拡張デジタルデスクを用いた ヒューマン・ロボットインタフェース

寺島 真秀^{*1} 坂根 茂幸^{*2}

A Human-Robot Interface Using an Extended Digital Desk Approach

Maho Terashima^{*1} and Shigeyuki Sakane^{*2}

Much attention has been paid recently to Augmented Reality (AR) which supports human's daily life by blending multimodal information into real world, typically on the objects being considered. Most of existing AR systems provide such information to human by using ordinary displays and see-through head mounted displays. However, monitoring and understanding of human actions is necessary for the system to achieve flexible and gentle interactions. The paper presents an attempt to make such interface using an extended digital desk for guiding and teaching robot system. A prototype system consists of a projector subsystem for information display and a realtime tracking vision subsystem for recognizing human's action. We implemented coordination of two levels of interactions using a virtual teaching panel and projected real image of the task environment on a desk for intuitive teaching. Experiments of teaching robot tasks demonstrate usefulness of the proposed system.

Key Words: Augmented Reality, Human-Robot Interface, Digital Desk, Tracking Vision, Projector

1. はじめに

社会が急速に高度化し複雑化していく時代を迎え、我々の身の周りの機械や情報機器も複雑になっていく一方である。そのような社会で人間がより快適に生活するために、日常的に直面する様々な状況に適応して、対象そのものが存在する場で多様な情報を提供し、我々の活動をサポートしてくれるAR (Augmented Reality) 技術に対する関心が近年高まっている [1][2]。

ARは、コンピュータの高い情報処理能力を積極的に我々の実世界に取り入れ、実世界での我々の環境を情動的に強化・拡張するものであるが、従来のARの研究では、普通のディスプレイや透過型ヘッドマウントディスプレイを用いた、システム側から人間への情報提示のシステムが多い。しかしこれと同時に、状況依存する人間の意図をシステム側がリアルタイムに理解し、それに対応する行動をとるのは、人間とシステムの間での柔軟なインタフェースを構成し、円滑なコミュニケーションを確立する上で不可欠である。

システム側が状況を判断し我々をサポートすることは、人と人との自然なコミュニケーション方式により近づくことであり、ロボットの家庭やオフィス等への進出が望まれている今、

このようなARによる人に優しいインタフェースの技術の開拓は重要な課題である。

本研究では、人とロボットが協調して作業を行う際の相互インタラクションを取り上げ、AR技術の一つとして提案された「デジタルデスク」をロボット教示向きに拡張し、より自然なロボットとのコミュニケーションを実現するシステムを構築した [3][4]。

作業環境内での双方向のインタラクションを実現するために、プロジェクタによる情報提示のサブシステムと、トラッキングビジョン装置による人間の行動理解のサブシステムを結合した。人が作業している様子や作業場を視覚で捉え、人間の行動理解を行い、必要な情報を作業者に提示する。

具体的には、我々が普段自然に使用している紙の上に仮想的なティーチングペンダント等を投影し、これに対する入力動作の認識を実現する下位レベルでの協調系と、机にロボットや作業場の画像を投影して、仮想的な作業空間に対する指示動作の認識をする上位レベルの協調系とを構成した。作業教示の実験を通してそれらの有効性を確認した。

以下、第2章では、AR技術の現状と実空間強化に基づくロボットインタフェースとしての位置づけについて述べる。第3章ではプロトタイプシステムの構成を述べ、第4章ではシステムの動作実験について述べる。第5章むすびでは、本プロトタイプシステム構成の特徴と評価をまとめ、今後の課題について述べる。

原稿受付 1997年11月14日

^{*1}セイコーエプソン(株)

^{*2}中央大学理工学部

^{*1}Seiko Epson Corporation

^{*2}Faculty of Science and Engineering, Chuo University

2. 実空間強化とヒューマン・ロボットインタフェース

2.1 Augmented Reality の技術的背景

従来の人間とコンピュータのインタフェースは、大きく3世代に分けることができる。第1世代はキーボードによるコマンド指向型のインタフェース。第2世代はマウスを取り入れ、アイコンをベースとしたGUI (Graphical User Interface) 型のインタフェースである。コマンド指向型インタフェースに比べて、GUIは操作やその結果をvisualに表示することでユーザの認知的な負担と操作量を減少した。しかし両者は我々にコンピュータの画面を凝視することを強い、我々の様々な仕事をオーバーラップして表示される均一なウインドウ環境に閉じ込めてしまう。

第3世代のインタフェースでは、様々なメディア技術を駆使したVRやARのようなマルチモーダルなインタフェースが登場した。ARは人間が日常的に直面する様々な状況を認識して人間をサポートするようなインタフェースであるが、もともとはVR (virtual reality) の研究から発展したものである。ARでは、VRのようにコンピュータの生成する情報世界に人間が入り込むのではなく、実世界に情報世界を融合し、実世界を情報的に拡張している。

AR技術では、例えば透過型ヘッドマウントディスプレイにより、ユーザの世界の視界とディスプレイの画像を重ね合わせて効果的なアドバイスを行う[1][5]。この技術は特に医療や工業分野で応用されており、胎児の超音波エコー像をあたかも妊婦の体が透けて体内から透けて見ているかのように表示する[6]。また、複雑な機械のメンテナンスや修理の際に、作業をどのように行うかを説明する三次元情報を実際の機械に重ね合わせて表示することにより、効果的なアドバイスを行うシステム[7]などが提案されている。さらに、デザイン等を行う際の支援としてプロジェクタを利用したシステムも構築されている。「デジタルデスク (digital desk)」[8]は、普段鉛筆やコーヒークップなどを置く現実の机にワークステーションの特性を得られるように強化したものである。コンピュータの画面が机の上に投影され、上から机を見下ろすように設置されたカメラが、ユーザの動作を認識する画像処理システムにデータを送ることで作業支援を可能にした。また、屋外活動を支援するAR技術には携帯型システムを利用したウォークナビシステム[9]等がある。

2.2 従来のロボットインタフェース

ロボットの分野においても、人間の意図を適切に伝えたり、補助したりするために、様々なインタフェースが存在する。従来の人間とロボットとのインタフェースでは、以下のようなものがオンライン教示、オフライン教示に用いられている。

● オンライン教示のインタフェース

オンライン教示時によく用いられているのは、ティーチングボックスやマスターアームを利用したシステムがあげられる。これらは、現実にはロボットを実行しつつ教示を行うため、使用法は簡単であり細かい教示も与えやすい、さらにセンサグローブ等を利用し、作業者に自分の腕のような操作感を与えるインタフェースや、スレーブロボットを直接操作し、その手先の位置データや指に加わる力のデー

タから操作状態の同定を行う方式[10]も存在している。これらのインタフェースはロボットに詳細な動きを指示するときには有効であるが、ある程度の手順をまとめたタスクレベルでの教示、エラー回復に関する教示、センサを用いたフィードバック機能等を教示するのは難しい。

● オフライン教示のインタフェース

オフライン教示では、仮想の作業環境の中でロボットを操作し、作業の教示を行う。仮想作業環境内では、多くのシミュレーションを行ったり、全体の作業の流れを何度でも組み立て直すことが可能である。しかし、作業者はコンピュータグラフィックスを利用することが多く、インタフェースはディスプレイの中のGUI環境となり、ディスプレイの束縛から逃れられない。仮想の作業環境を厳密に構築できる生産ラインなどでは実用化が進められているが、実ラインで動かすためには、微少な補正のために先のオンライン教示も必要となる。

また、ロボットへの作業教示の一つの理想として、人間がやって見せた作業をロボットが理解して行うというものが考えられる。Kuniyoshiらは、人間の実演による組み立て作業の画像に基づいて対象物や指の運動を解析し、作業の同定を行った[11]。

しかし、これらのインタフェースは、人間側からシステムへの一方向な情報入力を行うものが多く、システムが今どのような状態なのか次に何をすべきなのか、システム側からの情報提示が不十分と言える。そこでARを利用してシステム側からの情報を付加することで、システムの状態を知らせたり、作業者の負担を軽減したりするインタフェースが考えられている。

2.3 ARによるロボットインタフェース

ARを利用したロボットインタフェースは、システム側から積極的に情報提示を行い作業者を補助する。

松井ら[12]は、モデルに基づいたオフライン教示と高度な作業監視機能の実現のためにマルチメディアディスプレイをインタフェースに用いることで、環境モデルの作成、作業表示とシミュレーション、実行監視を統合化した環境を作り出し、操作者の負担を軽減した。Milgramらが開発したARGOS[13][14]では、作業場の三次元的な視野を提供するために、ステレオ視により取り込んだ作業場のディスプレイ画像を特別なスコープを装着することで立体視を可能にし、さらにその映像に仮想的な情報を融合させて表示し作業者の操作を支援する。しかし、これらの二つのシステムでは特殊な表示装置が必要である。また、佐藤らの製作活動を行う人間を支援する組み立て作業支援システム[15]では、製作活動を行う人間がどのような思考過程と意図の下にその作業を行っているかを理解し、有効な支援を行う。例えば、電子回路の組み立て作業を考えたとき、人がある部品をとろうとしたとき、その部品の機能的な役割を気にしていれば、その部品周辺の回路図を表示し、機構的な面に注意が向いていれば実体配線図的なアドバイスを行う。しかし、このシステムは人間の作業支援を目的としたもので、ロボットの作業教示には至っていない。

2.4 本研究での Augmented Reality

本研究では、人間とシステムの双方向の情報伝達を伴うデジタルデスク的なアプローチを採用し、ロボットインタフェースを指向して拡張を行った。すなわち、コンピュータにより仮想的な情報を付加した作業場の画像や仮想操作パネルを机上に投影すると同時に、それに対する人間の動作認識を行い、ロボットへの教示時に有効な情報提示やロボットへの指示を支援する多様なレベルのインタフェースを構成した。

実際には、ロボットへの教示の際に Fig. 3 のような画像上のカップをこちらに移動など、直接手で指示し作業全体の大きな流れを教示するより高いレベルでの作業教示、実際にカップをつかむ正確なロボットの姿勢の教示等が必要な際に、Fig. 2: A, B, C のような仮想的な操作パネルを投影し、ロボットの軸を直接教示するような低いレベルでの作業教示。また、作業が複雑な場合のメニュー表示による選択等を行う。

本システムによる特長は以下の2点である。

- (1) 投影範囲内での作業者の位置、方向、および作業状況に合わせての仮想操作パネルの提供を可能にする。
- (2) 実画像を利用したより高いタスクレベルでの直観的な作業の指示を可能にする。

システム側からの情報提示に本システムではプロジェクタを用いるが、同様のことを行う既存の方法には、(a) 透過型 HMD や (b) タッチスクリーンをもつ液晶デバイスなどの利用が挙げられる。これらに対するプロジェクタ方式を比較する。まず、(a) 透過型液晶 HMD に比べると、本方式は、

- 特別なセンサ類を体に装着する必要がなく、作業者の使用感を損なうことがない。また、頻繁に作業者が変わる環境でも利用できる。
- 複数の作業者が簡単に同じ情報を見ることができる。
- 高精度な頭部の位置、姿勢追跡を必要としない。

次に、(b) タッチスクリーンをもつ液晶デバイスに比べると、

- デバイスを手で持ち歩かなくてよい。
- 投影範囲は液晶画面の広さにくらべて十分大きい。
- 作業環境や作業対象への情報投影システムへ拡張が可能である。

などの利点がある。一方、欠点としては、情報の投影がプロジェクタの投影範囲内に限定されることがあげられる。しかしこの点は、プロジェクタを複数台用いることで改善が可能であり、さらに将来、既存の照明に代わってプロジェクション機能を有するインテリジェント照明が可能となれば、作業者の利用空間の拡大が期待できると考えられる。また、指の接触検出の信頼性の点では、タッチスクリーンの液晶デバイスの方が優位にあるが、本システムでは前述の作業環境や作業対象への情報投影への拡張性をより重視した。

3. プロトタイプシステムの構成

3.1 ハードウェア構成

本システムのハードウェア構成は、Fig. 1 のようである。システム側から人間への情報提示のためにプロジェクタ：EZ4000 (Sharp 社製) を用い、仮想的な操作パネルや作業場の画像を机の上に投影する。人間からシステムへ提示された情報の理解

Fig. 1 Prototype system of the human-robot interface

Fig. 2 Examples of the virtual teaching panel

のためには、CCD カメラ：EVI-G30 (Sony 社製) と高速画像追跡機能を持つトラッキングビジョン：TRV-CU (Fujitsu 社製) [16] を用いた。なお、実画像の投影のためにトラッキングビジョンの映像出力は、コンピュータ (Power Mac 7100/66AV) のビデオ入力に接続している。ロボットと作業場の画像の取得のためには、CCD カメラ：EVI-G20 (Sony 社製) を用いた。実際に作業を遂行するロボットとしては、6 自由度を持つマニピュレータ：Js2 (川崎重工社製) を動作させた。プログラムは電総研でロボット応用向きに開発されたオブジェクト指向言語 Euslisp [17] により記述している。

3.2 仮想操作パネルによるインタフェース

作業状況に合った操作パネルをその場で提供したり、作業者の位置に合わせて操作パネルを提供することは、非常に有意義である。本システムでは、任意の紙の上にプロジェクタによりソフトウェア的に操作パネルを投影することで柔軟性の高いインタフェースと場所的な制約の緩和を実現した。

例えば、ロボットが詳細な座標の指示を必要とするときに、作業者の持つ紙の上にボタン形状のパネルを投影すれば (Fig. 2: 左)、数値的な指示を行いやすいティーチングペンダントの代わりとして利用できる。以下では、これを仮想ティーチングペンダントと呼ぶことにする。また、大まかな動作の指示を必要とするときは、Fig. 2: 右のようにジョイスティック等を投影して作業者の指示を仰げば、より高いレベルでのロボット操作を可能にする。

類似の機能を実現させる方法としてタッチパネルを採用することが考えられるが、タッチパネルは作業者の方向を特定したり、大きな範囲での利用に向かない。本システムでは、作業台

周りの方向の自由とプロジェクタの投影内での自由が与えられている。また、システムをソフト的に Off にすれば、本来の机としての機能に戻ることができる。

構成した仮想操作パネルのアイテムとしては、button, slider, joystick, information window を用意した。各アイテムの使用例を以下にあげる。

- Fig. 2: A の仮想操作パネルは、button アイテムと information window (上段) を組み合わせた例である。ユーザのアクションにより、ロボットの xyz 座標を数値レベルで設定することができる。また、現在の座標などを information window に表示する。
- Fig. 2: B の仮想操作パネルは、slider, joystick, information window アイテムを組み合わせた例である。このパネルでは、joystick アイテム (中央左) でロボットの x-y 座標を、slider アイテム (中央右) でロボットの z 座標を設定する。設定された現在の座標が information window に表示される。
- Fig. 2: C の仮想操作パネルは、button アイテムと information window を組み合わせたメニューパネルの例である。作業が複雑な場合の選択等を行う。

指の動きをトラッキングした結果を投影するパネル側にフィードバックすることにより、パネル上の bar を移動したり、ボタンを押された状態にビジュアルに変化させることが可能になっている。

人間同士がコミュニケーションを行うとき、その人が何をしたいのか身振りなどからある程度推測することができる。近年、人間の身振りを機械に認識させる研究が多くなされているが、三次元の人間の動きを認識させるのは一般に難しい。本研究では、人間の動作を紙や机など二次元平面に拘束するので自由度は減るが、認識の信頼性を優先した。

3.3 実画像によるインタフェース

実画像上のある対象物を直接「ここへ移動」など指示することは人間の直観にも合致し、人間同士のコミュニケーション方法に近い。作業指示の点ではより高いタスクレベルでの指示につながる。さらに、実画像で幾何学的な環境モデル情報やその他の情報を重畳することは、従来の AR 技術の適用を図ることも可能になる。そこで、本プロトタイプシステムでは、Fig. 3 に示すようなカメラからの実画像をプロジェクタにより卓上に

投影し、作業者は指や手を使い実画像上でロボットに対する指示を与える。

また、教示をする際に、その操作が遠隔操縦になる場合は、作業場を見渡したり、特定の対象物をよく見たい場合がある。このようなときのために、作業場に設置したカメラを同時に操作することができる。作業者は、画像上の左端か下側を手で上下左右になぞることによりカメラの pan, tilt を行う。人間の動作検出には、トラッキングビジョンの optical flow を利用した。flow の検出場所を画像の左端と下側に 2×8 列張っており、その総和がある閾値を越えた時点で、カメラを動作させる。作業場の対象物の細部を見たいときは、画像上の物体を指さす。システム側は、指された位置の検出と計算を行いカメラの pan, tilt により対象物を中心に表示した後 zoom を行う。そして決められた時間後再び元の視野に戻る。

4. 実 験

本プロトタイプシステムの有効性を確認するために、(1) 仮想操作パネルに対する動作の認識、(2) 仮想操作パネルの投影位置誤差の評価、(3) 実画像の投影を用いた作業教示の実験を行った。

4.1 仮想操作パネルに対する動作の認識

仮想操作パネルに用いられるすべてのアイテムは、button アイテムが基本となっており、その動作や位置の検出はすべて同じ認識方法がとられている。よって以下の実験では、button アイテムに対して行った。

作業者の仮想パネルに対する操作を認識するために、Fig. 4 に示されるような、二つの画像の差を利用する。ユーザがパネルを押しているときの画像は、Fig. 4: 左に見られるように指の先と影が一致している画像である。これに対して、ユーザがパネルを押そうとしているときの画像が、Fig. 4: 右である。ユーザがパネルを押しているときの画像をテンプレートとして覚えさせ、トラッキングビジョンによりそのディストーション値 (実画像とテンプレート画像間の誤差値) を計算する。そのディストーション値により、ある閾値と比較して判断している。

指と指の影を効果的に区別するために CCD カメラを真上からではなく約 45 度の角度を持たせて監視した。トラッキングの方法としては、指先を常にトラッキングするのではなく、各パネルのアイテム上にトラッキングの検索領域をはりつけ、監

Fig. 3 Projected real image of the task environment

Fig. 4 Real image of fingers and the shadows

Fig. 5 Distortion values with motion of the fingers

Table 1 Distortion values

	touch	shadow only	finger only
average	52811.3	110023.7	100577.6
std	8293.6	7235.4	8352.8
number of data	27159	120	120
avg-2std		95552.9	83872.0

std: standard deviation, avg-2std = (average - 2 × std)

視している。検索領域は、各パネルの左上端から相対的に決定される。

作業者がパネルをクリックしている状態の判断基準の決定のために、実際にトラッキングビジョンのディストーション値が作業者の指の動きに合わせてどのように変化するかを実験した。Fig. 5 にそのグラフを示す。横軸に時間軸、縦軸にディストーション値をとっている。

このグラフでは、180,000 付近がパネル上に何も検出していない定常状態を表している。これに対して、深く落ち込んで 60,000 付近で安定しているところがあるが、それはユーザがパネルをクリックしている状態に対応する。また、110,000 付近で変動している箇所があるが、それは指のみであるか、もしくは指の影だけを検出しているときである。このような実験を Fig. 4 で利用したすべてのアイテム上について 5 秒間隔で 80 秒繰り返し、まとめたのが Table 1 である。特に平均については、押した状態ではデータ群の平均をとったが、影のみ検出、指のみ検出の状態では、それぞれのデータ群の中で最小値を取り、その最小値の平均値をとった。

両者の状態を区別する適切な閾値設定のために、押した状態と判断する値に余裕を持たせた (平均値 + 2 × 標準偏差) と、押されていない状態と判断する値に指のみ検出された場合の (平均値 - 2 × 標準偏差) を考慮した。

すなわち、押した状態の (平均値 + 2 × 標準偏差) = 69,398.6 と指のみ検出された場合の (平均値 - 2 × 標準偏差) = 83,872.0 とを比較して、作業者がクリックしたと判断する閾値として 70,000 程度の値を用い、良好な認識結果を得ることができた。

なお、視野内における紙の位置によるディストーション値の変化については、より厳しい条件下である、実験 4.3 の実画像の投影を用いた作業教示で行った。

Fig. 6 Detection of the paper and the virtual panel item

Fig. 7 Window for the rotated image

4.2 仮想操作パネルの投影位置誤差の評価

投影画像の生成には、弱透視変換 (weakperspective) [18] とアフィン変換を利用した。作業者が紙を任意の位置に移動させたり回転させても紙のある場所に正しくパネルを投影するため、紙の左上角、右上角、左下角の位置を常にトラッキングしている (Fig. 6)。

実際の投影画像生成の流れは以下のようなものである。

- (1) カメラ、プロジェクタ、投影面の位置が決定した時点で紙の座標系と投影用 window の関係を弱透視変換により求めておく (この行列を prj-matrix とする)。
 - (2) 同様に紙の座標系とカメラ座標系の関係を弱透視変換により求めておく (この行列を cam-matrix とする)。
 - (3) カメラより観測された紙の 3 点の座標を cam-matrix を用いて紙の座標系に変換する。
 - (4) 3 で変換された 3 点から、紙の座標系での平行移動、回転の関係をアフィン変換により求める (この行列を afin-matrix とする)。
 - (5) 4 で求めた afin-matrix によりパネルの画像 (Fig. 7: 左) を平行移動、回転させる。
 - (6) 5 で得られた画像を prj-matrix の逆変換により、紙の座標系から投影用 window の座標系 (Fig. 7: 右) に変換する。
- 投影面の各場所において、どの程度の誤差が発生するかを測定した。実験環境は、プロジェクタと投影面の距離を 275 [cm] カメラは約 30 度の向きから監視した。観測地点は、Fig. 8 のようにそれぞれの位置 (p1~5) で 500 回計測した。

Table 2 は位置 p2 で計測した値をまとめたものである。Table 3 は、それぞれの観測地点での誤差 (pixel) をまとめたものである。プロジェクタの投影面が 98.5 × 68.1 [cm] であり、

Fig. 8 Projection of the virtual panel

Table 2 Position p2

		true coords.	average	std	error
upper-left	x1	45	44.35	0.925	0.645
	y1	37	37.86	0.669	-0.867
lower-left	x2	46	47.93	0.344	-1.937
	y2	174	176.23	0.787	-2.238
upper-right	x4	172	173.67	0.949	-1.675
	y4	36	33.89	0.890	2.100

error = (true coordinates - average)

Table 3 Error of position

	upper-left		upper-right		lower-left		lower-right	
	x1	y1	x2	y2	x3	y3	x4	y4
p1	1	1.97	0	-0.52	-2.42	3.89	-0.42	3.37
p2	0.65	-0.87	-1.94	-2.24	-5.03	0.66	-1.68	2.10
p3	-1	2.88	-1.21	1.12	-3.29	2.85	-3.53	5.61
p4	-1.02	1.55	-4.0	3.23	-3.53	3.65	-1.55	3.06
p5	-2	2.71	4.04	4.59	1.52	4.94	-2.03	5.06

p1:center, p2:upper-left, p3:upper-right p4:lower-left, p5:lower-right (with respect to coordinates of projection)

投影のための画面が 600×400 [pixel] であるので, 1 [pixel] 当たり $98.5/600 = 0.164$ [cm] であるから, Table 3 より最大誤差値である 5.61 [pixel] の誤差が生じた場合でも投影面で 0.92 [cm] のずれとなり良好な結果を得た。

4.3 実画像の投影を用いた作業教示

机の上に投影された作業場の実画像 (Fig. 9) を直接指し示すことでロボットとのインタラクションを行う。指された対象物の同定のために, トラッキングビジョンによるテンプレートマッチングを利用している。システム側の知識ベース部に対象物のテンプレートと名前があらかじめ関連付けられている。

投影された実画像上の対象物に対して指示を行うためには, 作業者が指した位置をシステムが正しく認識する必要がある。しかし, 指した位置の対象物の画像の影響により指のテンプレート画像との誤差が大きくなり対象物の色によって指示できないものが生じた。そこで本システムでは, 投影した実画像の上での指さし動作を観測するカメラのアイリス値を大きくして背景の画像の影響を抑えた。

実際に今回使用した対象物に対して, 背景の影響がどの程度抑えられるかそれぞれの対象物に対して 2,000 回計測した結果

Fig. 9 Sample view of the teaching robot task

Table 4 Distortion values of the items

		Desk	Coffe	Cup	Pot	Robot
no touch	avg	65812.0	65803.6	64638.1	65380.9	65722.6
	std	0	16.13	305.74	118.71	70.15
touch	avg	26224.9	30974.3	33204.0	27582.5	28352.2
	std	610.78	623.20	8390.68	1115.97	716.09
	a+2s	27446.5	32220.7	34883.3	29814.5	29784.4

avg:average, std:standard deviation, a+2s=(avg + 2 × std)

Table 5 Distortion values of the cup

		center	UL	UR	LL	LR
no touch	avg	64638.1	65363.2	65713.6	64322.5	64995.8
	std	305.74	233.07	62.39	393.21	636.30
touch	avg	33204.0	37650.0	38876.8	39244.9	40490.8
	std	839.68	608.75	576.35	1199.00	800.32
	a+2s	34883.3	38867.1	40029.5	41642.9	42091.4

UL:upper-left, UR:upper-right, LL:lower-left LR:lower-right, avg:average, std:standard deviation, a+2s=(avg + 2 × std)

Fig. 10 Distortion values

を Table 4 に示す, さらに移動する頻度の高いカップについては, 作業台上でそれぞれ中心, 左上, 右上, 左下, 右下においてディストーション値の計測を 2,000 回行った (Table 5).

作業者が対象物を押した状態と判断する値に先の操作パネルの場合と同じく (平均値 + 2 × 標準偏差) を用いると, Table 4, 5 より, 最も誤認識しやすい値としてカップを右下に移動したときの, 42,091.4 が閾値設定の目安になる。

Fig. 10 は実験データの一例で, 縦軸にディストーション値, 横軸に時間をとったグラフである。右側の 3 レベルのデータ

Fig. 11 Overview of the experiment

は、上段が机上で約3[cm]指を離したとき、中段が右下にあるカップを指したとき、下段が机上を指したときである。時間軸で2,000を越えたあたりからグラフが落ち込むのは、その時間に指さしが行われたことを表している。

特に指を画像上から約3[cm]離して2,000回計測したときのディストーション値の平均が52,225.5、約1[cm]離して計測した平均が44,855.2であるので、今回は43,000を閾値として設け良好な結果を得た。

基本的な作業教示の流れは以下になる。

- (1) 作業者が指示したい対象物を指し示すか、その画像のある領域を指定する (Fig. 9 のロボット上の四角領域等)。
- (2) システムがトラッキングビジョンによりその対象物を同定し、結果を画像上に表示する (Fig. 9 の Hand や Pot の文字)。
- (3) 画像上に提示された情報を作業者が確認する (Fig. 9 右下)。
- (4) 指により対象物を移動したい場所まで指示する。このときシステム側は、指の軌跡を画像上に重ね合わせて作業者を補助する (Fig. 9 のロボットからポットへの線)。
- (5) 指を離した位置が、例えば机であればロボットは対象物をその場所に移動させる。

今回実際に教示した作業は、1. 黄色いカップをポットの下に移動 2. ポットを押す 3. 黄色いカップを中心に移動 4. スプーンをつかむ 5. 砂糖を黄色いカップに入れる 6. スプーンを置く 7. 出来上がったコーヒーを差し出す という一連の作業教示を行った。

実験の結果、投影された実画像で操作者の指さしによる作業教示の有効性を確認した。

Fig. 11 は、実際に実験を行っている様子を表したものである。今回は、プロジェクタを天井につるす代わりに、通常の設置姿勢で置き、机ではなくホワイトボードにティーチングペンダントを表示している。プロジェクタとカメラが写っていないが、写真の手前に位置している。

5. む す び

本研究では、人とロボットが協調して作業を行う際の相互インタラクションを取り上げ、AR技術の一つとして提案された「デジタルデスク」をロボット教示向きに拡張し、より自然な

ロボットとのコミュニケーションを実現するプロトタイプシステムを構築した。具体的には、我々が普段自然に使用している紙の上に仮想的なティーチングペンダント等を投影し、これに対する入力動作の認識を実現する下位レベルでの協調系と、机にロボットや作業場の画像を投影して、仮想的な作業空間に対する指示動作の認識をする上位レベルの協調系とを構成した。仮想的な操作パネル上での指による指示動作の認識では、指の影情報を積極的に利用することにより信頼性を高めることが可能になった。

今後の課題としては、実画像の投影に加えて、作業環境の実物体の上での直接的な指示動作をも可能にする拡張が考えられ、これらの多様なレベルでのインタフェースを統合していく必要がある。

なお本研究は平成8～10年度文部省科学研究費基盤研究(C)「能動的視覚センサプランニングに関する研究」の補助を受けた。

参 考 文 献

- [1] 坂村健 訳：電腦強化環境。パーソナルメディア、1995。
- [2] R.T. Azuma: "A Survey of Augmented Reality," PRESENCE, (MIT Press), vol.6, no.4, pp.355-358, 1997。
- [3] 寺島、坂根: "実空間強化とヒューマンロボットインタフェース—デジタルデスクを用いるインタフェース—", 第14回日本ロボット学会学術講演会, pp.691-692, 1996。
- [4] 寺島、坂根: "実空間強化とヒューマンロボットインタフェース—実画像の投影と教示—", 第15回日本ロボット学会学術講演会, pp.79-80, 1997。
- [5] J. Rekimoto: "Augmented Interaction: Toward a New Human-Computer Interaction Style Based on Situation Awareness," Interactive Systems and Software 2 (WISS '94 Proceedings), pp.9-17, 1994。
- [6] M. Bajura, H. Fuchs and R. Ohbuchi: "Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient," SIGGRAPH, pp.203-210, 1992。
- [7] S. Feiner, B. MacIntyre and D. Seligmann: "Knowledge-based Augmented Reality," Communications of the ACM 36, pp.52-62, 1993。
- [8] P. Wellner: "The DigitalDesk Calculator: Tactile Manipulation on a Desktop Display," In Proceedings of the ACM Symposium on User Interface Software and Technology, UIST '91, 1991。
- [9] 長尾 確: "エージェント拡張現実感—エージェントによる実世界と情報世界の統合—", 情報処理学会誌, vol.38, no.4, pp.257-266, 1997。
- [10] G. Hirzinger: "Robot Systems Completely Based on Sensory Feedback," Proc. of IECON '84, pp.46-51, 1984。
- [11] Y. Kuniyoshi, M. Inaba and H. Inoue: "Learnig by Showing," IEEE Trans. on Robotics and Automation, vol.10, no.6, pp.799-822, 1994。
- [12] 松井、塚本: "マルチメディアディスプレイを用いた統合型遠隔ロボット制御法", 日本ロボット学会誌, vol.6, no.4, pp.301-310, 1988。
- [13] P. Milgram, et al: "Application of Augmented Reality for Human-Robot Communication," Proc. of Int. Conf. on Intelligent Robots and Systems, pp.1467-1472, 1993。
- [14] D. Drascic, et al: "ARGOS: A Display System for Augmented Reality," ACM Conference on Human Factors in Computing Systems, p.521, 1993。
- [15] 佐藤知正 他: "ものづくり作業のその場支援", 日本機械学会ロボティクス・メカトロニクス講演会95講演論文集 (vol.A), pp.535-538, 1995。
- [16] 森田、沢崎、内山、佐藤: "カラートラッキングビジョン", 第14回日本ロボット学会学術講演会, pp.279-280, 1996。
- [17] 松井: "幾何モデリング機能を備えたマルチスレッド並列オブジェクト指向言語 Euslisp", 日本ロボット学会誌, vol.14, no.5, pp.650-654,

1996.

[18] R. Cipolla: "Uncalibrated Stereo Vision with Pointing for a

Man-Machine Interface," In Proc. IAPR Workshop on Machine Vision Applications, Kawasaki, pp.163-166, 1994.

寺島真秀 (Maho Terashima)

1971 年 5 月 3 日生. 1996 年, 中央大学理工学部管理工学科卒業. 1998 年, 同大学大学院修士課程修了. 同年, セイコーエプソン (株) 入社.

(日本ロボット学会正会員)

坂根茂幸 (Shigeyuki Sakane)

1949 年 8 月 20 日生. 1972 年, 東京工業大学工学部制御工学科卒業. 1974 年, 同大学大学院修士課程修了. 同年, 通産省工業技術院電子技術総合研究所入所. 1995 年, 中央大学理工学部教授. 知能ロボットシステムの研究に従事. 工学博士. 計測自動制御学会, 人工知能学会, 電気学会の各会

員.

(日本ロボット学会正会員)