



リアルでライブ配信可能な テレグジスタンス VTuber システム

野村 証貴¹⁾, 脇田 航¹⁾

Masaki NOMURA and Wataru WAKITA

1) 広島市立大学大学院 情報科学研究科 (〒731-3194 広島市安佐南区大塚東 3-4-1, nomura@ics.info.hiroshima-cu.ac.jp)

概要: 本研究では任意の実地にいる視聴者 (AR アプリケーション) 側に対して遠隔地にいる VTuber 配信者 (VR アプリケーション) 側がリアルでライブ配信可能なテレグジスタンス VTuber システムを提案する. 具体的には配信者側の位置を視聴者側の GPS に合わせることで視聴者側が実地でスマートフォンをかざすと VTuber が出現し, 配信者側は視聴者側のカメラ映像から第三人称視点としてテレグジスタンス可能である.

キーワード: ライブ配信, VTuber, 実世界指向メタバース, テレグジスタンス

1. はじめに

近年, VR・AR 技術はますます発展してきており, ライブストーリー配信や動画投稿等において 3D アバターや 2D イラストによる VTuber と呼ばれるようなコンテンツが生まれてきている. 昨今, 配信活動だけでなく VTuber を使用して地域の魅力を発信していくご当地 VTuber も存在してきている. ご当地 VTuber は地域の魅力を伝えるために, 現実の映像に VTuber を重畳した AR 映像を用いて観光地や食べ物のレポート動画を作成している. このような AR 映像によって, VTuber があたかも現地に実際に赴き, 現地から地域の魅力を直接伝えているような動画を作ることができる.

しかしながら, このような AR 映像は視聴者とリアルタイムにコミュニケーションが取れない動画形式であり, 実際に現地に行って VTuber に会えるような体験はすることができない. そこで本研究では任意の実地にいる視聴者 (AR アプリケーション) 側に対して遠隔地にいる VTuber 配信者 (VR アプリケーション) 側がリアルでライブ配信可能なテレグジスタンス VTuber システムを提案する. 具体的には, VTuber の位置を視聴者側の GPS に合わせることで, 例えば遠隔地にいる VTuber 配信者が指定した時間と場所に現れることをあらかじめ視聴者に伝えておき, 視聴者は指定された時間と場所に行きスマートフォンをかざすと現実世界に VTuber がリアルタイムに重畳表示され, 配信者側は視聴者側のカメラ映像から第三人称視点としてテレグジスタンス可能となり, 視聴者と VTuber との間でコミュニケーションがとれるシステムを構築する. これにより, VTuber がリアルタイムで視聴者とコミュニ

ケーションを取りながら現地の観光案内や記念撮影等を行うことができ, VTuber と近い距離間で実世界指向でのバーチャルなライブ配信が可能となる. また, VTuber を活用した地域の魅力を伝えるための地域活性化や話題作りのきっかけ等に応用することが期待できる.

2. 関連研究

遠隔地のユーザと現地のユーザが同じ時間・場所を共有するようなシステムとして Takeuchi らは GIBSON と呼ばれるシステムを提案している[1]. GIBSON では現地にいるユーザはスマートフォンを持ち, AR グラスをかけている. また, 遠隔地のユーザは HMD と VR コントローラを身につけている. 現地のユーザの AR グラスからは遠隔地にいるユーザのアバタが重畳表示されて見えており, スマートフォンで撮影された映像は遠隔地のユーザへストーリーミングされる. 遠隔地のユーザは現地のユーザがいる場所を 3D 都市モデルとして疑似的に同じ場所に存在することができる. また, 遠隔地のユーザは現地のユーザのスマートフォンからの映像を見ることができ, 現地のユーザが見ている風景と現地のユーザのアバタを仮想空間上で見ることができる. このようにして, GIBSON では現地のユーザは AR グラスを通して現実世界に重畳された遠隔地ユーザのアバタを見ることができ, 遠隔地のユーザは 3D 都市モデルで構成された仮想空間内に現地のユーザのアバタとスマートフォンから見える風景をみながら, お互いにコミュニケーションをとることができるシステムとなっている. しかしながら GIBSON システムでは, 遠隔地のユーザが見ることができるスマートフォンで撮影された映像は

現地の様子を映し出したそのままの映像であり、遠隔地のユーザの視点は常に1人称視点であるため、お互いがアバタを介してのコミュニケーションとなってしまう。

これに対し、本研究では現地のユーザがスマートフォンをかざすと、遠隔地のユーザを表すアバタが重畳表示され、遠隔地のユーザはバーチャル空間内で現地のユーザが撮影した映像そのままではなく自身が重畳表示された映像を見ることが可能なシステムを構築することで、遠隔地のユーザは現地のユーザの映像を鏡のように使うことができ、また、現地のユーザが撮影している映像を視点移動により第三人称視点のように見ることによって実世界指向でレイグジスタンス可能となり、現地にいるリアルユーザと遠隔地にいるバーチャルユーザ（アバタ）側との実世界指向のリアルタイムコミュニケーションの実現が期待できる。

3. システム概要

本システムの概要を図1に示す。

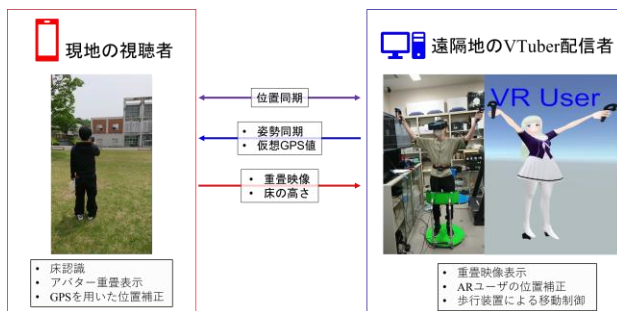


図1: システム概要

本システムは現地にいる視聴者と遠隔地にいるVTuberが同じ場所でコミュニケーションをとることが可能なシステムとするため、現地にいるユーザはスマートフォンによるARアプリケーション、遠隔地にいるユーザはPCによるVRアプリケーションとしてUnityを用いてそれぞれ実装を行う。また、それぞれのアプリケーションで位置同期、アバタの姿勢同期を行うためにMirrorと呼ばれるUnityアプリケーション間で通信を行うパッケージを使用し、スマートフォンで撮影した重畳映像の送受信にはUnity Render Streamingと呼ばれるUnityでレンダリングした映像を送受信可能とするパッケージを共通で使用する。

現地の視聴者は現地にVTuberが実際に現れているかのような体験を可能とするには、VTuberをスマートフォンで撮影した現地の映像に重畳表示させる必要がある。そのため、現地の視聴者が使用するARアプリケーションにはUnityのAR Foundation機能とGoogle社が提供しているAR Coreを使用する。AR Foundation機能はARアプリケーションを構築するための基本機能がそろっており、スマートフォンで撮影した映像にVTuberとしての3Dキャラクターを重畳表示させることや床の識別を行う。VTuberの重畳表示を行う際、遠隔地のPC上のVTuberアバタの位置

と現地位置を同期する必要がある。しかしながら、ARアプリケーション上での原点はスマートフォンでARアプリケーションを起動した位置と向きによって決定され、遠隔地のPC上のVTuberは高さ $y=0$ を床として立っているため、そのまま位置同期を行ってしまうとスマートフォンを起動した位置の高さにVTuberが表れてしまう。そのため、VTuberの位置同期の際にはARアプリケーション側で検出した一番低い床の位置情報を床としてVTuberを重畳させる補正処理を行う。また、AR CoreではGPSを用いてARアプリケーションの緯度と経度が取得でき、緯度と経度を指定してARアプリケーションの空間上にオブジェクトを固定するアンカ機能が備わっている。VTuberは緯度と経度を指定して現地の視聴者と会うことになるため、VTuberが出現する初期位置をアンカとして固定し、そこから位置同期によりVTuberは遠隔地からの操作により移動することとなる。現段階ではARアプリケーションとVRアプリケーションとの間は方向補正を行っていないため今後の課題とする。また、ARアプリケーションで生成されたVTuberが重畳表示された映像はUnity Render StreamingによってVTuberのVRアプリケーション側に送信される。

遠隔地にいるVTuberは現地とは異なる場所から配信を行うため、視聴者と現地で出会うためには、どこに出現するのかということを決め、現地に現れる必要がある。そのため、VTuberはUnity上では原点に現れるが、仮想GPSの値（緯度と経度）を指定して現れる場所を決め、視聴者のARアプリケーション側に緯度と経度情報を送信する。また、VTuber側のVRアプリケーションには視聴者のスマートフォンに映し出されているVTuberが重畳表示された現地の映像の受信を行い、重畳映像表示用オブジェクトとして仮想空間上に出現させる。この際、ARアプリケーションの原点をもとに位置の同期が行われているため、ARアプリケーションで検出された床の高さをもとに、VRアプリケーション上での重畳映像表示用オブジェクトの高さを決定する。また、遠隔地にいるユーザはVTuberとして現地の視聴者と同じ場所でコミュニケーションとっているような体験を可能とするために、自身の動きをトラッキングしてアバタを制御する。そのため、遠隔地にいるユーザはHMD、両手にVIVEコントローラ、腰・両足にVIVE Trackerを装着して6点でのフルボディトラッキングを行う。これらのトラッキング情報をもとに、VTuberとしてのアバタの姿勢を制御するため、SA FullBody IKと呼ばれるUnityアセットを利用し、取得したトラッキング情報からアバタの姿勢を決定する。これらの6点のトラッキング情報は現地の視聴者側に重畳表示されているアバタの姿勢としても送信される。また、アバタの移動操作には、現地のユーザと同じ場所を共有しているため、現地のユーザと同じように歩行できる必要がある。そのため、本システムでは、我々が開発している歩行装置[2]を使用し、アバタの移動操作を行う。

4. 実装結果

VTuber 側と視聴者側でコミュニケーションをとっている様子を図 2 に示す。



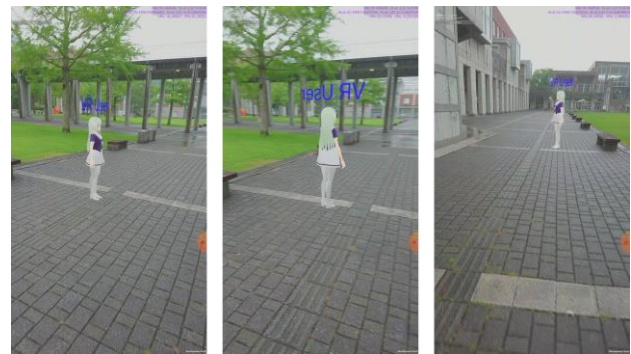
(a) VTuber 視点

(b) 視聴者視点

図 2: VTuber と視聴者のコミュニケーションの様子

図 2 のように、VTuber はフルボディトラッキングと歩行装置によりアバタを制御し、視聴者とコミュニケーションをとることができ、視聴者は重畳表示された VTuber の姿を見ることができる。また、VTuber は視聴者のスマートフォンで撮影された VTuber (自分自身) の重畳表示映像を見て、自分自身がどのように映っているのか・視聴者に見えているのかを確認しながらコミュニケーションとることができる。

図 3 に任意の指定場所 (仮想 GPS 値) に VTuber を出現させ、実地側のユーザは異なるスタート地点 (3 地点) で AR アプリケーションを起動し、実際に指定場所まで歩いた際のスマートフォンから見た結果を示す。3 地点ともに指定した場所まで行くと VTuber が現れることが確認できる。今回、VTuber の出現位置の経度と緯度の値を直接入力して指定したが、今後は Google マップでピンを刺すような直感的な方法で出現場所を指定し、任意の位置に出現・移動できることを考えている。また、図 2 の結果では、GPS を用いた位置同期を行っておらず、図 3 の結果では GPS の位置同期は行っているが、視聴者側のスマートフォンでレンダリングされた重畳映像の送受信機能を行っていない。このため、今後は GPS の位置情報に基づいた位置同期・姿勢同期や重畳映像の送受信機能を VTuber 側・視聴者側ですべて有効にした状態で安定した動作を行えるよう改良を行っていく。



(a) A 地点

(b) B 地点

(c) C 地点

図 3: 任意の位置に VTuber を出現させた様子

5. むすび

本研究では、遠隔地の VTuber が視聴者のいる現地に第三者人称視点でテレイグジスタンス可能にし、リアル側の視聴者はスマートフォンを介して VTuber とコミュニケーション可能な実世界指向のライブ配信システムを提案・構築した。

現状、VTuber 側は実世界側の任意の緯度・経度に出現、移動することができ、視聴者のスマートフォンの映像 (VTuber 自身が重畳表示された映像) を見ながら視聴者とジェスチャによるコミュニケーションをとることが可能となっている。VTuber 側の今後の課題として、現地のユーザと同じ場所を表す仮想空間の構築、Google マップを使用した直感的に出現場所を指定することができるシステムの実装を行うこと等があげられる。

また、現状、視聴者側は VTuber が指定した任意の場所にスマートフォンをかざすと VTuber が出現することが可能となっている。視聴者側の今後の課題として、緯度と経度だけでなく AR アプリケーション起動時の方向を考慮した位置同期を行うことや、障害物等の奥行情報をもとにした重畳表示を行うこと等があげられる。

参考文献

- [1] S. Takeuchi, K. Hashiguchi, Y. Homma, K. Kajitani, and S. Meguro: "GIBSON: AR/VR synchronized city walking system," SIGGRAPH Asia 2021 XR (SA '21 XR), No 9, pp.1-2, 2021
- [2] 脇田 航: 簡易没入型全方位 VR 歩行プラットフォームの改良 (第二報), 知覚情報研究会, マルチモーダル応用及び一般, PI-19-16, 2019.