

# 案内ロボットの遠隔操作のための GUI の提案

村上 友樹

大阪大学 工学部 応用理工学科

中西 英之

大阪大学大学院 工学研究科 知能・機能創成工学専攻

ロボットを遠隔地から操作して人を誘導して案内する場合，操作するロボットの位置や案内すべき人々の位置を，操作者が把握する必要がある．そこで，天井に取り付けた俯瞰視点のカメラの映像を利用して案内ロボットを操作することができるインタフェースを提案する．具体的には，GUIでファイルを操作するように操作者が俯瞰カメラの映像内の人や物の上に表示されたアイコンを操作することで，遠隔地から案内ロボットを操作できるようなインタフェースである．このようなインタフェースを利用する操作者は，対象とする人や物（オブジェクト）とそれに対して行う動作（メソッド）を指定するというオブジェクト指向の入力方式で直観的にロボットを操作することができる．

## Proposing GUI for Remotely Controlling Guide Robots

Yuki Murakami

Division of Mechanical, Materials and Manufacturing Science, Osaka University

Hideyuki Nakanishi

Department of Adaptive Machine Systems, Osaka University

When an operator controls a robot and guides the people from the remote site, he needs to grasp the position of the robot and the people. We propose an interface using a bird's-eye view for controlling the guide robots. The operator can control the robot by clicking or dragging the icon which is displayed on the camera image. This interface is object-oriented interface for controlling robot in the sense that the operator chooses the "object" and "method" that is a robot action to the object.

### 1. はじめに

遠隔地間で身体性を用いたインタラクションを実現するシステムとして，遠隔操作ロボットに関する研究が現在盛んになりつつある[1][2][3]．しかし，従来の遠隔操作ロボットのシステムを用いて，ロボットを移動させて何らかの仕事を

行う場合，次のような問題がある．まず，ロボットを移動させる際は遠隔地の様子やロボットの位置を操作者が把握する必要があるが，ロボットに装備したカメラやセンサを用いて操作者がそれらを十分に把握するのは困難である．また，ジョイスティックなどを用いてロボットを遠隔操作する場合，ロボットの動作を逐次操作

者がコントロールする必要がある、操作者の負担は大きい。そこで本論文では具体的なタスクとしてロボットを操作して遠隔地にいる人々を誘導、案内することを考え、上記の問題を解決するような遠隔操作ロボットのインタフェースの提案を行う。具体的には、俯瞰視点のネットワークカメラの映像を利用して、GUIを用いたオブジェクト指向の操作で案内ロボットを遠隔操作することのできるインタフェースを提案する。以下、2節では関連研究との比較を行い、3節では設計内容を示す。4節では予備的に行った実験内容を示し、5節でまとめる。

## 2. 関連研究

Jouppi の遠隔操作ロボット[1]や GestureMan[2]は現地の様子を操作者が把握しやすいように広画角のカメラや複数代のカメラをロボットに装備し、操作者の前に大きなスクリーンを配置していた。しかしながらロボットのカメラを用いた一人称視点では、ロボットの後ろ側や、人や障害物の後ろにいる人を確認することができず、操作者が遠隔地の様子を十分に把握できるとは言いがたい。また、周りの物体との距離が把握しづらくロボットが現在部屋のどこにいるのかを正確に操作者が把握することは難しかった。Roussou らはこの問題に対し、レーザを用いて人や壁などの位置を取得して三次元表示をしたり、地図の上にロボットの位置を表示したりすることで解決を試みた[3]。しかしこの方法でも操作者は人や障害物の後ろにいる人を確認することはできず、また、位置情報のみでは人々が何に注目しているかなどを操作者が把握することはできない。そこで、本論文では天井に設置したネットワークカメラを用い、三人称視点を基にしてロボットを操作するインタフェースを提案する。ロボットに搭載されたカメラによる一人称視点ではなく、天井に設置したカメラの三人称視点をを用いることで、部屋全体の様子やロボットの位置を把握しつつ、ロボットを遠隔操作することができる。

また、操作者のボタンやジョイスティックの操作に同期してロボットが動作するというシス

テム[1][2]の場合、操作者が逐次ロボットに動作命令を送る必要がある、操作者に負担がかかる。また、ネットワークの遅延の問題も発生する[4][5]。さらに施設の案内などの仕事をロボットで行う場合を考えると、操作者は同じような操作を何度もすることになる。この問題に対し、あらかじめロボットによく行われる動作パターンを設定しておき、操作者がそれを選択するという解決法が考えられるが[3]、そのような手法の場合、動作パターンの増加にしたがって操作コマンドの選択肢が増加し、操作が煩雑になる恐れがある。本論文ではこれを解決するために、ロボットの動作対象を”オブジェクト”、オブジェクトに対するロボットの動作を”メソッド”として捉えるオブジェクト指向型の操作を提案する。操作者はカメラ映像の中から動作対象となるオブジェクトを選択し、その後オブジェクトに対する動作をメニューから選択することができる。これにより、操作者はすばやく直観的にロボットの動作パターンを呼び出すことができると考えられる。

## 3. 設計内容

### 3.1 システム概要

本システムでは、ネットワークカメラ及びウェブインタフェースで遠隔操作可能なロボットを用いる(図1)。ネットワークカメラは天井に設置し、操作者はネットワークカメラの映像をウェブクライアント上で見ることによって環境全体の様子を把握することができる。また、ネットワークカメラの上にアイコンを重ねて表示し、操作者はGUIでコンピュータ内のファイルを操作するように、アイコンのクリックやドラッグなどの操作でロボットに動作指示を与えることができる。

### 3.2 三人称視点をを用いたインタフェース

本システムでは、ネットワークカメラの三人称視点の映像を用いる。三人称視点をを用いることにより、操作者は部屋全体の様子を見ることができ、人やロボットの位置関係を容易に把握

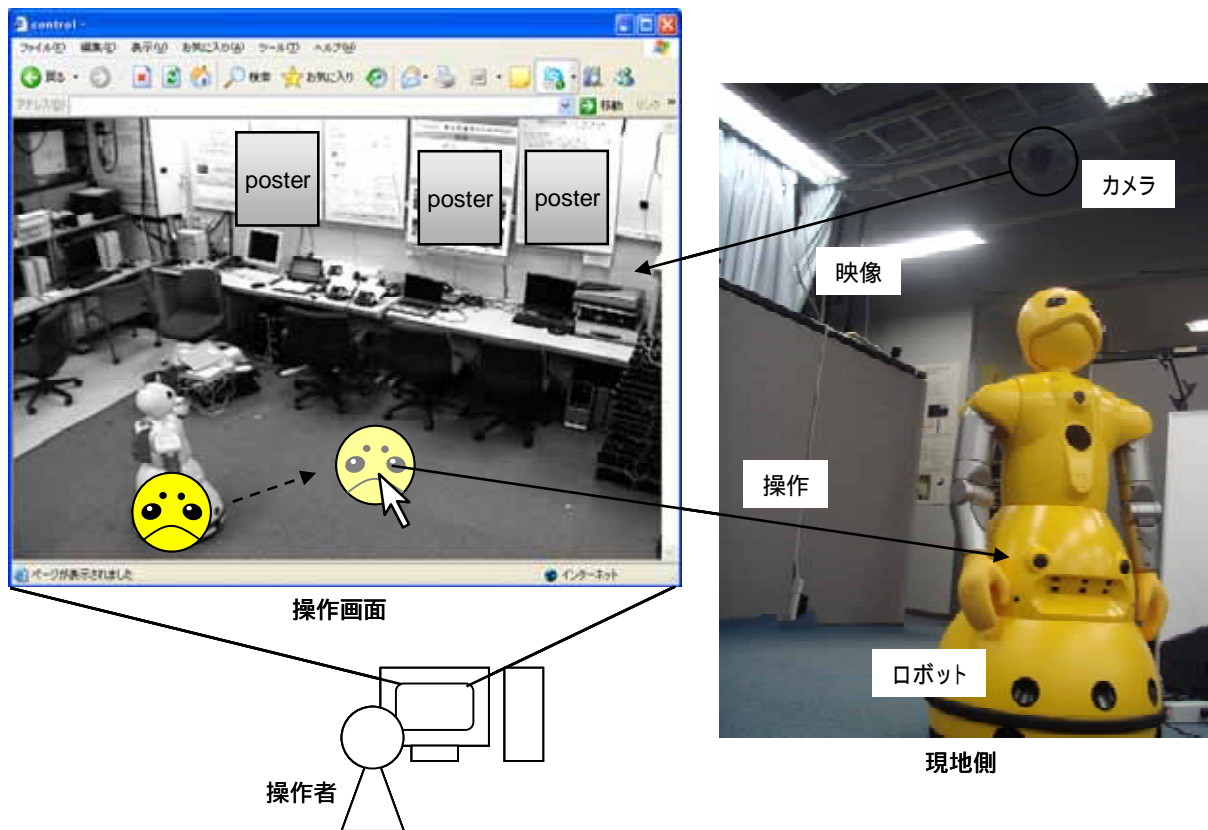


図1. システム全体図

することができる。ロボットの後ろに人を連れて案内をする場合においては、操作者は目的地、ロボットの現在位置、案内される人々の位置を把握する必要があり、それらを同時に確認できる三人称視点が適切であると考えられる。ロボットの移動操作に関しては、一人称視点の場合は、たとえば左にジョイスティックを倒せばロボットが左に回転するというのは直観的であるが、三人称視点の場合はロボットの左は操作者の視点では必ずしも左ではなく操作者が混乱する可能性がある。そこで本研究では三人称視点のカメラ映像の上で操作者が指定した位置にロボットが移動するという操作方法をとる。三人称視点の映像をそのまま移動先の入力に用いることで、直観的な操作が可能となる。

### 3.3 オブジェクト指向型操作

ロボットの動作対象を”オブジェクト”，オブジェクトに対するロボットの動作を”メソッド”として捉えるオブジェクト指向型の操作を提案

する。例えば、博物館の展示物をロボットの音声で紹介する場合、各展示物が”オブジェクト”であり、「この展示物を説明する」、「ここに移動する」などのロボットの動作が”メソッド”となる。操作者がカメラ映像内のオブジェクトを指定するとそのオブジェクトに対するメソッドの一覧が表示され、操作者がさらにそこからメソッドを選択することで、ロボットの動作が実行される。ロボットに実装されている動作パターンをすべて一覧として表示するのではなく、オブジェクトごとに分けることで操作者が選択する際の負荷を軽減する。また、オブジェクトの選択は実際のカメラ映像を使って行われるので、オブジェクトの一覧表などを用いるよりも直観的に選択できる。

### 3.4 GUIを用いたインタフェース

上記の三人称視点、オブジェクト指向型操作を実現するインタフェースデザインとして、GUIを用いることを提案する。ここで言う GUI

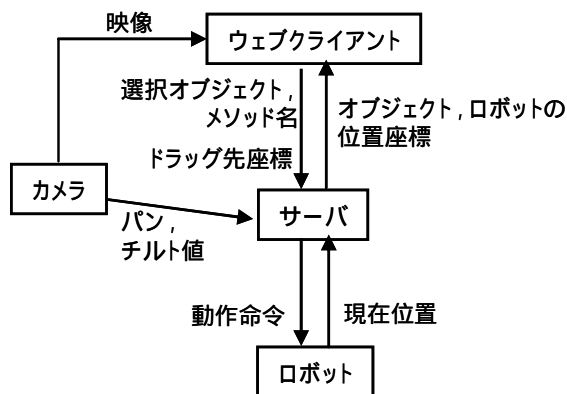


図2. システム構成図

とは、アイコンのクリック、ドラッグで操作するパーソナルコンピュータで主流となっているインタフェースのことである。ネットワークカメラの映像をもとにし、映像内の各オブジェクトとロボットの上にアイコンを表示する。そして映像内のロボットのアイコンをドラッグすると、そこにロボットが移動するというデザインにすることで、三人称視点のロボット操作を実現する。また、オブジェクト指向型操作に関しては、操作者がカメラ映像の中のオブジェクトのアイコンをクリックし、表示されたプルダウンメニューからメソッドを選択するというかたちで実現する。以上のように、全体のインタフェースデザインをすでにコンピュータのインタフェースとして普及している GUI をもとにすることで、操作者がロボットの操作に慣れるまでの時間を短縮することができると思う。

### 3.5 詳細設計

システムはネットワークカメラ、ロボット、サーバ、クライアントから構成される（図2）。以下、それぞれについて説明する。

#### 3.5.1 ネットワークカメラ

ネットワークカメラは天井に取り付け、部屋全体を見渡すことができるようにする。ネットワークカメラはウェブブラウザを用いて映像を見たり、パン、チルトの操作を行ってカメラの向きを変更したり、要求に応じて現在のパン、チルトの値を出力したりすることができるもの



図3. ウェブクライアント設計案

を用いる。本研究では AXIS 215 PTZ ネットワークカメラを使用することにした。

#### 3.5.2 ロボット

ロボットは移動可能なもので、音声をつかって展示物などの説明を行うことができるものを使用する。また、外部からの通信によって動作を制御できるものである必要がある。本研究では三菱重工の wakamaru[6]を使用することにした。wakamaru には TCP 通信で外部から命令が送られるとその命令に従って動作するように行動パターンを記述した。

#### 3.5.3 ウェブクライアント

ウェブクライアントはネットワークカメラの映像を表示し、映像内の展示物やポスターなどのオブジェクトの上に、映像に重ねるかたちでアイコンを表示する。また、ロボットの位置にもアイコンを表示する。カメラはパン、チルトの値によって向きが変わるので、アイコンをカメラ映像の上に重ねて表示するためにはカメラのパン、チルトの値とカメラの位置をもとに、オブジェクトやロボットの実空間上での位置座標を座標変換する必要がある。それらの座標変換などの作業はサーバが行うことにし、クライアントはサーバに対してオブジェクトやロボットのカメラ映像上での位置座標を定期的に要求し、サーバがそれに対して各オブジェクトとロボットの座標変換後のカメラ映像上での座標を

返すという設計にした。ウェブクライアントはサーバから送られてきた位置座標にしたがってカメラ映像の上にアイコンを描画する。操作者がオブジェクトのアイコンをクリックすると、プルダウンメニューが現れ、そのオブジェクトに対して、ロボットがどのような動作を行うかを選択することができる。例えば、ポスターをクリックすると「ここに移動する」、「説明する」などのメニューが現れる(図3)。すなわち、操作者はクリックすることでロボットの動作対象(オブジェクト)を選択し、プルダウンメニューからそのオブジェクトに対する動作(メソッド)を選択することになる。クリックされたオブジェクト名と、選択されたメソッド名はサーバに http で送信する。また、操作者はロボットのアイコンをドラッグすることでロボットの移動先を指示することができる。移動先をカメラ映像内で指示することにより、三人称視点での直観的な操作が可能となる。ドラッグ先の位置座標もサーバに http で送信する。以上のように、現在パーソナルコンピュータのインタフェースとして主流となっている GUI を模したインタフェースデザインにし、操作者に受け入れられやすくしている。

### 3.5.4 サーバ

サーバは大きく分けて2つの機能がある。1つ目はウェブクライアントからの要求に対してカメラ映像上でのオブジェクトやロボットの位置座標を返す機能である。サーバには各オブジェクトの名前、位置、メソッド名とメソッドの内容、すなわちそのメソッドが選択されたときにロボットがどのような動作をするかを記述したファイルを保持しておく。そして、クライアントからの要求を受け取ると、オブジェクト情報ファイルから位置座標を読み出す。さらに、ロボットからロボットの現在の位置座標を取得し、ネットワークカメラからパン、チルトの値を取得したうえで、各オブジェクトとロボットの実空間上での位置座標をカメラ映像内での座標、すなわちカメラ映像の左上の点からのピクセル数へと変換し、ウェブクライアントへ送信

する。実空間座標を $(x\ y\ z)$ 、カメラ座標を $(u\ v\ w)$ 、実空間座標原点とカメラとの相対距離を  $c_x, c_y, c_z$ 、カメラのパン値を  $\varphi$ 、チルト値を  $\theta$  とすると、座標変換式は以下ようになる。

$$(u\ v\ w\ 1) = (x\ y\ z\ 1)MPT \quad (1)$$

ただし、

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -c_x & -c_y & -c_z & 1 \end{pmatrix} \quad (2)$$

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(-\varphi) & \sin(-\varphi) & 0 \\ 0 & -\sin(-\varphi) & \cos(-\varphi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

$$P = \begin{pmatrix} \cos(-\theta) & 0 & -\sin(-\theta) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(-\theta) & 0 & \cos(-\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

となる。上式に加えて、ある  $w = w_0$  での  $u$  方向の可視範囲  $w$  と  $v$  方向の可視範囲  $h$ 、カメラ映像の横幅  $w_p$  と縦幅  $h_p$  を測定し用いると、カメラ映像上での左上からのピクセル数  $x_p, y_p$  は下記のようなになる。

$$x_p = d_p \left( \frac{1}{2} + \frac{w_0}{d} \frac{u}{w} \right) \quad (5)$$

$$y_p = h_p \left( \frac{1}{2} + \frac{w_0}{h} \frac{v}{w} \right) \quad (6)$$

2つ目は操作者のウェブクライアント上での操作を受信してロボットに命令を出す機能である。操作者の操作はロボットアイコンのドラッグと、オブジェクトアイコンのクリック及びプルダウンメニューの選択がある。操作者がロボットアイコンをドラッグした場合、ウェブクライアントから送信されたドラッグ先の座標をサーバが受信する。座標はカメラ映像内での座標であるため、受信したカメラ上の座標及び、 $z =$



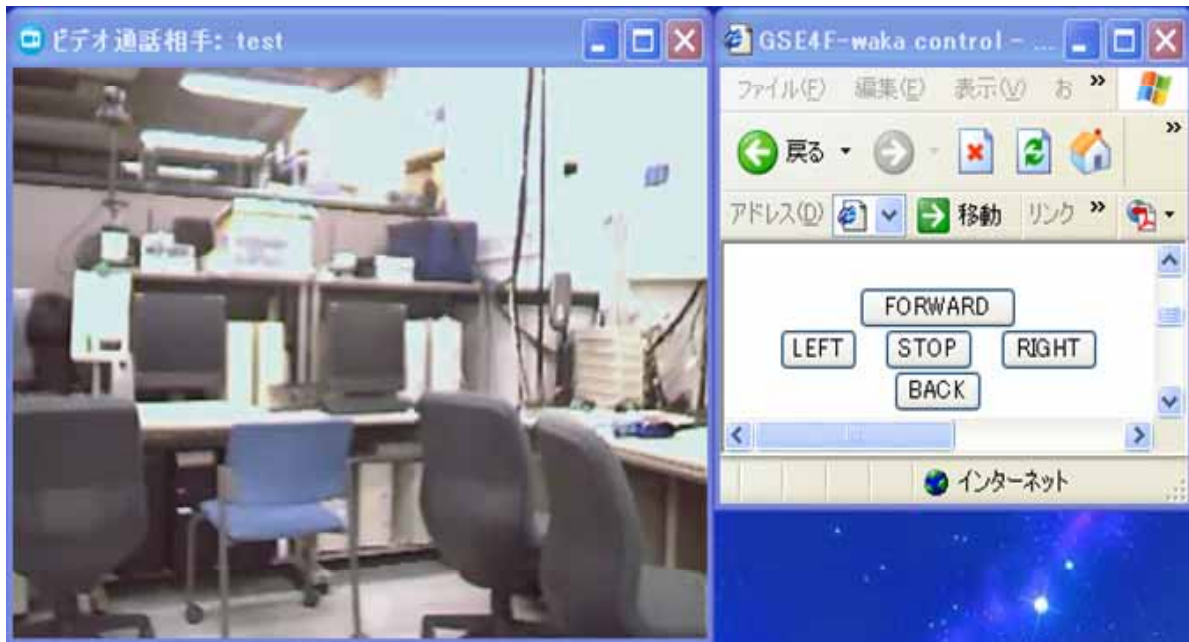


図4. 一人称視点

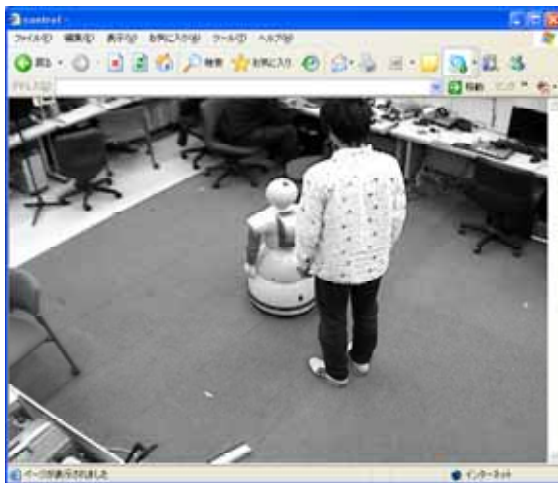


図5. 三人称視点

0 の条件を用いて上記の式(1)~(6)の座標変換の逆変換を行い、実空間上での座標に変換し、ロボットに移動先の座標として TCP 通信で送信する。ロボットは現在地の座標と、サーバから送られた移動先の座標を元に回転角と移動距離を計算し、移動する。操作者がオブジェクトをクリックし、プルダウンメニューからメソッドを選択した場合は、サーバはそのオブジェクト名とメソッド名を受信し、サーバに保持してあるオブジェクト情報ファイルを参照して、オブジェクトの位置座標とメソッドの内容を取得する。メソッドの内容とはそのメソッドが選択された

ときにロボットにどのような命令セットを送信するかを記述したものである。メソッドが「ここに移動」など、ロボットの移動を伴う場合は、ロボットアイコンがドラッグされたときと同様に移動命令をロボットに送信する。音声の再生や、身振り手振りなどのその他の動作命令の場合もロボットに命令セットを TCP 通信で送信し、ロボットはそれを受信してあらかじめ設定された動作を、動作命令に従って実行する。

#### 4. 実験

##### 4.1 実験内容

本論文で提案したインタフェースはネットワークカメラを用いた三人称視点、オブジェクト指向型操作を特徴としているが、今回は三人称視点を用いた操作の有用性について調べるための予備的な実験を行った。被験者に遠隔操作で指定した軌道に沿ってロボットを移動させて一人の人間を案内してもらった。案内される側の人間は実験者が行った。案内される側の人間がロボットに付いてこない場合はその場で一時停止して待つように被験者に指示したが、実験者は常にロボットの後方約 500mm の位置をロボット

に付いて歩いた。被験者は以下の一人称視点，三人称視点それぞれにおいて案内を行った。

#### 1. 一人称視点

ロボットに取り付けたウェブカメラの映像を操作者側の画面に映した。ウェブカメラはELECOM UCAM-E1D30MDSVを使用した。またウェブブラウザ上の前進，後退，左右回転，停止のボタンでロボットを操作することができるようにした（図4）。

#### 2. 三人称視点

天井に取り付けたネットワークカメラの映像をウェブブラウザに表示した。そしてその映像上をクリックすると，その地点にロボットが移動するという機能を実装した（図5）。

### 4.2 実験結果

一人称視点の場合，被験者は案内される人間がロボットの後ろに付いてきているかを確認するために，定期的にロボットを回転させて振り向かせていたが，三人称視点の場合はロボットと人との位置関係を常に把握できるのでそのような操作を行うことはなかった。また，一人称視点において，壁に向かって垂直に移動する場合に壁との距離をつかめず壁に接触する場面があったが，三人称視点の場合は適切な軌道でロボットを移動させることができていた。また，最後まで案内するのにかかった時間も一人称視点より三人称視点のほうが短かった。

### 4.3 考察

一人称視点よりも三人称視点を用いた方が正確に短時間で案内を行えていたので，案内などの移動を伴うロボットの遠隔操作においては三人称視点を用いるのが概ね有効であると思われるが，今回の実験で用いた一人称視点のウェブカメラの画角は47°と狭く，従来研究[1][2]のような広い視野を操作者に提供できていなかったため，三人称視点の優位性を示すにはより広い視野の一人称視点と三人称視点と比較し，検証する必要があると考えられる。また実験の設計においては，案内する軌道や部屋の状況，案

内される人数などもより厳密に定める必要がある。

また，今回の実験で用いた三人称視点のインタフェースには次のような問題点があったので，今後の実装でこれらを解決していく必要がある。

#### 1. カメラのひずみの影響

今回カメラ映像をクリックするという操作方法を実装したが，カメラのひずみによる座標変換のずれは考慮しておらず，画面上をクリックする操作方法では細かい操作ができなかった。今後はより正確な座標変換を行い，カメラ映像上でも細かい操作を可能とするインタフェースにする必要がある。

#### 2. カメラから遠い場所での操作

カメラを使った三人称視点の操作では，カメラから遠い地点であるほど単位ピクセルあたりの距離が大きくなり，細かい位置指定ができなくなる。この問題を解決するためには，部屋が広い場合には部屋の中に複数台のカメラを設置し，容易に切り替えることのできるようになるなどといったような工夫が必要である。

#### 3. ロボットの向きの指定

今回実装したカメラ映像をクリックするという操作方法では向きのみを変えることは不可能であり，向きを変える場合は現在地のすぐ近くをクリックして短い距離を移動して向きを変える必要があったが，上述の通り座標変換のずれにより細かい操作が不可能であったため，ロボットの向きのみを変えることが困難だった。これを解決するためには三人称型操作においても向きを変える操作を実装する必要がある。しかしながら，今回の実験のように画面上をクリックして移動させる場合，shiftキーを押しながらクリック，あるいは右クリックなどを「移動せずに向きのみを変更」に対応させることで解決できるが，3節でのウェブクライアント設計案のように，GUIを意識してドラッグの操作を割り当てた場合にどのような手法が適切であるかは容易に決定できず，向きを変える操作のデザインは今後の課題である。

## 5. おわりに

案内ロボットの遠隔操作インタフェースとして、コンピュータの GUI の操作を模した三人称視点、オブジェクト指向型のインタフェースを提案した。三人称視点のロボット操作として、天井に設置したカメラの映像内をクリックしてロボットを移動させるという操作方法を実装し、人をロボットの後ろに連れて案内する場合において一人称視点と比較した。その結果三人称視点の方がロボットや人の位置を把握しやすく、案内をしやすいと言えたが、今回の実験では比較対象の一人称視点の視野が狭かったことなどが問題であるとも考えられるので、三人称視点の有用性を証明するにはより厳密な比較実験を行っていく必要がある。また、ロボットに様々な動作を実装し、オブジェクト指向型の操作の有用性を調べる実験を行っていくことも今後の課題である。

## 謝辞

本研究の一部は、文部科学省科学技術振興調整費「先端融合領域イノベーション創出拠点の形成：ゆらぎプロジェクト」の研究助成によるものである。ここに記して謝意を表す。また、実験環境の構築に協力していただいた、河村竜幸氏、野上大輔氏、黒川高弘氏に感謝する。

## 参考文献

- [1] Jouppi, N.P.: First steps towards mutually-immersive mobile telepresence, Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 354-363 (2002)
- [2] Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y., Heath, C.: Dual ecologies of robot as communication media: Thoughts on coordinating orientations and projectability, Conference on Human Factors in Computing Systems - Proceedings, pp. 183-190 (2004)
- [3] Roussou, M., Trahanias, P., Giannoulis, G., Kamarinos, G., Argyros, A., Tsakiris, D., Georgiadis, P., Burgard, W., Haehnel, D., Cremers, A., Schulz, D., Moors, M., Spirtounias E., Marianthi, M., Savvaides, V., Reitelman, A., Konstantios, D., Katselaki, A.: Experiences from the use of a Robotic Avatar in a Museum Setting, Proceedings VAST 2001 Virtual Reality, Archeology, and Cultural Heritage, pp.153-160 (2001)
- [4] Luo, R.C., Chen, T.M.: Development of a multibehavior-based mobile robot for remote supervisory control through the Internet, IEEE/ASME Transactions on Mechatronics, Vol.5, No.4, pp.376-385 (2000)
- [5] Hu, H., Yu, L., Tsui, P.W., Zhou, Q.: Internet-based robotic systems for teleoperation, Assembly Automation, Vol.21, No.2, pp.143-151 (2001)
- [6] Kawachi, N., Koyuu, Y., Nagashima, K., Ohnishi, K., Hiura, R.: Home-use Robot "wakamaru," Mitsubishi Juko Giho, Vol.40, No.5, pp.270-273 (2003)