



ソーシャル VR におけるリアルタイムモーションキャプチャ とプロシージャル制御の融合による 一対多インタラクション手法

One to Many Body Interaction Techniques through The Fusion of
Real-Time Motion Capture and Procedural Control in Social VR

山岡凌¹⁾, 三武裕玄²⁾

Ryo YAMAOKA, Hironori Mitake

- 1) 明治大学 先端数理科学研究科 (〒164-8525 東京都中野区中野 4-21-1, cs232045@meiji.ac.jp)
2) 明治大学 先端メディアサイエンス学科 (〒164-8525 東京都中野区中野 4-21-1, mitake@meiji.ac.jp)

概要: VTuber と視聴者の交流として, ソーシャル VR 空間上での一対一の対面イベントや, 生放送でのコメントに対するリアクションが存在する. しかし前者は濃密なコミュニケーションが図れる代わりに一度に対話できる人数は少なく, 逆に後者は視聴者のコメントの一方的なコミュニケーションになりやすい. そこで本研究ではリアルタイムにモーションキャプチャしながらも身体動作の一部をプロシージャル制御したアバタと参加者で同時並列的に身体コミュニケーションを図れるシステムを構築した.

キーワード: プロシージャルアニメーション, VTuber, モーションキャプチャ, 躍度最小軌道

1. はじめに

ソーシャル VR 空間上ではしばしば VTuber によるお話し会や握手会などのイベントが開催され, 来場者とのコミュニケーションを楽しんでいる.

一方で VTuber の活動の多くを占める YouTube 上などの配信では, 数人~数十万人程度の視聴者がコメントを投稿し VTuber がコメントに反応している.

前者は参加者と一対一の濃密な会話や身体インタラクションを行えるのに対し, 後者は VTuber が一部のコメントに適宜反応していく. 特に人気な配信ではコメントの流速が早く, 自分のコメントを読んでもらえる可能性は低い. 従って配信では視聴者からの一方的で希薄なコミュニケーションになってしまうことがある.

そこで本研究では, 演者のモーションキャプチャと, 来場者の動きに反応するプロシージャルアニメーションを融合し単一の人型アバタを制御し, 複数人の来場者が VR 空間上で同時並列的に一対一で VTuber との身体インタラクションを行えるシステムを提案し実現する.

NPC ではなく人らしいインタラクションを実現するため, 演者の動作の一部分に来場者への追従動作を融合する方法をとり, 追従には人の手先軌道をモデル化した躍度最

小軌道を組み込んだ.

2. 実装

2.1 システム概要

本システムは, 演者のモーションキャプチャデータとプロシージャルアニメーションを融合して半自動制御するアバタ (以下, 制御用アバタ) とソーシャル VR 空間への参加者が, 一対一のインタラクションを同時並列的に行うシステムである (図 1).

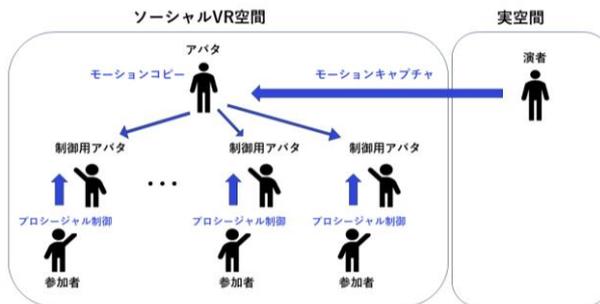


図 1: システム概要図

システムの実装は Unity で行い、ソーシャル VR 環境として VRChat を利用した。

VTuber の演者のモーションキャプチャはいわゆる 6 点トラッキング方式で行った。すなわち、mocopi や HTC Vive など普及価格帯のモーションキャプチャデバイスで計測した頭部、左手、右手、腰、左足、右足の位置を目標として、逆運動学 (IK) を用いてキャラクタモデルの各関節角度を算出する。

次に、制御用アバタにおいてインタラクションさせたい部位の位置を参加者の動きに応じてプロシージャルに変化させる。例えば、ハイタッチをさせる場合は制御用アバタの右手の x, y 座標は参加者の x, y 座標に追従するようにし、 z 座標はモーションキャプチャデータの z 座標をそのまま使用する(ここで本システムは y -up の左手系である)。ここで、人らしいモーションを演出するため、右手の x, y 座標の追従には躍度最小軌道を用いた。

2.2 躍度最小軌道

人らしいインタラクションを実現するため、制御用アバタの右手の追従に躍度最小軌道を採用した。

躍度最小軌道は Flash[1]らによって提案された人の手先軌道の特徴を予測し再現するモデルである。

始点・終点の速度が 0 であるという条件での躍度最小軌道は次のような五次式であらわされる

$$x(t) = x_0 + (x_0 - x_f)(15\tau^4 - 6\tau^5 - 10\tau^3) \quad (1)$$

y 座標も同様である。

ここで、 $\tau = \frac{t}{t_f}$ は運動時間で正規化した時間を表し、

(x_0, y_0) と (x_f, y_f) は運動の始点と終点を表す。

この軌道を時間区間を重複させて足し合わせることで刻々位置を変化させる目標への滑らかな追従動作を再現することができる。

Yeo ら[2]は直前の軌道の終点から現在の対象の位置に向けて 400ms で到達する躍度最小軌道を、200ms おきに生成して加算することで自然な捕球動作を再現している。本手法でも同様の手法で制御用アバタの右手の x, y 座標を参加者の手に追従させる。

3. 実装結果

実際に観客と制御用アバタがインタラクションを行っている様子が図 2 である。上図の右側が VTuber のアバタで、図左側の手前が参加者、奥側が制御用アバタである。

制御用アバタは一体のみ設置し、参加者ごとにインタラクションの軌跡を変更することで実質的に複数体の制御用アバタが存在するようになっている。ただしどの制御用アバタも同一の場所に存在してしまうと制御用アバタとプレイヤーが近接インタラクションを行うためにプレイヤー同士が同一個所に密集してしまうため、プレイヤーごとに制御用アバタの位置を変更している。

4. まとめと展望

演者のモーションと参加者の動きに応じたプロシージャルアニメーションを融合させた身体インタラクションが実現できた。

ユーザスタディは今後の課題である。現時点では筆者自身のみが体験できており、演者の動作もリアルタイムではなく、アニメーションデータによるものである。実際の演者や筆者以外の参加者に本システムを利用してもらい、動作の検証と印象評価を行うことは今後の課題である。特に、VTuber や VTuber の視聴者層など想定される利用者によるユーザスタディを行いたい。

動作の汎用化も課題である。今回用意したプロシージャルアニメーションは右手のハイタッチを想定したのみであるが、他にも頭を撫でたり肩を叩いたりするなど様々な身体インタラクションが考えられる。複数のインタラクションに対応できるようある程度システムを汎用化させていきたい。

今回は追従のためのプロシージャルアニメーションとして躍度最小軌道を採用したが、躍度最小軌道は身体の慣性を考慮しないため、腕を大きく振り回すような動作や上下方向など重力の影響が大きい動作では人らしい動作にならない可能性がある。トルク変化最小軌道[3]など身体の慣性を考慮した手法を用いることも検討したい。



図 2: ハイタッチの様子

参考文献

- [1] Flash, Tamar and Neville Hogan: The coordination of arm

movements: an experimentally confirmed mathematical model, *Journal of Neuroscience*, 1985.

- [2] Sang Hoon Yeo, Martin Lesmana : Debanga R. Neog, and Dinesh K. Pai. 2012. Eyecatch: simulating visuomotor coordination for object interception. *ACM Trans. Graph.* 31, 4, Article 42, 10 pages, 2012.
- [3] Uno Y, Kawato M, Suzuki R: Formation and control of optimal trajectory in human multijoint arm movement—Minimum torquechange model. *Biol Cybern* 61: 89-101, 1989