



フォトリアルな バーチャルヒューマンの 研究開発に携わって

加藤晴久[†]

1. まえがき

COVID-19の影響でzoomやteamsなどを活用したリモートワークが注目されていますが、対面との違いも議論されるようになってきました。その違いは同じ空間を共有している感覚が乏しいことに起因しているかもしれません。私が所属している部署ではこの違いを低減し、遠隔に居ながらにして同じ空間を共有するような通信を実現するための研究開発を行っています。目指す世界としては、バーチャルでありながら、リアルの旅行や店舗での買い物を楽しめるような「距離」を感じさせないゼロディスタンスな世界です。そのためには、表情や仕草などで情報伝達ができるフォトリアルなバーチャルヒューマンが自身のデジタルツインとして遠隔会議などで日常的に活躍するような世界観を想定しています。その取り組みの一環としてのバーチャルヒューマンに関する研究開発について紹介したいと思います。

2. 入社してからの経緯

学生のときは画像符号化等について研究しており、1999年にKDDIへ入社してからは希望通り研究所に配属され映像符号化の研究に10年程度従事しました。その後、直感的なユーザインタフェースの研究開発を経て、拡張現

実感 (AR) の研究開発を10年程度担当しています。映像の高効率符号化のために取り組んだ大域的動き補償等が、カメラワーク推定に繋がり映像理解に発展させてARの研究開発に活かされています。ARに関しては、主にスマートフォン向けARアプリ「SATC (サッチ)」等の研究開発を担当しました。SATCの名称は現実世界を認識しさまざまな状況を“察知”する意味合いを込めています。また、ARを扱ったアニメ「電脳コイル」の仮想キャラクターの愛称を連想させるものにもなっており、その世界観を実現する開発者を目指して研究開発を進めていたのを覚えています。2011年前後はSATCで手のひらの姿勢を推定し初音ミクなどの3D CGキャラクターを手の上にAR表示させる機能を開発するなど当時からARおよび3D CGの活用に関心を持っていました。現在、SATCはauのAndroid端末にプリインストールされAR等の研究成果が広く一般に利用されています。研究テーマの選定からアプリのリリース・サポートまで幅広い経験をさせていただいており、これまでの研究内容が結実していると感じています。現在は、AR/VRによるコミュニケーションの一環として過去の研究を応用しながらバーチャルヒューマンの利活用について研究開発を進めています。

3. バーチャルヒューマン の課題

バーチャルヒューマンとは、あたかも



図1 バーチャルヒューマンの一例

実在するかのような見栄えを実現する3D CGモデルのことを指します。図1にバーチャルヒューマンの一例を示します。自然な表情・仕草の再現、インタラクティブな動作を反映するバーチャルヒューマンは、遠隔コミュニケーション、遠隔カウンセリング、ナビゲーション、接客などさまざまなユースケースが想定されます。しかし、フォトリアルなバーチャルヒューマンを描画するにはスマートフォンやスマートグラスでは性能が充分ではないという課題があります。必然的に高性能なサーバで描画・配信することになりますが、サーバでフォトリアルなバーチャルヒューマンを描画しても映

[†]株式会社KDDI総合研究所
"Engaged in the Research and Development of Photo-realistic Virtual Humans" by Haruhisa Kato (KDDI Research Inc., Saitama)

像符号化して配信する場合、高品質を維持するにはネットワークコストの観点で課題がありました。特に、ARで利用者と同一空間に表示させようとするとバーチャルヒューマン以外の背景を透過領域にする必要がありますが、正確な透過領域の再現には数十Mビット/秒の帯域が必要になります。また、ネットワークコストの課題だけでなく、復号処理でもスマートグラスでの負荷が問題になりかねません。処理負荷が高むと伝送遅延に加え処理遅延によりリアルタイムのインタラクティブなコミュニケーションに支障が生じます。したがって、技術的ポイントとしては大容量3D CGの高品質かつリアルタイムな配信が求められます。

4. バーチャルヒューマンの効率的な配信

前述の課題を解決するため、ネットワークコストを抑えながら高品質なバーチャルヒューマンを描画する方法を模索しました。学生時代から入社後10年近くは高効率な映像符号化の研究開発に携わってきたので、映像符号化におけるスケーラブル符号化を応用できないかと考えました。スマートフォンでも比較的low品質ながらも3D CGを描画できることを利用し、サーバとスマートフォンとを協調させて分散レンダリングする方法を提案しています。具体的には、スマートフォンで描画しつつ、スマートフォンの性能では描画しきれない肌の透明感や潤い、瞳の映り込み、柔らかな陰影など高品質な質感をサーバで描画して、不足分だけを映像としてサーバからスマートフォンへ伝送します。スマートフォンでは伝送された映像と描画した画像とを合成することで図2のようなフォトリアルなバーチャルヒューマンを再現します。ネットワークで伝送される情報はスマートフォンでは表現できない不足分なので大幅な情報量削減が達成され、伝送コスト削減と3D CGの高品質化とを両立することを可能としています。特に、スマートフォン側



図2 著者そっくりのバーチャルヒューマン



図3 タブレット端末のカメラで認識した表情をリアルタイムに反映するバーチャルヒューマン

で透過領域を生成することでサーバから透過情報を伝送する必要がなくなるため、透過領域の境界での画質劣化を原理的に回避できるというメリットも併せ持ちます。

リアルタイムで利用する場合は、利用者の表情を画像認識し認識結果だけをサーバとスマートフォンに伝送します。サーバでは不足分だけを描画しスマートフォンへ伝送する一方、スマートフォンでもバーチャルヒューマンを描画し合成します。この場合でもネットワーク上には表情の認識結果と不足分の映像しか伝送されないののでバーチャルヒューマンの効率的な配信を実現できます(図3)。バーチャルヒュー

マンと映像符号化はまったく異なる技術分野ですが、異なる分野の知見を活用することで新たな方式を生み出すことに繋がっています。現在の技術分野の深化だけでなく新たな技術への取り組みと技術の水平展開が差別化には重要になってくると思います。

5. 有効性の評価とアウトプット

有効性を確認するためバーチャルヒューマンの画質や伝送量などの客観的な評価を行うだけでなく、バーチャルヒューマンを用いた遠隔カウンセリングを模擬した主観的な評価(図4)を行い、既存のテレビ会議と比較してど

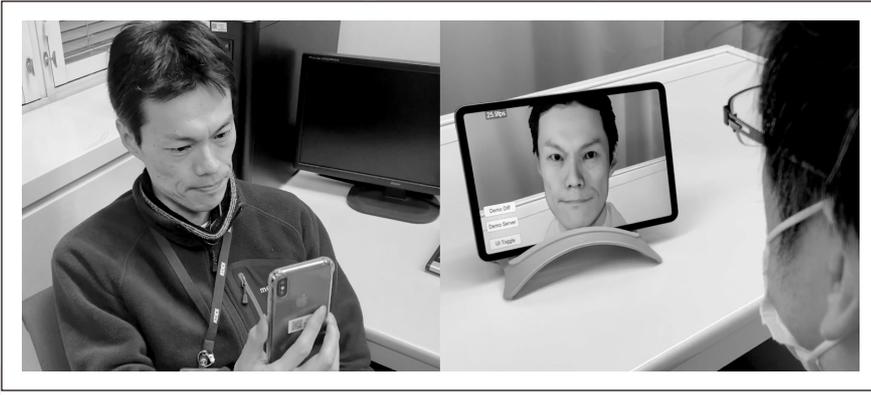


図4 バーチャルヒューマンを用いた遠隔カウンセリングの実験

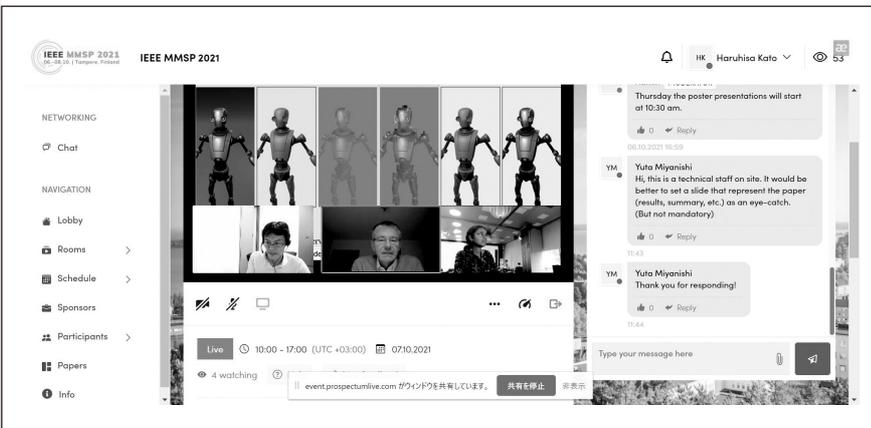


図5 国際学会会議でのオンライン発表

のような効果が得られるのかなどの調査を行っています。その過程で遠隔カウンセリングに取り組む企業との意見交換を通じてバーチャルヒューマンに求められる要件についてアドバイスを受たり、実運用に向けて解決すべき課題について指摘を受けたりするなど協力して研究開発を進めています。また、フォトリアルなバーチャルヒューマンを制作するには、フォトグラメトリーによる全身3Dスキャンや表情アニメーションのためのさまざまな顔形状の生成など専門の3Dモデル制作会社と一緒に品質を高める作業も協力して行っています。

このような方式を提案・実証し有効性を確認したら、特許明細書を執筆・出願し権利化を進めるとともに、論文として仕上げ国際学会(図5)やフルペーパーとして学術的な成果に繋げて

いきます。さらには、バーチャルヒューマンの事業における付加価値を付けた差別化要素としてサービスへの応用を提案していきます。すなわち、課題の発見から解決策の考案、学術的発表、事業化への貢献と全体を俯瞰して仕事に取り組む姿勢と実行が求められています。

6. 他分野との連携

バーチャルヒューマンは自身の動きを反映させるバーチャルYouTuber(VTuber)のような使い方だけでなく、自律的に動作する使い方にも注目が集まっています。自律的なバーチャルヒューマンは感情的なつながりを実現する次世代のチャットボットと位置づけることができます。本物の人間のような容姿と自動対話システムとの連携により自然な応答を実現することで

感情や知性を表現できたり、共感が得られるインタラクションを提供できたりするようになれば、店舗での接客や案内、教育・介護現場での応用や省人化が期待されます。さらには話し合える相手として人に寄り添える存在に成長していくことで社会に受容されていくと考えられます。その実現のためには、自然言語処理技術、自動対話技術、音声合成技術、感情推定技術などさまざまな技術を結集させる必要があります。幸い当研究所にはそれぞれの分野での専門家が在籍しているので、最新の知見を持ち寄りさまざまな技術を連携させながら研究開発を進めていくことができます。また、大学や他企業との共同研究により研究開発を加速していくこともあります。

7. バーチャルヒューマンの将来

2030年代の社会を見据えると、ライフスタイルは暮らしの変化や働き方の変化など遠隔コミュニケーションの高度化が求められてくると考えられます。このとき遠隔の人々が目の前に現れるような場の共有を実現することで、人と人を繋ぐ次世代のコミュニケーション基盤の構築が期待されます。フォトリアルなバーチャルヒューマンによりお互いに自分の容姿のみならず、表情・仕草・音声を空間的に重畳させることで相手の実在感、自分の臨場感、空間への没入感を実現できると考えています。また、自律的なバーチャルヒューマンが常時同行する身近なコンシェルジュとして生活をサポートしている未来が想像されます。このような未来を自分たちの手で創造していくことができる環境にいられるのは研究開発を進めていく上で非常に魅力的で、バーチャルヒューマンにより時間や空間の制約を取り払うゼロディスタンスが当たり前になる世界を築いていくのに学生の皆さんの新しい知見や柔軟な発想を取り入れて発展させていければと思います。(2021年11月30日受付)