

ジェスチャ操作を活用する広視野電子作業空間の設計と実装

木村朝子[†] 柴田史久[†] 鶴田剛史[†]
酒井理生[†] 鬼柳牧子[†] 田村秀行[†]

近未来社会のヒューマンインタフェースのあるべき姿として、映画『マイノリティ・レポート』に登場する広視野ディスプレイとジェスチャ操作の組み合わせを考える。本研究ではまず、広視野電子作業空間に適した作業の想定・分類から始め、これを実現する基幹システムとジェスチャ基本コマンドを開発した。その上で、このシステムに立体映像表示、実物体の導入、複合現実感へと発展する体系「MR キューブ」を提唱する。その試作実現例として、ビデオ映像編集と医用画像の閲覧・管理を題材としたシステムを開発し、その操作性と有用性について検証した。

Design and Implementation of Minority-Report-Style Gesture Interaction with Wide-view Electronic Working Space

ASAKO KIMURA,[†] FUMIHISA SHIBATA,[†] TAKESHI TSURUTA,[†]
TOSHIO SAKAI,[†] MAKIKO ONIYANAGI[†] and HIDEYUKI TAMURA[†]

Combination of a wide-view display and hand gesture operations that is presented in the feature film "Minority Report" is to be considered as what the computer-human in-teraction should be. Assumption and categorization of works suitable to Wide-view Electronic Working Space were the starting point of our study and we established a backbone system and a detection method for gesture commands. Now we propose "MR-Cube" system which will bloom to be a mixed reality system having included 3D image display and introduction of real objects based on the backbone system. We have developed a video playing and editing system and a system to browse, edit and manage medical images as prototypes in order to verify their operation and usability.

1. はじめに

20世紀末に急速に普及した WIMP 型ユーザインタフェースの全盛の中で、その限界を超える様々なユーザフレンドリーなヒューマンインタフェース (HI) の研究が活発化している¹⁾。とりわけ、実世界指向インタフェース、タンジブル・インタフェースの研究結果²⁾³⁾の中には、未来への可能性を感じさせるものが少なからず存在するが、まだ一般に広く普及する汎用的なものにはなっていない。最近、大型スクリーンにコンピュータの画面を投影し、そこに描かれたオブジェクトを操作する試みもいくつか見られるようになってきた⁴⁾⁵⁾。これは、PC で同時に開くウィンドウの数や作業対象が増え、もはや小さく狭いデスクトップ画面には収まりきらなくなったことを意味している。

筆者らは、広い電子作業スペースを確保できる広視

野ディスプレイに向かって、両手を大きく動かすジェスチャを用いる操作が、近未来の有望な HI の 1 つであると考えられる。映画『マイノリティ・レポート』(2002)の中では、この種のジェスチャ操作のシーンが再三登場する。もともと良質の SF 映画は「実現できるならば、こんなものが欲しい」という人間の自然な願望を忖度して描いているので、これは未来技術予測のある種の可視化結果だと考えられる。我々はこの視点から、古典的名作『2001年宇宙の旅』(1968)から最近の『アイ、ロボット』(2004)まで、約60本のSF映画を分析した⁶⁾。この調査分析の結果、『マイノリティ・レポート』中で主人公の捜査官が湾曲した半透明の広視野ディスプレイの前で行なうジェスチャ・アクションが、post-WIMP で最もあり得る好ましい HI と評価された。

指さし動作を HI とする試みとしては、約四半世紀前に MIT Media Lab. の「Put-That-There」⁷⁾が登場したが、本格的なジェスチャ操作には発展しなかった。最近、卓上の作業環境に対しては類した操作の実

[†] 立命館大学大学院理工学研究科
Graduate School of Science and Engineering,
Ritsumeikan University

現例⁸⁾⁹⁾¹⁰⁾が多数発表されているが、いずれも広視野領域を扱う上記映画のイメージとはかなり異なっている。我々は、もっと映画に似た形状の広視野電子作業空間とジェスチャ操作を組み合わせたシステムを、まず実現すべき達成目標とした。こうしたシステムを実際に開発して、操作性・有効性を HI 学の視点から評価する。続いてその様々な発展形を考え、広視野電子作業空間の技術体系を構築することを目指している。

本論文では、まず広視野電子作業空間に適した作業を想定して分析し、ジェスチャ・コマンドの体系を一般的に考える。次に、本研究の目指すインタフェースの基礎となる基本システム構成とジェスチャ操作の実装について述べる。さらに、この基本システムに立体映像表示、実世界指向、複合現実作業環境等を導入した発展型「MR キューブ」を提案する。この MR キューブの事例として、2 種類の試作システムを実現し、そこから得られた広視野電子作業空間の有効性と、ここで生じた問題点の解決法に関して報告する。

2. 広視野電子作業空間とジェスチャ入力に関する考察

2.1 広視野電子作業空間に適した作業

1970 年代のワークステーション研究に端を発するデスクトップ・メタファは、マルチウィンドウという革新的な情報提示手法を完成させた。しかし、PC で行ないたい作業が増加するにつれ、モニタの画面は次第に窮屈なものになって来た。ホワイトボードをメタファとした大型モニタにペン型入力を用いるシステムは、これまでも数多く研究され、市販製品も存在する¹¹⁾¹²⁾。スティック状の WAND デバイスを用いた研究例も報告されている¹³⁾。いずれも解像度的には、現存の PC の枠を超えていない。コンピュータ・モニタの高解像度化と大型化だけでは、この不満を解決できず、早晚もっと広くかつ使いやすい電子作業領域が求められると予想される。

利用者が没入感を感じる「イマーシブ・ディスプレイ」は、現在でもバーチャル・リアリティ(VR)分野で活用され、そのほとんどは 3 つ(あるいは、それ以上の)プロジェクトで映像を同時投影している。家庭用の大型 TV モニタの低価格化が進めば、この種の広視野ディスプレイを数々の日常業務で用いることも現実味を帯びてくると思われる。

我々はそうした近未来を想定し、広視野電子作業空間が存在すれば、そこで効果的に実行できると考えられる作業内容を検討した(表 1)。図 1 はその一例で、広視野電子作業空間を用いて、医師が医療画像を閲覧

表 1 広視野電子作業空間に適した作業の例
Table 1 Example of works suitable for wide-view electronic working space

作業 / システム名	内容
マルチメディア・プレイヤー	多様なマルチメディア・データを並べ、選択・再生する
データ整理システム	多数のデータを並べ、分類・整理する
地理情報システム	比較的広範囲の地図データを閲覧し、その上で各種作業する
安全監視センター	ビルに多数設置された監視カメラなどの情報を、切り替えて安全確認する
組み立て・解体・解剖の訓練システム	工場・建築・医療現場等での組み立て・解体、解剖の訓練を行う
戦術シミュレーション	団体スポーツにおける戦術パターンを視認し、協議する
動画編集システム	複数の動画データベースを閲覧しながら、整理・編集する
グループスケジュール管理システム	メンバーのスケジュールを一覧しながら、作業調整やプロジェクトの進捗管理を行う
ファッションコーディネート・システム	多数の洋服を並べ、ファッションコーディネートを行う

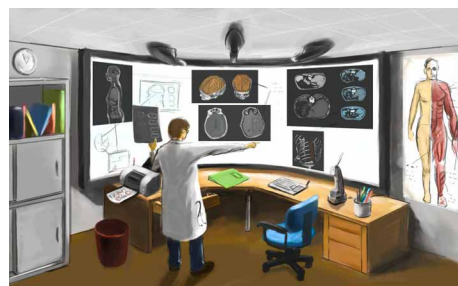


図 1 医用画像の閲覧・整理システム作業イメージ図
Fig. 1 Imaginary picture of the system to browse and manage medical images

したり、整理している風景である。この図は、実際の医師のアドバイスと要望を得て描いたイメージ図である。近未来、病院内の医師のオフィスではこのような光景が実現するものと考えられる。

表 1 に示した作業内容を整理すると、下記のような特徴を有している。

- 多種・多様なデータの一覧
- 多数のデータの整理・分類
- 大きいデータや、複雑なデータを大きく表示
- 沢山の補足資料を見ながらの作業

こうした広視野領域での作業の HI としてジェスチャ操作を導入しようというのは、極めて自然な考え方である。大きく両手を広げたり、時には歩き回るという身体行為を伴うことも、狭い画面内の作業からの解放感を与え、未来型 HI としての期待を抱かせる。我々が『マイノリティ・リポート』のシーンに魅力を感じるの、その身体的動作と直観的な操作方法に心地よ

さに共感を覚えたからだろう。

2.2 ジェスチャ入力のためのコマンドセットとその水準

ジェスチャ操作では、腕・手・指の動きのバリエーションにより、現在のマウス操作によるポインティングとクリックによる実行指示よりも多様なコマンドを表現することができる¹⁴⁾。しかしながら、ジェスチャの種類を多くするとユーザが覚えきれず、作業能率が低下するため、なるべく少数で覚えやすいコマンドセットにすべきだという報告がある¹⁵⁾。

これまでの研究では、HIとしてのジェスチャ操作は、識別率を上げるために、個別の応用に適した操作をその都度考えてきた傾向がある。例えば、Baudelら¹⁶⁾はプレゼンテーションでのスライド操作のためのジェスチャを、Segenら¹⁷⁾はVR空間内のFly Through操作を行うためのジェスチャを、Freemanら¹⁸⁾はテレビの操作を目的としたジェスチャを提案しているが、様々な業務に通用するコマンド体系を考えた上でジェスチャを対応づけているわけではない。過去の研究例は、信号レベルでのジェスチャ抽出方法は参考になるものの、本研究が対象とする広視野電子作業空間に適用できるジェスチャ操作体系に役立つものは存在しない。

本研究では、表1に示したような各種業務を類型化・抽象化して考え、それに適したジェスチャ操作も基本的な共通のコマンド体系をもつべきであると考えた。各応用システムの作業に依存したジェスチャ・コマンド群は、基本コマンドとは別に定義できる体系が望ましい。また、ジェスチャは人間の身体動作がそのまま反映されるので、操作者の体格による適性、覚えやすさ、操作に対する好みの違いもあると思われる。したがって、コマンド入力システムとしては、ユーザが自ら好みのジェスチャを選択・定義できるカスタマイズ機能を持たせることが得策である。

この考えに基づき、コマンドセットには以下の3つの水準を設ける。

- 「第1水準」は、最も共通性の高い基本コマンドセットである。この水準のコマンドは、応用には非依存で、ユーザによる対応づけの選択・変更を不可とし、全員が記憶しやすい操作を選ぶ。
- 「第2水準」もまた応用非依存の共通性の高いコマンドセットであるが、第1水準より利用頻度のやや落ちる操作をこの水準とする。この水準のコマンドには、ユーザ自身によるカスタマイズ機能を許す。
- 「第3水準」は、特定の応用領域に依存したコマ

表2 コマンドセット

Table 2 Command set

水準	操作例
第1水準	選択, 確定, 解除, 移動, 回転, 拡大/縮小
第2水準	複数選択, Undo, コピー, ペースト
第3水準	ムービーなどメディアデータの操作・編集

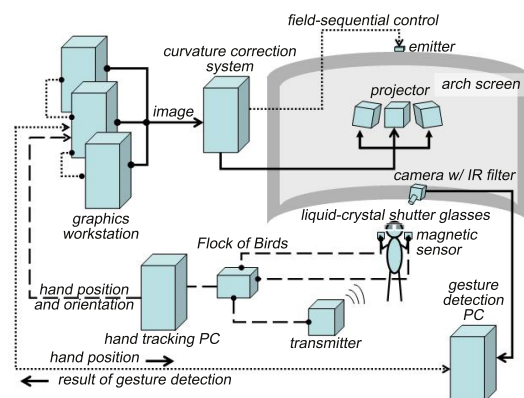


図2 システム構成図

Fig. 2 Hardware system configuration

ンドセットである。文字セットの「外字登録」に相当する概念で、熟練者が頻度高く利用することにより作業効率が向上するような操作を個別に定義する。

表2は、表1の作業内容を考慮して選んだ操作コマンドの例である。ジェスチャ操作には、手話のような言語的側面をもつものと、物理的に存在する対象物への動作をメタファとするものに大別できる。本研究が対象とするのは後者である。

本研究では、「各水準にどのような機能のコマンドを割り当てるか」と「各コマンドにどのようなジェスチャを対応づけるか」と「そのジェスチャ操作を実際に識別する方法」を切り離して独立に考えらえる枠組を採用した。こうした抽象化した概念設計を行なったことで、後述する基幹システムの設計が容易になり、それを様々な応用システムへ展開しやすくなった。

3. 広視野電子作業空間とジェスチャ入力の基本システム構成

3.1 アーチスクリーン型映像表示システム

広視野電子作業空間を実現するベースとなるシステムとして、図2に示すハードウェア構成のアーチスクリーン型映像表示システムを導入した。

広視野没入型ディスプレイとしては、高さ1.8m、弧長6.8mの大型アーチスクリーンを用いている。計3台のプロジェクタ(CDS社製のMirage6000 DLP)

を用いることで、アーチスクリーンの正面、左、右の部分を投影している。正面、左、右の映像が重なる部分は湾曲補正装置を用い、境目をブレンディング処理している。

Mirage6000 DLP は時分割立体視に対応している。このステレオ映像を、液晶シャッタ眼鏡 (CrystalEYES 社 CrystalEYES3) を用いて見ると、赤外線エミッタによる信号に応じて、液晶メガネのシャッタが開閉し、右目と左目用の映像が交互に映し出され、両眼立体視が可能である。

3 台のプロジェクタで出力される映像はそれぞれ右面、正面、左面用のグラフィックス・ワークステーションで同期を取りながら生成される。投影映像の開発環境は Visual C++ で、OpenGL ライブラリと、米イリノイ大学が開発した VR 空間を構築するための API である CAVELib を用いている。このライブラリを用いることで、アーチスクリーンに対する左右眼用の画像生成や座標変換、磁気センサからの信号処理をプログラミングすることができる。

後述のジェスチャ入力のために、両手の位置姿勢を検出するのに磁気センシング・システム (Ascension Technology 社製 Flock of Birds) を用いる。センサの計測範囲は半径 1.2m の半球内である。

3.2 ジェスチャ入力用手袋デバイス

ジェスチャ入力用には、伸縮性のあるピリヤード用のグローブを採用し、これを加工して手袋状のデバイスを試作した。このグローブは親指、人さし指、中指の 3 本の指のみを覆っている。これは映画『マイノリティ・リポート』に登場する手袋型デバイスでも同様であり、手の自然な動きを妨げることなくハンドアクションができる。

本研究でのジェスチャ認識では、磁気センサによる手の大まかな位置姿勢の検出と、指先につけた再帰性反射材をマーカとして用いた手指のアクションの識別という 2 つの方法を併用する。グローブで覆われた 3 本の指先の先端にそれぞれ再帰性反射材を装着し手の甲の部分に前述の磁気センサを取り付ける (図 3)。

またグローブ自体は黒色で環境光を反射させることなく、再帰性反射材の領域のみを検出できる。再帰性反射材を検出するために、アーチ型スクリーンの前面下部にビデオカメラ (Sony Network Handycam DCR-TRV70K) を設置した。また、ビデオカメラの夜間撮影機能を利用することで、被写体に赤外光を投影する。レンズには赤外線透過フィルタが取り付けられ、ジェスチャ認識用 PC で画像認識処理を行なう。

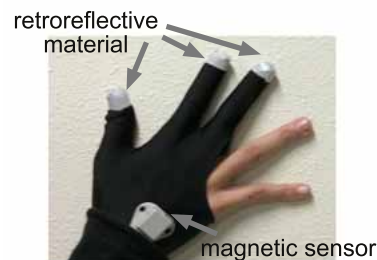


図 3 ジェスチャ入力のための手袋デバイス
Fig. 3 Glove device for gesture detection

表 3 第 1 水準のジェスチャ・コマンド
Table 3 First level gesture command

コマンド	ジェスチャ	コマンド	ジェスチャ
選択		確定	
解除		移動	
回転		拡大 / 縮小	

3.3 作業コマンドとジェスチャの対応づけ実験

ハードウェア・システムと対話デバイスの準備ができたので、作業コマンドとジェスチャを対応づける問題に取り組んだ (操作対象の)「選択」「移動」「複写」等の作業コマンド名にどのような手や指の動きを対応づけるかは、予め研究チーム内で候補を絞り、各コマンドの最有力候補を選定した。第 1 水準コマンドに関する対応は、表 3 に示す通りである (紙幅の制約から、他の候補や第 2 水準以上のジェスチャ・コマンドはここには明示しない)。

複数人が参加しかつ十分な議論の上で選んだものであるが、その妥当性を客観的に確認するために、多数の被験者に対して、以下の実験を行なった。

(1) 実験 1: 第 1 水準コマンドの選択実験

【実験目的】第 1 水準コマンドは強制的に覚えてもらう共通コマンドであるので、ジェスチャとの対応も慎重であるべきである。本実験は、設計側が選んだ各候補ジェスチャをユーザがどう評価するかを調べた。

【被験者】本研究チームに属さない学生 36 名。

【選択対象と候補】「選択 確定 解除」は一連の操作であるので、3 コマンドを一括りにし 3 つの候補を、「回転」「拡大 / 縮小」に関しては各々に 3 候補を提示した。「移動」は紛れがないので、他の選択肢は与えなかった。例えば、「回転」に対しては、「片手を握り、そ

の握り拳をオブジェクトになぞらえて回転」「両手でオブジェクトを挟むようにし、挟んでいるオブジェクトを回転」「親指、人差し指、中指を各々垂直に伸ばし、3次元座標の各軸になぞらえて回転」の3つ選択肢の中から、最も使いたいジェスチャを選ばせた。

【提示方法】手袋をはめない手・指の動きをビデオに収録し、格別な説明はつけずに被験者に映像を提示した。

【実験結果】「移動」を別にすれば、各コマンド（列）に対する最有力候補（即ち、表3のジェスチャ）の選択率は各々以下の通りであった。

「選択 確定 解除」= 97 %

「回転」= 53 %

「拡大 / 縮小」= 72 %

この結果は、各コマンドで集中度には差があるが、研究チームが予め選んだ最有力候補ジェスチャが過半数の被験者から支持されており、設計側の事前選定が妥当であったことが証明された。

(2) 実験2：第2水準コマンドの選択実験

【実験目的】第2水準コマンドはジェスチャ操作をユーザがカスタマイズできる機能を有しているため、ユーザの好みの偏りとそれがコマンド記憶実験にどれほど影響するかを予め記録する。

【選択対象と候補】第2水準「複数選択」「Undo」「コピー」「カット」「統合」「分割」「削除」「情報を見る」の8コマンドに対して、各々に3候補を提示し好みのジェスチャを選ぶ。例えば、「削除」は「オブジェクトを片手で後ろに放り投げる」「両手で破る」「両手で握り潰す」である。

【実験方法】実験1と同じ学生36名にビデオを提示。ただし、「グループA」17名には好みを聴いた上で指定ジェスチャを記憶することを求め、「グループB」19名にはその好みのジェスチャをそのまま記憶させた。

【実験結果】8コマンド中の6つに対して、かなり好みのジェスチャが分かれた。その選択理由は、「作業内容がイメージしやすいから」「片手だけでできるから」「あまり手を動かさずに済むから」などであった。第1水準に比べて作業内容がイメージしにくいものもあり、この実験では、ユーザ自身がジェスチャを選択できる余地を残すことは有効であると考えられた。

(3) 実験3：全コマンドの記憶・反復実験

【実験目的と方法】三者択一で自ら選んだものの影響度を調べることを目的とし、実験1・実験2で記憶したジェスチャを、作業内容が読み上げられた時に即座に動作できるかを試験した。正しいジェスチャを示せなかった被験者には、再度ビデオを見せて記憶させた。

【実験結果】全36名の被験者中、14名が2度目で、

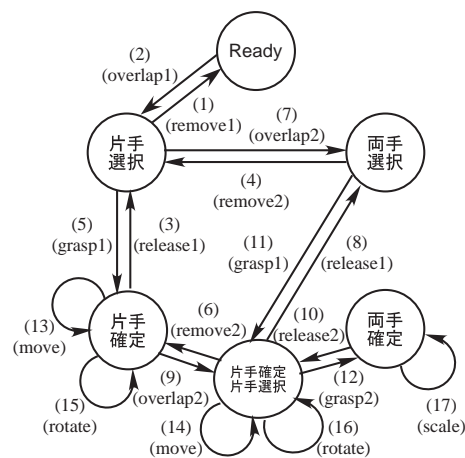


図4 ジェスチャ認識のための状態遷移図

Fig.4 State transition diagram for the gesture detection

16名が3度目で、5名が4度目で全問正解に達した。また、1度目の平均正解数が5.8種類、2度目では10種類と、多くの被験者は2回目で大半のジェスチャ・コマンドを記憶していた。誤回答を分析したところ、第2水準の正答率は第1水準よりも若干低かったが、好みのジェスチャである / なしは全く影響がなかった。グループAよりグループBの正答率が高いと想定したが、予想に反してほとんど差がなく、むしろ最も覚えが悪かった1名はグループBに属していた。また、約1週間後に数名に抜き打ち検査したが、全員ジェスチャ・コマンドを記憶していた。

【考察】この正答率の高さは、設計者らの予想を遙かに上回るものであった。ここから直ちに、ユーザ自身のカスタマイズ機能は一切不要と断定するのは危険であるが、ジェスチャ・コマンドは多少無理をしてデザインしても、ユーザは適応して覚えてくれることを示している。

3.4 ジェスチャ認識処理の実装

作業コマンドに対応づけたジェスチャの認識処理は、正面から見えている赤外光の領域数を片手毎にリアルタイムに検出し、それぞれの手の形状変化を検出する方法を採用した。具体的には、以下のような手順で安定した結果が得られた。

- (1) 手領域の検出：磁気センサから計測された右手・左手の位置・姿勢から画像内の右手・左手の領域を大まかに検出する。
- (2) 背景差分：入力画像の(1)で推測した領域に対して、事前に取得しておいた背景画像との差分処理を行い、環境光の影響を排除する。
- (3) 2値化：濃淡画像の2値化処理を行い、赤外光

表 4 ジェスチャ認識のための状態遷移表
Table 4 State transition table for the gesture detection

遷移	1つの物体に対する条件		コマンド
	手の位置姿勢	LEDの領域数	
(1)	物体と手が重なっていない	-	選択なし
(2)(3) (4)	ある物体と片手1が重なっている	2個以上(片手1)	片手選択
(5)(6)		1個(片手1)	片手確定
(7)(8) (9)(10) (11)	ある物体と両手が重なっている	2個以上(両手)	両手選択
(12)		1個(片手1)	片手確定
(13)		2個以上(片手2)	片手選択
(14)	手が動いている	1個(片手1)	移動
(15)		2個以上(片手2)	
(16)	手が回転している	1個(片手1)	回転
(17)		2個以上(片手2)	
(17)	手が動いている	1個(両手)	拡大・縮小

領域を検出する。

- (4) 連結成分のラベリング：ラベリングにより、得られた領域の数、その領域の重心を求める。
- (5) 手形状の識別：スクリーンに対して手を開いている場合、2個以上の赤外光領域が検出される。一方、ものを掴むように3本の指先を閉じている場合は、3つの赤外光が重なり合い1個の赤外光領域として検出される。

以上の処理で得られた手の形状（赤外光領域の数）と磁気センサから得られる手の位置姿勢情報により、図4の状態遷移図と表4の遷移表をもとにジェスチャコマンドを決定する。認識されたコマンドはグラフィックス用PCに伝えられ、コマンドに応じてCGデータが制御される。

4. MRキューブ：マイノリティ・リポート型HIとその発展形

映画『マイノリティ・リポート』に描かれた情景は、デザインの斬新さを無視すれば、アーチ型スクリーンに複数の2次元映像を表示し、これをジェスチャにより操作するものであると言える。ディスプレイが半透明であることに本質的な意味はなく、通常の前面もしくは背面投射型のスクリーンで同等の作業は達成できる。

これは前章のハードウェア構成で達成できたので、本研究では表示と操作空間に関して、次のような拡張・発展形を考える（表5）。

レベル1

基本となる広視野ディスプレイへの2次元映像の複

数表示を扱う。記録済み動画の再生だけでなく、仮想物体の映像の操作も考えられる。またネットワーク結合したLAN、WAN内に存在するオブジェクトも同等な対象になる。さらには、遠隔地のライブ映像を表示するのも、HI的にはこのレベル内でのバリエーションである。

レベル2

立体映像表示とイマーシブ・ディスプレイとの組み合わせは、既にVR分野でよく用いられている。偏光、液晶シャッタ方式が一般的で、これにより、スクリーン手前に飛び出した3Dオブジェクトを操作でき、眼前の空間も作業環境となる。

WIMP型がデスクトップ・メタファであるならば、この奥行きをもった広視野空間は部屋のメタファになり得る。大きな作業ボードであるスクリーンの他に、3D-CGで描いたごみ箱や戸棚を空間的に配した部屋を想定できる。その中で戸棚の中の仮想物体を取り出したり、不要な電子書類をごみ箱に向かって投げ捨てる、といった日常手慣れた身体的動作で円滑に操作する発展形が考えられる。

レベル3

電子的に配置して視認できるだけのCG製のごみ箱やファイルキャビネットではなく、実物の棚やごみ箱などの什器、プリンタ、スピーカといった出力機器を手の届く場所に配する発展形が考えられる。これによって、あたかもデータを実世界の棚に格納したりごみ箱へ捨てたりといった動作が可能になる。これは仮想世界と現実世界を融合する複合現実空間を構成することになる。複合現実感(Mixed Reality)¹⁹⁾²⁰⁾は、人工現実感の発展形として研究が活発な分野であるので、その研究成果も盛り込むことができる。作業者にとっては実物が良いか、仮想物の方で十分かの検証も研究対象となる。

実物体を配置するだけでなく、引き出しにセンサをつけて開閉を検出したり、スピーカなど機器のスイッチのON/OFFが現実世界にも仮想世界にも影響を及ぼすタンジブルI/Fへの発展も考えられる。日常使い慣れた実物体を使用し、そのアフォーダンスを活かすことは実世界指向の直観型インタフェース²¹⁾として興味深い研究対象となる。

レベル4

スクリーン型でなく、シースルーHMDを装着する本格的な複合現実感システムを導入して、広視野電子作業空間を構築する展開も考えられる。レベル3までは、スクリーンの正面に立つことが前提であるので、比較的大きなジェスチャをするとはいえ、操作の体験

場所は限定されている。HMD 装着型の体験であれば、もう少し移動の自由があり、それだけ広い作業領域を取ることができ、直感的な動作が可能になる。複合現実感システムの可搬化、ウェアラブル化が進めば、自由に歩き回った空間すべてが作業空間となる。複数人が向かい合って共同作業を行なうのにも適している。

以上の全レベルに対して音声コマンド、レベル 2 以降には 3D サウンド、触覚入出力の導入も考えられる。こうした拡張・発展形は表 5 のように整理でき、この種のシステムの総称を「MR キューブ」(MR3 = Minority Report-style Movement and Reaction in Mixed Reality space)と呼ぶことにした。

5. 応用システムの試作事例

5.1 ビデオ MR キューブ:映像再生・編集システム

我々の目指す広視野電子作業空間でのジェスチャ操作の有用性・操作性を確認・検証する第 1 ステップとして、映画『マイノリティ・リポート』中に見られる映像再生・編集操作に類したシステムを試作した。このシステムは 4 章で述べた MR キューブのレベル 1 機能に相当するもので、「ビデオ MR キューブ」と名付けた。

[機能概要]

本システムの操作風景を図 5 に、操作画面例を図 6 に示す。作業空間はアークスクリーン上の 2 次元面のみとし、立体映像は利用しない。中央画面が「再生・編集領域」で対象となる映像が大きく表示され、その下に映像の再生時間、および再生中の前後数コマが表示される。左右両サイドの画面は、それぞれ「動画データの一覧表示領域」兼「作業中ファイルの一時退避領域」で、サムネイルが表示される。また、スクリーン外部下方の見えない部分を、映像を捨て去る「ゴミ箱領域」と設定した。本システムで採用した映像の編集の操作コマンドを表 6 に列挙した。また、操作補助として、両手が指し示している位置にカーソルを表示する視覚フィードバックと、作業に応じたサウンドを発生する聴覚フィードバックを実現している。

[ジェスチャ・コマンドの実装]

ジェスチャ操作の基本セットは、表 3 の第 1 水準コマンドであるが、本システムでは「選択」「確定」「移動」「解除」のみを採用し、「回転」「拡大・縮小」コマンドは利用していない。

ビデオ MR キューブで実装した第 2 水準以上のジェスチャ・コマンドを表 6 に示す。第 2 水準コマンドセットからは「分割」「統合」「削除」が選ばれ、表 6 の上半分の「再生」「一時停止」「早送り/巻戻し」は、ビ



図 5 ビデオ MR キューブの操作風景
Fig. 5 The scene of the Video MR-Cube

表 6 動画再生・編集用ジェスチャ・コマンド

Table 6 Gesture commands for playing and editing video

コマンド	ジェスチャを利用した操作
再生	一覧表示領域に表示されている動画のサムネイルを作業領域に「移動」、作業領域に何もなければ再生が始まる
一時停止	再生中の動画、または作業領域下の再生中の動画のコマ送り画像を「選択」「確定」する
早送り 巻戻し	作業領域下のコマ送り画像を左右に「移動」すると、移動速度に応じて早送り、巻き戻しされる
分割	再生中の動画、または作業領域下のコマ送り画像の左右両端を両手で「選択」「確定」し、手を左右に離すと、一時停止中のコマの前後で分割される
統合	一時退避領域にある動画を、作業領域で再生中の動画の上に重ねるように「移動」すると再生中の動画の後ろに重ね合わせた動画が統合される
削除	作業領域または一時退避領域の動画を、スクリーン外下方のゴミ箱領域に「移動」する

デオ操作ならではの第 3 水準コマンドで、ビデオデッキにあるボタンに相当する。この第 3 水準コマンドのジェスチャは、本試作システムでは専用の複雑なジェスチャを定義せずに、第 1 水準コマンドをベースに、その組み合わせ、移動速度、移動先の領域の意味付けによって実現した。

直観的に実物体を操作するようなジェスチャといっても、画面上の映像オブジェクトは手が届く範囲にある訳ではない。人間が画面に近づくという選択肢もあったが、ここではあたかも差し出した手の方向に手が伸びているかのような操作性を与え、差し出した手の延長線とスクリーンが交わる点にカーソルを表示した。また、「確定」「一時停止」が行われたときには、効果音を出力した。本システムは、15~20fps の動作速度で稼働している。

[実体験結果と考察]

本システムの開発段階では、ジェスチャ入力 of 認識処理の実装を並行して行なった。まだ誤認識が多い段

表 5 MR キューブシステムのレベル分け
Table 5 Levels of MR-Cube Systems

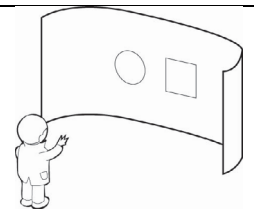
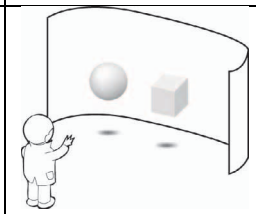
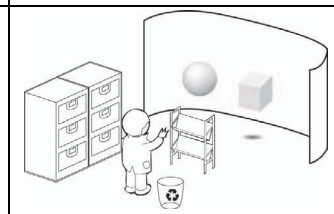
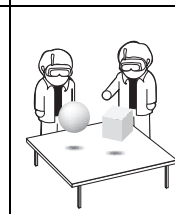
レベル 1	レベル 2	レベル 3	レベル 4
			



図 6 ビデオ MR キューブの画面例
Fig. 6 Screen image example of Video MR-Cube

階では、操作性が悪く、実感としてジャスチャ入力は耐えられるものではなかった。対話デバイスや照明環境の改善、パラメータ調整後、安定した認識結果が得られるようになってからは、ジェスチャ操作は快適に感じるようになり、ジャスチャのデザインも円滑に進行した。

しかる後に、本研究チーム以外の人間（含む、外部からの見学者）約 40 名にこのシステムを実体験してもらい、操作性・有効性に関する感想を求めた。この実験から得られた主な知見は、以下の通りである。

- 最初に操作方法を提示し、その後実際に手袋をはめて映像編集操作を実行してもらった。その結果、すべての被験者が問題なく編集操作を行うことができた。
- 体験後にコメントを求めたところ、ほぼすべての体験者から、直観的で操作感が高いとの評価を得た。
- ジェスチャ操作は腕が疲れるのではないかと危惧されたが、疲れるという意見はなかった。これは、腕を持ち上げなくても、肘から手までの動きでオブジェクトを操作できるためであると考えられる。
- 第 2 水準コマンドの「分割」は「両手で引き裂いてちぎる」もしくは「手刀で切る」が好ましく、「統合」は両手で粘土をくっつけるジャスチャの方が直観的で好ましいとの意見もあった。この意見を尊重するなら、ユーザのカスタマイズ機能を

許すことで解決できる。

- また、使用頻度の高い第 3 水準コマンドは、複数の基本コマンドの組み合わせではなく、専用のジェスチャを設けた方が便利だという意見もあった。
- 本試作システムでは、すべての体験者が円滑に編集操作を実行できることが確認できた。ここで、5 名の被験者には、動作速度を故意に 15fps 未満に設定したところ、オブジェクトが思い通りに操作できず、操作感が著しく低下することが確認できた。

以上のシステム開発と部外者の実体験結果から判断するならば、映画並みの快適さで「広視野電子作業空間」を実現することを否定する要因は見当たらない。ただし、本試作システムでは、有効性を確認するためまず安定したジェスチャ入力が達成できる環境を整えたが、もっと悪環境でも円滑に動作させるためにはジャスチャ認識方法の工夫がさらに必要かと思われる。

5.2 メディカル MR キューブ：各種医用画像の閲覧・整理システム

もう 1 つの試作システムは、MR キューブのレベル 3 までを目指すものである。広視野電子作業空間に立体映像表示、実世界のオブジェクトを導入するシステムを構築し、これらの問題点・有用性・操作性を確認する。

多種多様な画像データが存在し、立体映像表示が重要な意味をもち、かつ専門家の意見を得やすいという

理由から、ここでは対象として「医用画像」を選んだ。既に図1に示したのがそのイメージであり、本システムを「メディカル MR キューブ」と呼ぶ。具体的対象は、病院内で利用される各種医療用画像で、日々増加するデータを能率よく閲覧・整理し、診断所見も管理できるシステムを目指す。同一病院内の医局間を結ぶだけでなく、近未来には PACS(Picture Archiving and Communication System) が実用化されていて、遠隔地にある他病院の保管データも高速ネットワークでアクセス可能と考えられる。そうした統合管理システムの一部として本システムが活用されることを想定する。

各種検査結果のデータ(血液検査・生理機能検査・画像検査)、専門医による手術現場の映像、3D ボリューム・データの他に、VR/MR による手術シミュレーション、カルテ、投薬履歴なども表示対象となる。

利用者は、データの整理・分類を担当する若手医師や検査技師を想定している。日々の多忙な医療業務の中で、多種多様なデータを手際よく扱えることが望まれている。一般オフィス等に比べて、コストは大きな要因でなく、操作性・安定性・作業能率が重視される。

[拡張機能の概要]

(1) 3D 空間の利用

レベル1からレベル2への拡張として、スクリーンから体験者までの3D空間の任意の場所に、オブジェクトを配置可能とした。ユーザは、液晶シャッター眼鏡をかけることで、オブジェクトを両眼立体視することができる。対象としたデータは、2Dの静止画及び動画、立体視差をもったステレオ画像対(3D-CGで生成したのも含む)である。

ジェスチャ・コマンドセットは「ビデオ MR キューブ」を踏襲したが、先の尖った円錐形状のカーソルで3D空間内のオブジェクトの位置を指し示す機能を加えた。3D空間内でのポインティングは、手を最大限に伸ばした時に、最も奥のスクリーン面(利用者からの距離2.7m)を指し、手を最も縮めたときにユーザの目の前のオブジェクト(利用者からの距離0.4m)を指し示すものとした。

(2) 実物体の導入

さらにレベル3のシステムへの拡張として、ユーザが実世界での作業に日常用いるごみ箱・棚などの什器を導入した。これは「データを捨てる場所」「格納しておく場所」の分かりやすいメタファであると同時に、作業スペースをスクリーン上から日頃使い慣れた実世界へ拡張する働きをする。両眼立体視できる3D表示領域には限界があるので、広い作業領域を確保すると

いう目的からは、足下の実空間も活用することは意味がある。

この種の実物体を VPPE(Visually Perceivable Physical Equipment) と名付けたが、本試作システムでの VPPE としては、以下のようなものを配置した。

- データ長期保存用：木製キャビネット
- データ一時保管用：ファイルワゴン
- 不要データ削除用：ごみ箱

「木製キャビネット」は、デフォルトとして操作者の左手の不可視領域に配置するものとし、データをキャビネットまで移動し、手を放すとキャビネットに保存される。「ファイルワゴン」はスクリーン下の可視領域にあり、データを一時保管しておきたい場合に、サムネイル化して保存できる。「ごみ箱」は操作者の足下であり、データをごみ箱まで移動して手を放すとデータが削除される。ここで、実世界での利用でも頻りに動かす「ごみ箱」に関しては、磁気センサを取り付け、位置を自動検出することで、どこに置いても作動するようにした。

この他の VPPE として、電子データとしてのドキュメントやサウンドを物理的に出力する「プリンタ」や「スピーカ」なども考えられる。

[問題点と解決策]

上記の拡張機能を導入したが、早速いくつかの問題点に遭遇したので、次の解決策をシステムに常備することにした。

(1) 3D 空間での層状配置

データ・オブジェクトを3D配置することで空間の有効利用は図れたが、その一方でオブジェクトをランダムに配置すると、ものが散らかった部屋のようにオブジェクトへのアクセスが悪くなるとともに、どのオブジェクトが重要なのか整理が難しくなった。

この解決策として、オブジェクトを自在に配置できる3D領域ではなく、奥から手前に向けて多層の作業領域を設け、関心度・重要度が高いものを順に手前に配置する方式を採用した。ここで、ジェスチャ操作の利点を活かすのに、手が届く範囲に特別な意味を持たせることにした。

本システムでは3つの層を設けた。最前列の「フロントエンド(FE)層」は「ビデオ MR キューブ」の中央画面に相当する作業領域を手前に飛び出させ、軽く肘を曲げた楽な姿勢で操作できる距離(0.3-0.5m:可変)に配置した。多数のサムネイルが並ぶスクリーン面は「バックエンド(BE)層」で、利用者から2.7mの位置にあり、この層からデータを引き寄せFE層で作業する。さらに、FE層で頻りに使用するデータの

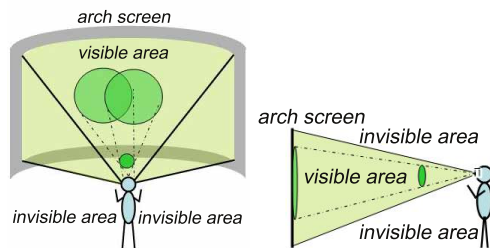


図 9 立体映像表示における可視領域
Fig.9 Visible area on stereoscopic display

待機場所 (1.0-1.7m : 可変) として「インタメディアート (IM) 層」を設けた。

それぞれの層に配置されたデータにアクセスするときは、FE 層は操作者が直接その位置に手を伸ばすことでアクセスできるが、IM 層、BE 層は距離があるので、図 7 に示すように、腕を伸ばすと BE 層、肘を少し曲げると IM 層にアクセスできるようにした。

本システムでの作業時のスクリーンイメージ例を図 8 に示す (ここでは便宜上単眼視データで示したが、実際には両眼視データが時分割で表示される)。

(2) サウンド機能による補完

スクリーンから体験者までの 3D 空間を活用する拡張は行なったが、視差のある左右の画像を両眼立体視する方式であるため、頭を頂点、スクリーンを底面とする四角錐領域内しか視認できないという制約がある (図 9)。広い作業領域を確保したいという目的からは、この四角錐の表示可能領域の外にごみ箱等の VPPE を配置するのは理にかなっていない。しかし、立体表示領域と足下の VPPE の間は連続していないので、データをごみ箱に移動する途中から手で掴んだはずのデータ・オブジェクトが視認できなくなるという問題が生じた (図 10)。実際に約 10 名の被験者に体験してもらったところ、約 9 割が「データをごみ箱に入ったという実感がない」と回答した。

この解決策として、データをごみ箱 (と想定する領域) に入った瞬間、同期して効果音を出力するようにしたところ、全員から「データをごみ箱を捨てた感じがする」との回答が得られた。そこで、本システムでは、ごみ箱以外の VPPE への操作にも効果音を付すことにした。それぞれの効果音の ON/OFF は、ユーザが選択できる。

〔実体験結果と考察〕

上記の改善策を講じた後に、研究チーム外の約 20 名に本システムを体験してもらった。「ビデオ MR キューブ」と同様に、最初に操作方法を提示し、手袋をはめて体験してもらった結果、以下の回答を得た。

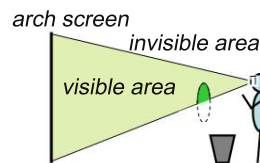


図 10 VPPE の不可視領域への配置
Fig.10 Placement of VPPE into the invisible

- 3つの層を利用したデータ配置は、重要度が一目で分かるという概ね好評であった。しかし、BE 層のデータにアクセスする際、腕を伸ばす必要があるため、疲れるという声があった。また、BE 層に対しては、まっすぐ伸ばした手の先にカーソルを表示するよりも、指の先からレーザービームを模した光線を CG 表示することを望む声もあった。
- 当然のことながら、FE 層は IM 層を、IM 層は BE 層を覆い隠してしまう。FE 層、IM 層を半透明表示すべきかどうかは、意見が分かれた。この隠された部分まで活用するのか、それをアクセスする場合、FE、IM 層の側面に回り込んで覗き込んだり、あるいは手前の層を一旦非表示にする機能を持たせるのかも、今後検討すべき課題である。
- VPPE が不可視領域に配置されている場合は、移動する途中から手で掴んだはずのデータ・オブジェクトが視認できなくなるが、その半面、VPPE には十分な大きさがあるので、周辺視でその存在と位置を視認できる。慣れ親しめば、VPPE の位置を意識しなくても操作できるようになり、作業効率が上がること確認されている。VPPE の導入には概ね好意的で、これまでにない体験に興味を示す被験者が少なくなかった。
- 「ビデオ MR キューブ」と異なり、立体映像表示を常用するために、「長時間利用する場合は目が疲れる」という指摘があった。その半面、奥行き感を誇張した立体映像は用いていないので、「テーマパーク等で体験する 3D 映像ほどは疲れない」という意見もあった。

以上の試作開発と評価実験を総合すると、「広視野電子作業空間」の実現には多々興味深い問題が内在しており、HI 分野として今後大いに研究を進める価値があると考えられる。

6. む す び

ポスト WIMP の実現性のある提案として、広視野電子作業空間を有効に活用するシステム「MR キューブ」の構想とその実現例について述べた。また、この

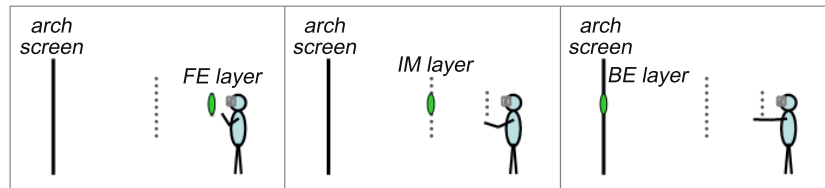


図 7 各層のデータへのアクセス方法
Fig. 7 Access method to each layer

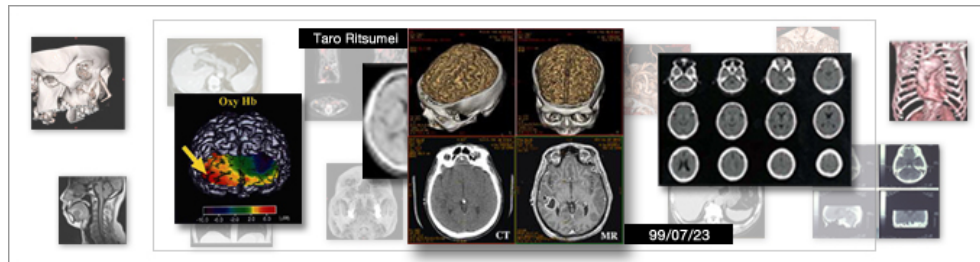


図 8 メディカル MR キューブの画面例
Fig. 8 Screen image example of Medical MR-Cube

ような作業空間を操作するためのジェスチャ・コマンドセットについて検討した。2つの応用システムを試作開発し、広視野電子作業空間をジェスチャ操作する上での操作性・有効性を評価した。

本研究は、SF映画での描写を引き合いに出したように、「広視野電子作業空間」が近未来社会で必ず求められるとの前提に立ち、これをいかに実現するか、矛盾なく素直に実現できるのかを検討し、そこから生じる問題の解決や発展形を探ることを目的としている。

これまでもジェスチャ認識に関する研究は数多いが、こうした広視野領域を対象としてジェスチャ・コマンド体系を考え、本格的に作業効率を論じた研究はない。本研究実験（レベル1）の結果は上々で、ジェスチャ入力の識別率が許容範囲で応答速度が15~20fps程度あれば、人は円滑にジェスチャ操作を実行でき、ジェスチャ・コマンドの種類も予想以上に覚えられることが確認できた。映画並みの未来型システムの実現を阻む大きな障害要因はないと言える。

広視野電子作業空間のレベル2, 3, 4への拡張は、我々の独自のアイデアであるが、既にレベル2, 3で遭遇した問題からHI研究にとって興味深い課題が見つかった。多層表示の視認性、実物体の導入時のサウンド補完機能などの解決策は、他分野にも適用できると考えられる。ただし、まだ一般化して結論づけられるものではないので、今後様々な種類の広視野空間向きの作業を想定し、ヒューマンインタフェース学としての評価実験を行って行きたい。

謝辞 「メディカルMRキューブ」の近未来の利用

価値や外科医としての要望に関して、関西電力病院西藤隆太博士に貴重なご助言を賜った。ここに深甚の感謝の意を表します。本研究の一部は、(財)国際コミュニケーション基金の助成による。

参考文献

- 1) Dam, A.v.: Post-WIMP User Interfaces. Comm. ACM, Vol. 40, No. 2, pp. 63 - 67, 1997.
- 2) 暦本純一: 実世界指向インタフェース?実空間に拡張された直接操作環境, 情報処理, Vol. 43, No. 3, pp. 217 - 221 (2002).
- 3) 石井裕: タンジブル・ビット?情報と物理世界を融合する, 新しいユーザ・インタフェース・デザイン?, 同上, pp. 222 - 229 (2002).
- 4) Bae, S., Kobayash, T., Kijima, R., and Kim, W.: Tangible NURBS-Curve Manipulation Techniques Using Graspable Handles on a Large Display. Proc. UIST 2004, pp. 81 - 90 (2004).
- 5) Khan, A., Fitzmaurice, G., Almeida, D., Burtynyk, N., and Kurtenbach, G.: A Remote Control Interface for Large Displays, ibid., pp. 127 - 136 (2004).
- 6) 田村, 柴田, 木村: 未来創像学?SF映画に学ぶIT機器とHIの未来形, 電子情報通信学会フェロー&マスターズ未来技術研究会, FM05-1-2, pp.5 - 10 (2005)
- 7) Bolt, R.: Put-That-There: Voice and Gesture at the Graphics Interface, Proc. SIGGRAPH 80, pp.262 - 270 (1980).
- 8) Koike, H., Sato, Y., and Kobayashi, Y.: Integrating Paper and Digital Information on En-

- hancedDesk: A Method for Realtime Finger Tracking on an Augmented Desk System, ACM Trans. on Computer-Human Interaction, Vol. 8, No. 4, pp. 307 - 322 (2001).
- 9) 石井, 中西, 小池, 岡, 佐藤: EnhancedMovie: 机型インタフェースを用いた動画編集システム, WISS2003 論文集, pp. 41 - 46 (2003).
 - 10) Wu, M. and Balakrishnan, R.: Multi-finger and Whole Hand Gestural Interaction Techniques for Multi-user Tabletop Displays, Proc. UIST 2003, pp.193 - 202 (2003).
 - 11) Elrod, S. et al.: Liveboard: A Large Interactive Display Supporting Group Meetings, Presentations, and Remote Collaboration, Proc. CHI 92, pp. 599 - 607 (1992).
 - 12) Ishii, H. and Kobayashi, M.: ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact, Proc. CHI 92, pp. 525 - 532 (1992).
 - 13) Cao, X. and Balakrishnan, R.: VisionWand: Interaction Techniques for Large Displays Using a Passive Wand Tracked in 3D, Proc. UIST 2003, pp. 173 - 182 (2003).
 - 14) Pavlovic, V., Sharma, R. and Huang, T.: Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review, IEEE Trans. on PAMI, Vol. 19, No. 7, pp. 677 - 695 (1997).
 - 15) Long, J., Landay, J., and Rowe, L.: Implications for a Gesture Design Tool, Proc. CHI '99, pp. 40 - 47 (1999).
 - 16) Baudel, T. and Baudouin-Lafon, M.: Charade: Remote Control of Objects Using Free-Hand Gestures, Comm. ACM, Vol. 36, No. 7, pp. 28-35 (1993).
 - 17) Segen, J. and Kumar, S.: Gesture VR: Vision-Based 3D Hand Interface for Spatial Interaction, Proc. 6th ACM Int. Conf. on Multimedia, pp. 455 - 464 (1998).
 - 18) Freeman, W. and Weissman, C.: Television Control by Hand Gestures, Proc. Int. Workshop on Automatic Face and Gesture Recognition, pp. 179 - 183 (1995).
 - 19) 田村, 大田: 複合現実感, 映像情報メディア学会誌, Vol. 52, No. 3, pp. 266 - 272 (1997).
 - 20) Feiner, S. K. (田村秀行訳): 複合現実感がひらく第3の視界, 日経サイエンス, 2002年7月号, pp.40 - 49 (2002).
 - 21) 池田, 木村, 佐藤: 道具の持つアフォーダンスを利用した触覚フィードバックデバイス, 日本VR学会論文誌, Vol. 7, No. 3, pp. 339 - 345 (2002).



木村 朝子 (正会員)

1996年大阪大学基礎工学部卒。1998年同大学院基礎工学研究科修士了。同大学助手を経て、立命館大学理工学部助教授。現在、同情報理工学部メディア情報学科助教授。

博士(工学)。実世界指向インタフェース、複合現実感、ハプテックインタフェースの研究に従事。2001年より2002年まで Mayo Clinic にて Special Project Associate。電子情報通信学会、情報処理学会、ヒューマンインタフェース学会、日本バーチャルリアリティ学会、ACM、IEEE 各会員。



柴田 史久 (正会員)

1996年大阪大学大学院基礎工学研究科博士前期課程修了。1999年同研究科博士後期課程修了。大阪大学産業科学研究所助手を経て、2003年4月より立命館大学理工学部助教授。

現在、同情報理工学部情報コミュニケーション学科助教授。博士(工学)。モバイルコンピューティング、複合現実感等の研究に従事。電子情報通信学会、日本ロボット学会、日本バーチャルリアリティ学会、IEEE 各会員。2004年日本バーチャルリアリティ学会学術奨励賞受賞。



鶴田 剛史

2004年立命館大学理工学部情報学科卒。現在、同大学院理工学研究科博士前期課程在学中。???の研究に従事。



酒井 理生

2005年立命館大学理工学部情報学科卒。現在、同大学院理工学研究科博士前期課程在学中。???の研究に従事。



鬼柳 牧子

2005 年立命館大学工学部情報
学科卒。現在，同大学院理工学研究
科博士前期課程在学中。??? の研究
に従事。



田村 秀行（正会員）

1970 年京都大学工学部電気工学
卒。工業技術院電子技術総合研究所，
キヤノン（株）等を経て，2003 年 4
月より立命館大学工学部教授。現
在，同 情報理工学部メディア情報学
科教授。工学博士。パターン認識，画像情報処理，マル
チメディア，バーチャルリアリティ等の研究推進と実
用化に従事。本学会論文賞，人工知能学会功労賞等受
賞。編著書：「コンピュータ画像処理」（オーム社）、「デ
ジタル映像」（日本経済新聞社）など。IEEE，ACM，
電子情報通信学会，人工知能学会，映像情報メディア
学会，日本バーチャルリアリティ学会等の会員。
