

# 擬人化エージェントを利用した複合現実空間でのインタラクション

穴吹まほろ 山本裕之 田村秀行  
(株) MR システム研究所

現実世界と仮想世界を融合した複合現実 (Mixed Reality; MR) 空間とのインタラクションに擬人化エージェントを導入することを試みた。リビングルームでのインテリア・シミュレーションを題材に、音声対話機能を持つロボット型エージェントに仮想家具の操作代行を行わせた。この種のエージェントの持つべき機能とユーザに与えるヒューマン・ファクタについて考察した。

## Interacting with Embodied Interface Agent in Mixed Reality Space

Mahoro Anabuki, Hiroyuki Yamamoto, and Hideyuki Tamura  
Mixed Reality Systems Laboratory Inc.

This paper describes an embodied interface agent introduced in Mixed Reality (MR) space, where the real and virtual worlds are merged together. As a case study of interior simulation in a living room, a robot type conversational agent is implemented to operate virtual furniture for the users. Functions necessary for this type agent and human factors given to the users are principally studied.

### 1. はじめに

複合現実感 (Mixed Reality; MR) は、現実世界と仮想世界を融合し体験する技術である[1]。従来の人工現実感 (VR) は主としてコンピュータ内に人為的に生成した仮想環境を扱ってきたが、われわれ人間が生活し行動する現実世界に立脚したMR空間体験は、VR技術の限界を打破するものとして大きな期待が寄せられている。現在、現実・仮想両空間の幾何学的・光学的な合成手法が活発に研究されている[2][3]。

VRは視覚・聴覚・触覚に訴える空間型メディアであり、ユーザが身体性を保ってリアルタイム対話できる新しいヒューマン・インタフェース技術として注目を集めた。この特質をすべて継承し、目の前の現実世界とのインタラクションも可能なMR技術は、HCI研究にとってさらに興味深い題材を提供する。

一方、HCI分野では、視覚化された擬人化エージェントが活発に研究されている[4]。この2つを組み合わせる、即ち、MR空間に身体的表示をもつインタフェース・エージェントを存在さ

せたら何が起こるだろうか、というのが本研究の発想のもとである。MRシステムで物理世界に融合される仮想世界は、必ずしも重畳表示されるCGデータだけでなく、そこから広大なサイバースペースにリンクされていてもいい。この場合、MR空間に置かれた擬人化エージェント (以下、MRエージェントと呼ぶ) は、サイバースペースの入口としてのインタフェース・エージェントと考えられる。また、MR空間中を自由に動き回れるエージェントの振る舞いや存在感は、コンピュータ・モニタ上のウィンドウの中に登場する擬人化エージェントとは大いに様相が異なるに違いない。

こうした観点から、我々は音声対話機能をもつMRエージェントの具現化を思い立ち、リビングルームでのインテリア・シミュレーションを支援するロボット型エージェントWelboを開発した。Welboを含むMRシステムの技術的側面は文献[5]で報告したので、本稿ではシステム構成の記述は最小限に留め、HCI研究の観点からMRエージェントの果たす役割と運用から得たヒューマン・ファクタ面での知見を述べる。

## 2. 複合現実空間における擬人化エージェント

### 2.1 MR エージェントの役割と特質

現実の物理空間に仮想空間の情報を融合するMRシステムでは、通常シースルー型HMD(Head Mounted Display)を用いて仮想データを重畳表示する。例えば医療分野では、患者の体表面に体内の映像を重ね合わせて見ることにより、診断や手術計画に役立つ。同様にCADデータがあれば、屋内では床下配線を、屋外では地中に埋蔵されたガス・水道等の配管を透視することができる。また、建築物の景観シミュレーションでは、ビデオやコンピュータ画面で完成予想図を見るのではなく、直接現場でHMDをかけて合成シーンを体験できる。いずれの場合も、ユーザの主観的視点でMR空間を実時間体験できるシステムが前提となっている。

こうしたMR空間に登場させるエージェントに期待する役割は、

- ・利用体験の少ないユーザに対応し、利用法や背景知識を教示する
- ・仮想物の選択や操作を代行する
- ・ユーザの移動に随行し、体験地点に応じた案内をする

などである。

これまで研究されてきた擬人化エージェントの目的が、利用者にとっての対話の自然さ、身体表示を持つことによる表現力の向上、存在感や親しみやすさ等であるなら、これらはMRエージェントにも当てはまる。一方、MRエージェントのならではの長は、MR空間という独自の環境に存在し、その中を自律的に移動する行動範囲の自由度である。

従来の擬人化エージェントのほとんどは、コンピュータ・モニタ上に存在してTVニュースのアナウンサのようにユーザに正対し、行動の範囲が

四角い2次元画面に限定されていた。南カリフォルニア大のSteve [6]は3次元仮想空間中を移動できるが、VRとMRの関係通り、現実空間と関わりを持つことはできない。電総研のジョニー [7]は、ユーザの質問に答え、現実空間中の事物を指さすことができるが、その存在はモニター画面上に限定されている。MITメディアラボのALIVEシステム[8]の自律エージェントは、ユーザと同一の空間内に存在するが、その様は客観視点の合成映像で表示され、ユーザ自身の主観視点で見ることができない。

これに対して我々の考えるMRエージェントは、以下のような長を有している。

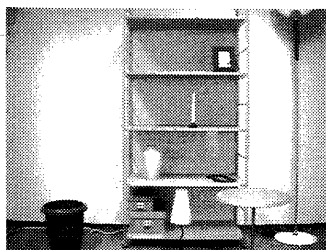
- (a) 立体視、立体音響を用いることで、眼前の空間に高レベルの実体感を描出することができる。
- (b) ユーザと同一空間内で、現実と整合のとれた実寸の仮想物を扱い、自らも実空間のスケールでユーザと接することができる。
- (c) ユーザの周り360°自由な空間位置に存在でき、現実物体・仮想物体との3次元遮蔽関係も表現できる(即ち、物体の陰に隠れたり、現れたりできる)。

### 2.2 MR エージェントの機能と振る舞い

筆者らは、複数人が現実空間と仮想空間を同時共有し、かつ実時間対話性を実感できる事例としてゲーム型のMRシステム[9][10]を開発してきた。しかし、MRエージェントの具現化対象としては、図1に示すようなインテリア・シミュレーションを選んだ。これは、リビングルームを想定した2.8m×4.3mの物理的に存在する空間に実物の家具や調度品が配されていて、ここに電子カタログに存在する仮想物を追加配置するというビジュアル・シミュレーションである[11]。スコアを競う対戦型のゲームや、ユーザ自らが手を動か



(a) リビングルームの外観



(b) 現実空間の光景



(c) 複合現実空間の光景

図1 MRリビングルーム・システム

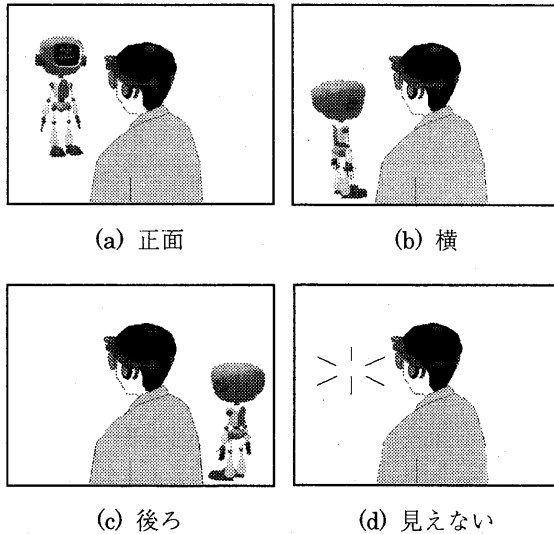


図2 エージェントの対応位置

す組立て作業よりも、MR エージェントの存在感や動作の代行を試すのに適しているからである。

我々が、このMR エージェントに与えた機能の概要は以下の通りである。

- ・ユーザの発話を音声認識し、合成音声で答える音声対話機能をもつ
- ・擬人化された身体をもち、関節を動かすことで動作を表現できる。
- ・ユーザの意図により、仮想物を選択し配置する動作を代行する。
- ・ユーザの指示がない場合も、自律的存在としてMR空間内で行動できる。
- ・ユーザの頭部位置・視線方向を知り、それに応じた行動が取れる。

この種のMR エージェントの存在や振る舞いとして特に興味深いのは、どのような大きさのキャラクターが、空間内のどの位置で対応することが好まれるかである。ユーザの視野の正面で対応するか、エレベータ・ガールのように斜め前に存在するか、VIPの通訳や大相撲協会の理事長のように後方耳元で待機するか、時として透明人間のようになら声だけの存在が望ましいのか、等々である(図2)。また、ユーザの視点移動に対して、常に視野内にいるのか、さりげなく見えないところにするべきかも検討に値する。個人差があることも想定して、こうしたヒューマン・ファクタを分析できるよう実装することにした。

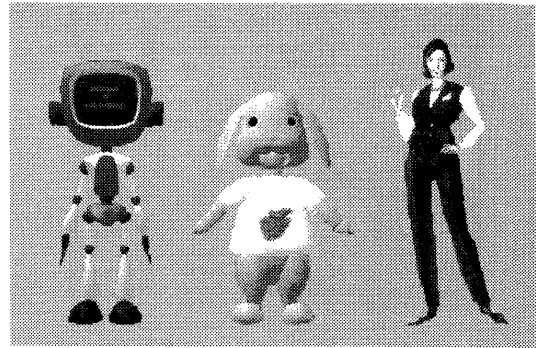


図3 エージェントのキャラクタ・デザイン

### 2.3 キャラクタのデザインと選定

過去の研究例を参考にして、MR エージェントのキャラクターを3種類デザインした。図3に示すロボット型、小動物型、(ニュースキャスタ風)人物型である。想定ユーザ10数名の意見、好みを反映して、キャラクターの選定と動作の決定をすることにした。

まず、現時点でのCG技術の能力を考慮して、人物型キャラクターが脱落した。CGアニメ、ゲーム、ウェブサイト上には、若い女性CGキャラクターも珍しく無くなったが、写実的な人物像を描くのは易しくない。実時間での描画にはポリゴン数の制約が大きいため、「目の前で対応する存在としては、中途半端なりアリティは却って気味が悪い」という意見が圧倒的であった。MR用途の場合には、現実空間との位置合わせや本来の仮想物(この場合、家具や調度品)のレンダリングにリソースを取られ、エージェントは付属物扱いであることは止むを得ない。

残る2体には、関節を持たせ動きや音声を与えて、MR空間内オーディションを行った。その結果、ロボット型が選ばれた。理由として、関節の動きが分かりやすく動作が視認しやすかったのと、合成音声との相性が挙げられた。現時点では最高レベルの規則合成音声を採用したが、やはりロボットの外観との親和性が高かったようである。

このロボット型MR エージェントを Welbo (Welcome Robot の意)と名づけた。Welboには、64関節をもつ約2000ポリゴンの幾何形状が与られている。フロア面に着地し歩く動作も実装され

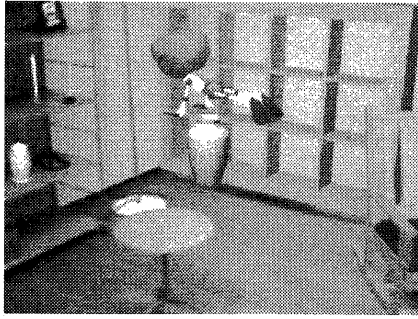


図4 飛行しながら花瓶を運ぶ Welbo

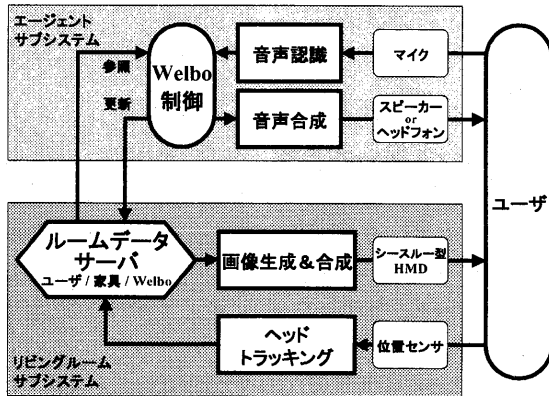


図5 システム構成図

たが、結果的には採用されなかった。ここでも中途半端な動きのリアリティは敬遠され、空間中に浮遊し飛び回って家具を運ぶ様の方が好まれた(図4)。

### 3. Welboのシステム構成

MR エージェント Welbo を実現するシステムは、MR 空間を描出するリビングルーム・サブシステムと、Welbo を制御するエージェント・サブシステムから構成されている(図5)。

#### 3.1 リビングルーム・サブシステム

リビングルーム・サブシステムは、さらに次の3つのモジュールに分かれている。

##### (1) ヘッドトラッキング部

ビデオシースルー型 HMD (自家製) に取り付けられたマーカを光学式センサ (OPTOTRAK 3020) で計測し、ユーザーの視点位置・姿勢を算出する。この測定誤差を実時間で補正するため、リビングルーム内に配された補正用マーカを画像認識する方法を併用している[12]。

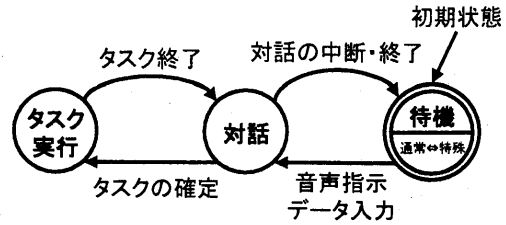


図6 状態遷移図

##### (2) ルームデータ・サーバ

ユーザの位置・姿勢、家具の形状・位置・姿勢と、擬人化エージェントの位置・姿勢・ポーズのデータ(ルームデータ)を管理する。仮想家具だけでなく、実物の家具の幾何モデルデータも管理されている。

##### (3) 画像生成・合成部

ルームデータに基づいて、仮想の家具とエージェントのCG画像を生成し、ユーザ視点の実写映像と合成する。合成結果は、ビデオシースルー型HMD(92万画素)に立体映像として表示される。

コンピュータは、画像生成・合成部に SGI 社 Onyx 2, ヘッドトラッキング部とルームデータ・サーバに SGI 社 O2, さらに補正用マーカ認識専用画像認識ボード (IP5005, 日立製) を搭載した PC を使用している。

### 3.2 エージェント・サブシステム

エージェント・サブシステムは、下記の3つのモジュールからなる。音声合成部用に PC1 台、他の2つのモジュール用に SGI 社 O2 を用いている。

#### (1) 音声認識部

ジョニーと同様の対話モジュール(電総研製)[13]を利用して音声認識と文章理解処理を行う。認識対象文は、挨拶と家具の配置変えに関する文章のみに限定している。認識結果は音声指示データに変換し、Welbo 制御部へ出力する。ここで、音声指示データとは [内容+対象+場所+方向] なる文法1つで全認識対象文を表現するデータである[5]。

#### (2) Welbo 制御部

音声認識部からの音声指示データ入力やユーザとの対話の経過に応じて状態を遷移させながら、エージェントの振る舞いを制御する(図6)。

「待機状態」では、エージェントがユーザ付近を浮遊するようにその位置と姿勢データを制御する。待機状態時に音声指示データが入力されると「対話状態」となり、発声と身振りを制御してユーザとエージェントの対話を実現する。対話が中断・終了すれば再び待機状態へ戻り、対話によってタスクが確定すればタスク実行状態となる。タスク実行状態では、仮想家具とエージェントの位置と姿勢、およびエージェントの身振りを制御することによって、家具の配置変えや場所指示を実現する。

仮想家具とエージェントの位置と姿勢の制御は、ルームデータ・サーバで管理されているデータの更新によって行われる。エージェントの身振り制御はエージェントのポーズデータ(キャラクターの関節角度データ)を、連続的に更新することによって行われる。発声の制御は、音声合成部へテキストデータを送ることによって行われる。

### (3) 音声合成部

Microsoft Speech API およびキャノン製日本語音声合成エンジン PureTalk を利用している。入力されるテキストデータを、部屋内に置かれた1対のスピーカー、もしくはユーザが装着するステレオ・ヘッドフォンから出力する。このとき、エージェントの存在位置から声が聞こえるように、左右の音源の音量が調節される。

## 4. MRエージェントの与えるヒューマン・ファクタ

### 4.1 エージェントの大きさと位置に関する考察

フルバーチャルの仮想空間のスケールや空間移動は、相対的・定性的なもので差し支えないが、目の前の現実空間と仮想空間を重ね合わせるMR空間では、すべての事物は絶対的な大きさを保つ必要がある。こうした臨場感が保たれた空間中では、エージェントの身体も実物スケールで考える必要がある。

そこまでの実在感があるなら、人間とほぼ等身大の存在が望ましいかと思われたが、この予想は見事に外れた。人間よりかなり小さめのキャラクターが好まれる傾向にある。以下、10数名の被験者による実験結果と考察は以下の通りである。

- (a) Welboのデザインは頭部が大きいこともあるが、他のキャラクターの場合でも人間大は好まれない。本システムのHMDの視野角は $51^\circ$ と

限られているため、視野に比して大きなエージェントは圧迫感を与えてしまうようである。

- (b) 同じWelboをコンピュータ画面で見ると比べ、約半数はHMD利用の場合にエージェントが大きく感じると答えた。ここでも見かけの視野角が影響していると考えられる。
- (c) 視野内の大きさは、絶対的なサイズとユーザからの距離の両方で制御できる。しかし、立体視可能なHMDの場合、ユーザからの絶対的な距離での存在感があるので、小さく見せるために遠くにエージェントを位置すると対話のしやすさが損なわれる(図7)。
- (d) Welboの身長は20cmから160cm、対応位置は30cmから3mまで可変とし、ユーザの好みで選択させたところ、身長40cm弱、待機位置90cm近辺が最も好まれた。個人差はあるが、総じて全身が画面内に入る程度の大きさが好ましいようである。
- (e) 正対して会話する場合に比べ、遠くで家具を運んでいる場合には、少し大きい目の方が動作は印象的になる。しかし、絶対的な大きさを状況に応じて変更するのは不自然なので、対話状態で望まれる大きさをWelboの標準の大きさと考えた。
- (f) こうした実験だけでなく、ユーザの体験時にも大きさを調整できる機能を解放することが考えられる。しかし、エージェントを欲するのは面倒な動作を代行して欲しいからであり、実用上は上記のようなデフォルト値に設定して置くのが望ましい。

### 4.2 エージェントの待機状態に関する考察

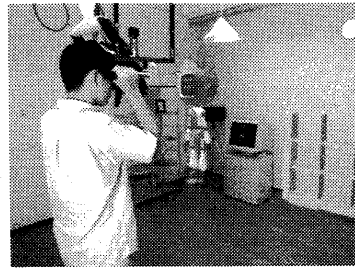
HMDの利用によって見回しが可能な3次元MR空間は、その広さに対して一度に観察できる領域に限られる。そこで、Welboの位置を認識しやすいよう、指示を待つ待機状態で、ユーザの動きに合わせて積極的に位置を変えるモードと追跡機能を設定した。

#### (1) 通常待機モードと特殊待機モード

初期状態とタスクの終了時に「通常待機モード」に入る。ユーザを追跡して視点の斜め前方の位置へ移動し、HMDの視界の端に浮遊して、ユーザが話し掛けやすい状況を作る。指示がないまま、待機状態が通算で一定時間(デフォルト値

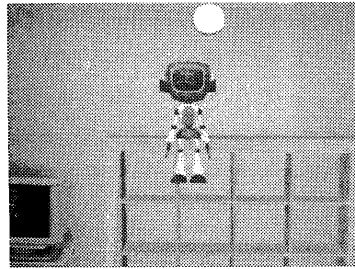


ユーザ視点画像

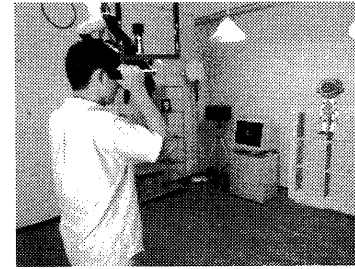


客観視点画像

(a) 距離が近いとき (20cm)



ユーザ視点画像



客観視点画像

(b) 距離が遠いとき (150cm)

図7 ユーザからの距離による視野内の大きさの違い (身長 36cm 時)

は 90 秒) 経過すると、ユーザの関心を引くために、Welbo は身振り状態となって、何らかの動作 (例：手を振る) を行う。動作後は再び待機状態となる。さらに指示がないまま、待機状態が長時間 (デフォルト値は 6 分) 経過すると、しばらく仮想の家具移動は行われないと判断し「特殊待機モード」に切り替わる。このモードで Welbo は室内を散歩したり、ソファで休息を取ったりする (図 8)。ユーザからの発話があれば、対話状態に戻る。2 つのモードの切り替え、および各モードの内部状態を図 9 の状態遷移図に示す。

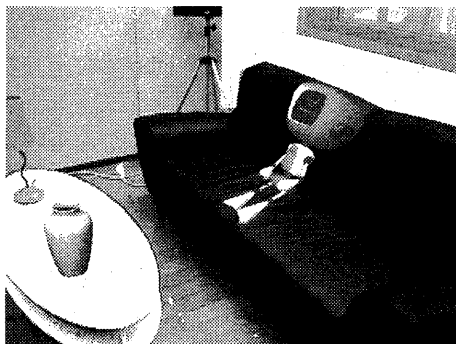


図 8 休息中の Welbo

(2) 追跡設定と非追跡設定

待機状態で「追跡設定」と「非追跡設定」を選択することができる。追跡設定時には、ユーザの動きによって視界から外れると追跡を開始し、視野内に入ると待機状態に戻る。非追跡設定では、状態を変えずにそのままの位置で浮遊を続ける。

様々な設定で実験し、体験者の反応を調査したところ、半数以上から「呼んだときだけ現れて、それ以外は邪魔にならないところに待機してい

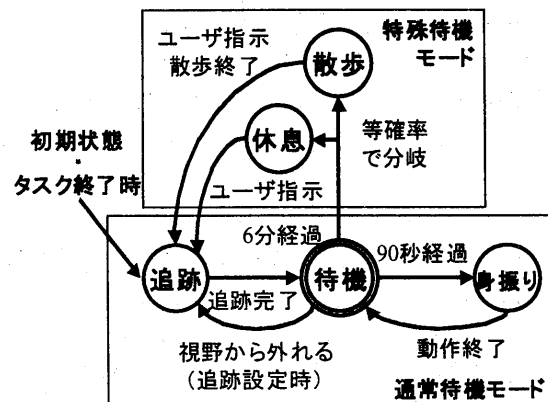


図 9 指示を待つ間の状態遷移図

てもらいたい」との反応があった。かなり小さなエージェントを用いない限り、しばしば視界の中に入ってくる追跡設定は好まれていない。ユーザは対話や指示した仮想物の移動時には Welbo を注視しているの、主体的にリビングルーム内を見回しているときは、動きのある追跡は煩わしい存在と感じられてしまう。

タスクにも依存するので、まだこの実験結果が一般的であるとは言い難い。現在の MR リビングルームは狭く、どこにいてもすぐに Welbo の居場所は確認できる。このため、元の位置で待機するか、決まった位置に待機していれば十分である。しかし、もっと広い空間が対象の場合、自分のエージェントを見失ってしまう恐れがある。そうすると「追跡はするが、視界には入らず、さりげなく近くに待機する」という慎み深いエージェントが欲しくなってくることだろう。

筆者らはかつて、多地点から複数人が仮想空間を共有し、参加者それぞれの存在をアバタで表示しながら、バーチャルモールでショッピングするシステムを開発した[14]。この場合は、同伴者や応対してくれる店員が、仮想空間内でどこにいるのかを確認するのに、アバタ表示はかなり有効であった。常に会話相手を視野内においておこうとする傾向があった。このシステムでは、本物の人間同士の会話が主であったからと考えられる。用がないときには現れるなどというのは、まだまだエージェントとの対話が自然でなく、本当に役立つ存在になっていないからであろう。

図 2(c)で想定したような後方耳元での待機は、実装はできるが、今回のシチュエーションでは余り意味がなかった。そうした解説モードが必要な状況になかったからである。また、今回の実装ではステレオ音響のみを利用したが、本格的には 3 次元音場を導入して聴覚の 3 次元空間特性も利用するのが望ましい。

いずれにせよ、待機状態に関しては、もっと MR エージェントの機能を充実させ、実用性のあるタスクに遭遇したとき改めて議論すべきことと思われる。

#### 4.3 MR エージェントのもつ娯楽性と遊び心

前項の特殊待機モードで述べた「散歩」や「休息」は、必要不可欠な機能ではなく、多少の遊び心から持たせた機能である。浮遊した待機状態

の挙動や Welbo の発言にも、ある種の遊びが埋め込まれている。「ちょっと重いなあ」「疲れるなあ」といった類いの発言である。

これらは当初目的とした HCI 研究の本質ではないが、開発者の意図以上に好評であった。音声認識のミスもまた愛嬌として好まれる傾向がある。それは、本システムは実運用を目的としたものでなく、研究目的での試作・実験に過ぎないので、見学者の体験が大半であるためと考えられる。これは、ショールームや娯楽を目的としたアトラクションで、MR エージェントのような存在が観客を和ませる働きをすることを意味している。無味乾燥になりがちな科学技術教育への応用でも、演出効果として大いに有望だろう。

Welbo の TV モニタ風顔面は、せいぜい関連事項の文字情報を表示できるだけで、顔の表情を描くことは想定していない。その文字情報もまだ十分に使いこなしていないが、エージェントの感情・表情を明示的に表現するには別のキャラクターが必要だろう。これは、通常の擬人化エージェントの研究でもなされているが、MR エージェントの場合存在感が大きいだけに、いろいろ工夫する必要はあるだろう。

## 5. むすび

現実世界と仮想世界を融合した MR 空間に擬人化エージェントを登場させる試みを、インテリア・シミュレーションを題材としたシステムで実現した。第 1 段階として、Welbo と名づけたロボット型のエージェントに音声対話機能と仮想物操作代行機能を持たせ、この種のエージェントがユーザに与える影響・効果を確認した。

MR 空間中の擬人化エージェントの大きさと位置については、事前の予想よりも小さめ存在が好まれ、必要のない限りユーザの視界の現れない方が好ましいことが判明した。他の例のない実施例だけに、まだこれ以上の結論を導き出すのは難しいが、MR エージェントという存在は今後も HCI 研究に有益な知見を与えるものと思われる。実際、エージェントの空間中の移動速度や 3 次元音響効果など、その設計がインタラクションに影響を与える要素を多く持っている。

また、MR 空間では複数のユーザと複数の擬人化エージェントが同一空間に存在できる。複数のユーザのインタラクションを支援するエージェ

ントや、複数のエージェントによるインタラクション支援の実現も今後の課題である。

## 謝辞

音声認識および言語処理モジュールの利用をご快諾いただいた電子技術総合研究所秋葉友良氏、伊藤克亘氏、長谷川修氏、速水悟氏に深謝します。日本語音声合成エンジン PureTalk は、キャノン（株）情報技術研究所麻生隆氏の協力を得てシステムに実装しました。ここに記して謝意を表します。

また、（株）MR システム研究所で本研究の基本方針の立案を担当された若月裕子氏、有益な議論および御指導を頂いた佐藤清秀氏、内山晋二氏、大島登志一氏、CG キャラクタ制作を担当された角田弘幸氏、秋山由希子氏に感謝します。

## 参考文献

- [1] 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, 1998.
- [2] Y. Ohta and H. Tamura (eds.): *Mixed Reality – Merging Real and Virtual Worlds*, Ohmsha & Springer-Verlag, p.418, 1999.
- [3] 田村秀行 (編): 「複合現実感」特集, 日本 VR 学会論文誌, Vol.4, No.4, 1999.
- [4] E. André (ed.): Special Issue on “Animated Interface Agents: Making them Intelligent,” *Applied Artificial Intelligence*, Vol.13, No.4-5, 1999.
- [5] 穴吹, 山本, 田村: “Welbo: 複合現実空間に存在する擬人化エージェント”, 第 5 回知能情報メディアシンポジウム論文集, pp.191-198, 1999.
- [6] J. Rickel and W. L. Johnson: “Animated agents for procedural training in virtual reality: perception, cognition, and motor control,” in [4], pp.343-382, 1999.
- [7] 長谷川, 坂上, 速水: “実世界知能のための人間型ソフトウェアロボットの試作”, 第 4 回知能情報メディアシンポジウム論文集, pp.69-76, 1998.
- [8] P. Maes, T. Darrell, B. Blumberg, and A. Pentland: “The ALIVE system: wireless, full-body interaction with autonomous agents,” *ACM Multimedia Systems*, Vol.5, No.2, pp.105-112, 1997.
- [9] 大島, 佐藤, 山本, 田村: “AR<sup>2</sup> ホッケー: 協調型複合現実感システムの実現”, 日本 VR 学会論文誌, Vol.3, No.2, pp.55-60, 1998.
- [10] 大島, 佐藤, 山本, 田村: “RV-Border Guards: 複数人参加型複合現実感ゲーム”, in [3], pp.699-706, 1999.
- [11] 大島, 佐藤, 山本, 田村: “MR リビングルーム – MR 空間の幾何的・画質的整合性に関する考察”, 日本 VR 学会第 3 回大会論文集, pp.309-312, 1998.
- [12] K. Satoh, T. Ohshima, H. Yamamoto, and H. Tamura: “Case studies of see-through augmentation in mixed reality projects,” in *Augmented Reality – Placing Artificial Objects in Real Scenes (Proc. IWAR'98)*, A K Peters, Ltd., pp.3-18, 1999.
- [13] K. Itou, S. Hayamizu, K. Tanaka, and H. Tanaka: “System design, data collection and evaluation of a speech dialogue system,” *IEICE Trans. on Information and Systems*, Vol. E76-D, No.1, pp.121-127, 1993.
- [14] 片山, 内山, 熊谷, 田村, 苗村, 金子, 原島: “Collaborative CyberMirage: リアリティと相互アウェアネスを追求した仮想空間共有体験システム”, 情処学論, Vol.39, No.5, pp.1484-1493, 1998.