

## 遠隔会議でのアイコンタクト実現手法の提案と評価

小峯 隆宏<sup>1,2</sup> 勝本 道哲<sup>1</sup> 丹 康雄<sup>2</sup>

**あらまし** 会議参加者の一部が遠隔地から参加する遠隔会議において対面会議と同様の活発な議論を展開するためには、遠隔地間とのアイコンタクトによるスムーズな話者交替の実現が重要である。我々は、遠隔参加者の分身として、参加者の目となるビデオカメラ、耳となるマイクロフォン、口となるスピーカ及び顔映像を出力する液晶ディスプレイを一体として上下左右旋回する機能をもつ代理人ロボットを試作した。本論文では、試作システムによるアイコンタクトおよび話者交替の実現に関する評価実験を実施し、同試作システムのアイコンタクト実現手法の有効性の評価について報告する。

## Proposal and evaluation for the interactive video teleconferencing system with the eye contact

Takahiro Komine<sup>1,2</sup>, Michiaki Katsumoto<sup>1</sup>, Yasuo Tan<sup>2</sup>

**Abstract** The realization of smooth changing of speakers by the eye contact among multi-sites is very important for successful interactive video teleconferencing as well as the ordinary conference in one site. We have developed the attorney terminal that a set of a camera, a monitor, a microphone and a speaker acts for remote participant's eyes/face/ear/mouth can turn round freely as the attorney of the remote participant with other real participants in the same real space. This paper evaluates the usability of the proto-type system of the attorney terminal and considers about the requirements for establishing the eye contact with concentrating the interactive teleconferencing.

### 1. はじめに

インターネットの広帯域化および H. 323[1]等の遠隔会議システム標準プロトコルの普及により、近年、遠隔会議の実用利用の機会が増加してきている。特に H. 323 を活用したシステムでは、インターネット等を経由して遠隔参加者がカメラを操作することが容易に実現できるため、遠隔参加者が会議室内の注目したい映

像を自分の意思で選ぶ等会議への積極的な行動がある程度可能になる。しかし、会議室側の参加者には小さな遠隔カメラが動作するのを確認できるものの遠隔参加者の映像は固定のままであるため、今どこに注目しているのか等の同一実空間内にいる参加者同士ではアイコンタクトや顔の方向によって瞬時に判断できる情報が把握できない。このため、現在の遠隔会議システムは、単一会議室内で行われる通常の対面会議では可能なスムーズな話者交替の実現を難しく、遠隔参加者との議論が極端に少ない会議になる傾向にある。

遠隔会議システムにおいてスムーズな話者交替を実現するには、遠隔参加者が会話をした

- 
- 1 独立行政法人情報通信研究機構  
National Institute of Information and Communications Technology  
E-mail: komine@nict.go.jp
  - 2 北陸先端科学技術大学院大学  
Japan Advanced Institute of Science and Technology

いと思う相手の方向に顔を向けアイコンタクトをおこなうことと、会議室内の参加者が遠隔参加者の視線方向や表情を把握できることが重要である。この実現手法としては、仮想空間上にアバタと呼ばれる自分の分身を配置してアバタ同士による自然なコミュニケーションを実現させようとする手法[2]と、遠隔参加者の分身として他の参加者と並んで同一実空間上の会場内に筐体（以後、代理人ロボットと記述する）を置き遠隔参加者が遠隔操作する手法[3]が存在する。

これに対し、本研究ではシステムの汎用化と技術的課題に対する簡便な手法での解決を目標とし、後者の手法を試みることにした。これにより、仮想空間を表現する大型スクリーンや特殊な表示装置等を使用せず、また参加者の実写映像の表示モニタを実際に振向かせるという極めて簡便な手法でアイコンタクトの実現をはかれる。ただし、同手法を遠隔会議に用いる場合、遠隔参加者が何らかの入力デバイスで遠隔操作を自ら行うため、操作が簡便でかつ会議への集中力を維持できる工夫が必要である。遠隔カメラの簡便な遠隔操作を実現するための入力デバイスと遠隔制御手法に関しては、これまでも研究[4]されているが、アイコンタクトの実現という視点で研究は少ない。本論文では、この部分に焦点をあてた議論と考察を行う。

本論文では、まず代理人ロボットの試作システムについて説明する。次に、この試作システムが、遠隔地間の参加者同士のアイコンタクトおよび話者交替の実現の面で有効に機能していかどうかの評価を、遠隔会議模擬実験での被験者の行動分析から試みる。

## 2. 代理人ロボットの試作システム

本研究で想定する代理人ロボットを利用した遠隔会議システム環境を図1に示す。会議室の参加者と遠隔会議室にいる遠隔参加者はインターネットを介して会議を行う。会議室には、参加者のほかに遠隔参加者の分身として他の参加者と並んで実際に会議に出席する代理人ロボットが設置される。音声・映像伝送装置により代理人ロボットのモニタ画面上に遠隔参加者の顔の実写映像が表示されると共に、頭部カメラにより会議室の映像が遠隔会議室内のモニタに表示される。また、代理人ロボット制御クライアントおよびサーバを介することで、

遠隔会議室内の遠隔参加者は、入力デバイス操作によって代理人ロボットの頭部旋回動作等の遠隔操作を行うと想定する。

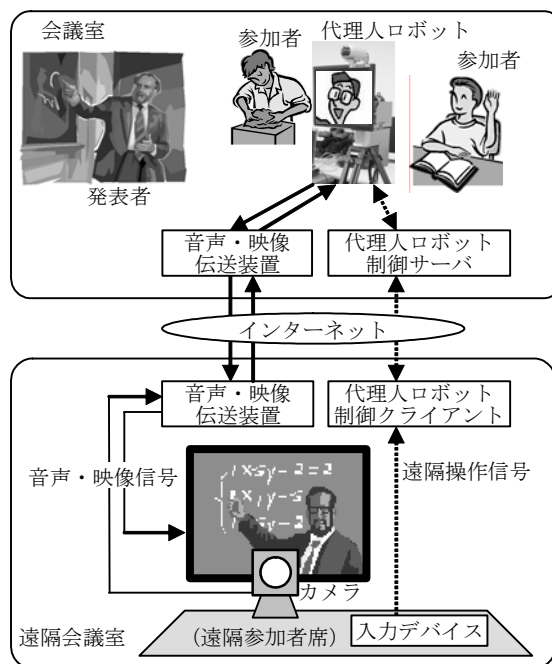


図1. 想定する遠隔会議システム

### 2.1 システム構成

図2に試作システムの構成を示す。代理人ロボットの本体は、遠隔カメラ、マイクロフォン、モニタ、スピーカから構成される頭部とその頭部を旋回動作させる駆動部とキャスター付の大型三脚から構成される胴体部分にわけられる。また、遠隔参加者の目に相当する遠隔カメラでは、旋回動作やズーム機能によって上述の頭部旋回では調整できない微小な視線移動や視野の拡大・縮小という行為が可能である。代理人ロボット制御サーバで稼動する制御プログラムは、インターネットを経由して届く遠隔会議室での遠隔操作の情報を取得して遠隔カメラや頭部旋回制御装置に対応する動作命令を送信する。

### 2.2 特長

試作システムの3つの特長を説明する。

1 点目は、代理人ロボット本体の全長が等身大スケールであり、かつ液晶モニタに表示される遠隔参加者の顔映像を実物大の大きさに調整して表示することが可能というものである。

人は普段同じ実空間上にいる実物の人同士で会話をすることに慣れているため、音声と映像を用いた遠隔での会話を行う場合、表示される顔映像が実物に比べて小さいとその存在感が低下してしまい、逆に実物に比べて大きいと威圧感や違和感が増加[5]してしまう。このため試作した代理人ロボットは、等身大スケールの全長および顔寸法にするよう設計した。

2点目は、遠隔参加者の目と耳の役割をする遠隔カメラと顔映像表示と口の役割をする液晶モニタを一体とした頭部を、遠隔操作により上下左右旋回等の動作を行うというものである。これにより、遠隔参加者は会話をしたいと思う相手の方向に顔を向けアイコンタクトをとることができる。また、会議室内の他の参加者からは遠隔参加者の視線方向や表情を把握できる。

3点目は、遠隔カメラと液晶モニタの取り付けマウント部分の標準規格対応、頭部旋回動作や遠隔カメラ制御等の動作命令関数のライブラリ化等による汎用性の確保である。遠隔カメラ用の取り付けマウントは DIN 規格、液晶モニタの取り付けマウントは VESA 規格に対応しており、遠隔カメラと液晶モニタの選択面での汎用性を高めた。また、代理人ロボットへの動作命令を関数ライブラリ化すると共に、機種毎に異なる遠隔カメラの制御命令パラメータを記述したプロファイル情報を関数ライブラリが参照することで、動作命令の一元化を実現した。これにより、遠隔カメラの機種に依存しない制御プログラムで動作し、新機種の遠隔カメラに交換する場合もプロファイル情報にパラメータ情報等を追加するだけで稼働可能となる。

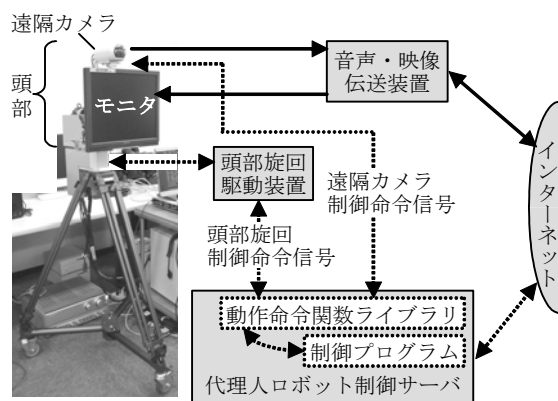


図2. 代理人ロボット試作システム構成

## 2.3 遠隔操作システム

今回試作した遠隔操作システムの概要を図3に示す。制御クライアントにおいてジョイスティック情報を100ms間隔のポーリングで取得し、その情報が頭部旋回制御命令もしくは遠隔カメラ制御命令の場合に、その情報を送信する。一方、制御サーバでは受信したジョイスティック情報に対応する動作命令関数を呼び出すことで代理人ロボットを動作させる。

頭部旋回は、図4に示すように現在の頭部の位置を基準とする相対移動とし、上・下・左・右の各旋回動作はジョイスティックのスティック部分を前方・手前・左・右にそれぞれ傾けることで動作する。なお、スティックを握る手のわずかな動きによる誤動作を防止するため、スティックのセンター位置からある一定傾き角度以内を無反応領域に設定した。これにより、移動感知領域にスティックがある場合のみ、センターからの相対座標(x, y)に対応した頭部の相対旋回動作コマンドを送信する。また、ジョイスティックのフィードバック機能は有効にしてあり、スティックから手を離すと中心位置に戻るようにした。

遠隔カメラのズームおよび旋回は、制御サーバの動作命令関数ライブラリに何種類かの市販の遠隔カメラの旋回制御コマンドおよびパラメータ特性を登録しておくことで、登録済の市販の遠隔カメラであれば同じジョイスティック操作で統一した遠隔操作を実現できるようにした。図5にジョイスティックでの操作命令割当を示す。

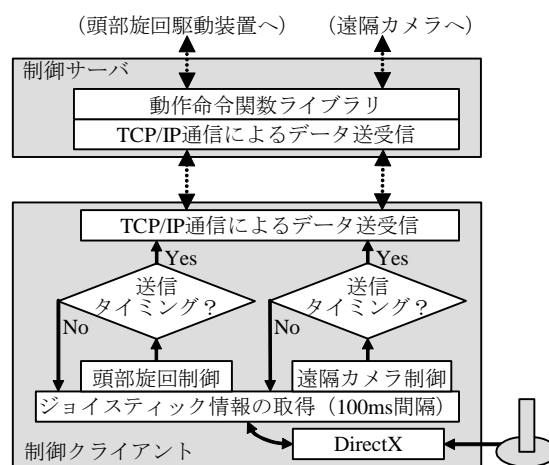


図3. 遠隔操作システム概要

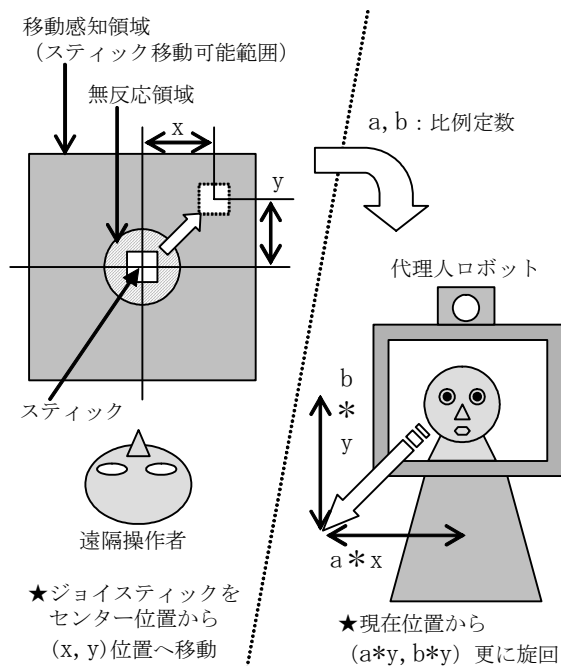


図4. スティック移動と頭部旋回の対応

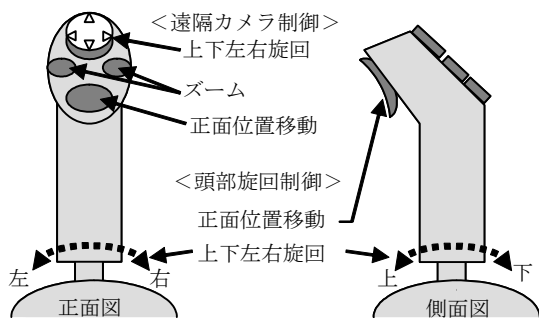


図5. ジョイスティックでの操作命令割当

### 3. 評価実験

試作システムが遠隔地間の参加者同士のアイコンタクトおよび話者交替の実現の面で有効に機能しているかどうかの評価を行う。被験者がタスクを実行する過程を観察し、遠隔会議模擬実験での被験者の行動分析[6]を行うことにした。

#### 3.1 実験構成

被験者は、情報通信研究機構の事務系職員20名である。被験者は全員、正常視力を有している。今回の実験では被験者4名ずつの計5グループに分けて評価実験を実施した。

評価実験のシステム構成を図6に示す。代理

人ロボットと3名の被験者W、X、Yを会議室内の1つの机を囲む形で配置し、被験者Zは遠隔会議室から代理人ロボットを介して参加する。被験者Zは必要に応じて会議室内の代理人ロボットの旋回動作、遠隔カメラのズームや旋回動作等をジョイスティック操作により実行する。音声・映像のリアルタイム伝送には、DVTs[7]を活用した。ネットワーク伝送遅延時間は、遅延発生器を用いて研究報告[8]にある東京・大阪間のインターネット模擬環境と同じ片道7msに設定した。

なお、後述する評価解析を行うために、4台のカメラ映像をマルチ合成映像として評価解析用記録テープに録画した。

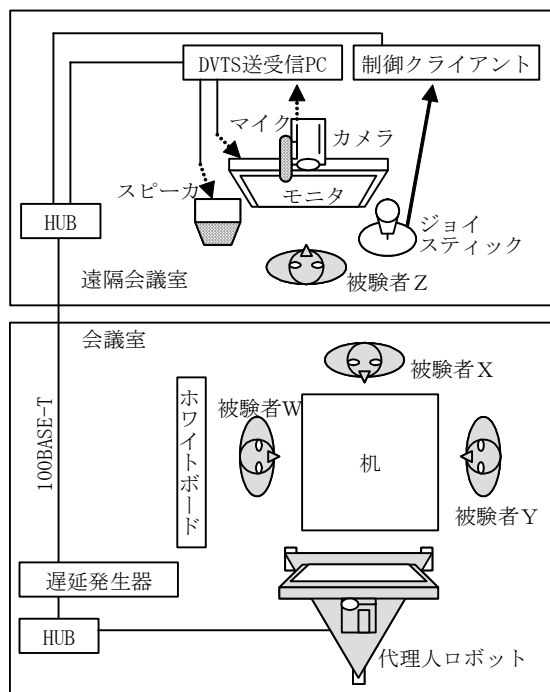


図6. 評価実験構成

#### 3.2 実験手順

今回被験者に与えたタスクは、会議室内にいる3名の被験者と代理人ロボットを介して別室から参加する1名の被験者の計4名に、「日本でのサマータイム制導入の是非について」というテーマについて議論してもらい、会議室内の被験者Wの近くに置いてあるホワイトボード上に要点をまとめてもらう作業とした。議論開始からまとめるまでの時間は30分とした。議論を進行させ会議室内のホワイトボード上

に要点をまとめさせる作業を担当する司会役は遠隔会議室から参加する被験者Zとし、実際にホワイトボード上に書き込む作業を行う書記役は会議室内にいる被験者Wとした。

各グループの被験者4名に対して、まず抽選で被験者W, X, Y, Zの役割を決めてもらい、各自所定の位置に着席後、代理人ロボットの遠隔操作方法の説明を行った。次に、取り組んでもらうタスクをテーマ名のみ伏せた形で説明した後、3分間システムを自由に利用してもらった。その後、会議室内において、テーマ内容を提示し議論を開始してもらった。

### 3.3 評価実験結果

話者交替が行われるトリガーおよびそのトリガーから実際に話者交替が完了するまでの被験者もしくは代理人ロボットの行動回数を分析することで評価を行った。

話者交替時の行動としては、遠隔会議室から参加している被験者Zの遠隔操作による代理人ロボットの旋回動作（以下、「旋回動作」と記述）、これから会話を開始しようとする2名の被験者間のアイコンタクト（以下、「アイコ

ンタクト」と記述）、呼びかけやジェスチャ等アイコンタクト以外で話者交替のトリガーとなり得る行動（以下、「発声ほか」と記述）の大きく3つの行動に分類した。

一方、話者交替時の話者の位置関係の変化については、以下の4種類に分類して行動分析を行った。代理人ロボットの行動が関与しないものとして、会議室内にいる3名の被験者間での話者交替（以下、「会議室内→会議室内」と記述）と、代理人ロボットと会議室内のある被験者の会話から会議室内の会話への話者交替（以下、「遠隔地間→会議室内」と記述）がある。代理人ロボット又はその遠隔操作者である被験者Zの行動が関与するものとして、会議室内の2名の会話から代理人ロボットと会議室内のある被験者との会話への話者交替（以下、「会議室内→遠隔地間」と記述）と、代理人ロボットと会話する会議室内の被験者の交替（以下、「遠隔地間→遠隔地間」と記述）がある。

表1は、各グループのタスク実行中での話者交替時の行動回数とその統計結果を、話者交替時の行動と話者の位置関係の変化で分類し、5つのグループ全体でまとめた行動分析結果で

表1. 被験者の話者交代時の行動回数

(5グループの総計)		話者交替時の話者の位置関係の変化			
		会議室内 ↓ 会議室内	遠隔地間 ↓ 会議室内	会議室内 ↓ 遠隔地間	遠隔地間 ↓ 遠隔地間
話者交替時の行動パターン	旋回動作 → アイコンタクト → 発声ほか	-	-	10	13
	旋回動作 → 発声ほか → アイコンタクト	-	-	44	100
	旋回動作 → アイコンタクト	-	-	0	11
	旋回動作 → 発声ほか	-	-	0	4
	旋回動作	-	-	0	0
	アイコンタクト → 旋回動作 → 発声ほか	-	-	10	1
	アイコンタクト → 発声ほか → 旋回動作	-	-	3	1
	アイコンタクト → 旋回動作	-	-	0	2
	アイコンタクト → 発声ほか	51	13	4	12
	アイコンタクト	0	0	0	0
	発声ほか → 旋回動作 → アイコンタクト	-	-	36	86
	発声ほか → アイコンタクト → 旋回動作	-	-	4	24
	発声ほか → 旋回動作	-	-	4	0
	発声ほか → アイコンタクト	88	41	10	7
発声ほか	8	3	0	0	
統計値	話者交替数：T	147	57	125	261
	話者交替時のアイコンタクト実施数：I	139	54	121	257
	話者交替時の旋回動作実施数：R	-	-	111	242
	旋回動作→アイコンタクト実施数：α	-	-	90	210
	話者交替時のアイコンタクト実施率：I/T	0.95	0.95	0.97	0.98
	話者交替時の旋回動作実施率：R/T	-	-	0.89	0.93
旋回動作後のアイコンタクト実施率：α/I	-	-	0.74	0.82	

ある。本評価実験では、話者交替時の行動を、トリガーとなる最初の行動から実際に話者交替が完了するまでの行動内容として同表にある15パターンの行動に分類した。そして、評価実験録画記録を再生・解析することでそれぞれの行動回数をカウントした。この統計結果から、まず話者交替時の話者の位置関係の変化の違いにより話者交替数が大きく異なっていることがわかる。これは司会を務める遠隔会議室にいる被験者Zと会議室内にいる被験者とのやりとりが多くなっているためと考えられる。このため、本評価実験での評価は行動回数の絶対値ではなく比率により行うこととする。話者交替時のアイコンタクト実施率については、会議室内の被験者間での話者交替時で95%、代理人ロボットが関与する遠隔地間での話者交替時で97~98%という極めて高い実施率を示している。代理人ロボットの有無に関わらず話者交替におけるアイコンタクトの重要性が非常に高いことが確認できる。また、遠隔会議での代理人ロボット利用と話者交替の実現との関係には、話者交替時の代理人ロボット旋回動作実施率が約90%を示していることから、代理人ロボットの利用は話者交替の実現の面で有効に機能していると考ええる。更に、遠隔会議での代理人ロボット利用とアイコンタクト実現との関係についても、旋回動作後のアイコンタクト実施率が80%近い値を示していることから、代理人ロボットの利用はアイコンタクト実現の面でも有効に機能していると考ええる。

#### 4. まとめ

本論文では、遠隔会議において対面会議と同様の活発な議論を展開するために、遠隔地間とのアイコンタクトによるスムーズな話者交替の実現について議論した。まず、他の参加者と実際に並んで会議に参加することを目的として、ほぼ実物大の遠隔参加者の実写顔映像が表示されるモニターや目となる遠隔カメラ等が一体となっている頭部が遠隔操作により実際に上下左右に振り向く動作が可能な代理人ロボットと遠隔操作システムを試作した。そして、同システムが遠隔地間の参加者同士のアイコンタクトおよび話者交替の実現の面で有効に機能しているかどうかの評価を行った。行動分析評価の結果、話者交替におけるアイコンタクトの重要性が非常に高いことが確認されると

共に、代理人ロボットの旋回動作がアイコンタクトおよび話者交替と強い相関性を持っていることが分かった。このため、代理人ロボットの旋回動作が遠隔地間の参加者同士のアイコンタクトおよび話者交替の実現の面で有効に機能しているという考察結果を得た。

なお、本研究は文部科学省の研究開発委託事業である「ITプログラム」[9]により行われた。

#### 参考文献

- [1] ITU-T Recommendation H. 323: Packet Based Multimedia Communication Systems (2003).
- [2] 伊藤京子ほか: キャラクターエージェントをアバタとナビゲータとして利用したネットワークコミュニティの実験, 情報処理学会論文誌, Vol. 44, No. 7, pp. 1812-1827.
- [3] 小山慎哉ほか: 実空間を対象とした遠隔コミュニケーション支援システムの設計要件提案と開発, 情報処理学会論文誌, Vol. 45, No. 1, pp. 178-187.
- [4] 郷健太郎ほか: ズーム情報を利用した適応型遠隔カメラ制御法, 情報処理学会論文誌, Vol. 43, No. 2, pp. 585-592.
- [5] Reeves, B.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places, The University of Chicago Press (1996).
- [6] 海保博之: プロトコル分析入門, 新曜社
- [7] Digital Video Transport System.  
<http://www.sfc.wide.ad.jp/DVTS/>
- [8] 町澤朗彦ほか: GbE によるネイティブ IPv6 広域実験網の設計と構築, マルチメディア, 分散, 協調とモバイルシンポジウム, pp. 85-90.
- [9] VizGrid Project. <http://www.vizgrid.org/>