階とローカル段階）を切り替えることができる. グローバルビューとローカルビューのパネルは、生成されたモーションデータを異なるカメラ視点から確認することができる. グローバルビューは、ユーザがズームやパンができる通常のカメラを設定している. その一方、ローカルビューは、モーションデータのルート部を画面の中央に表示するように設定した.

1) **グローバル段階**: キャラクターの大域的な動き（ルート部の動き）を指定する. 具体的な手順としては、まずユーザはルート部の軌跡（動き）を表すストロークをキャンバス上に描く. 但し、キャンバスはグローバルビューでのカメラ視点と連動している. 次に、提案システムは、キャンバスに描かれたストロークを基にデータベースの中からモーションデータを検索する. 最も類似度の高いモーションデータをグローバルビューとローカルビューのパネルに（一時的に）表示するに加え、上位五位までのモーションデータのルート部の軌跡を基に影ガイダンス（半透明のストローク群）を自動生成し、キャンバスの背景に表示する. 表示された影ガイダンスを直接マウスクリックすることで、関連性の高いモーションデータを選択することができる. 但し、グローバル段階でモーション選択を行った場合、キャラクターの四肢の動き（ローカル段階の編集結果）は初期化される.

2) **ローカル段階**: キャラクターの大域的な動き（ルート部）を設定した後、キャラクターの詳細部の動きの検索及び編集を行う. 具体的には、グローバル段階での操作方法と同様、

ユーザはキャンバスにストロークを描くことで、局所的な部位（頭部や手）の動きの軌跡を検索する. 但し、ローカルビューで設定されたカメラ視点に連動しているため、キャラクターのルート部を基準とした相対的な動きを描く必要がある. 更に、検索されたモーションデータを基に影ガイダンスを生成し、キャンバスに表示および選択機能を用いることで、直感的にモーション検索ができる.

最後に、得られた局所的な動き（頭部や手の動き）をグローバル段階でのモーション結果に合成する. 但し、足の動きを編集する場合、両足が床から浮かぶ（or 足が床にめり込む）等の不自然なモーションを生成してしまう可能性がある. このような結果を防ぐためには、足の位置を修正するための制約条件を別途追加する必要があるものの、編集結果を（後処理で）修正した場合、ユーザが描いた軌跡と生成された結果（軌跡）が大幅にずれてしまい、本研究の有用性の検証に影響が出てしまう可能性がある. そこで今回は、足部に対する制約条件は設けず、ローカル段階での編集可能な部位を頭部、左手部、右手部に限定する.

4. 実装

本システムは、Python を用いた Qt ソフトウェアとして実装している. 検索（または生成）されたモーションデータを簡単に確認するために、トップビュー、サイドビュー（左右）、フロントビューの四つのカメラ視点を指定する機能を追加した. また、ユーザが描いたストロークとデータベース内のモーションデータの類似度を計算する方法として、各関節の軌跡とのフレッシュ距離を用いた.

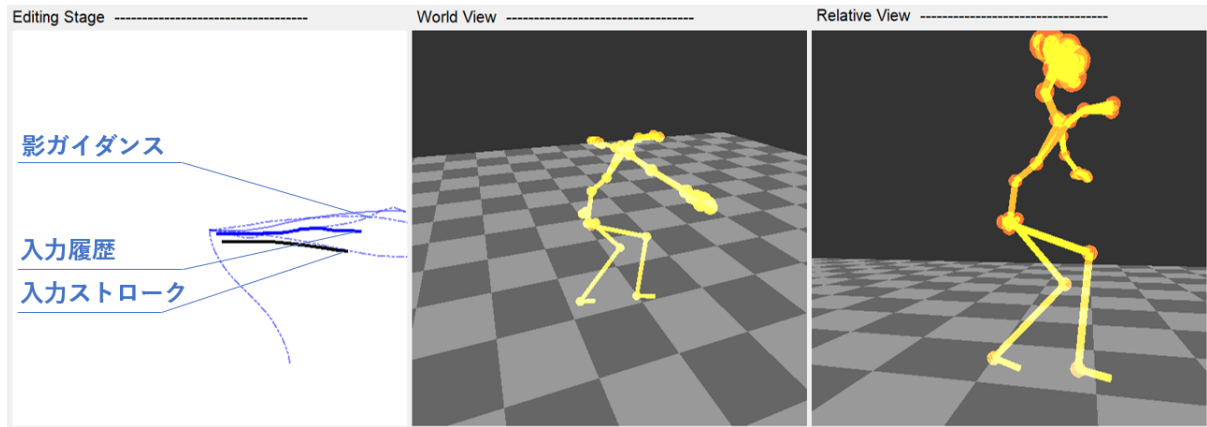


図 3: 提案インターフェース. (左) ユーザが動きの軌跡を描くためのスケッチキャンパス, (中央) キャラクタの全体的な動きを表示するためのグローバルビュー, (右) キャラクタの局所的な動きを表示するためのローカルビュー (キャラクタのルートが常に中央に表示される).

5. ユーザテスト

本手法の有用性を検証するために, 6 名の被験者 (= 男性 4 名 + 女性 2 名, 25 歳前後) を対象とするユーザテストを行った.

5.1 手順

被験者は計 55 個の歩行モーションデータ (例: 早歩き, 遅歩き, アヒル歩き, ゾンビ歩き) からの検索及び編集を行う. 具体的には, データベース中の複数のモーションシーケンスの四肢の動きを組み合わせた三種類のモーションデータ (図 4(上部)) を事前に用意し, 同様のモーション結果を被験者に再現してもらうものである. 更に本研究では, 一つのカメラ (グローバルビュー) のみで検索と編集を行う一段階のみのスケッチインターフェース (以下, 「一段階編集法」とする) を作成し, 二段階編集法 (提案手法) と一段階編集法 (ベースライン) の比較検証を行う. 但し, 各インタフェースの使用方法を覚えるために, 各被験者に約 5 分程度の練習時間を設けている.

本実験では, 定量的な比較検証と定性的な比較検証の二種類を行った. 定量的な比較としては, 編集に要する時間とインターフェース操作回数 (マウスクリックの回数) を記録し, 数値比較を行う. 定性的な比較としては, 二種類のユーザインターフェースを基に生成されたモーションデータの各関節の動きを記録し, 参照モーションとの類似性を評価する. 更に, 実験後に 4 つの項目に対するアンケート (5 段階リッカート尺度) に回答してもらい, 使いやすさや満足度に関する被験者の主観的な印象を収集した.

5.2 結果

被験者が作成したモーションデータの一例を図 4 と図 5 に示す. この結果から, 提案手法を用いて編集された右手の動きは, ベースラインと比較して参照モーションに近い

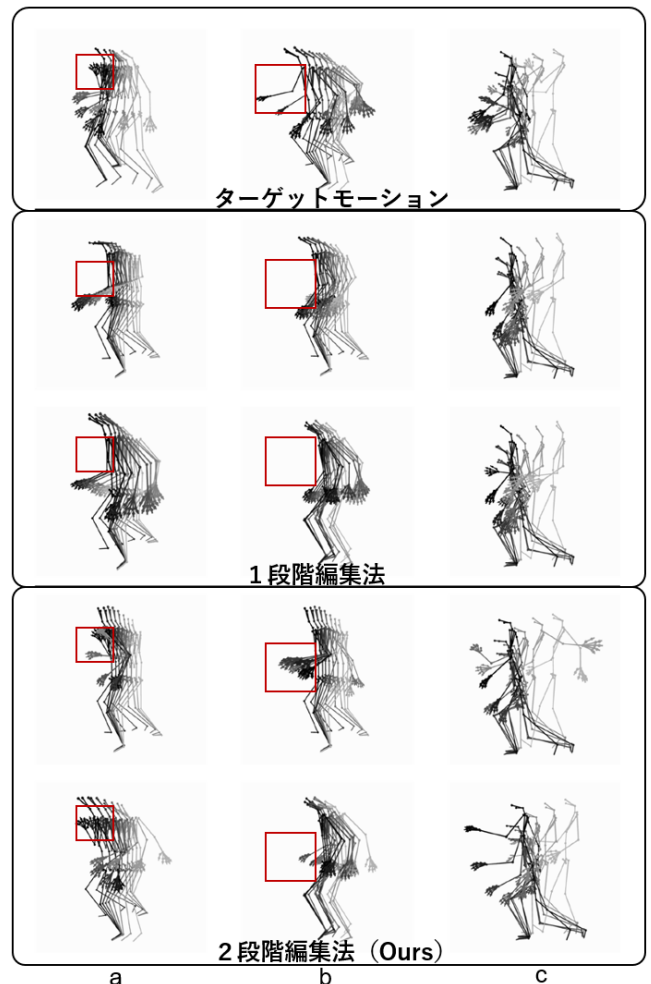


図 4: 編集結果の一例. (上部) 参照モーション, (中央) 一段階編集法 (ベースライン), (下部) 二段階編集法 (提案法).

ことが確認できる.

図 6(a) に被験者が編集に要した時間, 図 6(b) に操作回数 (マウスクリック数) を示す. この結果から, ベースライン (一段階編集) と比較し, 提案法を用いることで編集タスクを 10 秒程度早く完了し, 操作回数も約 17 回少な

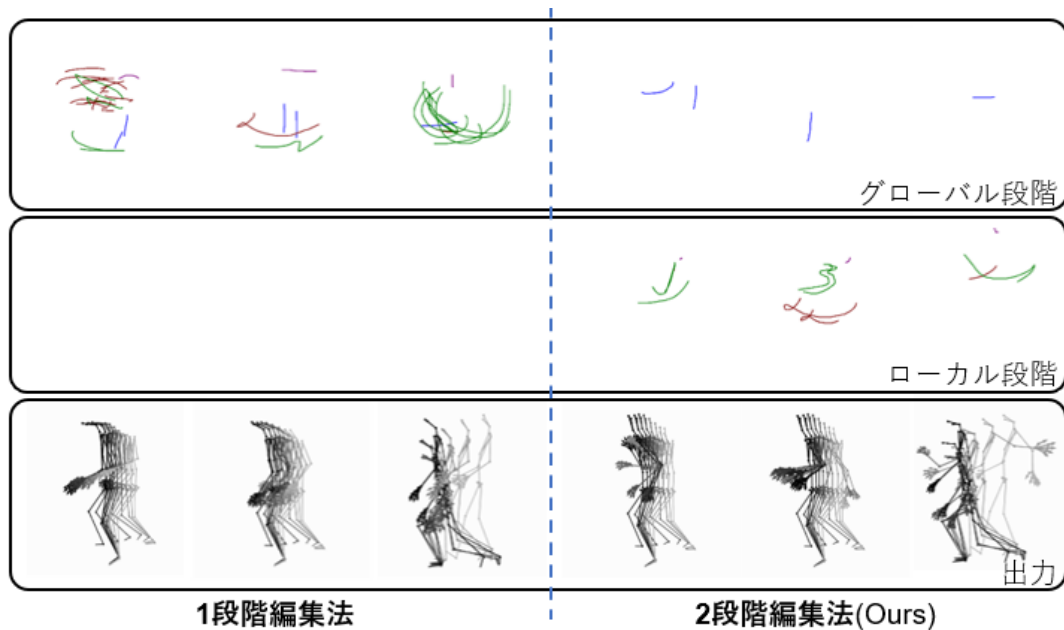


図 5: 各手法におけるユーザ入力の一例. (上部) グローバル段階, (中央) ローカル段階, (下部) 編集結果. 但し, 各部位に異なるストロークの色を割り当てている.

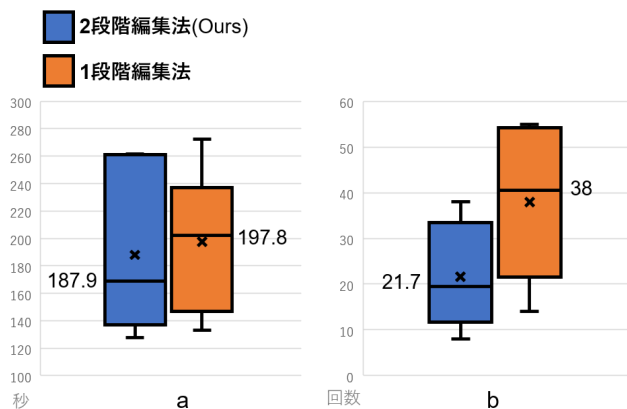


図 6: ユーザテスト結果. (a) 平均時間コスト, (b) 操作 (マウスクリック) の平均回数.

いことがわかる. これらの結果から, 提案手法によるモーション検索・編集はベースラインより簡潔な入力で, より直感的な出力を得ることができる.

表 1 にアンケートの結果を示す. この結果から, 被験者がベースラインに比べて, 提案手法によるモーション編集に満足していることが分かる. 特に, グローバル段階では, 影ガイダンス機能やスケッチ機能によるキャラクタの動きの設計が非常に快適であるといった回答が得られた (> 4.0). その一方, ローカル段階では, 一部の被験者からは「キャラクタのルートノードを基準とする相対的な動きを想像しなければならず, 多少の違和感を感じた」という意見が得られ, 比較的低いスコアとなった (< 4.0).

6. まとめ

本研究では, 二段階でモーションデータを編集を行うス

表 1: アンケートの結果.

#	質問項目	平均点	標準偏差
1	グローバルビュー UI の有効性	4.67	0.52
2	ローカルビュー UI の有効性	3.83	0.75
3	影ガイダンスの有効性	4.00	0.82
4	システム全体の満足度	4.17	0.41

ケッチインターフェースを提案した. また, ユーザが効率的にモーションの検索や編集を行うために, ユーザが描いた軌跡 (ストローク線) に類似するモーションの軌跡を表示する影ガイダンス機能を実装した. ユーザテストでは, 一段階で編集を行う方法 (ベースライン) と比較し, 特に初心者ユーザから高い評価が得られた.

現状のインターフェースでは, スケッチキャンバスと可視化パネルが分かれているため, ユーザはビューポートと仮想カメラの関係を理解するのに多少の時間を要することがわかった. 今後の課題として, これらのパネルを一つに統合することが挙げられる.

更に, ローカル段階では, ユーザが描いた局所的な動きをグローバル段階で設定したモーションデータに直接合成してしまうため, 生成された結果が不自然な結果になってしまう可能性がある (例: 床や壁とのめり込み). 将来的にはラフスケッチによって生成された手足の位置を補正するアルゴリズムも検討している.

謝辞 本研究の一部は, 公益財団法人 立石科学技術振興財団, JSPS 科研費 JP20K19845 と 19K20316 の支援を受けて実施されたものである.

参考文献

- [1] : CMU Graphics Lab Motion Capture Database, [http://mocap.cs.cmu.edu./](http://mocap.cs.cmu.edu/) (2021).
- [2] Chao, M.-W., Lin, C.-H., Assa, J. and Lee, T.-Y.: Human Motion Retrieval from Hand-Drawn Sketch, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 18, No. 5, pp. 729–740 (online), DOI: 10.1109/TVCG.2011.53 (2012).
- [3] Choi, M. G., Yang, K., Igarashi, T., Mitani, J. and Lee, J.: Retrieval and Visualization of Human Motion Data via Stick Figures, *Comput. Graph. Forum*, Vol. 31, No. 7pt1, p. 2057–2065 (online), DOI: 10.1111/j.1467-8659.2012.03198.x (2012).
- [4] He, Z., Xie, H. and Miyata, K.: Interactive Projection System for Calligraphy Practice, *2020 Nicograph International (NicoInt)*, pp. 55–61 (online), DOI: 10.1109/NicoInt50878.2020.00018 (2020).
- [5] Huang, Z., Peng, Y., Hibino, T., Zhao, C., Xie, H., Fukusato, T. and Miyata, K.: dualFace: Two-Stage Drawing Guidance for Freehand Portrait Sketching, *CoRR*, Vol. abs/2104.12297 (online), available from <https://arxiv.org/abs/2104.12297> (2021).
- [6] Lee, Y. J., Zitnick, C. L. and Cohen, M. F.: Shadow-Draw: Real-Time User Guidance for Freehand Drawing, *ACM Trans. Graph.*, Vol. 30, No. 4 (online), DOI: 10.1145/2010324.1964922 (2011).
- [7] Peng, Y., Huang, Z., Zhao, C., Xie, H., Fukusato, T. and Miyata, K.: Sketch-based Human Motion Retrieval via Shadow Guidance, *2021 Nicograph International (NicoInt)*, pp. 42–45 (online), DOI: 10.1109/NICOINT52941.2021.00015 (2021).
- [8] Peng, Y., Zhao, C., Huang, Z., Fukusato, T., Xie, H. and Miyata, K.: Two-Stage Motion Editing Interface for Character Animation, *The ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, SCA '21, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3475946.3480960 (2021).
- [9] Thorne, M., Burke, D. and van de Panne, M.: Motion Doodles: An Interface for Sketching Character Motion, Vol. 23, No. 3, p. 424–431 (online), DOI: 10.1145/1015706.1015740 (2004).