

3D アバターによるバーチャル講義コンテンツ収録のための単眼カメラを用いた全身トラッキングとオブジェクト操作 UI の提案

松本 拓未[†] 佐々木 茂^{††} 田中 誠一^{†††}

帝京大学大学院理工学研究科総合理工学専攻[†] 帝京大学工学部情報電子工学科^{††}

文星芸術大学美術学部美術学科マンガ専攻^{†††}

1. はじめに

近年、動画配信サイトなどで、2D/3D アバターで演じての配信が広く行われている。3D モデルのアバターを演じる場合は、ヘッドマウントディスプレイ (HMD) とコントローラ、トラッカーを身につけて全身のモーションをキャプチャする手法や、Web カメラ等で顔の表情と上半身、場合によっては手の仕草や全身の姿勢をキャプチャする手法が用いられる。

また、3D アバターを纏いバーチャル空間でさまざまな体験をするメタバース上のコンテンツ等も数多く提供されるようになってきており、メタバースや VR の教育分野での活用も多く見られるようになるとともに反転授業の実践も見受けられるようになった。

メタバース・バーチャル空間を反転授業向けの事前学習教材として活用する実践の例として、矢野の例がある[1][2][3]。

VR の特徴として没入感が高いという点があり、バーチャル空間内ではアバターにより表情やジェスチャーが伝わりやすくなるとおもわれることから、反転授業の事前学習コンテンツをバーチャル空間内で、3D のアバターが教員の代わりに説明することで、学生が興味を持ちやすくなることや講義コンテンツをより注視することを促し、結果として講義内容についての理解を深める要因になるのではないかと考えられる。事前学習は学生が好きな時間に個々に取り組むものであるため、教員がリアルタイムで説明することはできない。そのため教員役のアバターはノンプレイヤーキャラクター (Non Player Character, NPC) として実装する必要がある。NPC の 3D アバターがスライド等を用いて説明する講義コンテンツの収録には、アバターの演者の全身および表情のトラッキングに加え、バー

チャル空間内の移動・スライドやポインタの操作なども行えるようにする必要がある。

本研究では、単眼カメラを用いた全身及び表情トラッキング、バーチャル空間内でのジェスチャ認識による移動や、スライド・ポインタオブジェクト等を操作するための UI を統合した機能を提案し、試験的に実装した。

2. バーチャル講義コンテンツの概要

本研究で作成するバーチャル講義コンテンツは、バーチャル空間で NPC のアバターがスライドを使い説明するものである。NPC のアバターの動作については、全身あるいは上半身の動きと、顔の表情、両手の仕草を単眼カメラでキャプチャする。キャプチャしたアニメーションは、全身の動作と手の仕草については BVH 形式でファイルに記録する。顔の表情については、独自の形式でテキストファイルに記録する。アバターの動きに加えて、スライドの切り替えについても記録する。音声は、合成音声で記録したものを、リップシンクによりアバターの口の動きと同期させた。

記録したバーチャル講義コンテンツは、独自に開発した VR アプリ上でユーザが再生する。

バーチャル講義コンテンツ収録システムでは、スライドの説明をする教員役の演者が収録の操作も行う。収録システムは、コンテンツを再生する VR アプリをベースに作成しており、教員役の演者は、バーチャル空間で動く自身のアバターを見ながらスライドの説明をする。本研究の VR アプリは、PC のディスプレイに表示するものであり、Web カメラと組み合わせて使用する。HMD とコントローラは用いない。スライドの切り替えやポインタオブジェクトの操作は、バーチャル空間内に配置したボタンオブジェクト等で行う。

このような機能と UI を備えた、バーチャル講義コンテンツ収録システムと、収録したコンテンツを再生して閲覧できる VR アプリを開発した。

3. バーチャル講義コンテンツ収録システム

本研究で開発したバーチャル講義コンテンツ収録システムでは、全身と手、および顔の表情

Proposal of whole-body tracking and object control UI with a monocular camera for recording virtual lecture contents by 3D avatar

[†]Matsumoto Takumi, ^{††}Sasaki Shigeru · Teikyo University

^{†††}Tanaka Seiichi · Bunsei University of Art

のキャプチャについては著者らの開発したシステム[4]を用いている。本システムでは、これらのキャプチャに加えて、腕を振るジェスチャーによってアバターが移動する UI[5]も備えている。全身と手の動作の記録は BVH Tools の BVHRecorder[6]を用いているが、表情およびスライド切り替えについては独自のプログラムで記録する。リップシンクには uLipSync[7]を用いている。

バーチャル講義コンテンツ収録システムは Unity2022.3.12f1 で作成している。本システムは現状では Unity エディタ上で実行する。図 1 にコンテンツ収録システムの実行画面を示す。



図1 バーチャル講義コンテンツ収録システムの実行画面

画面に配置されているボタン等は、収録開始前の設定等に用いるものであり、収録の開始と終了、スライドの切り替えやポイントオブジェクトの操作は、バーチャル空間内のボタン等で行う。これは、収録される教員役の演者の動作が、バーチャル空間内で自然なものとなることを意図している。

4. バーチャル講義コンテンツを閲覧する VR アプリ

記録したバーチャル講義コンテンツは、独自に開発した VR アプリ上でユーザが再生する。本研究で作成した VR アプリでは、アバターおよびスライドの動作を記録した BVH ファイルおよび独自形式のファイルを読み込んで、アプリ内に用意してあるアバターとスライドの動作を再生する。ユーザは VR アプリ内で説明をするアバター・切り替わるスライド・それらに合わせて再生される説明の音声により、講義コンテンツを閲覧する。図 2 に講義コンテンツを閲覧する VR アプリの実行画面を示す。ユーザは VR アプリ上でスライドと教員役のアバターの近くに配置されたボタンを押すことで、講義コンテンツを再生できる。現状では、この VR アプリも PC のディスプレイを用いている。

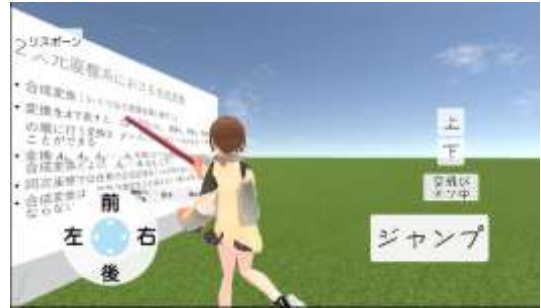


図2 講義コンテンツを閲覧する VR アプリの実行画面

5. 結果と考察

本研究で作成したバーチャル講義コンテンツ収録システムで、講義を行う教員役のアバターの全身と表情の動作を収録することができた。また、VR アプリ上で収録した講義コンテンツを再生することができた。松田らは仮想空間内を案内・紹介するキャラクターエージェントの開発手法として HMD を用いる方法を提案している[8]。アバターの動作を記録する点は本研究と同じだが、松田らの手法では顔の表情が記録されない。一方、本研究の手法では顔の表情も記録されるが、演者は PC のディスプレイでモニターするため得られるバーチャル空間内の情報が HMD を用いる場合よりも限定的である。この点はどちらを優先するかトレードオフであると考えられる。

参考文献

- [1] 矢野浩二郎, 反転学習むけ事前学習教材としてのバーチャル空間の活用, 2021 PC カンファレンス 論文集, pp. 73--75, 2021
- [2] Kojiro Yano, Production of instructional videos using a virtual presentation room on a mobile head-mounted display, ICAT-EGVE 2019, 2019
- [3] 「素人でもわかる VR シリーズ 第6回 - VR を活用したオンライン授業配信 - 」, <https://www.youtube.com/watch?v=fiZprs0d7zg&t=247s>, (2024年1月7日閲覧)
- [4] 松本拓未, 佐々木茂, 田中誠一, 一般的な単眼カメラによる表情・骨格の認識とアバターへの反映に関する研究, 情報処理学会第84回全国大会, 6ZM-04, pp. 4-1041 - 4-1042, 2022
- [5] 松本拓未, 佐々木茂, 田中誠一, 単眼カメラによる全身トラッキングを用いた 3D アバターの姿勢の同期と UI を包括的に扱う手法に関する研究, 情報処理学会第85回全国大会, 4ZK-08, pp. 4-799 - 4-800 (2023)
- [6] "BVH Tools for Unity", <https://github.com/emilianavt/BVHTools>, (2024年1月7日閲覧)
- [7] "uLipSync", <https://github.com/hecomi/uLipSync>, (2024年1月7日閲覧)
- [8] 松田 康生, 矢野 浩二郎, 自分の分身として仮想空間内を案内・紹介するキャラクターエージェントの開発手法の提案, 第24回日本バーチャルリアリティ学会大会論文集, 2019