

# 人型入力デバイスを用いた遠隔動作指示手法

田山 友紀<sup>1,a)</sup> 安藤 禎晃<sup>2,b)</sup> 萩野 実咲<sup>2,c)</sup> 岡田 謙一<sup>1,d)</sup>

**概要:** 遠隔地にいる人へ体の動きや姿勢を指示するには音声通話、ビデオ映像を送るなどいくつかの方法が存在する。しかし従来の遠隔コミュニケーション手法において指示者は被指示者の間違いを直感的に正すことができない、被指示者にとっても自分の行っている動きが正しいのか判断するすべがない、といった様々な問題点がある。そこで我々は、人型入力デバイス QUMARION を用いた遠隔動作指示手法を提案する。指示者は QUMARION を操作して指示を出し、被指示者の動きは Kinect を用いてトラッキングされる。双方のデバイスから得たデータを比較し、その結果をもとに被指示者の姿勢がシステムによって評価される。これらの手順によって得られた被指示者の姿勢の正誤を指示者および被指示者にリアルタイムに伝えることが可能になる。本システムによって指示者は遠隔地にいる人にも正確で容易な指示が行えることと同時に、被指示者の高い習熟度が期待される。

**キーワード:** 遠隔, コミュニケーション, インタフェース, Kinect

## Using a Human-Shaped Input Device for Remote Pose Instruction

YUKI TAYAMA<sup>1,a)</sup> YOSHIAKI ANDO<sup>2,b)</sup> MISAKI HAGINO<sup>2,c)</sup> KEN-ICHI OKADA<sup>1,d)</sup>

**Abstract:** There are several ways to instruct poses or motions to someone in a remote area, such as phone calls using voice and video messaging using images. However, these methods do not allow the instructors to correct the student's poses intuitively. Moreover, the students have no way to know whether the pose they are making is correct or not. In this study, we propose a method of instructing poses to a person in a remote area using a human-shaped input device. The instructor uses QUMARION as an input device to instruct a pose, while Kinect tracks the student's motion. By comparing the data acquired by both devices, the system evaluates the student's pose. The result is processed and shown on both the instructor's and the student's displays in real-time. By using our system, more effective instructions by the instructors and quicker training of the students is expected.

**Keywords:** Remote, Communication, Interface, Kinect

### 1. はじめに

人が他人に指示をする時、言葉や身振りを使うことが多い。しかし指示する側である指示者と指示される側である被指示者が、お互いの声が聞こえず姿が見えない2地点に

いる場合はコミュニケーションをとるのが難しい。現在では音声通信を使用した電話や音声と映像を使用したビデオチャットといった技術が発達している。しかしダンスのように動きを伴う動作や、体操やリハビリテーションのように指示通りの正確な姿勢をする必要がある行動において、既存の伝達手法では伝えきれない場合がある。同じ空間にいれば指示者は被指示者に目の前で動きを見せることが出来ると同時に、間違っていた場合は直接被指示者の体に触れて腕の角度を正すといった直感的な指示が出来るが、両者が遠隔地にいる場合指示者は直感的な指示がしづらい。また、被指示者も目の前で細部まで指示者の姿勢を見るこ

<sup>1</sup> 慶應義塾大学理工学部  
Faculty of Science and Technology, Keio University  
<sup>2</sup> 慶應義塾大学大学院理工学研究科  
Graduate School of Science and Technology, Keio University  
a) tayama@mos.ics.keio.ac.jp  
b) ando@mos.ics.keio.ac.jp  
c) misaki@mos.ics.keio.ac.jp  
d) okada@z2.keio.ac.jp

とが出来なくなるため、自身の姿勢が正しいのかどうか判断し難い。さらに、人間同士のコミュニケーションにおいて周囲の環境から得られる情報は重要である [1]。例えば「あれ」や「それ」といった指示語の内容を共有出来たり、「腕をここまで挙げて」と言えるのは、同一環境にいるからこそである。環境の共有が難しい遠隔コミュニケーションにおいては、支持者の意図を直感的に伝えるための媒介が存在しないため、お互いの意思疎通に問題が生じる場合が多く存在する。

このような背景から本研究では、人型入力デバイスを用いた遠隔動作指示手法を提案した。本提案システムにより指示者はデバイスを手で操作しながら直感的に指示をすることが出来、そのデータはリアルタイムで被指示者と共有される。さらに被指示者の状況もセンサを使ってトラッキングされ、指示者と共有される。双方のデータをシステムが自動的に比較検討する機能を実装したので、被指示者は自身が指示通りの姿勢になっているのかどうか瞬時に理解することが出来る。以下、2章では関連研究と問題点について述べ、3章ではその問題点を解決するための本研究の提案を述べ、4章でそれに基づく実装について述べる。そして5章では実験概要について記述し、最後6章を本研究のまとめとする。

## 2. 関連研究

### 2.1 ユーザインタフェース

入力インタフェースとしては、CLI (コマンドラインインタフェース) に分類されるキーボードや GUI (グラフィックカルユーザインタフェース) にあたるマウスが一般的である。CLI では、基本的なコマンドをある程度覚えないと使いづらいと言った短所がある一方で、一度習得してしまえば手慣れたユーザにとっては作業をしやすい [2]。GUI ではシステムのリソースを消費してしまう短所があるが、視覚的情報が豊富なため初心者には操作しやすい。また、両者とも人間の動きのような煩雑な情報の入力を行うことは非常に困難という欠点がある。

最近では視線やジェスチャといった、NUI (ナチュラルユーザインタフェース) もよく耳にするようになった。岡田ら [3] はハンドジェスチャによってコンピュータを操作する手法を作成した。人間にとってより自然な動作でコンピュータを操作出来るのが特徴であり、より直接的に入力を行うことが可能である。しかし現段階では認識が可能なジェスチャの種類に限りがあり、複雑な指示を伝達するには不適切である。

一方で、TUI (タンジブルユーザインタフェース) も研究されている。触れて感知することが出来る、つまり実体のあるデバイスに物理的に触れることで、直感的に操作することが可能なインタフェースである。図 1 に示す関口ら [4] の RobotPHONE は、ネットワークを介して2人の



図 1 RobotPHONE

ユーザをつなぐ。ユーザの持つぬいぐるみの関節を動かして操作すると、もう一方のユーザが持つ同形のぬいぐるみも同様に動く。人や物の動きを表すには CLI や GUI よりも優れていると考えられるが、しかしながらぬいぐるみでは可動範囲が狭いため、人間の動作を伝えるには不十分である。この様に特にロボットをインタフェースとして使用する手法を RUI (ロボットユーザインタフェース) とも呼ぶ。

### 2.2 実映像を用いたシステム

近年は遠隔にいる人とコミュニケーションをとったり、相手の様子を知りたいといった時にはビデオチャットのような手法を利用することが多くなってきた。しかしこれらは実映像を用いたものが多く、プライバシーが守れないといった問題点がある。

Dufaux ら [5] や Zhang ら [6] は監視カメラの利用が、プライバシー侵害とならないようシステムを作成した。監視カメラには不特定多数の人間が写ってしまう恐れがあり、万が一映像データが漏洩した時に個人データが拡散しないためのシステムとなっている。しかしこれらの研究では人物にスクランブルをかけたかばやかすとといった映像を加工する手法がとられている。さらに、人物の動きに合わせて画像にフィルタをかけてしまうため、実際の人物の詳細な動きが視聴者に伝わりにくくなる恐れがある。

### 2.3 遠隔指示手法の例とその問題点

指示者と被指示者が異なる空間にしながら、ネットワークを介してデータを共有し指示や指導を行う研究は多くなされてきた。遠隔で動作を指示する例として挙げられるのが、リハビリテーション指導やダンスである。医療などに関する専門知識や、特定のノウハウを持った指示者が被指示者と離れた遠隔地にいる場合は直接目の前で指導に当たることができず、遠隔で指示者にとってわかりやすい指示を出すシステムが必要になる。

飯野ら [7] の研究では、ストリートダンスにおいてモーションキャプチャシステムを用いて各関節に関する動きの分析と指導を行った。しかしこのシステムではモーション

キャプチャにカメラが18台必要であり、環境を整えるのにコストと時間がかかってしまう。

筋野ら [8] の研究ではダンス指導を目的としたシステムを作成した。指示者の動きは Kinect を用いてモーションキャプチャし、被指示者の動きも Kinect を用いてトラッキングされる。しかし、このシステムでは予め指示者のモーションデータを取得し、入力しておく必要がある。一連の流れのモーションデータ取得に時間と手間がかかると同時に、指示者はリアルタイムで直接指導にあたる事が出来ず、システム使用後に再度直接指導に当たるなどさらに余計な手間となってしまう可能性もある。

Chiri ら [9] は遠隔で手のリハビリテーションを行うシステムを作成した。手袋型の装置を手にはめて使用することによって各関節の細かい動きも指示出来る。手袋型の装置以外にはカメラやセンサといったものは必要としない。しかし指示者、被指示者共にこの手袋型の特殊な装置が必要であり、このシステムの使用は手の動きに限定されてしまっている。よって、体の他の部位のリハビリテーションには使えず、利用出来る場面が少なく汎用性が低い。

### 3. 人型入力デバイスを用いた遠隔動作指示手法

遠隔地にいる人へ動きや姿勢を指示する場合、効率よく指示者が指示を行うためには伝えやすさが、被指示者が素早く学習するためにはわかりやすさが重要となる。しかしながら現在の音声のみを使用する手法では、姿勢を言葉にして伝えるのが難しい。一方で、映像を使用した伝達手法では伝えやすさは向上するが実映像を使用した場合はプライバシーが守れないといった様々な問題点が指摘されている。また、映像を使っても現在主流である CLI や GUI といったインタフェースを用いた入力手法では、指示者が直感的に被指示者の間違いを正せない、といった問題は改善されない。モーションキャプチャを用いた入力も研究されているが、こちらの場合正確な入力を行うには複数のカメラが必要であったり、環境整備に時間やコストがかかってしまう。さらにビデオチャットのような形式ではなく録画映像を見る形式ならば、リアルタイム性の問題も依然として残る。

そこで、本研究では人型入力デバイスを用いた遠隔動作指示が行えるシステムを構築した。指示者に対しては人型入力デバイスを手で操作することでより直感的な指示を出すことが可能となる。被指示者は Kinect の前で指示された姿勢をとるだけでシステムが姿勢の正誤を自動判断するため、自身で姿勢の細かな修正が出来るようになる。

#### 3.1 想定環境

今回提案するシステムでは指示者と被指示者が異なる空間、つまり遠隔地にいることを想定する。使用用途として



図 2 人型入力デバイス QUMARION

は、一つ一つの姿勢を正確に指示する必要があるリハビリテーションの姿勢指導などを想定する。

#### 3.2 人型入力デバイスの使用

遠隔地にいる人へ姿勢を正確に伝達するには、指示者の伝達のしやすさが重要である。姿勢をただ見せる場合は自身はその姿勢をとり、その映像を送ることも一つの手法である。しかし、この手法ではその姿勢をとる場所があることが条件となる。例えば、病院内の狭いオフィスから被指示者へ指示を出そうと思っても、十分な場所が確保出来ず、狭い中では正確な姿勢が作れない。また、自身の姿勢や動きをキャプチャする設備を設置する必要もあり、決して楽な手法とは言えない。

そこで本研究では、図 2 に示す人型入力デバイス QUMARION を使用することで指示者の環境を問わず指示を行うことを可能とした。このデバイスは体長約 30cm で、机の上で楽に操作が出来る。さらに、指示者は実際にデバイスを触りながら操作することが出来るので直感的に指示を出すことが出来、被指示者の間違いを正すことが出来る。また、センサがデバイスに組み込まれているため座標値に誤差が生じることはなく、より正確な指示を行うことが可能である。

#### 3.3 システムによる自動姿勢正誤判断

他人から何かを教わる場合において、フィードバックがあるかないかは結果に大きな差をもたらす。フィードバックが細かければ細かい程、より正確な姿勢へ修正することが出来る。リハビリテーションの場合、正確な姿勢や動きをすることが最大限の効果を得る方法となるため、フィードバックをわかりやすく提示することは非常に重要である。1人で目視による判断をただけでは体の部位全てに目が行き届かず、正確な指示が行えない可能性が高い。

そこで、指示者からの指示を妨げずに姿勢の正誤判断を自動的に行う機能が必要である。本研究では、指示者は



図 3 Kinect

QUMARION, 被指示者は図 3 に示す深度センサを搭載した Kinect を使用する。QUMARION, Kinect 双方のデバイスから取得した座標値をそれぞれ比較し、姿勢が一致しているかどうかを判断する。一致していなければ、一致していない関節点を画面に列挙する。このことにより、被指示者は画像および文字を見ながら姿勢の修正を行うことが可能となる。

### 3.4 状況のリアルタイム共有

姿勢や動作を指示する状況において、指示者と被指示者がお互いの様子をリアルタイムで見ることが出来る環境は不可欠である。例えば非同期ビデオ映像を使用して動作指示を行う場合、指示者はどのタイミングで次の姿勢へ移ってよいかわからない。一方でリアルタイムでお互いの様子がわかる場合、指示者は次の姿勢指示へ移ってよいのか、あるいは現在の姿勢が出来ていないためまだ指示する必要があるのかを瞬時に判断することが出来る。また、被指示者も非同期の場合は自分で映像を操作し早送りや一時停止をする必要があるが、リアルタイムに状況が共有されることによってこのような作業が不要となる。さらにフィードバックを即時に受け取ることが出来るためより効率の良い習熟につながる。

## 4. 提案システムの実装

### 4.1 実装環境

本研究に置ける実装環境は、Windows 上で Visual Studio 2010 を用いて実装されている。また、プログラミング言語は Visual C++ を使用した。

指示者は人型入力デバイス QUMARION をパーソナルコンピュータに接続し、それを手で操作することで指示を出す。被指示者は Kinect を指示者とは別のパーソナルコンピュータへ接続し、その前に立って使用する。

### 4.2 システム構成

図 4 に本システムの全体構成を示す。まず指示者は QUMARION を起動し、それぞれの関節の座標値を取得する。得られた関節の座標値をもとに QUMARION の様子をリアルタイムに描画し、表示する。被指示者は Kinect を起動した後、Kinect の前に立つことで関節の座標値取得を開始し描画する。指示者のコンピュータをサーバ、被指

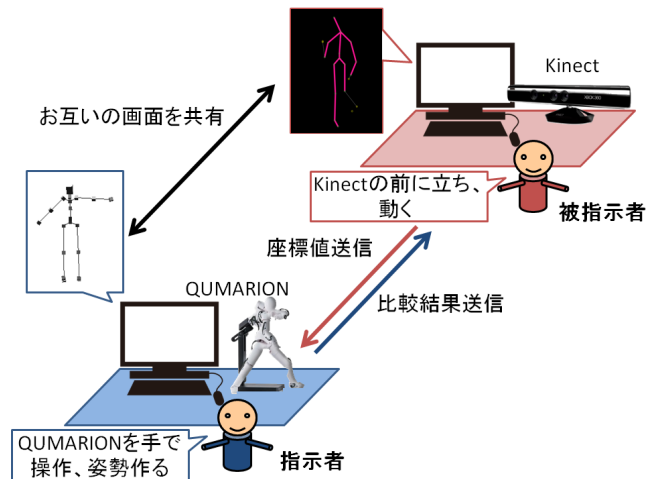


図 4 システム構成

示者のコンピュータをクライアントとした通信を確立し、Kinect で得た座標値情報を送信し、双方の値を比較する。比較した結果は指示者のコンピュータに格納されると同時に、被指示者へ送信される。また、座標値の通信と同時に双方で描画したデータもお互いに送信しあう。指示者、被指示者双方の画面には QUMARION による描画、Kinect による描画、および比較結果が文章で表示される。

### 4.3 指示情報および被指示者情報の取得

指示者は QUMARION を使用して被指示者へ指示を送る。被指示者へ伝えたいポーズや動きを QUMARION を手で操作することで作成する。図 5 のように、QUMARION には座標値をとれる関節点が 20 カ所あり(頭、右肩、左ひざなど)、それぞれに角度センサが内蔵されている。角度センサから得た情報をもとに実際の座標値を計算し、取得する。次に、それぞれの関節点において立方体を描画し、関節同士を線で結ぶことで人型アバターを完成させ、画面上に表示する。

指示者の画面に表示されたアバター画像はビットマップファイルとしてキャプチャされ、ソケット通信を通じて被指示者へと送信する。被指示者側では得られたビットマップファイルをそのまま画面に表示させ、被指示者はそれを見ながらポーズをとる。

一方で、被指示者は Kinect を使用して自身の様子を指示者に送り、システムにポーズの正誤を判断してもらう。Kinect の前に立ち人間を検出した時点から、自動的に関節点の座標値を取得し始める。図 6 のように Kinect も検出可能な関節点が 20 カ所あるため、取得した座標値をもとに人型アバターを表示させる。

被指示者の画面に表示されたアバター画像はビットマップファイルとしてキャプチャされ、指示者へ送信する。また、それと同時に実際の関節の座標値も全て指示者へ送信する。

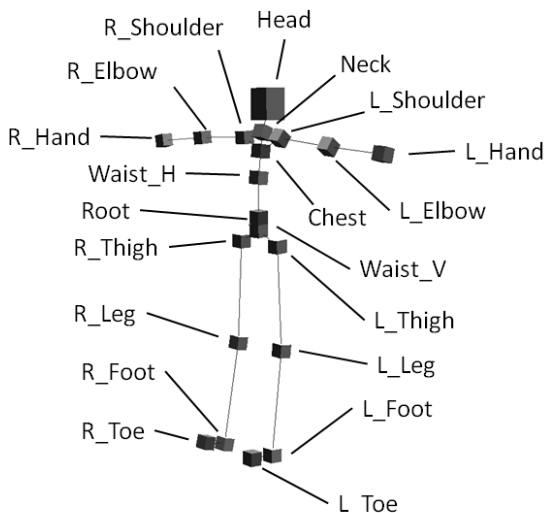


図 5 QUMARION で取得可能な関節

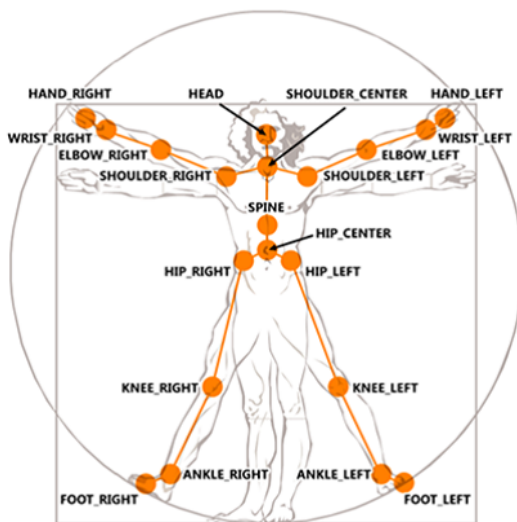


図 6 Kinect で取得可能な関節

#### 4.4 姿勢の正誤の決定

指示者のコンピュータに、QUMARION より取得した座標値と Kinect より取得した座標値が格納された後、それらを比較する。今回座標値を比較するためのパラメータとして角度を用いた。ある関節 A において隣接関節 B と C が存在する場合、まず点 A と点 B の 2 点を通る直線  $\alpha$  と、点 A と点 C を通る直線  $\beta$  の式を求める。次に、直線  $\alpha$  と直線  $\beta$  の 2 直線がなす角度を計算する。これを xy 平面、yz 平面、xz 平面においてそれぞれ行い、一つの関節につき計 3 つの角度を得る。足先や指先のように隣接関節が 1 つしか存在しない場合は、角度を 0 として入力しておく。

角度計算を QUMARION と Kinect の関節点全てに行った後、それぞれに対応する関節点の角度を比較する。この時、差異が 15 度以内であれば一致、それ以上であれば不一致と判定することにした。角度が一致しなかった関節はリストにして保存した。15 度とした理由の一つとして、お辞

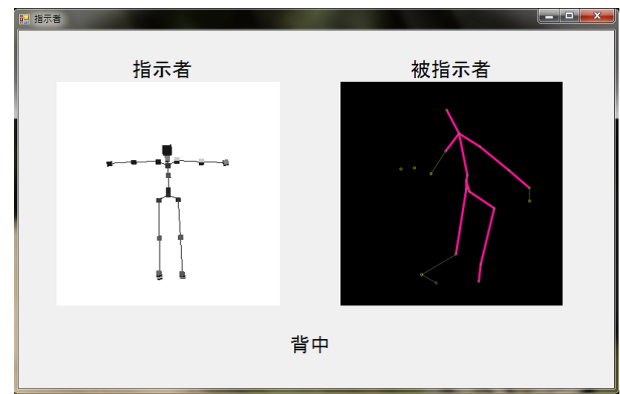


図 7 表示画面

儀が挙げられる。一般的に会釈をするには体を 15 度傾け、敬礼をするには 30 度傾けると言われている。よって 15 度傾くと違う姿勢をとっていると見なされるため、本システムでは閾値を 15 度とした。

#### 4.5 表示画面

QUMARION と Kinect の座標値を取得し、比較をした後は、指示者と被指示者の使用端末であるコンピュータの画面上に両者の様子と結果をあわせて表示する。図 7 にその表示画面を示す。表示画面は指示者と被指示者で同一のものとした。左側に指示者が操作する QUMARION のアバター、右側に被指示者の様子をリアルタイムで更新し表示する。画面下部には被指示者のどの関節が指示と異なっているのかを文字で示す。

文字表示について、位置が異なっている関節のリストをそのまま全て読み込んで表示するのではなく、一部の関節のみ表示することとした。これは、姿勢が全く一致していない場合に全関節が羅列されるのを防ぐためである。図 8 のように、体の中心部分の関節をルートとし関節の親子関係を構築した。体の中心関節を背中として設定したため、上半身と下半身を別々に走査していく。親関節が一致していれば子関節の走査をし、一致していなければその親関節名のみを表示する。これにより、姿勢が全く一致していなかった場合は「背中」のみを表示し、背中と下半身は一致しているが首が一致しない場合は「首」を表示するようになる。体の中心部から走査を始め、文字を表示することでおおまかな姿勢をまず正すことができる。その後子関節への走査が進んでいくことで、指先や足先といった細部への指示を行うことが可能となる。また、被指示者は文字で示された関節を中心に、QUMARION の映像も見ながら自身の映像が一致するように動き、姿勢を正していく。図 7 では背中の中の関節が一致していなかったが、一致すると図 9 のように表示される文字が変化する。

指示者、被指示者共に画面上で特に操作する必要はなく、指示をすること、指示を受けることだけに集中出来るようにした。

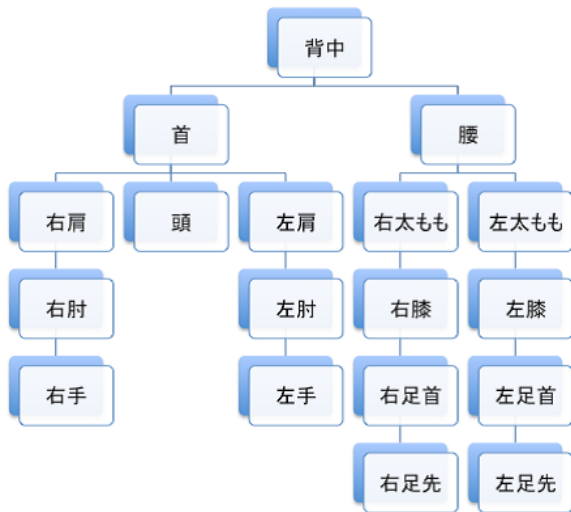


図 8 関節の親子関係

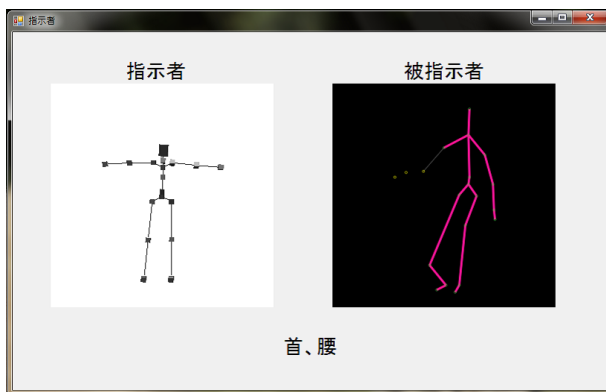


図 9 背中関節が一致した場合の表示画面

## 5. 実験案

### 5.1 実験目的

被指示者は指示者によって揭示された情報をもとに自身の姿勢を作り、学習をする。そのため、指示者が揭示した指示が被指示者に正確かつわかりやすく伝わっているかが大きな要素となる。そこで本実験では遠隔動作指示において本提案手法を用いることで、どの程度正確に指示を伝えることが出来、被指示者の学習に貢献したかを検証する。

### 5.2 実験内容

まず被験者を2グループに分け、一方のグループには本提案システムを、もう一方のグループには従来手法の一つである音声通話を使用してもらう。それぞれのグループの中でペアを作り一人は指示者役、もう一人は被指示者役とする。実際のリハビリテーションで使用されている姿勢をいくつか指示者に指示してもらい、被指示者は指示された通りに姿勢をとる。この時、被指示者が一連の姿勢の習熟にかかった時間を両グループで計測し比較する。さらに、本提案手法を使用した指示者を対象として、意図した姿勢

を指示することが出来たかどうかを検証することで、本提案手法の有用性を評価する。

## 6. おわりに

お互いの声が聞こえず姿が見えない遠隔空間にいる2人がコミュニケーションをとるのは非常に困難である。特に姿勢や動きを指示する、といった視覚情報が重要となるコミュニケーションにおいてはその伝達方法が要となる。しかしながら従来の音声のみの伝達方法や映像をそのまま表示するだけの手法では、指示者は直感的に被指示者の姿勢の間違いを正すことが出来ず、被指示者も自身のとっている姿勢が正しいのか判断出来ないといった問題点があった。遠隔地にいるため同じ空間を共有することができず、「この腕」や「この部分」などと言った指示語では伝わりにくく、コミュニケーションがとりづらい。姿勢や動きを指示する例として挙げられるダンスでは、キレのある動きや姿勢が見た目の美しさを左右するため正確な伝達が必要である。別の指示例として挙げられるリハビリテーションの姿勢指導では、一つ一つの姿勢や動きに意味があり、それらを正確に被指示者が習熟し実践することではじめてリハビリテーションの効果が得られる。

そこで本研究では、人型入力デバイス QUMARION を用いた遠隔動作指示手法を提案した。人型入力デバイスを用いることで指示者は実際にデバイスを触りながら伝達できるので、より直感的に指示を出すことが出来る。さらに、被指示者の姿勢情報を Kinect を用いて収集し、その情報をもとに姿勢の正誤を自動判断する機能によって、被指示者へ姿勢に関する的確な修正フィードバックが可能となった。さらにお互いの状況も送信しあい共有する。これらの情報収集、座標値比較、姿勢の正誤判断、画面共有は全てリアルタイムで行われる。実験では、リハビリテーションポーズを遠隔地にいる人へ指示することで、遠隔動作指示において本提案手法はどの程度指示者にとって伝えやすく、被指示者にとってどの程度わかりやすいのかを検証する。指示者の容易かつ正確な指示を支援することにより、被指示者の正確かつ効率の良い習熟につながると期待する。

**謝辞** この研究の一部は文部科学省科学研究費補助金 (B) 課題番号 23300049(2013 年) の支援により行われた。

### 参考文献

- [1] Christian Heath, Paul Luff, Hideaki Kuzuoka, Keiichi Yamazaki, and Shinya Oyama, "Creating Coherent Environments for Collaboration" in Proceedings of ECSCW'2001, Bonn, Germany, pp.119-38, 2001.
- [2] ThinkIT 「第3回 グラフィカルユーザインタフェースとコマンドラインインタフェース」  
<http://thinkit.co.jp/article/1120/1>
- [3] 岡田隆三, 風間久, 鈴木孝子, "PC 向けユーザインタフェース「ハンドジェスチャリモコン」", 映像情報メディア学会

- 誌, Vol. 64, No. 12, pp.1812-1815, 2010.
- [4] Dairoku Sekiguchi, Masahiko Inami and Susumu Tachi, "RobotPHONE: RUI for Interpersonal Communication" in Proceedings of CHI '01 Extended Abstracts on Human Factors in Computing Systems, New York, USA, pp.277-278, ACM Press, 2001.
  - [5] F. Dufaux and T. Ebrahimi, "Scrambling for video surveillance with privacy", in Proceedings of IEEE Computer Vision and Pattern Recognition Workshop on Privacy Research in Vision, 2006.
  - [6] W. Zhang, S. S. Cheung, and M. Chen, "Hiding privacy information in video surveillance system", in Proceedings of International Conference on Image Processing, 2005.
  - [7] 飯野友里恵, 高橋時市郎, "ストリートダンス動作の分析とダンス指導への応用", 映像情報メディア学会技術報告, Vol. 35, No. 14, pp.49-52, 2011.
  - [8] 筋野正太, 森谷友昭, 高橋時市郎, "NPR 機能を付加したダンスの動作解析・指導システム", 情報科学技術フォーラム講演論文集, Vol. 11, No. 3, pp.353-354, 2012.
  - [9] Azzurra Chiri, Mario Cortese, Paulo Rogerio de Almeida Riberio, Marco Cempini, Nicola Vitiello, Surjo R. Soekadar, and Maria Chiara Carrozza, "A Telerehabilitation System for Hand Functional Training", Converging Clinical and Engineering Research on Neurorehabilitation Biosystems & Biorobotics, Vol. 1, pp.1019-1023, 2013.