

全周囲裸眼立体ディスプレイを用いた サイバースペースへのテレイグジスタンス

渡邊 孝一^{*1}

南澤 孝太^{*1}

新居 英明^{*2}

舘 暲^{*1}

Telexistence into cyberspace using an immersive auto-stereoscopic display

Kouichi Watanabe^{*1} Kouta Minamizawa^{*1} Hideaki Nii^{*2} and Susumu Tachi^{*1}

Abstract – We propose a mutual communication system in which several users can join the same cyberspace by telexistence. We focused on immersion into space, face-to-face communication, physical interaction with the environment, and real-time communication to construct a system based on these factors in order to solve the problems of conventional systems. The system is based on TWISTER, which is an immersive auto-stereoscopic display, and it consists of a user's face image and systems for motion capture, sound and network communication. Two users in separate places participated in shared cyberspace by using CG avatars that had actual images of the users' faces, allowing them to see each other face-to-face and talk with gestures. Thus, two users can enter the same cyberspace and communicate face-to-face with detailed embodiment.

Keywords : telexistence, telepresence, TWISTER, cyberspace

1 はじめに

近年、遠隔コミュニケーション技術の進展により、遠隔地と相互にやり取りする感覚情報の質が向上し量が増大している。市販されている遠隔ビデオ会議システム [1] は、感覚情報として視覚情報と聴覚情報を相互に伝達するが、遠隔地から会議に参加する参加者は実際の会場にいる参加者と、その場にいる感覚で議論し合えるまでには至っていない。特に、遠隔からの参加者と実際の会場にいる参加者との間に空間的な隔りがあり、対等な条件で参加ができない。この不平等な状態は、コミュニケーションを行う空間を実空間と同等な質で共有できていないことから生じる。

遠隔コミュニケーションにおいて、遠隔参加者は実際の会場を映像や音声を利用して共有するが、平面ディスプレイやプロジェクタを利用した遠隔ビデオ会議ではそれらが空間的に広がりを持っておらず、会場へ視聴覚的に没入するための臨場感が足りない。そのため、遠隔参加者は実際の会場にいる参加者と共にその会場にいる、という感覚を持ってない。また、遠隔参加者は映像を通じて表情やジェスチャーを伝えることが可能ではあるが、それぞれが存在する空間が異なるため空間的位置関係が一致せず、相手と向き合って話しているかどうかかわからない。このため、相手の視線までを含めた対面コミュニケーションができてい

とは言えない。加えて、会場との空間的な隔りにより遠隔参加者は自身の身体性を実際の会場内で持たず、伝えられるジェスチャーに限られる。例えば、実際の会場内にある物を指し示すときは、遠くからの指差だけでは伝わらないことが多く、実際に物のそばに行き直接指差す動作が必要となるが、遠隔参加者はそれを行うことが出来ない。これらの問題を解決することで、実空間と同等なコミュニケーションの場の共有が成立し、参加者同士が対等な条件での遠隔コミュニケーションが実現できる。そこで、会議だけでなくパーティなども含めた多人数が参加する遠隔コミュニケーションを想定し、全ての参加者が対等な条件でコミュニケーション空間に参加し、空間を共有できる環境の構築を目指す。

遠隔コミュニケーションにおけるコミュニケーションを行う「場」の定義として、「場の移動」と「場の接続」の2つの分類が文献 [2] にて提起されている。この定義に基づき、「場の移動」によるコミュニケーションにおいて、場をバーチャル空間、即ちサイバースペースとして設けて、全ての参加者がサイバースペースを共有することで対等な条件においてコミュニケーションを行うことが可能となる。このサイバースペースに、臨場感と存在感と身体性を兼ね備えた遠隔作業/遠隔コミュニケーションのための技術概念であるテレイグジスタンスを適用することで、参加者が身体性をサイバースペース内に持つことが出来る。図1にテレイグジスタンスを利用した場の移動による、バーチャル空間を共有した対等な遠隔コミュニケーションの概念を

^{*1}慶應義塾大学

^{*2}National University of Singapore

^{*1}Keio University

^{*2}National University of Singapore

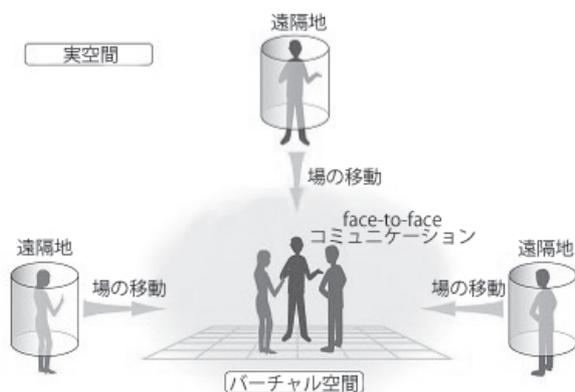


図1 バーチャル空間（サイバースペース）を共有する遠隔コミュニケーションの概念図 [2]
Fig.1 Concept of the remote communication system using VR space (cyberspace).

示す。全ての参加者がサイバースペース内に参加者の代理となるアバターを用意し、テレプレゼンスすることで、音声や映像だけでなく身体性を兼ね備えた高臨場感遠隔コミュニケーションが実現可能となる。即ち、多くの参加者がテレプレゼンスにより同一のサイバースペースに場を移動し、サイバースペースを共有することにより、多人数で対等高臨場感遠隔コミュニケーションを行うことが可能となる。

本研究では、多人数で対等高臨場感遠隔コミュニケーションを実現するため、サイバースペースをコミュニケーションの場として共有する遠隔コミュニケーションシステムを設計構築する。システムに必要な要件を検討し、全ての要件を満たす統合システムを構築し有用性を検証する。

2 サイバースペースを共有する遠隔コミュニケーションシステムの設計要件

サイバースペースを共有する遠隔コミュニケーションシステムを構築するため、システムの目標を掲げて設計要件を検討する。

2.1 提案する遠隔コミュニケーションシステムの目標

サイバースペースを共有し円滑にコミュニケーションを行うためには、まずサイバースペースに入り込んだ時に自身の周囲全方向に広がるサイバースペースに没入して空間内を自由に見渡し、サイバースペースの環境の音を実際の環境と同様に方向性を持って体感できる臨場感が必要である。

サイバースペース内で他の参加者とコミュニケーションをするときは、相手が誰であるかを知るための顔と声の情報が重要である。顔は人の考えや行動の心理状態が現れやすいと言われており [3]、コミュニケーションにおける意思伝達には顔から表出する表情が重要な要素である。表情だけでなく、相手と向き合い、

相手の向きに応じた顔が見え、相手の視線まで正確に伝えなければ円滑なコミュニケーションは行えない。視線はコミュニケーションに大きな影響を与えることが生理学的 [4]、心理学的 [5] に知られており、相手の視線と自分の視線が一致する対面コミュニケーションを満たす必要がある。

サイバースペース内にいることを体感するには、サイバースペース内で思うままに身体を動かし、その動かしした身体を視覚的に、体性感覚的に自身のものとして感じられなければならない。また、サイバースペースにおける臨場感という観点で見ると、臨場感とは個々の感覚が個別に作用するのではなく、身体と外界との相互作用における時間的変化、そこから生じる自分自身の身体認識やコミュニケーション対象に対する心理状態が総合的に作用して生じるものであると言われており [6]、身体の動きに応じた環境とのインタラクションが実現されなければならない。さらに、身体性は相手に意思を伝える際にも重要である。人とのコミュニケーションにおいては声による言語コミュニケーションだけでなく表情やジェスチャーによる意思伝達からなる非言語コミュニケーションが占める割合が多いとされており [7]、空間内で身体性を持つことはジェスチャー等の質を高めることにつながる。即ち、サイバースペースの中においても参加者が自身の身体を認識し、環境に働きかけ、相手に意思伝達するための身体性が確保されている必要がある。

これらの要望をまとめ、サイバースペースを共有する遠隔コミュニケーションにおける目標を掲げる。

目標 1 サイバースペースで参加者が視聴覚的に高い臨場感を有する

目標 2 参加者同士の視線が一致する対面コミュニケーションが可能

目標 3 サイバースペースで参加者が身体性を有する

2.2 提案する遠隔コミュニケーションシステムの設計要件

3つの目標に対して、システムを構築するための具体的な要件を検討する。

目標 1「サイバースペースで参加者が視聴覚的に高い臨場感を有する」に対する要件

コンピュータにおいて生成したサイバースペースに入り込むには、映像を3次元的な空間広がりを持って提示しなければならない（要件 1-1）。人が空間に没入するためには、人が持つ視野全体に空間を提示することが望ましい。人の視野は両眼で周辺視野を含めて水平視野で 200 度程度と言われており [8]、水平視野 200 度以上を満たす映像提示システムによる映像空間の提示が必要である（要件 1-2）。また、空間における

音場は、音源と参加者の位置関係に適切となるように形成される必要がある。よって、参加者の全周囲 360 度をカバーする音響提示システムによる音空間の提示が必要である（要件 1-3）。映像空間は、参加している参加者や環境の要因によって常に変化すべきであるが、その変化の更新レートはコミュニケーションを阻害しない値でなければならない。一般的に、30 fps の更新レートがあれば映像がなめらかに知覚できると言われており、本システムでも 30 fps 以上の更新レートを要件とする（要件 1-4）。以上より、目標 1 に対する要件は以下の 4 つとなる。

要件 1-1 空間が 3 次元映像として提示されること

要件 1-2 水平視野角 200 度以上の映像の提示が可能であること

要件 1-3 参加者の全周囲 360 度に音響の提示が可能であること

要件 1-4 映像の更新レートが 30 fps 以上であること

目標 2「参加者同士の視線が一致する対面コミュニケーションが可能」に対する要件

対面コミュニケーションを行う上で、相手を認識するための情報の一つである声の伝達は必須である。自身の声がコミュニケーションをする全ての相手に、コミュニケーションする全ての人の声が自身に伝わるよう、相互伝達できなければならない（要件 2-1）。相手を認識し、相手の表情から意思を汲み取るための顔の情報として、できるだけ実空間での対面コミュニケーションで見る顔と同等の顔を提示すべきである。アバターの顔として CG など人工的に作成された顔を用いる方法では、心理状態まで含めた意図を表現することは難しいため、本システムではカメラにより相手の目や口を含んだ裸眼の実映像を取得し、サイバースペースに反映させることを要件とする（要件 2-2）。実映像を用いることの 1 つの特徴として、目の動き、即ち視線が認識できる場所にある。しかし、自身と相手の視線を一致させるには、映像の目の動きが視認出来るだけでは不十分である。自身と相手の空間的位置関係に従って正しくアバターが配置され、アバターの目の位置からの顔の実映像が相手に伝わらなければ、視線の一致は満たされない。即ち、相手に送るための自身の顔の実映像を撮像するカメラは相手の目を代行するものであり、その位置が相手アバターの空間的位置と一致することを要件とする（要件 2-3）。以上より、目標 2 に対する要件は以下の 3 つとなる。

要件 2-1 声の相互伝達が可能であること

要件 2-2 表情を含む相手の顔の 3 次元実映像をカメラで取得できること（表情による意思伝達の実現）

要件 2-3 顔映像取得カメラと相手アバターの空間的位置が一致すること（視線が一致するコミュニケーションの実現）

目標 3「サイバースペースで参加者が身体性を有する」に対する要件

サイバースペースにおいて身体性を有するためには、参加者の代理となるアバターが人と同等なサイズで、人が持つ身体部位を有していなければならない。頭、胴、腕、脚の 4 部位は最低限必要であり、作業やジェスチャーにおいては腕の中でさらに手指も必要である。アバターはこの 5 部位を、人間と同等の大きさや自由度で有していなければならない（要件 3-1）。身体性によるジェスチャーや環境とのインタラクションを行うには、サイバースペース内を動きまわり、オブジェクトに対して働きかける必要がある。これを実現するには、アバターの腕と手指を参加者が直感的に思うままに動かせるべきであり、参加者の腕と手指の位置姿勢を取得しアバターに直接的に反映する必要がある（要件 3-2）。本システムにおいては、環境とのインタラクションとして物の把持と移動などを対象とするが、その際に実空間においては生じる触覚情報の再現までは対象としない。腕と手指の動きだけでなく、アバターの脚部は操作可能な形で提供され、参加者がアバターを空間内で動かすことが出来なければならない（要件 3-3）。移動は平面上の歩行移動を仮定し、ジャンプなど垂直方向への移動はここでは考慮しない。身体性を共有するサイバースペースに反映させる際に重要なことは、自身の動きがサイバースペースに遅れなく表現され、相手の動きも遅れなく同期することである。サイバースペースは要件 1-1 及び 1-2 に従い映像空間として提示されるため、映像の通信遅延の許容要件を定める必要がある。視聴覚の通信遅延として、運動視の表示遅れの許容時間が 120 ms であるという実験結果が挙げられているため [9]、本システムではアバターの位置姿勢情報が遅延 120 ms 以内で相互伝達できることを要件とする（要件 3-4）。以上より、目標 3 に対する要件は以下の 4 つとなる。

要件 3-1 アバターが人と同等な身体部位（頭、腕、手指、胴、脚）を有すること

要件 3-2 腕と手指の 3 次元位置姿勢が取得できアバターに反映できること（直感的な身体性の反映）

要件 3-3 アバターの 2 次元平面上の移動操作が可能であること

要件 3-4 相手アバターの位置姿勢が遅延 120 ms 以内で相互伝達できること

本研究では、要件 1-1 から要件 3-4 までの全ての要件を満たす、統合遠隔コミュニケーションシステムを

提案する。各要件を満たすよう要素システムの構築を行い、統合システムを構築し評価することで有用性を示す。

2.3 従来研究

サイバースペースを共有する遠隔コミュニケーションにおける従来研究として、例えば Second Life[10] や there[11] は、一般提供されているサイバースペースを共有する遠隔コミュニケーションシステムの一例である。参加者はアバターを介してサイバースペースを共有し、文字や発話によるコミュニケーションやアバターの動きを介したコミュニケーションを行うことが可能である。しかしながら、操作は基本的にキーボードやマウスによるものであり、アバターは身体を有しているものの、ユーザの身体行動を直接反映したものでないため要件 3-2 の直感的な身体性に基づく環境とのインタラクションが行えない。また、アバターの顔が CG であり、表情としては予め用意されたパターンしか表現できないため、要件 2-2 に基づく表情による意思疎通や要件 2-3 に基づく視線の一致を伴う対面コミュニケーションは実現できない。従って、本研究で示す要件のうち一部しか満足していない。また、没入型空間提示装置である CAVE[12] を用いたビデオアバターシステム [13] は、セカンドライフとは異なり要件 3-2 に基づく直感的な身体性の反映が行われており、行動に伴う環境の改変やジェスチャーなどを用いたコミュニケーションが可能である。しかし、CAVE における 3 次元映像を体験するためにはシャッターメガネを利用することが条件であり、相手の目の動きを含んだ表情の取得が行えない。そのため要件 2-2 に基づく表情による意思疎通が不十分であり、本研究で示す要件のうち一部が満たされていない。

TELESAR4 システム [2] は、本論文で示す要件を満たすシステムではあるが実空間をコミュニケーションの場として共有する手法であり、参加者間の対等性が損なわれる。

他にも文献 [14] の 3 章 3 節ビデオ会議システム及び 6 章 2 節共有仮想空間において、既存の遠隔コミュニケーションシステムがいくつか挙げられているが、各要件のうちいずれかのみが実現されるに留まっており、本論文で掲げる全ての要件を満たす多人数が対等な遠隔コミュニケーションシステムは無い。

3 サイバースペースを共有する遠隔コミュニケーションシステムの設計

3.1 全周囲 3 次元映像音響提示システムの設計

要件 1-1 及び要件 1-2 に対し、サイバースペースの空間映像を提示するディスプレイとして TWISTER (Telexistence Wide-angle Immersive Stereoscope) [15]

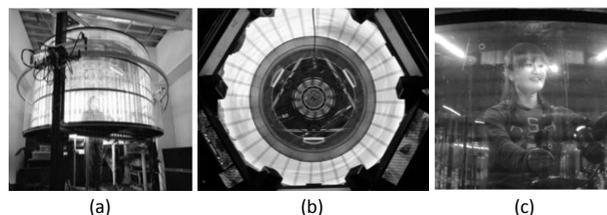


図 2 (a) TWISTER の概観, (b) TWISTER の下から見た全周囲映像, (c) TWISTER 外部から透けて見えるユーザの映像。

Fig. 2 (a) General view of TWISTER. (b) Auto-stereoscopic image of TWISTER (view from below). (c) User seen from outside TWISTER.

を採用した。TWISTER は右目用と左目用の LED 列を有する 36 枚の LED ユニットのユーザ中心の半径 1 m の円周上に等間隔に配置し、それらを回転させることで残像効果により 360 度全ての方向に 3 次元映像を提示する。回転機構により視差画像を提示するため、メガネ等を着用せずに裸眼で 3 次元映像を体験できる。TWISTER の提示解像度は 3162 x 600 pixel であり、更新レートは 60 fps である。また、ユーザは回転体内部において上半身を自由に動かすことが可能である。図 2(a) に TWISTER の外観及び (b) に TWISTER 内部を下から見た全周囲提示の様子を示す。高速回転による残像効果によって映像を見せているため、図 2(c) のように、LED を装着した円筒面が透けて内部のユーザが外側からはっきりと見える。

要件 1-3 に対し、TWISTER の回転体内部下方に全周囲音響を提示するためのスピーカを配置する。6 台のスピーカを 60 度間隔で配置し、サイバースペースの環境音及びユーザの発話が全周囲の空間音響となるように出力するスピーカと音量を調整して提示する。音響システムは文献 [16] と同様の構成となっており、音の方向提示が可能であることが同文献において確認済みである。これらの全周囲 3 次元映像及び音響提示システムにより、要件 1-1, 要件 1-2, 要件 1-3 が満たされる。

3.2 サイバースペース環境と身体性を有するアバターの設計

TWISTER で没入する環境であるサイバースペースとして、図 3(a) 及び (b) に示すような構成図に従うサイバースペースを構築した。近未来の装飾を施した会場に集う状況を想定し、サイバースペース内に構築したユーザの代理となるアバターの身体を借りてそこにテレイグジスタンスし、リアルタイムで変化する相手の表情と動作を見ながらコミュニケーションを実施する。サイバースペースの広さは移動可能な領域として 14 m x 10 m で構築し、動きまわるのに十分な広さを確保した。サイバースペースには近未来のバーを想

定し、オブジェクトとしてバーカウンターや机、人型キャラクタなどを配置している。バーカウンターにはコップのオブジェクトが配置されており、ユーザがアバターの腕を介して掴み取ることが出来るようになっている。サイバースペースの映像更新レートは 30 fps を満たすよう設計されており、要件 1-4 が満たされている。

アバターとして、文献 [13] のように映像として取得したユーザの全身像を利用する方法もあるが、アバターの全身映像を相互伝達すると通信遅延が大きくなる。そこで、アバターは CG で構築し、本提案において重要な顔の情報のみを実映像として提示する。単純に、CG の頭部に顔映像を貼り付けただけでは、首から上が実映像であるアバターとなってしまう気持ちが悪く不自然である。そこで、ユーザはサイバースペースに宇宙服を着込んで入り込むと仮定し、アバターの頭部を半透明なヘルメットとして構築してその内部に顔の実映像を提示する。こうすることで、相手ユーザが宇宙服を着て空間にいるように見え、CG と実映像の境界の不自然さがなくなる。また、要件 3-2 の身体性の反映については、ユーザから取得した腕と手指の位置姿勢情報に基づいて、CG で構築されたアバターのアームが動けばよいので、送りあう情報は位置姿勢情報のみで良い。そのため、ユーザはあらかじめサイバースペース内に配置されたアバターに入り込み、サイバースペース内を自由に移動することになる。図 3(c) にサイバースペース内でユーザの代理となるアバターの外観を示す。アバターは要件 3-1 を満たすように、一般的な成人男性と同等の背丈と肩幅を有し、片腕に 7 自由度（肩 2 自由度、肘 2 自由度、手首 3 自由度）、片手に 8 自由度（親指 3 自由度、人差～小指各 1 自由度、アブダクション 1 自由度）を有する。アバターの脚部には、車輪などの地面に平行に移動可能な機構を想定した移動機構を有する。人間の脚部のような機構では操縦が難しく、必要以上に映像空間が上下動してしまうことから平行移動機構を採用した。この移動機構により、要件 3-3 で求められる移動操作に伴う 2 次元平面上の移動が実現できる。顔映像を埋め込む顔部は透明なヘルメット体の内部に楕円形の平面板（以降、顔ボードと表記する）を有しており、それがビルボードとしてユーザに対して常に直面するように回転しつつ、位置関係に適切な相手ユーザの正面や横顔、後頭部などの画像が貼り付けられる。顔ボードに顔画像を貼る手法を用いることで、頭部サイズの異なるユーザ間でも作り変えることなくアバターを利用可能となり、様々なユーザに対応可能である。

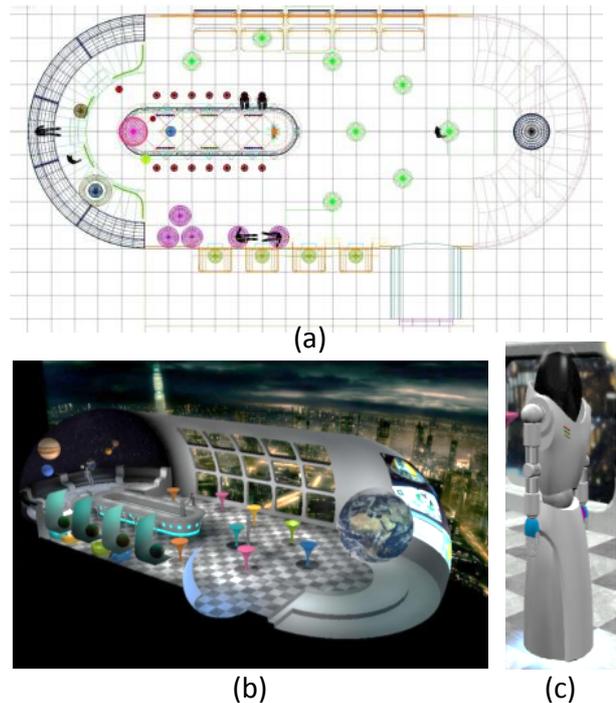


図 3 (a) サイバースペース全体の設計図面 (b) サイバースペースの概観 (c) CG アバターの概観

Fig. 3 (a) Design drawing of the cyberspace. (b) Overview of the cyberspace. (c) Overview of an avatar.

3.3 3次元顔映像取得システムの設計

TWISTER 内のユーザの顔を撮像するカメラシステムの詳細を図 4 に示す。TWISTER の特徴である内部が透けて見える効果を利用し、TWISTER の外周に自走が可能なステレオカメラを構築する。TWISTER は外側からは内部が透けて見えるが、内側からは提示映像しか見えないため、ユーザはカメラの存在を意識すること無くサイバースペースに没入できる。さらに TWISTER は顔部への装着デバイスを必要としない裸眼の立体映像提示装置のため、要件 2-2 で求められるユーザの表情を含む顔の実映像を取得できる。図 4(a) に示した自走可能なステレオカメラを、図 4(b) に示すように TWISTER の外側に円周状に設置したレール（半径 1.2 m）の上に TWISTER の中心を向くように複数台設置し、撮像する。レールは TWISTER 内のユーザの視線の高さになるように設置し、ユーザ間の背丈の違いは TWISTER 内の昇降台によって調整する。自走ステレオカメラは人間の平均眼間距離に合わせて 65 mm の間隔に置かれた 2 台のカメラ (Firefly MV, Point Grey Research 社) を有し、下部のモータによりレール上を自由に動くことができる。カメラの水平画角は 17 deg, 垂直画角は 13 度であり、1.2 m 離れた位置から撮影できる範囲は水平 0.36 m, 鉛直 0.27 m となる。日本人男性の顔の全頭高の平均は 0.24

m であり [17], この範囲を十分にカバーできる画角を有している. また, カメラの更新レートは 30 fps であり, 取得する画像のサイズは 640 x 480 pixel である.

自走ステレオカメラは, 配線など機構的な制約や複数ユーザへの対応のため, 円周上を 3 つのエリアに区切り, 各エリアに 2 台ずつ, 全体で 6 台を配置した. 自走ステレオカメラは, TWISTER 内ユーザに対する相手ユーザの相対位置関係と相手ユーザの向きから, TWISTER 内ユーザのどの方向から映像を取得すれば良いのかを判断し, その位置まで円周上を移動し撮像を行う. これにより, 映像提示されるサイバースペース内のアバターとカメラの方向が一致し, 常にカメラがアバターの目として働くことで, 相手の視線からの顔映像を取得できる. 取得した映像は相手側に送信され, 相手のサイバースペース内で自分の代わりとなるアバターの顔部に貼りつけられて提示される. アバターの設計において述べたとおり, 顔ボードはビルボードのように常に相手ユーザに向くように動くため, カメラで撮像した顔映像をそのまま顔ボードに貼り付けることで, ユーザからみた相手ユーザの顔映像を適切な位置関係で視認することが可能となる. 図 4(c) にアバターの位置関係と顔ボードの位置関係を示す. 相手ユーザが横に移動しながらユーザを眺めると, ユーザの横顔が撮像され, 顔ボードに貼り付けられることで立体に見ることが出来る. 少数のステレオカメラによりユーザのあらゆる方向からの映像を取得するため, 相手ユーザに 1 台の自走ステレオカメラを決め打ちで割り当ててのではなく, ユーザとの位置関係によってリアルタイムに適宜利用する自走ステレオカメラを切り替えながらユーザの撮像を行う. 切り替え時の映像がスムーズになるように撮像していないカメラは撮像しているカメラに出来るだけ近い位置にいるよう制御する. 以上の機構及び制御により, 要件 2-3 で求められる視線の一致が満たされる対面コミュニケーションが可能となる.

3 人以上のユーザが同一のサイバースペースに存在する場合は, 6 台のカメラのうち各ユーザの方向に一致した位置を撮像可能なカメラがそれぞれ担当する. サイバースペース内で自分以外の 2 人のユーザがすれ違う場合, 割り当てられるカメラが切り替わる必要がある. 図 5 に 2 人のユーザがすれ違う場合のユーザの移動とカメラの移動, 割り当てられるカメラの入れ替わりを示している. 左の図のような状態から 2 人のユーザが近づく方向へ移動すると, 中央の図のような状態となり, 2 台のカメラが物理的に衝突寸前の状態になる. その状態から 2 人がすれ違うと, カメラの空間位置はそのまま, ユーザ A の視点からユーザ C を撮像していたカメラはユーザ B の視点となり, ユーザ B

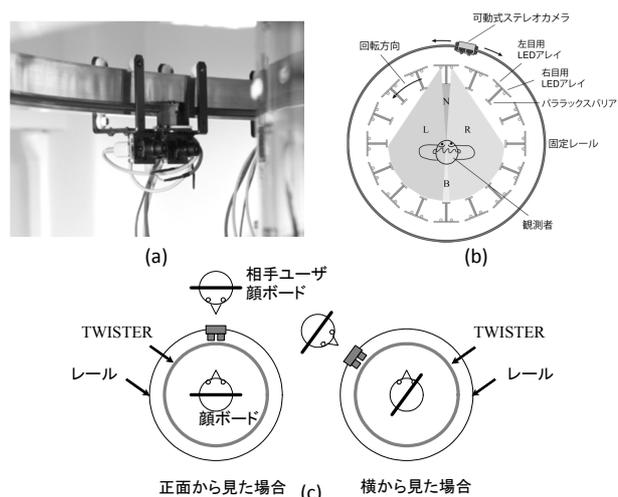


図 4 3次元顔映像取得システム (a) カメラの概観 (b) カメラの配置 (c) 顔ボードとカメラの位置関係
 Fig. 4 System for capturing user's face image. (a) Overview of the camera. (b) Layout of cameras. (c) Relationship between the face board of the avatar and the camera.

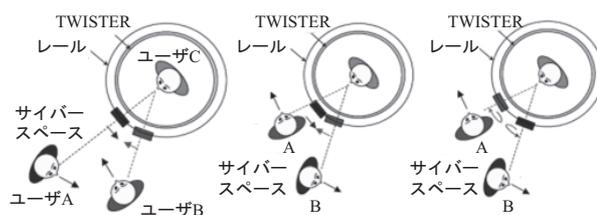


図 5 3次元顔映像取得システムにおける 2 ユーザの入れ替え
 Fig. 5 Two users change the algorithm of the image capturing system.

の視点からユーザ C を撮像していたカメラはユーザ A の視点となるように, 担当するカメラを電子的に交換する. その後は, 右の図のように入れ替わったカメラが各ユーザの担当をするよう振る舞う. このように制御することで, カメラの物理的な制約によらず, サイバースペース内の複数のユーザの移動に対応した 3次元顔映像の取得が可能となる. ユーザの切り替え手法に関する基礎評価は文献 [18] にて示されており, ユーザ切り替えが正しく機能することが確認されている.

要件 2-1 を満たすため, 声の取得デバイスとして小型のヘッドセットを利用する. 発話を取得する方法は幾つかあるが, ヘッドセットは顔の表情や目, 口の動きを阻害せず, 簡易に取得できるためこれを採用する. これらの設計により, 要件 2-1, 要件 2-2, 要件 2-3 が満たされる.

3.4 動作取得システムの設計

要件 3-2 で求められる直感的な身体性による物の把持と移動やジェスチャーを行うための腕と手指の位置姿勢を取得するシステムとして, 光学式モーションキャ

プチャとジョイスティックを合わせたシステムを構築する。モーションキャプチャの方式は接触式と非接触式の2手法があるが、ジェスチャーを実現することを考慮すると、計測の正確性よりも拘束性の低さが重要であるため非接触の光学式モーションキャプチャシステムを採用した。ただし、手指の動きについては非接触式ではオクルージョンの問題を解消しにくく、ある程度安定した把持や指によるジェスチャーが求められるため、装着式であるデータグローブを採用する。また、アバターの移動手法としてはジョイスティックを用いる方法を採用する。アバターの設計において、地面に対して車輪のようなものを想定した平行移動を前提として脚部を設計しており、人の足の歩行動作などの直接的な入力よりはジョイスティック入力による操作の方が容易である。

TWISTER 内のユーザの動作を取得するシステムの詳細を図6(a)に示す。赤外線を用いた光学式モーションキャプチャシステム (OptiTrack FLEX:V100R2, NaturalPoint 社) を右腕に適用するため、TWISTER の回転体内部の半径 1 m の空間全体を取得できるように TWISTER 内の上部から計測空間全体をカバーするように赤外線カメラを配置する。右腕の手首と肘、肩にそれぞれ赤外線マーカーを装着し、3点の空間的絶対位置を計測する。肩の位置を基準とした手首位置を算出し、肘の位置を参照しつつ逆運動学により関節角度を算出してアバターに与える。肩に対する手先の位置姿勢から得られる6自由度と肘の1自由度から、アバターの腕が持つ7自由度を完全に表現できる。マーカーは 10 ms 以内の遅延にて計測される。手指の動作を取得するため、5指の関節角度を取得可能なデータグローブ (5DT Data Glove Ultra, 5DT 社) を右手に装着し、各指の曲げ角を取得する。データグローブで取得可能な5自由度以外の自由度は、取得した情報から角度を推定して与える。データグローブは 13 ms 以内の遅延にて計測される。データグローブの把持動作によって指の関節角度情報が送られ、サイバースペース内に移動可能オブジェクトとして配置したコップなどのオブジェクトを把持し自由に移動させることが出来る。

左腕及び左手にはモーションキャプチャを装着せず、要件 3-3 に基づいてサイバースペース内を移動するためのジョイスティックの操作に充てる。ジョイスティックは前後左右の移動と左右回転の3自由度が割り当てられ、サイバースペース内を自由に動きまわることが出来る。サイバースペース内のアバターとユーザの動作の整合性を保つため、ジョイスティックを動かすユーザの左腕に対応したアバターの左腕は、CG のジョイスティックを把持し、ユーザの左腕の位置姿勢とジョ

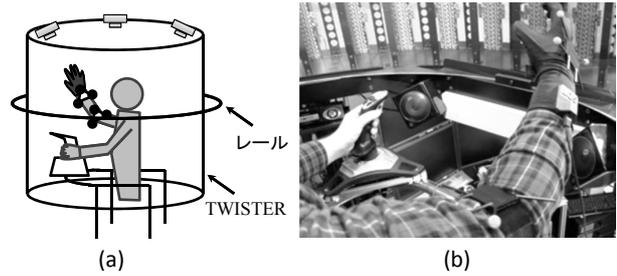


図6 動作取得システムの (a) マーカ及びジョイスティックの配置, (b) 実際の配置の様子
Fig.6 System for capturing user's motion.
(a) Layout of markers and joystick. (b) Overview of actual system layout.

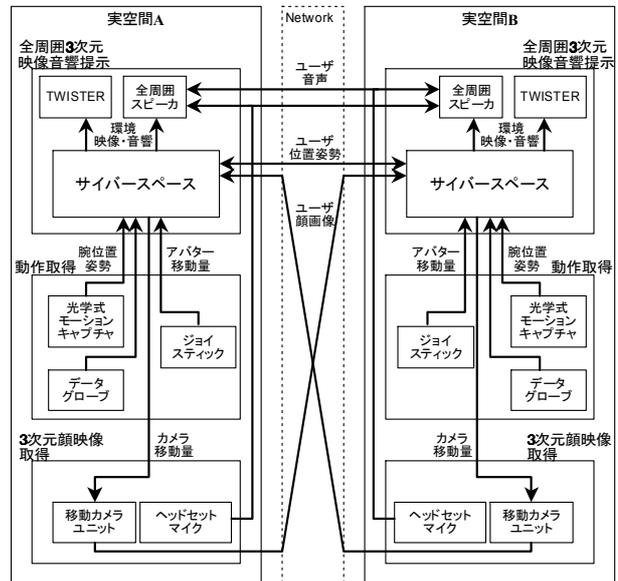


図7 提案する遠隔コミュニケーションシステムの構成図
Fig.7 Configuration of the proposed remote communication system.

イスティックの操作行動を再現する。これらの動作取得システムにより、要件 3-2, 要件 3-3 が満たされる。

3.5 統合システムの設計

図1の概念図に基づいた、2台の TWISTER を利用した遠隔コミュニケーションシステムの構成図を図7に示す。

全周囲 3 次元映像音響提示システムは、映像提示のための TWISTER と全周囲スピーカ、共有するサイバースペース情報を有し、サイバースペース環境の映像及び音響情報を TWISTER 及び全周囲スピーカに送る。動作取得システムは、光学式モーションキャプチャとデータグローブ、ジョイスティックを有し、ユーザの腕と手指の位置姿勢情報、アバターの移動量が計測され、全周囲 3 次元映像音響提示システム内のサイバースペースへ送られる。サイバースペースは、ユーザの顔映像とユーザ発話以外のユーザ位置姿勢情報を集約し、相手側のサイバースペースと UDP 通信によ

り差分情報を相互伝達し、受け取った情報を用いて空間情報の変更を行う。3次元顔映像取得システムは、移動カメラユニットとヘッドセットマイクを有する。移動カメラユニットは、サイバースペースからアバターの位置関係に基づくカメラ移動量を取得して、それに従って自走ステレオカメラを適切な位置へと動くよう制御する。撮像したユーザの顔画像は、ネットワークを介して相手側のサイバースペースへ直接送られる。ヘッドセットマイクで取得するユーザ発話情報は、自分側の全周囲スピーカーと相手側の全周囲スピーカーへネットワークを介して送られ、環境音響と合わせて提示される。

4 遠隔コミュニケーションシステムの検証

設計した遠隔コミュニケーションシステムの検証を行うための通信実験を試みた。設定した各要件が満たされているか否かの検証を行い、全ての要件が同時に満たされることでサイバースペースを共有した円滑なコミュニケーションが行えるかどうかを確認した。

4.1 検証実験の設定

遠隔コミュニケーションの検証条件として、今回は参加する遠隔ユーザを2人とし、1対1の遠隔コミュニケーションを行うとした。実施環境として、TWISTERを東京大学本郷キャンパス（東京都文京区本郷）と慶應義塾大学日吉キャンパス（神奈川県横浜市港北区日吉）に設置した。通信ネットワーク回線として、両TWISTER間に10ギガビット・イーサネットの専用回線を利用した。検証実験における体験者は20代6名であり、1人30分程度、体験者のうち2名は連続で2時間程度の体験を実施した。慶應義塾大学側で3名が1人ずつ交代で体験、東京大学側も3名が1人ずつ交代で体験した。体験の基本的な流れを示す。まず、東京大学側のユーザが慶應義塾大学側のユーザに近づき、コップの置いてあるカウンターに移動を促す。コップのあるカウンターに到着した段階で双方がコップを持ち上げ、そのままコップを持って窓側へ移動を行う。窓際に到着したあと、乾杯の動作をしてコップを飲む（ふりを）した上で簡単に雑談をする。その後は、自由に空間内を移動したり会話したりしながら体験を続ける。

4.2 各要件の検証

要件1-1から要件1-4について、サイバースペースの全周囲3次元映像提示についてはTWISTERの基本性能によって満たされており、統合システムにおいても損なわれずに実現されていることを確認した。図8にサイバースペースの投影映像の例を示す。

要件2-1から2-3について、サイバースペース内で言語・非言語コミュニケーションを行い、相手ユーザの

実映像から表情や視線を一致させた意思伝達が可能であることを確認した。図9(a)に投影されたアバターの顔部に相手ユーザの顔映像が提示されていることが確認できる。(a)は正面を向いている相手ユーザを見た場合の投影映像である。相手ユーザの横から相手ユーザの顔を眺めた時のアバターの顔映像を図9(b)に示す。これにより、相手の横顔が見えていることがわかり、(a)と合わせてユーザの位置関係に適切な3次元顔映像が提示されていることがわかる。また、ステレオカメラとアバターの位置を示すため、TWISTER外部の照明を調節しアバターとステレオカメラを同時に視認可能にした結果が図9(c)である。アバターの顔位置に一致するようにステレオカメラが重なって表示されており、その結果として視線が自然に一致するシステムが実現できたと言える。

要件3-1から要件3-3について、ユーザが人と同等な身体性を持つアバターを介し、サイバースペース内の移動やジェスチャー、オブジェクトであるコップを持ちながらの移動が可能であることを確認した。図10(a)と(b)にジェスチャーとして手を振ったり指差しをしたりする動作を行なっている様子を示す。また、図10(c)にコップを持った相手アバターの映像を示す。要件3-4の動作の許容遅延量を満たしているかどうかを、構築したシステム要素の性能から理論的に検証した。まず、腕や手指の計測にかかる遅延は、データグローブの計測遅延である13msに依存する。計測したデータをPCで取得しUDP送信処理を行うのにおよそ10ms程度かかるとし、送信先に送る。伝送路上での送信にかかる時間は、アバターの位置姿勢情報と移動可能なオブジェクトの位置姿勢情報のみであり、データのサイズは数100B程度と非常に小さいことから10ms程度とする。受け取り側でUDP受信処理に10ms程度、画像に反映させる際に最大30msかかると考える。これらを合わせると73msとなり、他の要因で通信遅延量が多少増えても遅延許容量120msは満たされるといえる。要件3-4は動作における運動視遅延に対する要件だが、相手の顔画像に対して同様に通信遅延を見積もってみる。かかる処理時間は、画像の取得(30ms)、PCでのキャプチャ(30ms)、エンコード処理(10ms程度)、通信と共有ファイル書込み(30ms程度)、読み出しとデコード処理(30ms程度)、画像反映(30ms)となり、160ms程度と見積もられる。運動視の遅れ条件である120msと比較すると許容値を超えるが、顔映像内の表情の多少の遅れは声やジェスチャーによって補完できると考えられるため、ここでは腕と手指の動作にかかる遅延が満たされていることで共有空間の同期が満たされているとみなす。なお、今回送り合っている情報は利用するネットワー

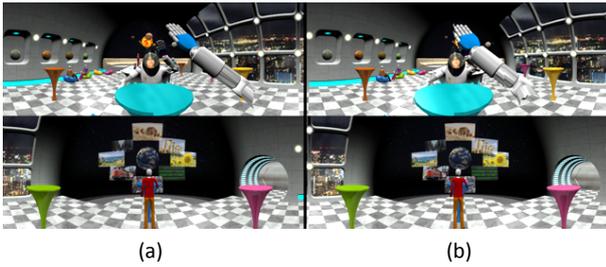


図 8 TWISTER に投影したサイバースペース映像の例. (a) 左視点映像 (b) 右視点映像
Fig. 8 Example of the cyberspace image projected on TWISTER. (a) Left-eye view. (b) Right-eye view.



図 9 (a) アバターの正面顔映像, (b) アバターの横顔映像, (c) アバターと重なるカメラ
Fig. 9 (a) Front view of an avatar face. (b) Side view of an avatar face. (c) Camera overlapping on an avatar.

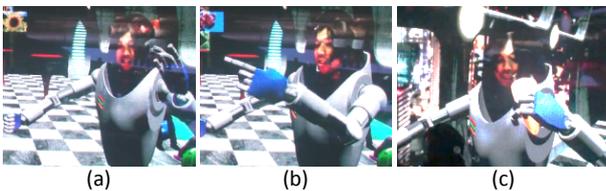


図 10 (a) アバターの手振り動作, (b) アバターの指差し動作, (c) コップを持ったアバター
Fig. 10 (a) Hand-waving motion of an avatar. (b) Pointing motion of an avatar. (c) Avatar with a cup.

ク伝送路の性能に対して少ないため、参加するユーザー数や移動オブジェクトの種類や数などを、通信伝送路の通信遅延を満たす範囲で自由に追加することが可能である。

4.3 遠隔コミュニケーションの主観的検証

統合システムにおいて遠隔コミュニケーションを行った結果を図 11 に示す。(a) が慶應義塾大学側の TWISTER ブース内の映像を示しており、(b) が東京大学側の TWISTER ブース内の映像を示している。体験の基本的な流れについては、滞りなく実施できていることを確認した。

体験者から得られた主観的評価として、すべての体験者より構築したサイバースペース内に実際に居るかのような感覚を得ることが出来たという意見を得た。ある体験者からは、現実の身体そのものがサイバースペース内に入ったというよりは TWISTER で構築さ



図 11 遠隔コミュニケーション実施の様子. (a) 慶應義塾大学側 (b) 東京大学側
Fig. 11 Results of remote communication. (a) Image result at Keio University. (b) Image result at The University of Tokyo.

れたブースとともに入り込んだように感じたという意見があった。そのことから、宇宙服のようなものを着てサイバースペースに入り込む状態に近い感覚を誘発できたと考える。また、「あっち」や「これ」といった現実空間で多用する指示代名詞が示す目的が自然と理解でき、コミュニケーションが行いやすかったという意見があり、感覚的にも対面コミュニケーションが実現できていることがわかる。これらの主観的な検証結果より、提案した全ての要件が満たされた統合システムは、身体性や意思伝達を含んだ遠隔コミュニケーションに有用であるといえる。ジョイスティックによる移動操作については、コミュニケーションしやすいと感じた人としにくいと感じた人の双方があり、これについては今後検討する必要がある。

5 おわりに

本研究では、多人数で対等な高臨場感遠隔コミュニケーションを実現するため、サイバースペースをコミュニケーションの場として共有する遠隔コミュニケーションシステムを設計構築した。サイバースペースを共有する遠隔コミュニケーションシステムを実現するための目標として、(1) サイバースペースで参加者が視覚的に高い臨場感を有する、(2) 参加者同士の視線が一致する対面コミュニケーションが可能、(3) サイバースペースで参加者が身体性を有する、の3目標を掲げ、各目標を達成するために必要なシステム要件を検討し、全ての要件を満たす統合システムを構築し有用性を検証した。全周囲3次元裸眼ディスプレイ TWISTER に、表情や視線を取得するための3次元顔映像取得システムとユーザの腕と手指の位置姿勢を取得するためのユーザ動作取得システムを加えて、サイバースペース内に構築したアバターを介してコミュニケーションを行う統合システムを構築した。その結果、本研究が提案する各要件を全て満たすシステムが実現できていることを確認した。また、実際に2者による遠隔コミュニケーションを実施した結果、お互いが対等な条

件で、サイバースペースを共有した遠隔コミュニケーションを行えることを主観的に確認した。これらの結果より、提案するシステムがサイバースペースを共有する遠隔コミュニケーションシステムとして有用であるといえる。

今後は、3者以上での遠隔コミュニケーションの検証実験を行うことで多人数でのコミュニケーションシステムの有効性を検証するとともに、触覚情報も含めた身体性を相互伝達することによるコミュニケーションの質の向上を図っていく。

謝辞

本研究は、総務省の戦略的情報通信研究開発推進制度 (SCOPE) 及び、科学研究費補助金基盤研究 (A) 課題番号:23240017 の一環として行った。

参考文献

- [1] RealPresence, <http://www.polycom.com/products/>.
- [2] 館暲, 渡邊孝一, 竹下佳佑, 南澤孝太, 吉田匠, 佐藤克成. 再帰性投影技術と全周囲裸眼 3D ディスプレイを用いて存在感と臨場感を実現する相互トレイグジスタンスシステム: TELESAR4. 日本バーチャルリアリティ学会論文誌, Vol. 17, No. 1, 2012.
- [3] 長谷川修, 森島繁生, 金子正秀. 「顔」の情報処理. 電子情報通信学会論文誌. D-II, 情報・システム, Vol. J80-D-2, No. 8, pp. 2047-2065, 1997.
- [4] K. Kampe, C.D. Frith, and U. Frith. "hey john": Signals conveying communicative intention toward the self activate brain regions associated with "mentalizing", regardless of modality. *The Journal of Neuroscience*, Vol. 23, No. 5, pp. 5258-5263, 2003.
- [5] M. Cook and J.M.C. Smith. The role of gaze in impression formation. *British Journal of Social Clinical Psychology*, Vol. 14, pp. 19-25, 1975.
- [6] D Friedman, A Brogni, C Guger, A Antley, A Steed, and M Slater. Sharing and analyzing data from presence experiments. *Presence*, Vol. 15, No. 5, pp. 599-610, 2006.
- [7] M.F. Vargas. Louder than words: an introduction to nonverbal communication. *Ioa State University Press*, 1986.
- [8] 増田千尋. 3次元ディスプレイ. 産業図書, 1990.
- [9] 佐藤隆宣, 伴野明, 岸野文郎. 運動視表示の遅延による仮想空間操作への影響. テレビジョン学会技術報告, Vol. 14, No. 68, pp. 1-6, 1990.
- [10] Second life, <http://secondlife.com>.
- [11] There, <http://www.there.com>.
- [12] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, R.V. Kenyon, and J.C. Hart. The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, Vol. 35, No. 6, pp. 64-72, 1992.
- [13] Tetsuro Ogi, Yoshisuke Tateyama, and Satoshi Oonuki. High presence collaboration using plug-in video avatar. In *Proceedings of the ASASME 2009 IDETC/CIE 2009*, Vol. DETC2009-86762, 2009.
- [14] 知識ベース「知識の森」S3群8編. 電子情報通信学会. <http://www.ieice-hbkb.org/portal>.

- [15] Susumu Tachi. TWISTER: Immersive omnidirectional autostereoscopic 3d booth for mutual telexistence. In *In Proceedings of ASIAGRAPH 2007*, Vol. 1, pp. 1-6, 2007.
- [16] 土場健太郎, 南澤孝太, 新居英明, 川上直樹, 館暲. TWISTER のための全周囲音響伝送システム. 日本バーチャルリアリティ学会第 14 回大会論文抄録集, pp. 2A3-2, 2009.
- [17] 勝浦哲夫, 栃原裕, 佐藤陽彦, 横山真太郎 (編). 人間工学基準数値数式便覧. 技報堂出版, 1992.
- [18] 佐野乾一, 等康平, 南澤孝太, 新居英明, 川上直樹, 館暲. TWISTER における多人数対面コミュニケーションのための自走式撮像システムの構築. 日本バーチャルリアリティ学会第 14 回大会論文抄録集, pp. 2A3-1, 2009.

(2011 年 12 月 19 日受付)

[著者紹介]

渡邊 孝一 (正会員)



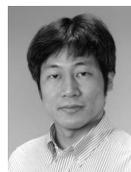
2004 年千葉大・工・情報画像工学科卒. 2006 年東大院・情報理工・システム情報学専攻修士課程修了. 2009 年同大博士課程修了. 情報理工学博士. 同年慶應大院・メディアデザイン研究科助教. 2011 年慶應大・メディアデザイン研究科リサーチャー. 2012 年木更津高専・情報工学科講師. トレイグジスタンスマスタスレーブシステムの研究に従事.

南澤 孝太 (正会員)



2005 年東大・工学部・計数工学科卒. 2010 年同大大学院・情報理工学系研究科博士課程修了. 博士 (情報理工学). 2007 年より日本学術振興会特別研究員 (DC1). 2010 年慶大大学院・メディアデザイン研究科・特別研究助教. 2011 年同研究科・特任講師. ハプティックインタフェース, 3D ディスプレイ, トレイグジスタンスの研究に従事.

新居 英明 (正会員)



2003 年から 2006 年まで電通大・電気通信学・機械制御博士課程在籍. 2006 年から 2009 年まで東大・情報理工学・システム情報学専攻助教. 2008 年博士 (工学) 取得. 2009 年より慶大メディアデザイン研究科講師. 2010 年よりシンガポール国立大学研究員.

館 暲 (正会員)



1968 年東大・工学部・計数工学科卒. 1973 年同大大学院博士課程修了. 工学博士. 東大・先端研・教授などを経て, 1994 年同大・工学部・計数工学科・教授, 2001 年同大大学院・情報理工・教授, 2009 年慶大大学院・メディアデザイン研究科・教授. 東京大学名誉教授. トレイグジスタンス, 人工現実感などの研究を行う. IMEKO ロボティクス会議議長, 計測自動制御学会第 46 期会長, 日本バーチャルリアリティ学会初代会長.