

# エージェントを介した遠隔地コミュニケーションの実現に向けたWOZシステムの開発

大橋 洋輝\*1  
Hiroki Ohashi

高橋 暁弘\*2  
Akihiro Takahashi

岡田 将吾\*1  
Shogo Okada

大本 義正\*1  
Yoshimasa Ohmoto

西田 豊明\*1  
Toyoaki Nishida

\*1 京都大学大学院情報学研究科知能情報学専攻

\*2 京都大学工学部情報学科

## 1. はじめに

本研究では遠隔地においてエージェントを介した協調タスクを行うことを目的とし、その基盤として、エージェントを介した円滑なコミュニケーションができる環境の実現を目指している。このような環境下で円滑なコミュニケーションを行うためには、次の二つのことが重要であると考えられる。まず一つ目は、エージェントを操作する際に、操作のために必要な動作によってコミュニケーションを阻害することがないようにすることである。もう一つは、エージェントのいる遠隔地での状況を知りえない操作者のために、エージェントがある程度の自律性を持って操作者に有用な情報を収集し提供することである。本論文では、これらの要件を満たすシステムを実現することを目標として、そのための基盤として必要となるWOZ(Wizard Of OZ)システムとそれを実現するための具体的なアーキテクチャについて提案を行う。

## 2. 研究の枠組み

### 2.1 研究の位置づけ

エージェントを介したコミュニケーションシステムについてはこれまでも多数研究がなされてきたが、その多くは2次元エージェントに関するものであった([Kubota 00]など)。ところが2次元エージェントを人間とのコミュニケーションに用いる場合には、エージェントが物理的な身体を持っていないことに起因するいくつかの問題が生じる。例えば、物理的な空間内を自由に動き回ることができないために、操作者に提供する映像は固定点からの映像(の組み合わせ、切り替え)になるという問題、ディスプレイに表示されるエージェントの体や顔、指差しなどの方向が正確に伝えられない(エージェントが右を向いている画像はどの場所にいる人から見ても自分の右を向いているように見える)という問題などである。

本研究ではエージェントを介して協調タスクを行うことを前提としているため、エージェントの視点の変化に応じて、操作者に示せる映像も動的に変化させることや、3次元空間内で正確に方向を伝達することが重要である。そこで、これらの要件を満たすために本研究ではエージェントとして3次元のロボットを採用する。

ロボット研究については[Kawakami 05]など多くの先行研究がなされているが、特に本研究ではロボットを介してコミュニケーションを行うことに研究の焦点を当てる。日常のコミュニケーションにおいては、人間は無意識に頷いたり、姿勢を変えたり、ふと何かを見たりということをしていると考えられるため、ロボットの操作インターフェースとしてはリモコンなどを

用いて明示的に指示を与えなければならないものよりも、人間の動作をそのままロボットの動作に反映させるようなものを開発する。

またロボット操作者は遠隔地の状況についてはシステムを通してしか知りえない。そこで、円滑なコミュニケーションを支援するために、有用な情報をロボットが自律的に収集して(必要ならば加工を加えた上で)操作者に提供できるようなシステムを目指す。

### 2.2 ロボットWOZシステム

本研究を進めるに当たって、その基盤としてまず我々はロボットWOZシステムを開発する。ここで、ロボットWOZシステムとは操作者の身体動作を自動認識し、同じ動作をロボットが再現するようなシステムを指している。

このような技術は、先に挙げた直感的な操作インターフェースとしての重要性に加えて、自律行動のできるコミュニケーションロボットを開発するに当たっても大変重要であると考えられる。

自律ロボットを開発するためには、ロボットに多種多様なインタラクションのパターンを教えなければならない。これに対し、多くの従来研究では計測を行い、分析をし、それを実装して評価するというサイクルを繰り返してインタラクションパターンを逐次実装していたため、多大な時間的なコストがかかっていた。

本研究で開発するWOZシステムを用いると、操作者はロボットのおかれた状況を直感的に把握でき、かつ(リモコン操作や視点変換といった)認知的負荷なしに容易に行動を取ることができる。そこで、人間がこのロボットWOZシステムを介して様々なインタラクションを行うことによって、ある状況の時にはどんな動作をするというような行動知識ベースを簡単に蓄積することができる。さらに、インタラクションデータをその場で学習器にかけることによって、知識ベースを逐次更新して特定の状況に適応していくことも考えられる。

## 3. システムの構成

図1に本研究で実現を目指すシステムの完成像を示す。

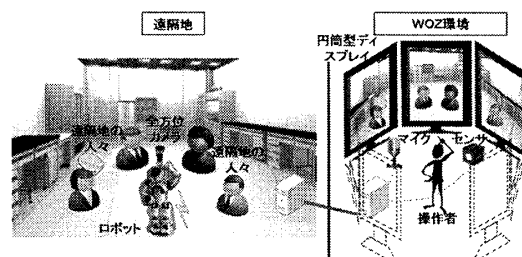


図1: システムの完成イメージ図

図に示したように、本研究では操作者にロボットの視点からの映像を提示するために、全方位カメラと円筒型ディスプレイを用いる。これらを用いることで、ロボットを中心とした周囲

Development of a WOZ System which Operates Using Human Pose Estimation for Agent-Mediated Telecommunication

\*1 Department of Intelligence Science and Technology, The Graduate School of Informatics, Kyoto University

\*2 The School of Information and Mathematical Science, Faculty of Engineering, Kyoto University

の状況をそのまま操作者にも提示することができるため、強い没入感・臨場感を与えることができると期待できる。さらに、カメラを中心とした特定の角度で映像が途切れるということがないため、周辺視野の映像についてももれることなく提示でき、対面会話に近い状況が再現できる。センサに関しては、コミュニケーションを行う操作者の行動をなるべく制約しないように、非接触のものを使用する。非接触センサを用いて人の動作認識を行うには、光学式のモーションセンサを使うと安定した結果を得やすいと考えられるが、光学式モーションセンサはマーカの隠れに弱く、この問題を解消するためには、多数のカメラを設置して様々な角度から計測を行う必要がある。本研究では、円筒型ディスプレイ内での使用を前提とするため、カメラを設置できる範囲には限界がある。このため本研究では操作者の動作認識のためのセンサとしてはレンジセンサを用いる。ロボットについては Aldebaran Robotics 社の NAO を採用する。NAO はワイヤレスで動作するプログラム可能なロボットである。首、肩、肘、手首、指、股関節、膝、足首を制御することができ、基本的なコミュニケーション動作を行うことができると考えられる。

これらの主なデバイスを図 2 に示した。



図 2: システムに使用する主なデバイス。左から、全方位カメラ、円筒型ディスプレイ、レンジセンサ、ロボット NAO

#### 4. ロボット制御のためのアーキテクチャ

本節では、操作者の動作をロボットに反映するための手法について述べる。尚、コミュニケーションを行う上では上半身の動作がより重要になると考えられるため、本論文では上半身部分を制御するアーキテクチャに限定して記述する。

アーキテクチャは図 3 のように構成する。まずレンジセンサにより距離画像、アンプリチュード (反射強度) 画像を取得する。レンジセンサを用いる利点は、これが一般のカメラと違い距離を含めた 3 次元情報が取得できるため、円筒型ディスプレイ内で操作者の背後に映る可能性がある他の人の画像との切り分けが簡単な方法でできると期待できることである。その後、前処理としてセンサデータのクラスタリングを行い、ノイズに該当するクラスを除去する。そして、そのノイズが除去されたデータに対し体部分の推定コンポーネントと顔方向推定コンポーネントで別々に推定の処理を行う。この時、特徴点としていくつかの点座標だけを用いるのではなく、多数の点群データの 3 次元座標から推定を行うことにより、隠れなどに強く、安定した推定結果を得ることができると期待できる。それぞれのコンポーネントは肩や肘の関節の角度、顔方向 (首) の角度を出力し、それをロボットに無線で送信する。

#### 5. システムプロトタイプ

我々は、プロトタイプの一部として操作者の上半身の姿勢や腕の動き、顔の方向などを自動認識してその結果をロボットの動作に反映させるようなシステムを開発した。

図 4 は体・腕の姿勢推定コンポーネントの認識の様子である。現在のところ図のように腕と胴体や頭などとの重なりが少ない

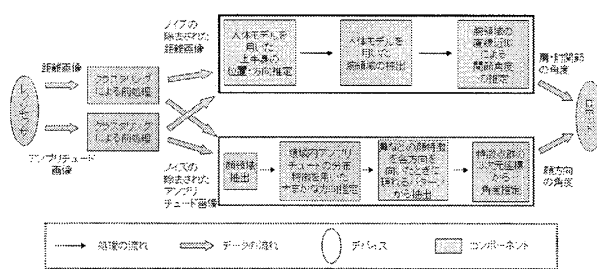


図 3: ロボット制御のアーキテクチャ

状況ならば、ポインティングなどの基本的な動作が取得できている。

図 5 には今回開発したプロトタイプの動作例を示す。

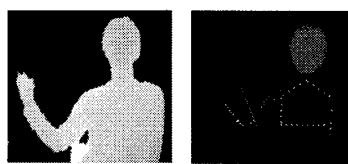


図 4: 姿勢推定と腕追跡の結果。左) レンジセンサで取った距離画像。右) 推定された上半身と腕をポリゴン表示した画像。五角形の部分の法線が推定された体の向きに相当し、灰色で表示された二本の直線が推定された腕に相当する。



図 5: プロトタイプの動作の様子

#### 6. まとめ

本論文では円滑な人間-ロボットコミュニケーションを行うための基盤として、ロボット WOZ 環境の構築方法について提案を行った。

今後は腕の隠れ、重なりなどにも対応できるように動作認識部の推定結果をより頑健にし、実際にロボットの行動知識ベースを学習させるという実験を行っていく予定である。

#### 参考文献

[Kawakami 05] Kawakami, N., Sekiguchi, D., Kajimoto, H., and Tachi, S.: TelesarPHONE - Communication Robot based on Next Generation Telexistence Technologies, *Proceedings of 36th International Symposium on Robotics (ISR2005)*, (2005)

[Kubota 00] Kubota, H. and Nishida, T.: EgoChat Agent: A Talking Virtualized Member for Supporting Community Knowledge Creation., *The Proceedings of the AAAI Fall Symposium on "Socially Intelligent Agents-The Human in the Loop"* (2000)