

## キネティックテレプレゼンス *MMSpace*

### 小規模グループ対グループ遠隔会話場の構成と非言語行動解析

大塚 和弘<sup>†</sup>

<sup>†</sup> 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所

〒 243-0198 神奈川県厚木市森の里若宮 3-1

E-mail: †otsuka.kazuhiro@lab.ntt.co.jp

**あらまし** 複数人-対-複数人の遠隔会話を対象とした「キネティック・テレプレゼンスシステム *MMSpace*」を提案する。*MMSpace* は、対面会話の場を模した双方向・対称形の遠隔会話場の構成、及び、そこでの人物行動解析を介したシステム評価とコミュニケーション理論の深化を射程とした研究プラットフォームである。複数のキネティック・アバタから構成され、各アバタは、遠隔参加者の顔画像を投影する平面パネル、及び、その位置・姿勢を動的に変化させるためのアクチュエータ群からなり、人物の頭部運動と連動してパネルの位置・姿勢が自動的に制御される。これにより頷きや首振り、前傾・後傾など様々な身体動作が表現可能である。この機械的運動の重畳により、遠隔地間において会話参加者が交わす非言語情報が補強され、対面状況により近い会話場が再現されるものと期待する。また、視線一致カメラ配置と動的カメラ切替によって、任意の参加者間において擬似的な相互凝視（アイコンタクト）を可能とする仕組みも特徴である。2人-対-2人の遠隔会話実験において、機械的運動の有り無しの二条件を比較し、機械的運動がもたらす効果を検証した結果、会話参加者の印象評定から、視線や相互凝視、ジェスチャなどの非言語行動の知覚、親近感、相互理解などにおいて、有意な正の効果が確認された。また、客観的な非言語行動分析の結果、平均注視時間の増加、視線遷移時間の減少、頭部ジェスチャの表出量・共起率の増加などの効果が確認された。これらの観点において、提案システムにて導入された機械的運動によって、非言語行動が対面状況により近くなり、遠隔地間の非言語コミュニケーションが促進されることが明らかになった。

**キーワード** 遠隔コミュニケーション、遠隔会議、テレプレゼンス、非言語行動、視線、相互凝視、頭部運動

## *MMSpace*: Kinetic Telepresence System for Enhancing Nonverbal Communications in Small Group-to-Group Settings

Kazuhiro OTSUKA<sup>†</sup>

<sup>†</sup> NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation

3-1 Morinosato-Wakamiya, Atsugi-shi, 243-0198 Japan

E-mail: †otsuka.kazuhiro@lab.ntt.co.jp

### 1. はじめに

今日、電子メール、ソーシャルネット、遠隔会議、テレプレゼンスシステムなどの電子情報通信技術は目覚ましい発展・普及を遂げており、生活や仕事など幅広い場面で活用されている。しかしその一方、実際に人と会って話をする「対面会話」は、依然、基本的かつ重要なコミュニケーションの形態と考えられており、重要な場面であればあるほど、貴重な時間や費用、労力を費やして遠方まで足を運ぶことも珍しくない。空間

的に離れた場所にいる人々があたかも互いに対面しているかのように会話を行うことができるコミュニケーション技術の開発は積年の課題である。この課題への取り組みは、人の移動の削減による環境負荷の低減、化石燃料消費抑制による持続的発展が可能な社会の構築、職住一致・近接によるQoL向上、年齢・性別・地域・国籍の壁を越えた多様な人材の社会参加促進など、多くの社会的課題の解決へつながるものと目される。近年、遠方にいる人物の代理として、物理的な実体を持つロボット端末（テレプレゼンスロボットと呼ばれる）の活用に期待

が寄せられている [1], [2]. 通常のディスプレイに表示される人物画像を超えた, 物理的実体をもたらす顕著な社会的存在感 (social telepresence) が利点とされるが, 現状では遠隔操作の負担や通信遅延, 運動性能の限界, 環境の非対称性など様々な課題があり, その設計論や評価法などより一層の研究進展が期待されている. その営みの一環として, 著者らは, 人のコミュニケーションにおける非言語行動の機能の側面から遠隔ロボットによる機械的な運動表出の方式を検討し, システムとして実装, 更にその運動表出が会話参加者の非言語行動に与える影響を解析することでシステムの洗練やコミュニケーション理論の深化へつなげるという一連の構成論的研究パラダイムの実践を進めている. 本稿では, そのプラットフォームとして開発されたキネティック・テレプレゼンスシステム *MMSpace* を紹介する [3], [4] (注1) (注2) (注3).

本研究では, キネティック・ディスプレイ (または, キネティック・アバタ) (注4) と呼ぶテレプレゼンス・ロボット (端末) に注目する [7]~[10]. キネティック・ディスプレイとは, 平面ディスプレイパネルを遠隔人物の顔として「見立て」, そのパネルの位置や姿勢を遠隔人物の動作 (や操作) に応じて動的に変化させることができる表示装置を指す. ディスプレイパネルに表示される顔画像に加えて, パネル自身の動きとして身体運動 (主に頭部の) を物理的に表出できることが特徴である. この運動チャンネルの追加によって, 通信により失われる人の非言語情報 [11] が補強・補完され, 複数人の遠隔会話がより対面状況に近くなるものと期待される.

ここで機械的に表出する運動として, 特に頭部運動に着目する理由としては, 頭部運動が非言語情報の表出・交換において重要な役割を担っていることによる. その代表例としては, 頷きや首振りなどの頭部ジェスチャや, 視覚的注意 (視線行動) があげられる. 頭部ジェスチャは, 話者によるリズムとり, 問い掛け, 強調など, また, 聞き手による相づち, 承諾, 同意/非同意, 態度, 興味や理解の度合いなど, 様々な非言語手がかりの表出・交換に用いられている [12]~[14]. また, 頭部の動きやその向きは, 視線行動や視線方向の手がかりとしても有用とされる [15], [16]. 視線行動は, コミュニケーションにおいて, 他者の反応や行動をモニタリングする機能の他, 態度や意図, 親密さの表出にも用いられている [17]~[19]. また, 話し掛けや傾聴のサインとしてグループ会話の組織化に寄与し [20], 発話交替の調整にも利用されている [18], [19].

従来, キネティック・ディスプレイの効果として, 他者の視線行動の知覚促進 [7], [9], [10], [21], [22], 感情や態度, 姿勢の伝達 [8], [10], [22], [23], 存在感の増強 [7], [22] などがあげられている. しかし一方で, 機械的運動のぎこちなさ, モーターの騒

音, 意図しない動きによる煩わしさ, 運動の解釈の曖昧さ, 遠隔ユーザの認知的負荷, 非対称環境による参加者間・地点間の不平等性など解決すべき問題が多く残されている [9], [22]. これまでの研究では, 主に一対一, または一対多の状況を対象とし, 典型的には1台のキネティック・ディスプレイが主となる会議室に設置され, 遠隔参加者は, 通常の端末 (デスクトップ型やタブレット型など) を用いて会議室の様子を観察しつつ, 端末の操作を通じて会話に参加する一方, 会議室にいる人物は通常の対面会話と同様のスタイルにて参加ができ, さらにキネティック・ディスプレイの表出情報も利用できるという, いわば「非対称形」のコミュニケーションに焦点を当てていた.

本研究では, 複数人物による対面会話の「場」を遠隔地間において再現することを究極の目標として据え, そのため通常の対面会話のようにユーザによる機器の操作を不要とし, また, 地点毎の環境や装置の差異を廃して, 全参加者が同一の条件で参加ができるような「対称形」の遠隔会話場の構築を試みた. この会話場の「再現」により, 対面状況の複数人会話で生じる非言語行動と遠隔状況とのそれとが直接的に対比でき, 新規に導入された仕組み (つまり本稿では機械的運動) による効果を浮き彫りにして評価することを可能とした. このような方法論の提案と実践も本研究の貢献と言える. さらに高性能な機構 (高精度・低遅延・低騒音) を導入することで, 従来システムの問題 (動きのぎこちなさや遅延・騒音など) を解決し, 非言語コミュニケーションの促進に資する水準の機械的運動表出を実現した点も貢献の一つとして数えられる.

また, 著者らの従来研究により, 複数台のキネティック・ディスプレイによって会話の場が近似的に再現されることが, 4人対話のオフライン再生実験により示唆されているが [21], [22], 本稿の *MMSpace* は, これを実時間かつ双方向の複数人会話システムへ拡張し, 会話実験により非言語行動を綿密に調査した点に新規性がある. 他の特徴としては, *MMSpace* にて導入したキネティック・ディスプレイは, 回転2自由度, 並進2自由度の計4自由度のアクチュエータ群により, 頷きや首振りに加えて, 前傾・後傾など様々な身体動作が表出できることがあげられる. また, 視線一致カメラ配置と動的カメラ切替により, 任意の参加者間での (擬似的な) 相互凝視 (アイコンタクト) を可能とする仕組みも新たな特徴といえる.

2人-対2人の遠隔会話実験を行い, 機械的運動の有り無しの二条件を比較し, 機械的運動をもたらす効果を検証した結果, 主観評価により, 他者の視線方向, アイコンタクト, 他者の反応, ジェスチャ, 顔表情, 姿勢など非言語行動の知覚, 相互理解, 親近感・存在感などにおいて, 有意な正の効果が確認された. また, 参加者の非言語行動について分析を行った結果, 機械的運動がある場合, 視線行動に関しては, 注視一回当たりの平均時間が増え, また, 視線遷移に要した合計の時間が減少することが分かった. また, 頭部ジェスチャの生成時間長や遠隔地の人物との共起率についても増加した. これらの観点において, *MMSpace* にて導入した機械的運動によって, 非言語行動が対面状況により近くなり, 遠隔地間の非言語コミュニケーションが促進されることが明らかになった.

(注1): 本稿は, 既出の英語論文 [3], [4] からの抜粋・加筆であり, 日本語の学術資料としては初出である.

(注2): YouTube にてデモムービーが視聴可能. 英語版 [5], 日本語版 [6].

(注3): 本システム, デモムービーとも著者らによる内製である.

(注4): 一般に「ロボット」から連想される, (擬人的な) 身体や顔の形態・様相ではなく, 純粋な物理的運動の表出とその対人コミュニケーションにおける機能に焦点を当てるといふ意図から, 本研究では, “kinetic” (運動学的) という用語をあえて選択して使用している.

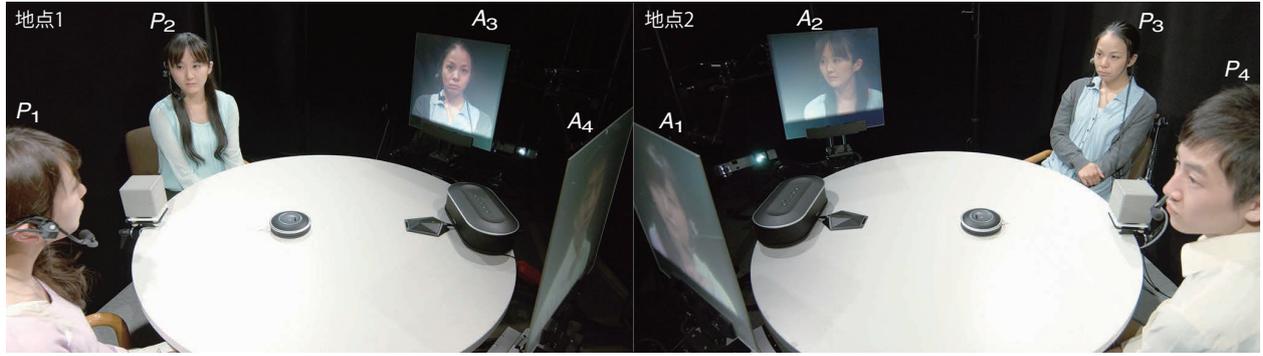


図1 2人-対-2人の対称形グループ会話を対象としたMMSpaceの全体像。会話参加者( $P_1, P_2, P_3, P_4$ )とそのアバタ( $A_1, A_2, A_3, A_4$ )。文献[4]より引用。



図2 対面会話の例。文献[4]より引用。

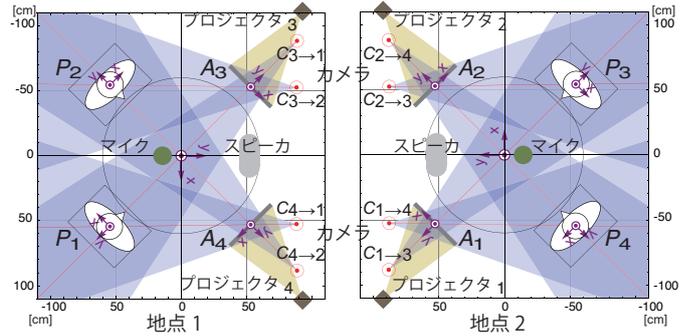


図3 MMSpaceの空間配置。会話参加者( $P_1, P_2, P_3, P_4$ )とそのアバタ( $A_1, A_2, A_3, A_4$ )の配置を示す。 $C_{i \rightarrow j}$ は、アバタ $A_i$ の視点から参加者 $P_j$ を捉えるカメラを示す。文献[4]より引用。

## 2. キネティックテレプレゼンス: MMSpace

MMSpaceは、小規模な対称形のグループ対グループの遠隔会話を対象としたテレプレゼンスシステムである。図1には、2人-対-2人のグループ会話を対象としたプロトタイプを示す。これは図2のような円卓を囲んだ対面会話の場を二地点間において再現することを意図している。図3には、空間配置図を示す。地点1の参加者 $P_1, P_2$ が、地点2においては、それぞれのアバタ $A_1, A_2$ として同じ相対位置に配置されている。地点2の参加者 $P_3, P_4$ についても同様な配置がなされる。

図4には、キネティックアバタの外観を示す。キネティックアバタは、平面スクリーンパネルとプロジェクタ、及び、パネルを支持するアクチュエータ群からなる。スクリーンパネルは、光線を拡散させる成分が配合されたアクリル板であり、高い透過性と広い視野角を有する。この透過性は、後述の視線一致カメラ配置のための必要条件となる<sup>(注5)</sup>。背後のプロジェクタによりパネル上に等身大の人物像(肩から上、背景抜き)が投影される。パネルの位置・姿勢に応じて動的に射影関数のパラメータを更新することで常に歪みの無い画像が投影される。アクチュエータ群は、図5下部よりXYステージ、回転ステージ、ゴニオステージの合計4軸で構成され、それぞれ左右・前後の並進、鉛直軸周りの回転(首振り方向)、傾き方向の回転の各運動成分を表出するために用いられる。全軸ダイレクトドライブモータを用いることで、静音性が実現されている(円卓中

央にて34.3 dBA [3])。

図4のパネルに表示された人物のように、参加者はヘッドセットマイク・センサを装着して会話に参加する。側頭部に配置された磁気式モーションキャプチャ(Fastrak)用のセンサにより頭部位置・姿勢が計測される。また、ヘッドセットマイクで集音される音声は後の行動解析に用いられる。なお、地点間の音声コミュニケーションには電話会議端末を用いた(図3中にてそのマイク・スピーカの位置を図示)。センサにより計測された各人の頭部位置・姿勢に応じて、別地点のアバタのパネルの位置・姿勢が実時間で制御される。なお、各軸のアクチュエータの可動範囲に対応すべく、シグモイド関数を用いた位置・姿勢角度の変換が行われ、制御目標値が設定される。

図3に示すように各アバタのパネルの背後にはカメラが設置される。カメラ $C_{i \rightarrow j}$ は、アバタ $A_i$ の目線から参加者 $P_j$ の顔を捉えるように、その光軸がパネル上の顔と参加者の顔(どちらも初期位置にておよそ両目の中心)を通るよう配置される。参加者 $P_j$ の顔画像は同室の複数のアバタのカメラより捉えられる。各時刻において、これらのアバタのカメラのうち一台が選択され、参加者 $P_j$ のアバタ $A_j$ のパネルに顔画像が投影される。ここでカメラ選択の方式として、注視対象(アバタ)に対して常に正視顔を向けるようにカメラを選択する方式を採用している。例えば、人物 $P_3$ がアバタ $A_2$ に顔と視線を向けるとき、そのアバタのカメラ $C_{2 \rightarrow 3}$ が選択され、地点1のアバタ $A_3$ のパネルに、人物 $P_2$ を見つめる人物 $P_3$ の顔画像が投影さ

(注5)：透過性パネルによってもたらされる、空間中に浮き上がるような人物像の表示によって、遠隔人物の存在感や共在感の向上に寄与することを期待しているが、このような表示法の直接的な効果は未評価である。

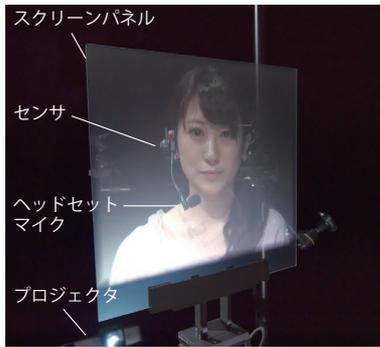


図4 キネティックアバタの外観. 文献[4]より引用.



図5 キネティックアバタの機構. 文献[4]より引用.

れる. 同時にこのアバタ  $A_3$  のパネルは, 人物  $P_2$  の方向に向けて回転する. このように視線を向けるという行為が, パネルの向きの変化, 及び, パネル上の画像の変化として表現される. このとき人物  $P_2$  がアバタ  $A_3$  に視線を向け返すと, 同様にアバタ  $A_3$  のカメラ  $C_{3 \rightarrow 2}$  が選択され, 地点2のアバタ  $A_2$  のパネルに人物  $P_3$  を正視する人物  $P_2$  の顔画像が表示される. このようにして別地点の二者間にて擬似的にアイコンタクトが成立することが期待される. ここでカメラ切替に用いられる人物の視線方向は, 頭部の向きで代用され, 頭部回転角の閾値処理により推定される.

なお, 本稿で提案する *MMSpace* は, 機械的運動の影響調査を主目的としたため, 他の要素については極力単純化を図った. 黒背景の利用により人物像切り出しのための画像処理を簡略化した. また, 通信遅延最小化のため地点1と地点2は隣り合う防音室とし, 各モダリティの信号伝達にはそれぞれ専用の伝送路を使用した. なお, 計測のサンプリングレートは, 画像, 音声, 頭部運動それぞれ各センサ当たり, 30Hz, 48kHz, 60Hzであった. 単一の基準信号源 (10MHz) より分周・通倍・分配されるクロックにより各センサを駆動することで, 各信号間・モダリティ間にて同期がとれた計測・収録を行った. 映像の遅延量 (カメラのシャッター開口からプロジェクタによる投影開始まで) は約 150msec であった. 電話会議端末本体の遅延量は約 35msec であったが, 映像の遅延量と一致させるため意図的にさらに 115msec の遅延を加えた. また, 機械的運動の遅延量は, 57msec~93msec と報告されている [3].

### 3. 実験

#### 3.1 実験条件

16名の実験参加者を排他的に4つのグループに分け, グループ毎に複数回のディスカッション (合意形成課題) を実施した. 参加者は20台~40台の女性であった (全員初対面). 各グループにおいて, 対面会話 (図2の配置) 2セッション, 及び, *MMSpace* にて4セッションのディスカッションを実施した. *MMSpace* を用いたセッションには, 機械的運動のあり・なしにより二つの条件を設定し, 動きあり条件 ( $M$  と記す) にて2セッション, 及び, 静止条件 ( $S$  と記す) にて2セッションを行った.  $M$  と  $S$  の順番は交互かつグループ毎に先行する条件を入れ替えた. 全ての  $M$ ,  $S$  セッションを挟む形で前後

に対面セッション ( $F2F$  と記す) を行った. ここで静止条件  $S$  は, パネルを動かさない他は,  $M$  と同条件とした. セッションの長さには一応の制限時間を設けたが, 強制終了はせず, 実験者は経過を観察しつつベル音により早期終結を促した ( $M$ ,  $S$  の平均会話長は 9.6 分).

#### 3.2 主観評価

$S$ ,  $M$  条件の各セッション直後に参加者の主観的な印象を質問紙法により調査した. 調査対象として, 会話の自然さ・楽しさなどの体験の質, 発話交替, 相互理解, 視線知覚, 非言語行動知覚などの容易さ, 存在感, 行動伝染, 映像音声品質, パネルの動きの自然さ ( $M$  のみ) などから構成される設問 ( $S$ : 33項目,  $M$ : 38項目) を用意し, 各々, 7段階のリカート尺度 (1: 全くそう思わない~7: とてもそう思う) にて回答させた. その結果に基づき, 参加者各人において条件毎の印象の差を分析するため (被験者内実験計画), 各設問に対して, 人ごとに同条件2セッションの平均スコアを計算し, それらについて対応のある  $t$  検定を行い統計的有意差 (両側 5%水準) を調べた.

図6には, 設問のうち6項目を抜粋した結果を示す. 図6(a)~(f)には, それぞれ視線方向知覚 (「別室の人が誰を見ているのか容易に判別がついた」), アイコンタクト知覚 (「別室にいる相手と目が合ったと感じることがあった」), ジェスチャ知覚 (「別室にいる人の仕草やジェスチャがよく分かった」), 他者理解 (「別室にいる人の言いたいことが容易に理解できた」), 被理解 (「自分の言いたいことが, 別室の相手にもよく伝わった, 理解されたと感じた」), 親近感 (「別室の会話相手に親近感を感じた」) についての結果を示す. 図6(a),(b)より,  $M$  条件においては, 相手の視線方向がより分かりやすく, また, よりアイコンタクトを感じることが示唆された. また, 図6(c)より,  $M$  条件において, 他者のジェスチャがより分かりやすく, 他項目の結果 (文献[4]参照) とあわせて, 視線やジェスチャ, 顔表情など含めた非言語行動の知覚が促進されたことが示唆された.

また, 図6(d)より, 他者理解については条件間で差はなかったが, 平均スコア ( $M$  条件にて 5.75) より他者の発言は比較的容易に理解できたことが読み取れる. 一方, 図6(e)より,  $M$  条件において, 自身が他者によって理解されたとの印象がより強かった. これら二項目をあわせると,  $M$  条件において, より地点間での相互理解が促進されたとの見方が可能である. さ

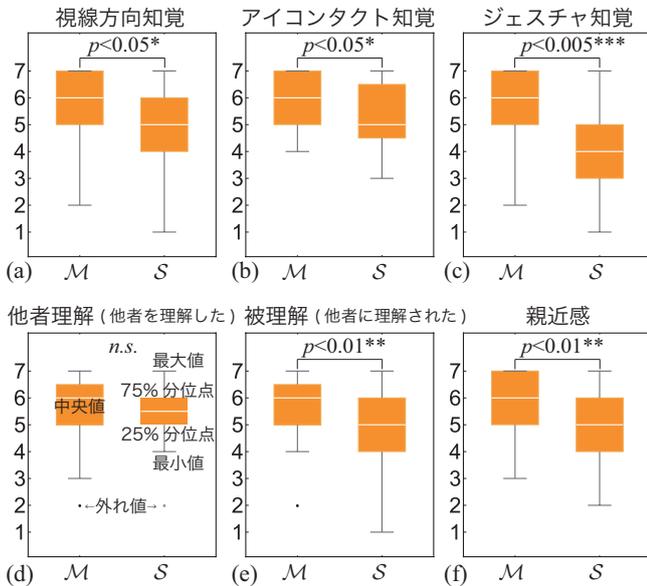


図 6 Box-Whiskert 図による主観的印象の条件間比較. 縦軸は 7 段階リカート尺度 (1:全くそう思わない~7:とてもそう思う).

らに, 図 6(f) より,  $M$  条件において遠隔の人物に対してより強い親近感を抱いたことが分かった. これら結果より, 機械的運動表出は, 視線やジェスチャなど非言語情報の伝達に留まらず, 会話参加者間のラポールの形成など, より豊かなコミュニケーションの実現に寄与しうることが示唆された.

### 3.3 非言語行動解析

$MMSpace$  ( $M$  条件) にて導入した機械的運動が, 会話参加者の非言語行動に及ぼす影響を客観的に明らかにするため, 収録された映像・音声・運動などのデータを用いて, 発話パターン, ターン, 話し掛け/応答, 視線, 頭部ジェスチャについて分析を行った. 本稿では, その内, 最も効果が顕著であった視線, 及び, 頭部ジェスチャに関する結果を紹介する. 分析に先立ち, 1 名の外部観察者によって, 視線行動, 及び, 頭部ジェスチャに関するラベリング作業が行われた. 視線方向のカテゴリとして, 他の参加者 (人, 及び, アバタ) の顔, 視線遷移中, 及び, その他を用いた. また, 頭部ジェスチャに関しては, 頷き, 首振り, 傾げ, その他 (ジェスチャ無しを含む) のカテゴリに分類したが, 後述の分析では, これら三種のジェスチャを統合してジェスチャの有無のみを分析対象とした. これらラベルの時間単位は, 毎ビデオフレーム毎 (30 fps) とした.

図 7 には, 平均注視時間, 視線遷移時間長, 頭部ジェスチャ表出量, 及び, 頭部ジェスチャ共起率について, それぞれ三つの条件を比較した結果を示す. 各項目について, 各参加者の同一条件 2 セッションの平均値を計算し, 条件間での差異の有無を調査するため, 反復測定 1 元配置分散分析を行った. そこで有意差があった場合, 各条件間の事後比較 (Holm 法による  $p$  値補正あり) を行った. まず, 図 7(a) より,  $S$  における平均注視時間は,  $M$ ,  $F2F$  と比較して, 統計的に有意に短いことが分かった. 一方,  $M$  と  $F2F$  の間には有意差が見られなかった. 次に, 図 7(b) には, 視線遷移時間の合計が全体の会話時間に占める割合を比較した結果を示す. この図 7(b) より,  $S$  に

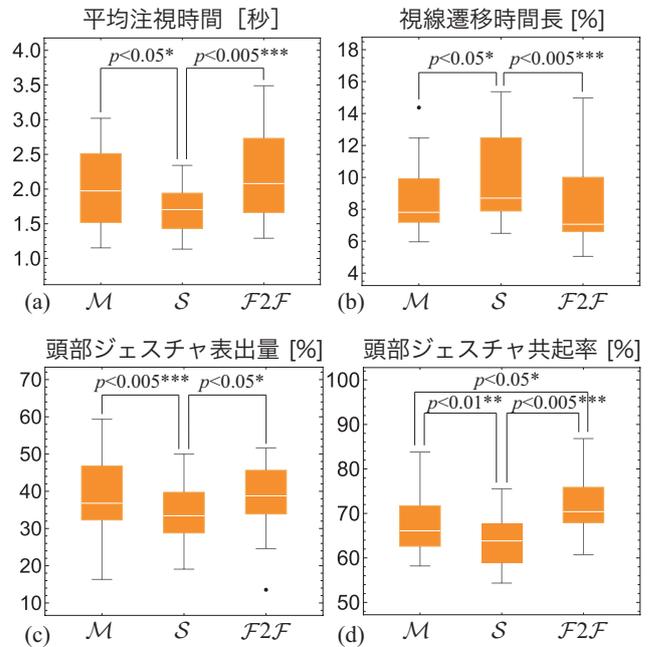


図 7 Box-Whiskert 図による非言語行動の条件間比較.

おける視線遷移時間長は,  $M$ ,  $F2F$  と比較して, 統計的に有意に長いことが分かった. 一方,  $M$  と  $F2F$  の間には有意差が見られなかった. また, 各人の平均相互凝視時間についても,  $S$  が他条件より有意に短いとの結果が得られている [4].

以上の結果より,  $S$  においては, 参加者は一つの対象について長い時間, 注意を注がず, また, 注視対象を探索するためにより多くの時間を費やしたことが読み取れる. 一方, これら尺度において,  $M$  における視線行動は,  $F2F$  のそれと類似していると言え, より対面状況に近い自然な視線行動が促されたことが示唆される. その原因を考察すると, ヒトの周辺視野における運動知覚の感受性 [24] に起因して,  $M$  条件においては, 直視している人物の周辺を含めた「会話の場」を直感的に把握でき, それにより「次に誰を見るべきか」が容易かつ的確に判断できるようになり, 迅速な視線遷移に結びついたものと推察される. これと表裏一体の関係として, 注視時間の増加は, 他者に対する注意の深まりを示唆し, 前述の印象評価で述べた非言語行動の知覚促進につながったものとも考えられる.

次に頭部ジェスチャに関する結果として, 図 7(c) には, 頭部ジェスチャの表出量 (会話の時間長に対してジェスチャ表出の時間割合) を示す. 図 7(c) に示すように,  $M$ ,  $F2F$  と比較して,  $S$  において頭部ジェスチャの表出量が有意に少ないことが分かった. 一方,  $M$  と  $F2F$  の間には有意差が見られなかった. また, 図 7(d) には, 各人の頭部ジェスチャが遠隔人物 (二人の内いずれか一人以上) と共起した割合を示す. なお,  $F2F$  における「遠隔」人物は,  $M$ ,  $S$  で遠隔に位置する人物を指す (例.  $P_1$  に対する  $P_3, P_4$ ). 図 7(d) に示すように,  $S$  では,  $M$ ,  $F2F$  と比較して, 頭部ジェスチャの共起率が有意に低いことが分かった. 一方,  $M$  より  $F2F$  の方が有意に共起率が高いことも判明した. 人物間において身体動作が共起・同期する度合いは, コミュニケーションの活発さなどを推し量る指標とし

て知られており [25], これら頭部ジェスチャ表出量や共起の観点から, パネルの機械的運動には, 頭部ジェスチャを介した非言語コミュニケーションを促進する効果があることが示唆される. 一方,  $F2F$  と  $M$  との共起率の差をみると, 依然として, 非言語情報の交換の場として対面状況の優位性が伺える.

#### 4. むすびと今後の課題

複数人-対-複数人の遠隔会話を対象とした「キネティック・テレプレゼンスシステム *MMSpace*」を紹介した. *MMSpace* は, 人物の頭部運動を遠隔アバタの機械的な頭部運動として, 顔画像とともに提示すること, 及び, 対面会話場を模した双方向・対称形の遠隔会話場を構成することに特徴がある. 会話中の視線や頭部ジェスチャなどの非言語行動を定量的に解析することで, 遠隔コミュニケーションの遂行・促進に対する提案方式の効果を明らかにした. 今後の課題としては以下があげられる.

*MMSpace* の特徴の一つとしてアイコンタクトの成立を謳っている. また, 主観評価によりアイコンタクト知覚の促進が確認されている. しかし, モナリザ効果 [26] によって過剰にアイコンタクト知覚が催されている可能性に留意が必要である. モナリザ効果の改善にはパネル回転角の最適化や, (より顔形状に近い) 立体パネルの利用などが要検討である. また, 実験において, 頭部運動よりも眼球運動を主とした視線遷移が多くみられ, 頭部方向のみに依存したカメラ切替方式には限界があることが分かった. 今後, 眼球運動計測などの導入が望まれる.

また, 提案法の効果は, 会話参加者の性別や性格, 母語などにも依存するものと思われる. 一般に女性の方が男性よりも他者の非言語行動に敏感であると言われており [17], [19], その点からは, 女性に対してより効果が大きいことが予想される. 一方, *MMSpace* を非言語情報の増幅器として捉えれば, 個々人の知覚能力を補強し, 能力差を埋める手段としても活路を見いだせる. また, 対人恐怖症や視線恐怖症などを有する参加者にとっては, 他者の存在感や非言語情報の緩衝材, あるいは適応のための訓練装置としても *MMSpace* を活用できる可能性がある.

さらに *MMSpace* をより多様な状況 (人数や地点数) に展開することも今後の課題である. また, 今回は触れていない通信遅延の影響についても調査が待たれる. さらに *MMSpace* 自体の発展以外にも, 一連の非言語行動解析を通じたコミュニケーション理論研究への貢献や, そこで得られた知見を更なる新方式の開発へとつなげていくことも期待される.

#### 文 献

- [1] E. Guizzo, “When my avatar went to work,” *IEEE Spectrum*, vol.47, pp.26–50, 2010.
- [2] J.P. Conti, “Technology telepresence - I see live people,” *Eng Manag J*, vol.17, no.3, pp.12–15, 2007.
- [3] K. Otsuka, “MMSpace: Kinetically-augmented telepresence for small group-to-group conversations,” *Proc. IEEE Virtual Reality 2016*, pp.19–28, 2016. [Link to PDF].
- [4] K. Otsuka, “Behavioral analysis of kinetic telepresence for small symmetric group-to-group meetings,” *IEEE Transactions on Multimedia*, vol.20, no.6, pp.1432–1447, June 2018. [OpenAccess]. <https://ieeexplore.ieee.org/document/8101001/>
- [5] K. Otsuka, “MMSpace: Multimodal meeting space embodied by kinetic telepresence,” *Proc. CHI’17 EA*, pp.458–458, 2017. <https://youtu.be/MYUmD0qZSHk>
- [6] 大塚和弘, “キネティック・テレプレゼンス: MMSpace デモビデオ 2016(日本語版),” 2016. <https://youtu.be/LoeDXpxR1Uo>
- [7] N. Yankelovich, N. Simpson, J. Kaplan, and J. Provino, “Porta-person: Telepresence for the connected conference room,” *Proc. ACM CHI ’07 EA*, pp.2789–2794, 2007.
- [8] S.O. Adalgeirsson and C. Breazeal, “MeBot: A robotic platform for socially embodied presence,” *Proc. ACM/IEEE HRI ’10*, pp.15–22, 2010.
- [9] D. Sirkin, G. Venolia, J. Tang, G. Robertson, T. Kim, K. Inkpen, M. Sedlins, B. Lee, and M. Sinclair, “Motion and attention in a kinetic videoconferencing proxy,” *Proc. INTERACT’11*, pp.162–180, 2011.
- [10] D. Sirkin and W. Ju, “Consistency in physical and on-screen action improves perceptions of telepresence robots,” *Proc. ACM/IEEE HRI ’12*, pp.57–64, 2012.
- [11] J. Short, E. Williams, and B. Christie, *The social psychology of telecommunications*, John Wiley & Sons, 1976.
- [12] S.K. Maynard, “Interactional functions of a nonverbal sign: Head movement in Japanese dyadic casual conversation,” *J. Pragmatics*, vol.11, pp.589–606, 1987.
- [13] D. Heylen, “Challenges ahead: head movements and other social acts during conversations,” *Proc. J. Symp. Virtual Social Agents*, pp.45–52, 2005.
- [14] S. Dancan, “Some signals and rules for taking speaking turns in conversations,” *Journal of Personality and Social Psychology*, vol.23, no.2, pp.283–292, 1972.
- [15] R. Stiefelwagen, J. Yang, and A. Waibel, “Modeling focus of attention for meeting index based on multiple cues,” *IEEE Trans. Neural Networks*, vol.13, no.4, pp.928–938, 2002.
- [16] K. Otsuka, Y. Takemae, J. Yamato, and H. Murase, “A probabilistic inference of multiparty-conversation structure based on Markov-switching models of gaze patterns, head directions, and utterances,” *Proc. ICMI’05*, pp.191–198, 2005.
- [17] M. Argyle, *Bodily Communication – 2nd ed.*, Routledge, London and New York, 1988.
- [18] A. Kendon, “Some functions of gaze-direction in social interaction,” *Acta Psychologica*, vol.26, pp.22–63, 1967.
- [19] M. Argyle and M. Cook, *Gaze and Mutual Gaze*, Cambridge University Press Cambridge, Eng. ; New York, 1976.
- [20] C. Goodwin, *Conversational Organization: Interaction between Speakers and Hearers*, Academic Press, 1981.
- [21] K. Otsuka, S. Kumano, D. Mikami, M. Matsuda, and J. Yamato, “Reconstructing multiparty conversation field by augmenting human head motions via dynamic displays,” *Proc. ACM CHI ’12 EA*, pp.2243–2248, 2012.
- [22] K. Otsuka, S. Kumano, R. Ishii, M. Zbogor, and J. Yamato, “MM+Space:  $N \times 4$  degree-of-freedom kinetic display for recreating multiparty conversation spaces,” *Proc. ACM ICMI ’13*, pp.389–396, 2013.
- [23] C. Breazeal, A. Wang, and R. Picard, “Experiments with a robotic computer: Body, affect and cognition interactions,” *Proc. ACM/IEEE HRI ’07*, pp.153–160, 2007.
- [24] K.L. Shapiro and T.L. Johnson, “Effects of arousal on attention to central and peripheral visual stimuli,” *Acta Psychologica*, vol.66, no.2, pp.157–172, 1987.
- [25] E. Delaherche, M. Chetouani, A. Mahdhaoui, C. Saint-Georges, S. Viaux, and D. Cohen, “Interpersonal synchrony : A survey of evaluation methods across disciplines,” *IEEE Trans. Affective Computing*, vol.3, no.3, pp.349–365, 2012.
- [26] H. Heiko, B. Evgenia, and K. Akiyoshi, “The Mona Lisa effect: Testing the limits of perceptual robustness vis-à-vis slanted images,” *PSIHLOGIJA*, vol.47, no.3, pp.287–301, 2014.