

2-5 実在感を伴う遠隔コミュニケーションの探究とその基盤技術の研究開発

2-5 *Investigating Human Cognitive Mechanisms and Developing Key Technologies Towards Remote Communication Conveying the Sense of Reality*

安藤 広志 Juan Liu Michal Joachimczak 内元 清貴

ANDO Hiroshi, LIU Juan, JOACHIMCZAK Michal, and UCHIMOTO Kiyotaka

本稿では、我々が取り組んでいるヒトが感じる実在感の本質の探究と実在感を伴う遠隔コミュニケーションを実現するための基盤技術の研究開発について概説する。当研究室では、空間・時間・身体の制約を超えたコミュニケーションの実現を目指し、実世界の人物や環境をデジタル化してサイバー空間に再構築しそれらを遠隔の人々に多感覚の情報で伝えることにより相互理解の深化やインタラクションの促進を図る技術の研究開発を推進している。本研究開発においては、心理行動・脳機能解析によるヒトの知覚認知の仕組みの解明、実世界の人物・環境のモデル化・理解を行う AI 技術の開発、多感覚の情報をリアルかつ自然に伝える XR (VR/AR/MR) インタフェースの設計・開発を統合して進めている。本稿では特に、我々が現在開発しているリアルで表情豊かな 3D アバターの構築・再現技術 REXR と今後期待される将来の応用展開について述べる。

This paper overviews our investigation on the nature of reality that humans perceive from multisensory information and the development of key technologies towards remote communication with the sense of reality. Our aim is to create communications that transcend the constraints of space, time, and body in order to deepen mutual understanding and promote interactions among distant people. We are conducting human behavioral and brain functional imaging studies to explore human perceptual and cognitive mechanisms, developing AI technologies to digitalize and understand people and environments in the physical world, and designing XR (VR/AR/MR) interfaces to naturally convey reality to remote people via image, sound, touch and other sensory modalities. This paper particularly describes REXR, a technology for constructing and reproducing realistic and expressive 3D avatars, and its expected future applications.

1 まえがき

次世代の情報通信技術 Beyond 5G / 6G の研究開発が世界の国々で始まっている。将来、これまで以上の超高速・大容量、超低遅延、超多接続、超高信頼のネットワークが利用可能になれば、サイバネティック・アバター社会が到来すると期待されている。サイバネティック・アバターとは、自分の 3D 映像やロボットによる分身を意味しており、このようなアバターを遠隔から自在に制御することで、自分自身をまさにどこにでも「瞬間移動」させて、遠隔からでも実空間と同等の活動が可能になると考えられている。このような未来社会では、リアルとバーチャルの情報が高度に融合されて、人々が空間・時間・身体の制約を超えて多様なつながりを生み出し、一人ひとりの活動領域が飛躍

的に拡大すると期待されているのである。NICT が 2021 年に公表した Beyond 5G / 6G ホワイトペーパーでは、このような技術の確立・普及が想定される 2035 年の人々の生活スタイルを具体的に描いている。例えば、ある技術者は、一日の中で「感触を含む五感による遠隔の製品開発会議に参加」、「海外のロボットを遠隔操作してトラブル対応」、「遠くに住む高齢の親と仮想空間の中で昼食を共にしながら食事の介助」、「XR (VR/AR/MR) を用いた遠隔の疑似登山により心身リフレッシュ」、「自動通訳とテレプレゼンス技術による海外との円滑なビジネス交渉」などを一人で楽々こなせるようになるイメージしている。

このようなリアルとバーチャルを高度に融合させる技術は、これまでになかった新たな社会的価値を生み出す。その一つは、レジリエントな社会の実現であり、コロ

ナ禍のような感染症拡大や震災などの自然災害により人々の移動の制約が生じて、サイバー空間を通じて平時と同等の活動を維持可能になれば、経済的・社会的ダメージを最小限に食い止められる可能性がある。もう一つは、労働生産性の飛躍的向上である。我が国をはじめ多くの先進国では、少子高齢化が急速に進み、労働人口の減少による経済・社会の弱体化の危機に直面している。アバターを活用することで一人ひとりの労働生産性を向上できれば、たとえ労働人口が減少しても社会全体の維持・発展が見込めるとともに、親の遠隔介護、子育て、地域活動を行いながらの遠隔就労も可能になり、ライフワークバランスの充実も期待できる。

NICTではこのような未来社会の構築に向けた研究開発を積極的に進めるために、2021年4月、先進的リアリティ技術総合研究室を立ち上げた。当研究室では、空間・時間・身体の制約を超えたコミュニケーションの実現を目指して、実世界のノンバーバルな情報や多感覚の情報を遠隔の人々に伝えることで相互理解を深め人々のインタラクティブな活動を促進するための技術開発を実施している。

本稿の2では、まず当研究室で進めているリアリティ(実在感)を対象とした研究の基本的アプローチについて述べる。次に、現在の研究は前身となる研究室・研究グループにおいて実施してきた研究開発の蓄積を踏まえて実施しているため、3においては、これまでの取組と主な成果について概説する。特に、視覚、聴覚、体性感覚、嗅覚等、ヒトの多感覚認知メカニズムとクロスモーダル知覚に関する研究及び超臨場感コミュニケーションの実現に向けて実施してきた技術開

発について述べる。4では、現中長期計画で実施している研究の狙いとこれまでに得られた主な成果について紹介する。本節では、まず我々が提案する実世界と仮想世界の融合の枠組みについて概説した後、現在進めているリアルで表情豊かな3Dアバターの構築・再現技術に関する研究開発について述べる。5では、今後期待されるこれらの技術の活用領域と将来の応用展開について述べる。最後に、6で本稿をまとめる。

2 リアリティ研究のアプローチ

先進的リアリティ技術総合研究室では、空間・時間・身体の制約を超えたコミュニケーションの実現を目指して、実世界の人や環境をデジタル化してサイバー空間に再構築し、それらを多感覚の情報で遠隔の人に伝えることにより、人々の相互理解の深化とインタラクティブな活動を促進を図る研究開発を進めている。

このような実世界の情報の取得・伝達・再現を実現するための研究開発においては、3つの研究手法を統合しながら進めるアプローチが重要と考える(図1)。第一は、ヒトが多感覚の情報から感じるリアリティ(実在感)の本質をヒトの行動解析や脳機能イメージングにより深く探る研究である。第二は、実世界の人・環境をモデル化し対象の理解や拡張を図るための人工知能(AI)の研究開発である。第三は、遠隔の人々にそれらを映像・音響・感触など多感覚の情報でリアルかつ自然に伝えるXR(VR/AR/MR)インタフェース技術の研究開発である。

第一の「ヒトの行動解析・脳機能イメージング」においては、ヒトはどのような情報からリアリティ(実在

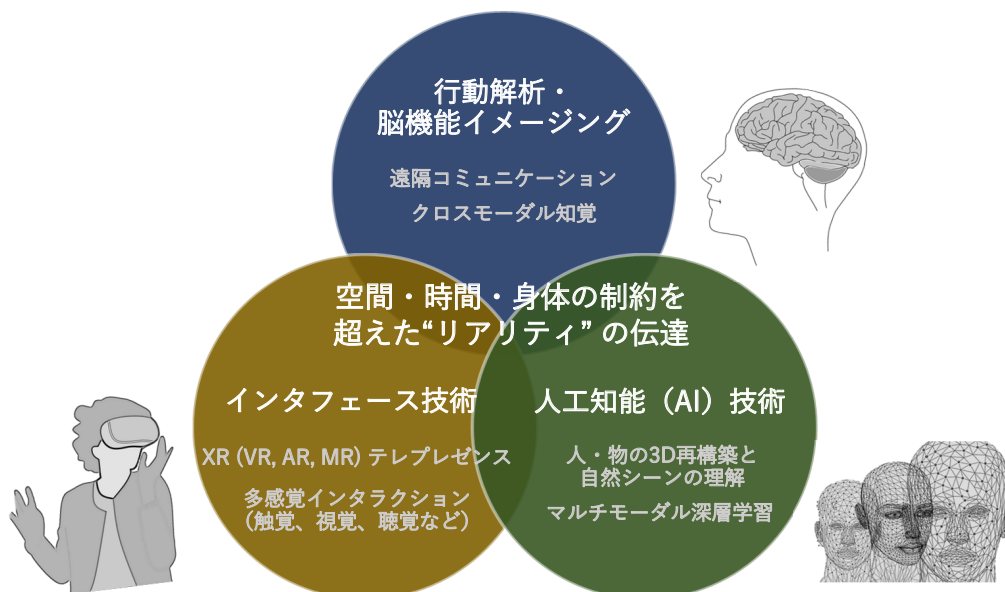


図1 リアリティ(実在感)の取得・伝達・再現を目指した研究開発のアプローチ

感)を感じ取っているのか、実在感をヒトに生じさせるメカニズムを科学的手法により明らかにする。具体的には、心理物理学・fMRI(機能的磁気共鳴撮像法)・脳波計測などの手法を用いて、ヒトの感覚・知覚に関する信頼性の高い計測データを取得・解析し、遠隔のコミュニケーションやインタラクションの促進を可能にするシステム設計に役立てる。特に、多感覚情報のクロスモーダル知覚(感覚間相互作用)に関する知見は、特定の感覚を他の感覚で補完・補強できる可能性を示しており、効率的で効果的なシステム設計の指針を与える。

第二の「AIに基づく実世界の人・環境のモデル化・理解・拡張」においては、実世界の人や環境の3D形状・内部構造、表面特性(テクスチャ・反射率関数)、動的特性等をモデル化し、サイバー空間やリアル空間における情報の再構築・理解・拡張に活用する。ヒトの脳内では、たとえ入力された感覚情報が断片的で不完全であっても、過去の経験で学習・獲得された対象の知識をトップダウン的に用いて入力情報の補完・再構築が行われることが知られている。実世界モデリングのシステム開発においても、機械学習により対象の知識を獲得し入力情報のモデル化に活用することは極めて有効と考える。特に、AIに基づく実世界モデリングは、設置・計測・解析に手間のかかるセンサ群に頼る必要がないため、低コストかつ実時間でモデル構築に威力を発揮すると考えられる。

第三の「XR(VR/AR/MR)インタフェース技術」においては、遠隔からでもリアルかつ自然な多感覚情報を伝え直感的なインタラクションを可能にするインタフェース技術を開発する。ここではヒトの脳や身体に違和感・疲労を与えない自然な情報提示の手法や長時間に渡るトレーニング等を必要とせず直感的に操作できるシステムの実現を目指す。特に、ヒトに対して心的負荷を与えない多感覚情報の時空間提示条件、仮想情報の実空間・実時間からのズレに対するヒトの許容範囲等の知見に基づいてシステムの最適化を図る。

これまでの取組と主な成果

3 (2006年4月～2021年3月)

本節では、現研究室の前身となる研究室・研究グループにおいて実施してきた研究開発における主な成果や社会展開に向けた取組について概説する。NICTにおける5年ごとの中長期計画・組織改編に伴い重点研究項目に変化はあるものの、一連の研究室・研究グループにおいて掲げてきた基本的理念と方法論は一貫しており、本節ではこのようなりアリティ研究のこれまでの取組の概要と主要な成果について述べる。

3.1 超臨場感システムグループにおける取組(2006年4月～2011年3月)

NICTの第2期中長期計画(2006年4月～2011年3月)においては、総務省の「UNS戦略プログラム」(平成17年情報通信審議会答申)において設定された「超臨場感コミュニケーション」(超高精細・立体映像、高臨場感音場再生、触覚・嗅覚を含めた五感通信等)の実現に向けた研究開発を進めるために、「超臨場感システムグループ」と「超臨場感基盤グループ」から成るユニバーサルメディア研究センターがNICT内に発足した。

現研究室のルーツに当たる超臨場感システムグループでは、ヒトが感じる臨場感の指標(構成要素・要因)を設定し、臨場感を客観的・定量的に解析する手法の開発を進めた[1]。特に、広視野3D映像によって生ずる包囲感・没入感等をfMRI脳活動計測に基づいて客観的に捉えるために、MRIの高磁場環境下で水平視角度100度の両眼立体映像を提示可能なシステムを世界で初めて開発し、後頭野から頭頂野の視覚関連領域の活動パターンを明瞭に捉えることに成功した[2]。また、モノの質感の主要な要素である光沢感が両眼視差情報・動的変化情報により増強して感じられることを心理物理学的手法により実証し、高級感・上質感等が求められる製品開発や商品提示においてこのような知見が有効であることを明らかにした[3]。

また、多感覚インタラクション技術の研究開発を進め[4]、モノの立体映像(視覚)、感触(触覚)、接触音(聴覚)、香り(嗅覚)の四感覚を同時に制御可能な多感覚インタラクションシステム(MSenS)を世界で初めて開発した[5]。特に、仮想物体の材質・表面特性、ユーザの触り方に応じた接触音を実時間で生成可能な技術及び香り提示の時空間制御が可能な嗅覚ディスプレイを新規に開発し、本システムに実装した。さらに、携帯型の把持感覚提示デバイスの開発に着手するとともに[6]、画像の臨場感を高める香りの選別手法の提案も行った[7]。多感覚インタラクションシステムに関しては、貴重な文化財のインタラクティブ体験等、多数の多感覚コンテンツを制作し、全米/全欧最大の放送機器展NAB/IBCや最先端IT総合展CEATECなど国内外のイベントに出展、本技術の効果を国内外の多数の人々に体験してもらい高い評価を得た。一方、立体映像に関しては、世界最大の200インチ裸眼立体ディスプレイ[8]及びテーブル型裸眼立体ディスプレイ[9]を開発し、多数のイベントに出展し体験デモを実施した。

さらに、産学官連携による研究開発・実証実験・標準化等の推進を図るために、2007年3月、総務省・NICTが中心となって超臨場感コミュニケーション産学官フォーラム(URCF)を設立し、3D映像や五感に

2 多言語コミュニケーション技術

関する国際ワークショップ・国際会議の主催や企業・大学等と連携した実証実験の実施に努めた [10]。

3.2 多感覚・評価研究室における取組 (2011 年 4 月～2016 年 3 月)

NICT の第 3 期中長期計画 (2011 年 4 月～2016 年 3 月) においては、ユニバーサルコミュニケーション研究所に多感覚・評価研究室が発足し、ヒトが多感覚情報から感じる臨場感を定量的・客観的に評価するための技術開発を更に進めるとともに、自然でリアルな映像・音響・感触・香り提示を達成するための技術要件を導出するために、心理物理・行動実験等によりヒトへのポジティブ／ネガティブ効果を検証した。

立体映像 (視覚) に関しては、成人 (20 歳～69 歳) 500 名及び未成年者 (12 歳～19 歳) 133 名を対象にした「眼鏡あり 3D 映像の視聴が人に与える疲労」に関する大規模評価実験を URCF の活動の一環として実施し、主観的・客観的評価の実験結果を報告書として取りまとめ一般公開した。また、国際標準化団体 ITU-R の WP6C に本実験結果を寄与文書として提出し 2014 年 11 月、レポート BT2293 に採択された [11]。また、立体映像が人に与える不快感・疲労の個人差要因の特定に向けて、眼の調節・輻輳機能等の個人特性が快適視差範囲に与える影響を明らかにする実験を実施し、詳細な実験データを取得した [12]。質感知覚に関しては、多視点立体映像の視点数とクロストーク (光線の重なり) を制御可能なシミュレータを開発して光沢感再現の最適条件を心理物理実験により特定し、多視点立体ディスプレイの設計指針を提案した。

立体音響 (聴覚) に関しては、MVP (Multiple-Vertical-Panning) 方式による大画面裸眼立体ディスプレイ用立体音響実験システムを構築し、ディスプレイの上下に配置したスピーカアレイからの音を制御することで画面上に音像が定位することを心理物理実験で実証した [13]。また、音の頭部伝達関数 (HRTF) におけるスペクトルの基本特性 (ピークとノッチ) を個人ごとに異なる耳介形状の特徴量 (複数の特徴点間の距離) から推定可能であることを音波伝播の計算機シミュレーションにより世界で初めて明らかにし、個々のヒトに最適な立体音響を生成するための個人適応化の基盤技術を確立した [14]。

感触 (触覚) に関しては、親指と人差し指で仮想物体を把持・操作できる力覚提示デバイスを開発し、多感覚インタラクションシステムに組み込んだ。また、3次元空間内で仮想物体の操作を行う作業において、映像と感触の空間的な一致・不一致の条件が作業効率に与える影響を心理物理実験により評価・分析した。さらに、2 台の多感覚システムをネットワークで接続し、

仮想物体の感触情報を通信して他者と共有することに成功した [15]。

香り (嗅覚) に関しては、6 種類の香りを瞬時に切り替えて映像と同期して提示可能な香り噴射装置アロマシューターを開発した。さらに、香りの微細な濃度調整を可能にする技術を開発し、デジタル方式の嗅覚検査の実験システムを構築した。本香り噴射装置に関しては、NICT 発のベンチャー企業を起ち上げ、製品化を行った [16][17]。現在、香りのデジタルサイネージ (ブランディング・集客等) やリラクゼーション空間構築・ストレス軽減などの分野において社会実装が進められている。

さらに、無人の建設機械による遠隔作業の操作性向上を目指して、高精細立体 (4K3D) 非圧縮映像の (光無線) 伝送システムの開発を進め、遠隔作業の作業効率の向上効果を土木研究所の実験施設及び実環境 (雲仙普賢岳の災害復興現場) において行動実験により実証し、遠隔操作の快適性に求められる技術要件をとりまとめた [18]。

3.3 多感覚認知グループにおける取組 (2016 年 4 月～2021 年 3 月)

NICT の第 4 期中長期計画 (2016 年 4 月～2021 年 3 月) においては、脳情報通信融合研究センター (CiNet)・脳機能解析研究室において多感覚認知グループを形成し、ヒトの解析・評価技術の研究開発を更に進めるとともに、脳内の多感覚認知メカニズムやクロスモーダル知覚に関して多くの有益な知見を得た (この期間の文献は [19] にも掲載)。

視覚に関しては、fMRI 脳活動計測実験により、光沢知覚の情報処理に関わるヒトの脳部位を世界で初めて特定するとともに、企業との共同研究により肌の光沢に由来する魅力度に関わる脳活動を捉えることに成功した [20]。これらの知見は、将来、感性価値の高い製品開発等への応用が期待されている。また、独自に開発した fMRI 用広視野立体映像提示装置を用いて、広視野の動的映像が引き起こす自己運動感覚に関わる脳部位として CSv (帯状溝視覚野) を同定し、VR 映像がヒトに与える正負の影響を客観的に評価するための基盤を築いた [21]。

聴覚に関しては、立体音の頭外への空間定位機能を担う脳部位として pSTG (上側頭頭後部) を同定するとともに、音声の定位が発話者の映像位置に引き付けられるクロスモーダル効果を pSTG の脳活動から実証した [22]。さらに、発話者の声質 (voice quality) を規定する声門流 (声帯における気流) を発話情報から逆推定する新たな数理手法を提案した [23]。この手法は、将来、ロボット・AI が声質からヒトの情動・意図を読み

解くための核となる技術として期待されている。

触覚に関しては、体性感覚と聴覚のクロスモーダル効果に関して、接触音が物体表面の剛性(硬さ)知覚に与える影響を心理物理実験により実証するとともに、音と感触に基づく硬さ判断においては、音の材質カテゴリではなく低次の音響特性(周波数・減衰特性)がより強く影響することを明らかにした [24]。また、視覚と触覚に基づいて身体近傍空間に位置する物体形状を認知する仕組みを行動実験により解析し、異種感覚情報は最終判断を行う感覚モダリティにおいて照合されることを明らかにした [25]。このような触覚に関わるクロスモーダル知覚の特性を活用することにより、将来、医療や遠隔作業等の分野において自然なインタラクション操作が可能になると期待される。

嗅覚に関しては、香りが感触知覚(物体表面の硬さ・ざらつき知覚)に与えるクロスモーダル効果を心理物理実験により実証するとともに、嗅覚情報が視覚的動きの知覚にも影響することを心理物理実験と fMRI 実験により示した [26]。また、頭の中だけで行った純粋なメンタル課題(自伝的記憶の想起)を全脳の fMRI データに対する機械学習に基づいて高い精度で推定可能であることを示し、その脳内機序を明らかにした [27]。このような脳活動に基づく心的推定は、将来の高齢者・障害者支援に特に有用であると考えられる。

さらに、MR (mixed reality: 複合現実) テレプレゼンス技術を開発し、その心理的効果を検証した [28][29]。この技術は、人の 3D 形状を奥行きセンサの情報から実時間で再構築して伝送し、遠隔の人が MR ヘッドセットをかけて見ると 3D 伝送映像が実空間の中に重畳して知覚され、相手との会話も実時間で円滑にできるというものである。頭部位置のセンシングにより 3 次元空間内に 3D 映像を高精度で定位させて表示できるため、実空間での対面に近いコミュニケーションが可能となる。一方、本システムでは実空間とは異なり、伝送される映像のサイズ等に変調をかけることも可能になり、それによりユーザの心的状態(緊張感等)を調整できることをヒトの行動実験と生体情報解析により実証した。

また、当研究グループは JST (科学技術推進機構) の地域発研究開発実証拠点プログラム「けいはんなリサーチコンプレックス (RC) (2015 年 12 月～2020 年 3 月)」を受託し、安藤副室長(当時)が研究推進リーダーとして、企業 29 社、研究機関 2 機関、大学・高専 11 機関、団体・自治体 5 機関による研究開発全体の統括を行った [30]。本プログラムでは「*i-Brain* × *ICT* による「超快適」スマート社会の創出」をテーマに人間中心の技術イノベーションの創出を目指し、ウェアラブル心電計測デバイスを用いた乳幼児の生体リズム解析、快

適性と省エネ効果を高度に実現する空調・五感統合制御システムの開発、潜在的な快適性を読み解く脳機能解析技術の開発等、15 件以上の産学連携プロジェクトを創出・実施し、高い評価を得ることができた。

4 現在の取組 (2021 年 4 月～) : 第 5 期中長期計画

第 5 期中長期計画の期間 (2021 年 4 月～2026 年 3 月) においては、ユニバーサルコミュニケーション研究所に先進的リアリティ技術総合研究室が発足し、リアリティ(実在感)の知覚認知の仕組みの解明とリアルとバーチャルの情報を高度に融合させる技術開発を進めている。今期は、前節で概説したこれまでの取組と成果を更に発展させて、遠隔地の人々とも円滑なインタラクションを実時間で行えるようにするための基盤技術の研究開発を行う。特に、現状のオンライン会議では十分に伝えられない非言語情報(表情・視線・ジェスチャー等)を効果的に伝え、複数の人々が 3D 仮想空間を共有して一体感のあるリモートの円卓会議を行えるようにするための基盤技術の研究開発を実施している。本研究では、本人のリアルで表情豊かな 3D アバターを Web カメラ一台の映像から構築しリアルタイムでサイバー空間に再現する技術を開発するとともに、このアバター構築・再現技術と NICT で現在開発が進められている同時通訳システムを連動させることで、海外の人とも同じサイバー空間内でそれぞれの母国語を用いてコミュニケーションが行えるシステムの実現を目指している。また、ヒトの心理・行動・脳機能解析に基づく 3D アバターの視線・姿勢・表情再現の効果検証や共有仮想空間の設計に求められる技術要件の導出等を目指した研究開発を進める。さらに、モノや環境の情報をデジタル化し遠隔の人々ともこれらのサイバー空間を共有し、触覚を含む多感覚情報を用いたインタラクションを直感的に行えるようにするための研究開発も実施していく。

4.1 バーチャル・リアル融合の枠組み

当研究室で目指しているバーチャル(仮想世界)・リアル(実世界)融合の枠組みを図 2 に示す。この枠組みでは、ユーザが用いる 3D アバター(分身)を 2 軸で分類している。縦軸は、実世界指向 (Real World Oriented) vs. 仮想世界指向 (Metaverse Oriented) を表す軸であり、横軸は不特定アバター (Unspecified Avatar) vs. 個人アバター (Personalized Avatar) を表す軸である。まず、これらの 2 軸で表現された 4 つの領域 (Field) A、B、C、D に特徴的なユーザ体験について概説する。

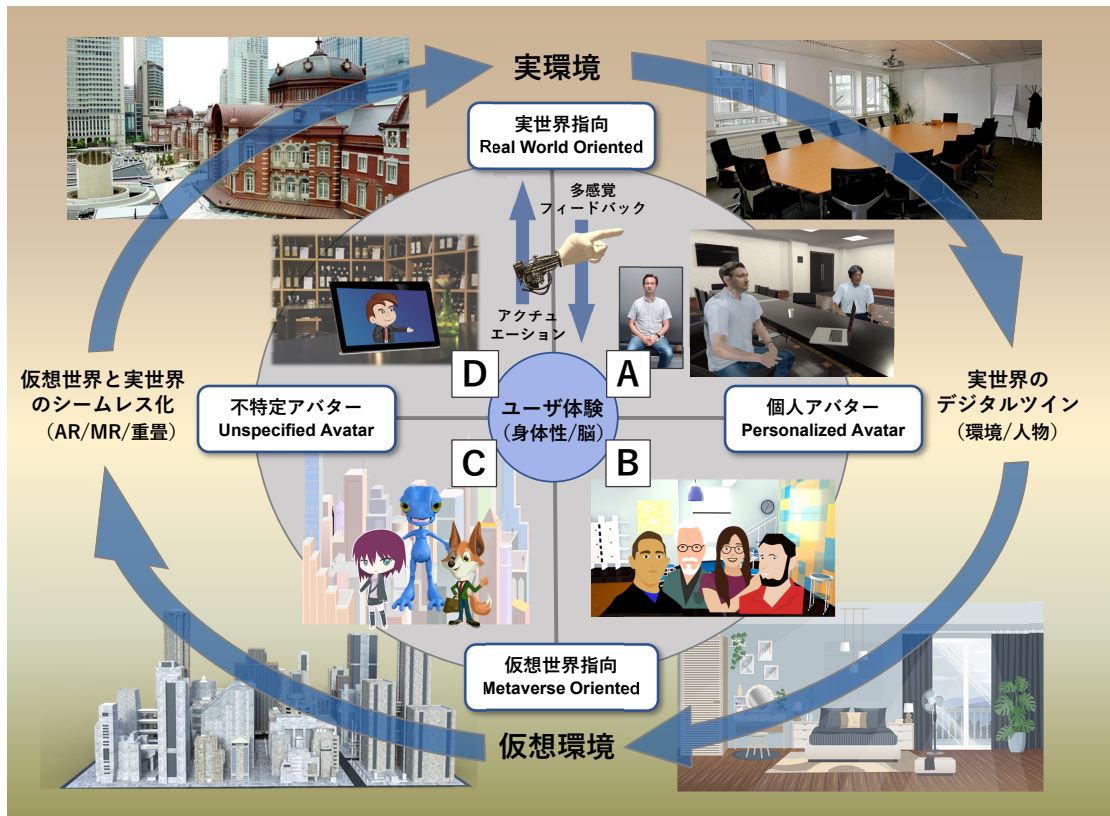


図2 パーチャル(仮想世界)・リアル(実世界)融合の枠組み

まず、領域Cでは、ネット上の仮想空間(イベント等)においてユーザが不特定アバター(CG/アニメの人物・キャラクタ)に変身し匿名で交流を楽しむといった利用形態が考えられ、既に商用化が進んでいる。現在メディアでよく取り上げられる「メタバース」は、このような仮想世界を指していることが多い。この領域には、実世界の制約から解放されて自由な体験や交流をゲーム感覚で楽しめるというメリットがある。本人の身体的特徴や人格とは全く異なるキャラクタで自由に交流できることは、特に身体的/心理的障害を持つ方々にとっては活動の場の拡大につながると考えられる。

領域Dは、本人とは容姿の異なる不特定の3Dアバターを実空間の中で利用するといった使われ方を意味している。例えば、実空間の店舗に設置したモニターの中の3Dアバターやアバター・ロボットを遠隔操作して接客サービスを行うといった利用形態が考えられる。接客では一般に本人の実際の年齢・容姿・人格は重要ではないため、より魅力的な容姿に変身して接客することにより客に好印象を与えられる可能性がある。また、本人とアバターを一対一に対応づける必要がないため、操作者が複数のアバターを遠隔から同時に操作するといった使われ方も想定できる。

領域Bは、ユーザが自分の個人アバターを指定し仮想空間の中で実名を用いて交流をするといった利用形

態を示している。現在商用化が行われているプラットフォームでは、ユーザはあらかじめ用意された3Dアバターの中から自分に容姿が似た3Dアバターを選択したり、3Dアバターの髪形などのパーツを更にカスタマイズしたりして利用することが行われている。将来的には、このような仮想空間においてNFT(非代替性トークン)を用いた新たな経済圏を形成していくことも想定されている。

最後の領域Aは、本人のフォトリアリスティックな3Dアバターを用いて実世界と近い環境の中で遠隔コミュニケーションや遠隔の共同作業を行うといった利用形態を示している。自分自身が別空間に言わば瞬間移動(teleportation)して実空間と同等の活動を行うといったイメージである。ここで用いられるリアルな3Dアバターは、無意識に表出される本人の顔の表情(目元・^{まぶた}視線・眉・口元・頬等の微細な変化)や非言語情報(手・頭部・姿勢の変化等)を豊かに再現できるため、相手の気持ち(関心、賛否、好意等)を理解しやすくなり、親密な意思疎通が可能になると考えられる。NICTの当研究室では、現在、このような3Dアバターの構築と活用を目指した技術の研究開発を進めており、次節でその詳細を述べる。

一方、図2の枠組みには、3Dアバターの異なる利用形態だけでなく、実環境と仮想環境の融合の流れを示している。その一つは、実世界(環境・人物)をデジタ

ル化・モデル化し、仮想環境の中で利用していく(上から下への)流れである。実世界のデジタルツインの構築・活用により実環境と同等の活動が仮想環境においても可能になると期待される。それに対し、仮想環境の情報を実環境の中にシームレスに融合していく(下から上への)流れもある。シームレス化の手法としては、AR(拡張現実)／MR(複合現実)の利用やプロジェクション映像の実空間への重畳等も考えられる。一方、ユーザが遠隔から実世界に対し直接的に働きかけるためには、遠隔操作型アバター・ロボット等のアクチュエータ(ハードウェア)が必要となる。それらを自分の身体の延長として直感的に操作可能にするためには、実環境から多感覚(感触／音響／映像等)のフィードバック情報をユーザに適切に与える必要がある。「脳」が実空間における様々な活動で獲得した「身体性」を壊すことなく、新たなバーチャル・リアル融合世界においても自然かつ直感的な活動を可能にするインタフェースの実現が求められる。当研究室では、このような世界の実現に向けた研究開発を進めていきたいと考えている。

4.2 リアルで表情豊かな 3D アバターの構築・再現技術 REXR

本節では、当研究室で研究開発を進めている 3D アバターの構築・再現技術について概説する [31][32]。現状のオンライン会議では、全員が別々のフレームの中で正面を向いているため、誰が誰に対してどのような気持ちを持っているのかが掴めず、親近感・共感・信頼関係などを形成するのが困難である。一方、実空間でのコミュニケーションにおいては、人々は 3 次元空間の中で向き合っており、表情・視線・動作・姿勢等の非言語情報を用いて互いの気持ちを自然に伝え合うことが可能で、その結果、相互理解を深めることができる。そこで、遠隔であっても実空間のような自然かつ親密なコミュニケーションを実現するために、当研究室では仮想空間を共有して本人の細やかな表情や動作を伝え合うことができる 3D アバターの構築・再現

技術 REXR (Realistic and EXpressive 3D avataR : レクサー) の研究開発を進めている。

REXR には 3 つの特徴がある。第一は、本人の細やかな顔の表情・視線・ジェスチャを入力映像と同程度に精細に 3D アバターで再現可能である点である。これに対し、現状、商用化されているアバターの顔の変化は、ランダムな瞬きやリップシンク(音声に合わせた口の開閉)程度であり、刻々と変化する本人の細やかな表情は十分に再現できてはいない。第二は、多数のカメラや特殊なセンサは不要で、Web カメラ 1 台の映像だけから 3D アバターを構築できる点である。一方、現状、本人の 3D アバターを構築するには、数十台から数百台のカメラを設置したスタジオや特殊な奥行き／位置センサ等が通常必要であり、一般ユーザがこれらを日常生活で利用することは難しい。第三は、構築した 3D アバターを仮想空間内の任意の位置・向きに配置して表情・動作・姿勢を滑らかな動きで自在に再現できる点である。単に身体全体の向きが変えられるだけでなく、各関節を動かして異なる動作や姿勢を取らせることも可能であり、将来の応用範囲は広い。

REXR を用いて正面の Web カメラの映像だけから本人のリアルな 3D アバターを構築・再現した例を図 3 に示す。入力映像だけから、細やかな顔の表情やジェスチャが 3D アバターで再現され、任意の方向から表示可能であることが分かる。さらに図 4 では、本人が表出する怒り、嫌悪、驚き、ほほ笑みなどの多様な感情も 3D アバターの表情で様々な方向から豊かに再現可能であることを示している。

このような諸機能が Web カメラ 1 台の映像だけから達成できるのは、複数の AI モジュール(ヒトに関する知識を機械学習により獲得したニューラルネットワーク)を連動させて用いているからである。すなわち、REXR は、本人の身体の基本モデル、顔の表情、身体の姿勢を映像から再構築するための AI から構成されている。具体的には、REXR は 2 つのステップから成り立っており、Step1 では、Web カメラの前で自ら一回転し異なる方向から身体の 2D 映像を取得し、そ

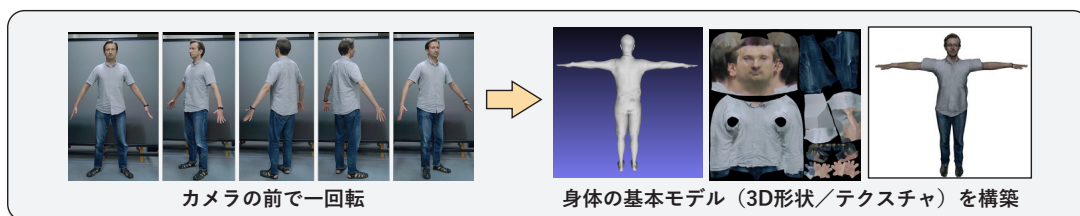


図 3 REXR で構築・再現された本人のリアルで表情豊かな 3D アバター



図4 REXRによる多様な感情表出の再現例(上から怒り、嫌悪、驚き、ほほ笑み)

【Step 1】



【Step 2】



図5 REXRにおける3Dアバター構築・再現の処理ステップ

これらの情報を基に本人の360度フルボディの基本モデルを再構築する(図5上部)。この基本モデルは、一万个以上の頂点から成るメッシュで表現された身体の3D形状、3Dテクスチャ画像(3D形状と対応づけられた服装・肌・顔パーツ等の画像)及び個人の形状の特徴や関節の3D位置等を表す200個近いパラメータから構成されている。次にStep2において、本人が様々

な動作や表情を見せると、入力映像のフレームごとに顔の表情(3D表情パラメータ)と身体の姿勢(3D骨格・関節角パラメータ)を推定し、あらかじめ構築しておいた身体の基本モデルをこれらの情報をもとに更新するとともに、各フレームで推定された顔の3Dテクスチャ画像と身体の3Dテクスチャ画像を統合する。この3Dアバターを仮想空間上に任意の向きで配置すれ

ば、刻々と変化する本人の表情や身体動作・姿勢をどの方向からでも表示・再生させることが可能となる(図5下部)。

5 将来の展開

前節で概説したREXRを用いると、本人が無意識に表出する細やかな表情(微表情: micro-expressions)やジェスチャをどの方向からでも入力映像と同程度に精細に3Dアバターで再現できるため、相手の心の機微(微妙な感情変化・意図など)も捉えやすくなると考えられる。特に、細やかな表情変化は相手に対する同意・誠実さ・好意等の社会的関係性の情報を豊富に含むため、複数の人の間の円滑なコミュニケーションを成立させるためには極めて重要な情報となり得る。

REXRの将来の応用としては、複数人が参加するオンライン会議での利用が挙げられる。REXRを用いるとリモートであっても仮想空間の中で各人が互いに顔を見合わせた対話が可能となり、現在のオンライン会議システムでは困難な相互理解の深化が図れると期待される(図6、動画は[32]参照)。例えば、初対面の遠隔ミーティングであっても、互いの細やかな表情が伝わり信頼関係を構築しやすくなると考えられる。また、微妙な表情変化や無意識の動作から相手の真意を汲み取ることができるため、シビアなビジネス交渉も円滑に進められると期待される。今後は、複数の人々が仮想空間内で「一体感」を感じられる自然な遠隔コミュニケーションの実現に向けて、3Dアバター構築の精度向上や処理の高速化(リアルタイム対応)の技術開発に取り組んでいく予定である。また、3Dアバターを用いたコミュニケーションが人々の一体感等の形成に与える効果を心理行動実験や脳活動計測実験により検証していく予定である。

一方、3Dアバターのリアリティが向上し本人の実映像と区別が付かなくなってくると、他人による「なりすまし」のリスクも生じてくる。本人が気づかないうちに、他人が自分のアバターになりすまして悪用されてしまうリスクを避けるためには、電子透かし技術の活用や本人認証の仕組みの導入等が必要になってくると考えられる。さらに、図2に示した仮想世界(メタバース)指向の領域も含めて考えると、アバターや仮想物体の著作権、肖像権、仮想空間内での傷害・事故に対する責任、ユーザの心身への影響・負荷等についても対策を検討していく必要がある。このような倫理的・法的・社会的課題(ELSI)に関しては、URCF(超臨場感コミュニケーション産学官フォーラム)のメンバー(企業・研究者)とも連携して検討を進め、本技術の社会展開に取り組んでいきたい。

6 おわりに

本稿では、先進的リアリティ技術総合研究室で実施している実在感を伴う遠隔コミュニケーションの実現に向けた研究開発について概説した。まず、本研究の基本的アプローチについて述べ、これまでに取り組んできたヒトの多感覚認知メカニズムの解明と超臨場感コミュニケーション技術の研究開発に関してその概要を述べた。次に、これまでの取組を踏まえて実世界と仮想世界を融合する新たな枠組みを提案するとともに、当研究室で現在開発を進めている「リアルで表情豊かな3Dアバターの構築・再現技術REXR」と今後期待される将来の応用展開について概説した。

新型コロナ禍は世界の人々に社会活動の制限を強いた一方、リモートでの活動の重要性と可能性を実体験として認識させた。今後は、世の中がリモート活動を積極的に活用したハイブリッド社会へと進む中で、「次

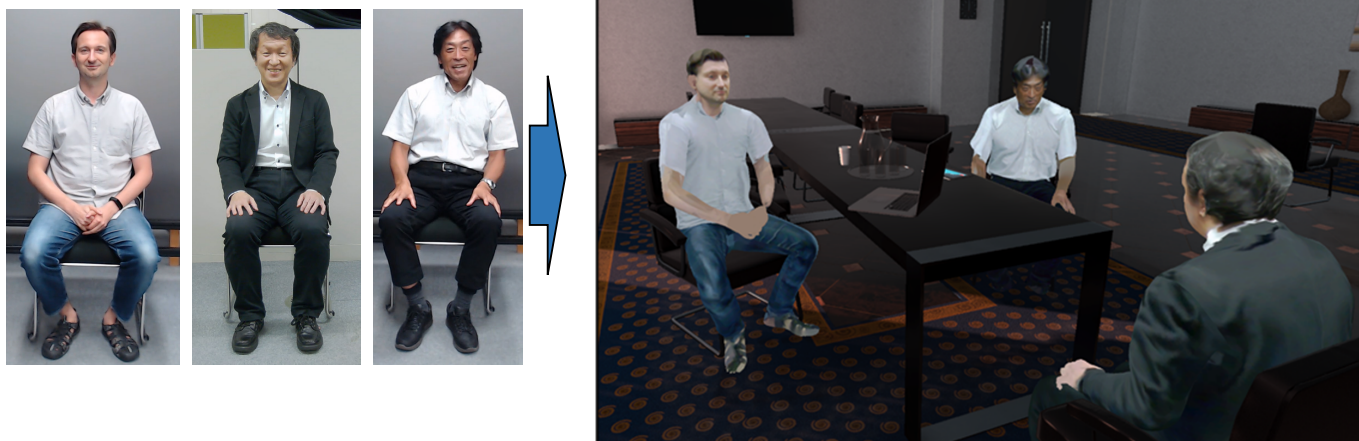


図6 将来のREXRの活用: 複数ユーザの3Dアバターによる遠隔ミーティング(左: 正面カメラからの入力映像、右: 仮想空間を共有したコミュニケーション)

世代の情報通信技術 Beyond 5G / 6」と「実世界と仮想世界の高度な融合技術」の実現に対する期待が一層高まると考えられる。我々はこのような未来社会の実現に寄与すべくリアリティ研究を更に推進していきたいと考えている。

【参考文献】

- 1 安藤 広志, カラン 明子, Norberto Eiji Nawa, 西野 友里恵, Juan Liu, 和田 充史, 坂野 雄一, “臨場感の知覚認知メカニズムと評価技術,” 情報通信研究機構季報, vol.56 nos.1/2, pp.157-165, 2010.
- 2 安藤 広志, 和田 充史, 坂野 雄一:情報通信研究機構プレスリリース「脳活動計測による 3D 映像評価装置を開発」2010年 11月 1日 <https://www.nict.go.jp/press/2010/11/01-1.html>
- 3 Y. Sakano and H. Ando, “Effects of head motion and stereo viewing on perceived glossiness,” *Journal of Vision*, 10 (9):15, pp.1-14, 2010.
- 4 安藤 広志, Juan Liu, Dong Wook Kim, “多感覚インタラクティブ技術とシステム応用,” 情報通信研究機構季報, vol.56 nos.1/2, pp.147-155, 2010.
- 5 安藤 広志, Dong Wook Kim, Juan Liu:情報通信研究機構プレスリリース「四つの感覚を統合した多感覚インタラクティブシステムの開発に成功」2009年 11月 4日 <https://www.nict.go.jp/press/2009/11/04-1.html>
- 6 中山 功一, 安藤 広志, “硬さ弁別能力の解明と携帯型力覚デバイスの研究開発,” 情報通信研究機構季報, vol.56 nos.1/2, pp.119-136, 2010.
- 7 大島 千佳, 中山 功一, 安藤 広志, “画像の臨場感を高める香りに関する研究,” 情報通信研究機構季報, vol.56 nos.1/2, pp.137-146, 2010.
- 8 安藤 広志, 河北 真宏, 岩澤 昭一郎:情報通信研究機構プレスリリース「世界初「200インチの自然な裸眼立体表示技術」の開発に成功」2011年 1月 25日 <https://www.nict.go.jp/press/2011/01/25-1.html>
- 9 吉田 俊介:情報通信研究機構プレスリリース「多人数で観察できるテール型裸眼立体ディスプレイの開発に成功」2010年 7月 1日
- 10 超臨場感コミュニケーション産学官フォーラム (URCF) <https://www.urcf.jp/>
- 11 Report ITU-R BT.2293-1 “Principles for the comfortable viewing of stereoscopic three-dimensional television (3DTV) images,” 2014.
- 12 H. Mizushima, J. Nakamura, Y. Takaki, H. Ando. Super multi-view 3D displays reduce conflict between accommodative and vergence responses. *Journal of the Society for Information Display*. 24/12 pp.747-756, 2016, <https://doi.org/10.1002/jsid.520>
- 13 木村 敏幸, 安藤 広志, “Multiple Vertical Panning を用いた立体音響システムにおいて音源位置の離散化が臨場感に及ぼす影響,” 日本バーチャルリアリティ学会論文誌 vol.20, no.3, 2015.
- 14 P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, “Frequency and amplitude estimation of the first peak of head-related transfer functions from pinna anthropometry,” *Journal of the Acoustical Society of America (JASA)*, vol.137 no.2, pp.690-701, 2015.
- 15 安藤 広志, “五感情報通信” 感覚デバイス開発～機器が担うヒト感覚の生成・拡張・代替技術～ 第2編 第2章 (NTS) ISBN 978-4-86469-064-5, 2014.
- 16 D.W. Kim and H. Ando, “Development of directional olfactory display,” *Proc. VRCAI (Virtual Reality Continuum and its Applications in Industry)*, ACM SIGGRAPH, 2010.
- 17 NICT 発ベンチャー (株)アロマジョイン (2012年 10月設立) <https://aromajoin.com/>
- 18 伊藤 禎宣, 坂野 雄一, 藤野 健一, 安藤 広志, “無人化施工において遠隔操作の映像環境が作業効率へ与える影響について” 土木学会論文集 F3 (土木情報学), vol.73, no.1, pp.15-24, 2017.
- 19 安藤 広志, 和田 充史, 坂野 雄一, カラン 明子, Parham Mokhtari, Juan Liu, 西野 由利恵, 對馬 淑亮, Norberto E. Nawa, Daniel Callan, Michal Joachimczak, “多感覚情報処理の脳・認知メカニズムの解明とその応用,” 情報通信研究機構研究報告, 64巻 1号, pp.39-49, 2018, https://doi.org/10.24812/nictkenkyuhoukouku.64.1_39
- 20 Y. Sakano, A. Wada, H. Ikeda, Y. Saheki, K. Tagai, and H. Ando, “Human brain activity reflecting facial attractiveness from skin reflection,” *Scientific Reports* 11:3412, 2021, <https://doi.org/10.1038/s41598-021-82601-w>
- 21 A. Wada, Y. Sakano, and H. Ando, “Differential responses to a visual self-motion signal in human medial cortical regions revealed by wide-view stimulation,” *Frontiers in Psychology*, 7:309, 2016, doi: 10.3389/fpsyg.2016.00309.
- 22 A. Callan, D. Callan, and H. Ando, “An fMRI study of the ventriloquism effect,” *Cerebral Cortex* 25, pp.4248-4258, 2015.
- 23 P. Mokhtari, B. Story, P. Alku, and H. Ando, “Estimation of the glottal flow from speech pressure signals: Evaluation of three variants of iterative adaptive inverse filtering using computational physical modelling of voice production,” *Speech Communication*.104, pp.24-38, 2018, <https://doi.org/10.1016/j.specom.2018.09.005>
- 24 J. Liu and H. Ando, “Metal sounds stiffer than drums for ears, but not always for hands: Low-level auditory features affect multisensory stiffness perception more than high-level categorical information,” *PLoS ONE* 11: e0167023, 2016, doi:10.1371/journal.pone.0167023.
- 25 J. Liu and H. Ando, “Response modality vs. target modality: sensory transformations and comparisons in cross-modal slant matching tasks,” *Scientific Reports* 8, Article number: 11068, 2018, doi: 10.1038/s41598-018-29375-w.
- 26 Y. Tsushima, Y. Nishino, and H. Ando, “Olfactory stimulation modulates visual perception without training,” *Frontiers in Neuroscience*, 2021, <https://doi.org/10.3389/fnins.2021.642584>
- 27 N. E. Nawa and H. Ando, “Effective connectivity within the ventromedial prefrontal cortex-hippocampus-amygdala network during the elaboration of emotional autobiographical memories,” *NeuroImage*, 189, pp.316-328, 2019, <https://doi.org/10.1016/j.neuroimage.2019.01.042>
- 28 M. Joachimczak, J. Liu, and H. Ando, “Real-time mixed-reality telepresence via 3D reconstruction with HoloLens and commodity depth sensors,” *Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*, 2017, <https://doi.org/10.1145/3136755.3143031>
- 29 M. Joachimczak, J. Liu, and H. Ando, “Effects of the size of mixed-reality person representations on stress and presence in telecommunication,” *International Journal of Semantic Computing* 13, no.03 :311-28, 2019, <https://doi.org/10.1142/S1793351X19400130>
- 30 <https://www.jst.go.jp/rc/index.html> , <https://keihanna-rc.jp/>
- 31 M. Joachimczak, J. Liu, and H. Ando, “Creating 3D personal avatars with high quality facial expressions for telecommunication and telepresence,” *Proceedings of IEEE Conference on Virtual Reality (IEEE VR 2022)*, 2022, <https://doi.org/10.1109/VRW55335.2022.00279>
- 32 安藤 広志, Juan Liu, Michal Joachimczak : 情報通信研究機構プレスリリース「カメラ 1 台の映像から自分の 3D アバターを構築し表情や動作を豊かに再現」2022年 3月 14日 <https://www.nict.go.jp/press/2022/03/14-1.html>
動画 : <https://youtu.be/LuqhKmkAEag>



安藤 広志 (あんどう ひろし)

ユニバーサルコミュニケーション研究所
先進的リアリティ技術総合研究室
上席研究員/
未来 ICT 研究所
脳情報通信融合研究センター
脳機能解析研究室
研究マネージャー
Ph.D.
認知脳科学、計算神経科学、多感覚インタフェース、機械学習
【受賞歴】
2016年 建設ロボット研究連絡協議会 2016年建設ロボットシンポジウム優秀論文賞
2013年 The 5th International Conference on 3D Systems and Applications (3DSA2013), Best Paper Award
2008年 IEEE Human System Interaction Conference 2008, The Best Paper Award



内元 清貴 (うちもと きよたか)

ユニバーサルコミュニケーション研究所
研究所長/
先進的音声翻訳研究開発推進センター
研究開発推進センター長/
先進的リアリティ技術総合研究室
室長
博士(情報学)
自然言語処理
【受賞歴】
2001年 情報処理学会 2001年度(平成13年度)山下記念研究賞



Juan Liu (ジュエン リュウ)

ユニバーサルコミュニケーション研究所
先進的リアリティ技術総合研究室
主任研究員/
未来 ICT 研究所
脳情報通信融合研究センター
脳機能解析研究室
主任研究員
Ph.D.
多感覚認知と行動、ヒューマンインタフェース、機械学習
【受賞歴】
2008年 IEEE Human System Interaction Conference 2008, The Best Paper Award



Michal Joachimczak (ミハウ ヨアヒムチャク)

ユニバーサルコミュニケーション研究所
先進的リアリティ技術総合研究室
主任研究技術員
Ph.D.
Mixed reality telepresence,
3D reconstruction, Deep learning
【受賞歴】
2016年 Genetic and Evolutionary Computation Conference (GECCO), Virtual Creatures Video Competition Winner
2015年 International Symposium on Artificial Life and Robotics (AROB), Young Author Award