

ソーシャルVRにおける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身体インタラクション手法

山岡 凌* 三武 裕玄*

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が図れるシステムを構築した。インタラクションの一例としてハイタッチシステムを構築し、人間的な動作を実現するため、人の手先軌道を上手く表現できるモデルである躍度最小軌道を採用した。更に頭部と体の向きを参加者のほうに向くようにした。参加者の右手位置にプロシージャル制御したアバタの右手位置を追従させ、参加者の位置に応じてプロシージャル制御したアバタの向きを動的に変化させた。

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、VTuberの活動の多くを占めるYouTube上などでの生配信では数人～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。従って配信では同時に多くの視聴者が参加できる代わりに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参加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、複数人の参加者が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参加者への追従動作を融合する手法を取り、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 提案手法

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）。

制御用アバタは演者が直接操作するアバタとは違い、参加者と相対する半自動制御された人型モデルのことである。

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などの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

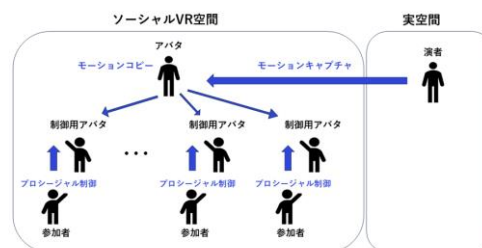


図1. システム概要図

Copyright is held by the author(s). This paper is nonrefereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

* 明治大学

キャラクター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したデータの 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。

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したデータの 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は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。本手法でも同様に、以下の加速度計算のプログラムから目標位置を算出している。

$$\mathbf{a} = \mathbf{a} - dt \left\{ \frac{60}{t^3} (\mathbf{p}_0 - \mathbf{p}_f) + \frac{36}{t^2} \mathbf{v} + \frac{9}{t} \mathbf{a} \right\} \quad (2)$$



図 2. インタラクションの様子

ここで、 \mathbf{a} を加速度ベクトルとし、 dt は 1 フレームに掛かった時間で、 $\mathbf{p}_0, \mathbf{p}_f$ はそれぞれ制御用アバタの現在位置と目標位置を表す 3 次元ベクトルである。

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。図右側のモデルが演者が直接操作するアバタであり、左側の奥のモデルが制御用アバタ、手前が参加者である。

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箇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際にプレイヤー同士が同一箇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 今後の課題

ユーザスタディは今後の課題である。現時点では筆者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。

また、ハイタッチ以外にも肩をたたいたり、頭をなでたり様々な身体インタラクションが考えられるため、プロシージャルアニメーションの種類を増やしたい。また、ハイタッチ以外にも肩をたたいたり、頭をなでたり様々な身体インタラクションが考えられるため、プロシージャルアニメーションの種類を増やしたい。

また、ハイタッチ以外にも肩をたたいたり、頭をなでたり様々な身体インタラクションが考えられるため、プロシージャルアニメーションの種類を増やしたい。

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 : Debang 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