情報処理学会論文誌　Vol. 52　No. 1　97–108 (Jan. 2011)

遠隔ユーザのジェスチャの可視性を向上させる手法の提案と評価

山下直美[†1]　梶　克彦[†2]　葛岡英明[†3]
平田圭二[†1]　青柳滋巳[†1]

対面協調作業においてテーブル上の作業領域の共有が有効であることが示されている．遠隔においても同様な環境を実現することが望まれており，これまでに遠隔地間でテーブル面を共有する様々なビデオシステムが提案されてきた．しかし，遠隔ジェスチャは二次元映像として表示されるために見逃されやすいうえ，見えていたとしてもジェスチャ映像が実物や人の手などに隠れて見えなくなってしまう状況が頻発する．本論文では，このような不可視性に関する問題を緩和するための手法として，Remote Lag を提案する．Remote Lag とは，一瞬前の遠隔ジェスチャの映像を現在の映像に重ねて表示する手法を指す．本研究では，Remote Lag の効果を検証するために，Remote Lag を表示した条件と表示しない条件における遠隔作業指示を比較する実験を実施した．実験データを分析した結果，ユーザが Remote Lag を用いることによって不可視の状態から遅延なく復帰できる場面が多数発見された．また，Remote Lag を表示させた条件では作業者による質問や指示者の指示発話が減少するなどの効果が認められた．

The Usability Evaluation of Showing the Immediate Past Remote Gestures

Naomi Yamashita,[†1] Katsuhiko Kaji,[†2]
Hideaki Kuzuoka,[†3] Keiji Hirata[†1]
and Shigemi Aoyagi[†1]

Distributed tabletop activities involving real objects are made difficult by some of the invisibility factors introduced in the workspace: Remote gestures often get occluded, transected, and slipped out of one's attention. In this study, we proposed a technique, "remote lag," that aims to mitigate the invisibility problems. The technique provides people with an instant playback of the remote gestures to recover from the invisible or missed context of coordination. To examine the effects of the proposed technique, we studied four-person groups engaging in two mentoring tasks using physical objects with and without remote lags. Our results show that remote lags are effective in alleviating the invisibility problems, resulting in fewer questions/confirmations and redundant instructions during collaboration. Although the technique seems to increase a helper's physical effort, it decreases the helper's overall effort in performing the tasks and alleviates both collaborators' temporal demands.

1. はじめに

対面協調作業において，テーブルを共有することの重要性が指摘されている[18]．これは，テーブルを共有した（少人数の）対面協調作業では，他者の動作と作業領域の関連性を把握したり他者の動作意図を汲むことができ，作業が進めやすいためである[1),22]．

遠隔においても同様な環境を実現することが望まれており，これまでに遠隔地間でテーブル面を共有する様々なビデオシステム（以後遠隔テーブル共有システムと呼ぶ）が提案されてきた[3),14),17),21),24]．近年は 3 人以上のユーザの支援[20] やユーザの柔軟な配置の支援が望まれている[18),25]．以上のシステムの評価結果より，遠隔テーブル共有システムを構築するうえで以下の知見が得られている．

- 遠隔地間のテーブル面（作業領域）を重畳表示させ，その作業領域上に遠隔ユーザのジェスチャ（本論文では「手振り」を意味する）を直接表示させることが有効である[4),6),11]．
- 遠隔ユーザのジェスチャの表示方法に関しては，多種多様なジェスチャ表現を可能とする「手の映像の直接投影」が最も有効とされている[11),20]．
- 遠隔ユーザの上半身映像を表示するビデオ会議システムとテーブル共有を可能とするシステムを組み合わせることによって遠隔ユーザの意図が把握しやすくなり，作業効率が向上する[16),26]．

これらの設計指針をもとに，我々はテーブル環境共有型システム t-Room を構築した（**図 1**）[26]．本システムの新規性は，ユーザの立ち位置を制限せず，テーブル周りの自由な移動を可能にした点にある．我々は，t-Room を用いることで対面協調作業と同程度に円滑

†1 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
NTT Communication Science Laboratories
†2 名古屋大学大学院工学研究科
Nagoya University
†3 筑波大学大学院情報工学研究科
University of Tsukuba

© 2011 Information Processing Society of Japan

図 1　t-Room を用いて遠隔指示が行われている様子
Fig. 1　People at remote sites working on a mentoring task using t-Room.

な遠隔協調作業が実現されることを期待したが，実際には期待に反して非効率な作業場面が散見された．たとえば，t-Room を用いた遠隔指示に関する予備実験では，作業者が遠隔指示者の指示内容を容易に把握できずに質問や確認に関する発話をする場面や，指示者が何度も冗長な指示を繰り返す場面が多く見られた．

本論文では，このような非効率性の一因と考えられる遠隔ジェスチャの不可視性の問題に注目し，この問題を軽減する方法を提案するとともにその評価結果を報告する．ここで不可視性の問題とは，遠隔ユーザの映像が見えない，あるいは見えにくくなることによって，遠隔ジェスチャを認識できなくなる問題を指す．我々は予備実験の観察や関連研究をもとに，ビデオシステムにおける遠隔ジェスチャの認識を妨げる要因を 2 つ特定した．ここであげる 2 要因は t-Room に限らず，従来の遠隔テーブル共有システムに共通に生じる現象であることに注意されたい．

- 注意の離脱：ユーザは作業に集中しているときや他の物事に注意を払っているとき，遠隔ユーザのジェスチャを見逃しやすい傾向にある．たとえば，遠隔指示において，作業者が手元に気をとられているときに遠隔指示者が次の指示に関連するジェスチャを行っても，作業者はそのジェスチャを見逃してしまうことが多い．
- オクルージョン：ユーザの手や物体がテーブル表面を覆うと，テーブル上面に表示された遠隔ジェスチャの映像がそれら実物体の下敷きになって隠れてしまうことがある．この現象（オクルージョンと呼ぶ）は，特に実物体が介在する遠隔協調作業で頻繁に観測され，作業効率を下げてしまう傾向がある．たとえば，遠隔指示において，遠隔指示者のジェスチャ映像の一部が物体の下敷きとなり，作業者が指示内容を正確に把握できないことがある．作業者が物体を持ち上げて指示者のジェスチャ（たとえば指差し位置）を確認しようとすることがあるが，持ち上げた時点ではたいていのジェスチャは終了しているため見えないことが多い．

我々はこれらの問題を軽減するための手法として Remote Lag を提案する[10]．Remote Lag とは，遠隔ユーザの数十ミリ秒〜数秒前のジェスチャ映像を現在の映像に重ねて表示する手法である．Remote Lag によって表示される遅延映像をうまく利用することによって，注意の離脱やオクルージョンによって遠隔ジェスチャを見逃した場合でも不可視状態から回復する効果が期待できる．

Remote Lag の効果を検証するため，指示者 2 名が遠隔にいる作業者 2 名に指示を出し，その際 Remote Lag を表示した条件と表示しない条件を比較する実験を実施した．実験では我々の期待したとおり，ユーザが Remote Lag を用いることによって不可視の状態から遅延なく復帰できる場面が多数発見された．また，実験データを分析した結果，Remote Lag を表示した条件では，それを表示しない条件と比較すると，作業者の指示者に対する質問や確認の発話数が有意に少なくなり，これにともなって指示者の発話数も有意に少なくなることが分かった．

本論文の構成は以下のとおりである．まず関連研究を紹介し，Remote Lag の特徴について説明する．次に，実物体が介在する遠隔指示作業を対象とした Remote Lag の評価実験を説明し，実験中に被験者が Remote Lag を利用していた様子を詳しく紹介する．さらに実験データを分析することによって，Remote Lag が協調作業に及ぼす効果を調べた結果を述べる．最後に，Remote Lag を表示することによる問題点をあげ，改善方法を議論する．

2. 関 連 研 究

2.1　ジェスチャの表示方法

システム上でジェスチャを表現する方法として，これまでにポインタ[6),7)] やペンによる描画[4)] など，様々な手法が取り入れられてきた．近年の遠隔テーブル共有システムでは，タンジブルインタフェースの利用が進んだことやペーパーワークのような実物を扱った協調作業が増加したことを受け[4),16),20)]，ジェスチャをそのまま表示することが重要視されている[11)]．特に，遠隔地間の作業領域を重ね合わせて 1 つの作業領域とし，この作業領域上に遠隔ユーザの手のジェスチャを直接表示させることが有効であると考えられている[11),15),26)]．

© 2011 Information Processing Society of Japan

2.2　遠隔ジェスチャの不可視性

遠隔協調作業の研究が進展するにつれ遠隔協調作業の効率も改善されてきたが[13),26)]，対面協調作業の効率にはまだ及ばない．その 1 つの理由として，従来の遠隔テーブル共有システムにおける不可視性の問題があげられる．実際にそのような不可視性のために遠隔ジェスチャ認識が妨げられているが，これまでほとんど議論されてこなかった．

本論文では，不可視性の問題を引き起こす要因として特に以下の 2 つの現象に注目する．ここで，t-Room を含む多くの遠隔テーブル共有システムは遠隔ジェスチャを作業領域の背面から投影（表示）する方式をとっていることに注意されたい[3),14),20),21)]．これは，マルチタッチ検出技術の実装[2),8)] や映像の明瞭性の確保など，様々な技術的要因によって背面投影（表示）方式を採用するテーブルトップディスプレイが多いことによる．したがって本論文ではこのような背面投影型のシステムに多く見られる不可視性の問題を議論する．

2.2.1　注意の離脱

一般に，ビデオシステムでは遠隔地の 3 次元的空間が 2 次元に縮退されるため，遠隔ユーザの動作へのアウェアネスが阻害されてしまう．Gaver はこの問題に対して以下のように述べている[5)]：*平面スクリーンは，我々が本来備える変化を見分ける能力や距離感覚を低下させてしまう．遠隔ユーザの様子をスクリーンに表示したところで，2 次元映像ではユーザの注意を十分に引くことができず，遠隔ユーザの様子や出来事に対するアウェアネスを伝えあうことは困難である．*

遠隔協調作業に取り組むユーザが 2 名の場合，ユーザが協調作業中に注意を向けるべき遠隔対話者は確定しているため，この現象はさほど大きな問題とならないケースが多い．しかし，我々が実施した実験のように（テーブル周りの）移動をともなう多人数遠隔協調作業では，だれがだれに対して話しかけようとしているのかを判別するためには，ユーザはスクリーン上に表示された複数名の遠隔ユーザの動きをつねに識別し続ける必要があり，これができなければ協調作業を円滑に遂行できないことが予想される．実際，我々が実施した予備実験では，この要因による遠隔ジェスチャの見逃しが多数観測され，円滑な協調作業の妨げとなっていた．

2.2.2　オクルージョン

オクルージョンとは，スクリーン上に置かれた実物体やユーザの体の一部によってスクリーンに表示された映像が隠れて見えなくなってしまう現象を指す（**図 2**）．遠隔協調作業において，情報伝達手段の比重が言葉よりもジェスチャに多く置かれた場合（たとえば言葉での表現が難しい内容をジェスチャで説明した場合），ジェスチャがオクルージョンによって隠れてしまうと遠隔ユーザ間の意思疎通が困難になる．なお，この問題は，ユーザが書類など比較的大きな遠隔物体に対してジェスチャを行う場合や遠隔ユーザがローカルユーザと同じ立ち位置で協調作業を行う場合（遠隔ユーザの腕の映像がローカルユーザの腕の下に表示される場合）に生じやすい．また，テーブルの上に立体物が置かれている場合，ユーザから見て立体物の裏側に表示されたジェスチャは見えにくくなる．これは上面投影方式のディスプレイを利用した場合にも生ずる問題である．

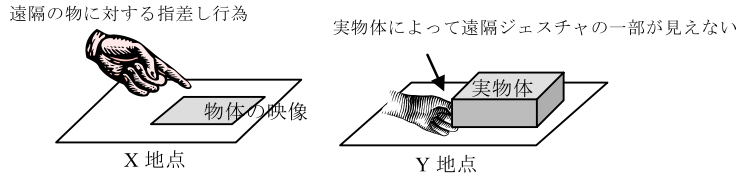

図 2　オクルージョンによって遠隔ジェスチャが見えない例
Fig. 2　An example of remote gestures partly occluded by a real object.

3.　Remote Lag

2 章であげた不可視性の問題を軽減する方法を考案するにあたり，我々は予備実験におけるユーザの動作，すなわちユーザが不可視性の問題に直面したときの反応に着目した．実験データにおいて不可視性の問題が生じた場面を調べたところ，ユーザが遠隔ジェスチャを見逃したとき，わずかの遅れによって見逃してしまう場面が多く見られた．なお，ジェスチャにはたいてい，「これ」や「あれ」といった発話がともなっていたため，これを遠隔ユーザが見逃した場合でも，自分が見逃したこと自体には気づいている様子がうかがえた．たとえば見逃した後に，見逃したと思われるスクリーンを凝視したり左右を見渡して探したりする仕草が観測された．また，ユーザから遠隔ジェスチャの映像が実物体の下敷きになって見えなかったときには，実物体を少し持ち上げて隠れている遠隔ジェスチャを確認しようとする動作が見られた．しかし実験では，指示を出す側（指示者）がこのユーザが実物体を持ち上げる動作を見て，ユーザが指示を理解したと解釈する場面が見られ，そのとき指示者は遠隔ジェスチャを終了してしまう．その結果，ユーザが指示内容を幾度も聞き返す場面が見られた．

以上の観察に基づき，我々は Remote Lag という支援手法を提案する．Remote Lag とは，一瞬前の遠隔ジェスチャの遅延映像を現在の映像に重ねて表示する方法を指す．本手法により，リアルタイムの遠隔ジェスチャを見逃してしまった場合でも，遅延映像を見ること

© 2011 Information Processing Society of Japan

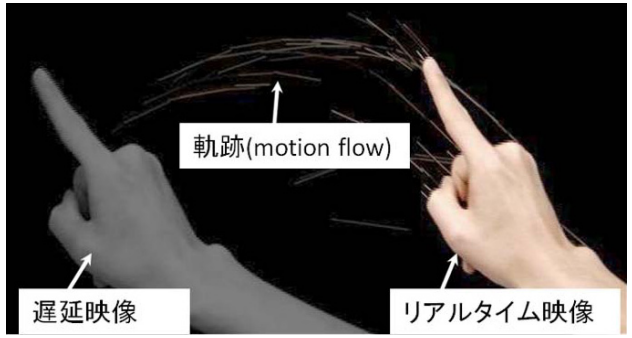

図 3　Remote Lag の表示例
Fig. 3　An example of Remote Lag.

によって協調作業の文脈に復帰する機会が与えられる．

Remote Lag は次のように実装されている．まず，撮影された映像の最新フレームから 1.1 秒前フレームまでのすべてをメモリ上に展開し，OpenCV ライブラリを用いて隣接フレーム間でのオプティカルフローを求める．それらを連結した線によって motion flow を構成する．表示される Remote Lag 映像は，一番奥にグレースケール化した 1.1 秒前フレーム（遅延映像），次に motion flow，一番手前に最新フレーム（リアルタイム映像）の順に重ねて作成する．

Remote Lag は，ユーザの一瞬前の動作を共有領域に表示するという点で Local Lag[7] と類似している．Local Lag は通信遅延による協調作業の問題を軽減するために考案された手法であり，ユーザのフィードバック映像にネットワーク通信時間分の遅延を導入することによって自身の動作が遠隔ユーザにどのように表示されているのかを確認可能にしている．Local Lag と Remote Lag は，実装方法こそ似ているが，遅延を導入する対象と使途が大きく異なる．まず Local Lag がユーザ自身のフィードバック映像に対して遅延を導入するのに対し，Remote Lag は遠隔ユーザの映像に対して遅延を導入する．また，Local Lag が通信遅延による遠隔協調作業の問題を軽減することを目的としているのに対し，Remote Lag は遠隔ジェスチャの不可視性による問題の軽減を目的とする．

Gutwin はカーソルの遅延映像の表示方法についていくつかのタイプを比較し，表示領域を必要以上に煩雑にさせないために，軌跡の表現は比較的短く，低いコントラストで，フェードアウトさせることが好ましいことを見い出した[6]．本研究では，Gutwin の方法を参考にしてユーザの指先の Remote Lag 表示方法を検討するため，t-Room[9] においていくつかの Remote Lag 表示方法を比較する予備実験を実施し，以下の表示法を採用することとした（**図 3**）[10]．

- リアルタイム映像と遅延映像の間の時間幅は 1.1 秒とする．
- リアルタイム映像と遅延映像を直観的に区別できるように，遅延映像をグレースケールとする．
- 静止しているリアルタイム映像を遅延映像と間違えないようにするために遅延映像はリアルタイム映像の下に重ねる．
- スクリーンを煩雑にしすぎないために遅延映像を複数枚表示せず，代わりにリアルタイム映像と遅延映像の間にフェードアウトする軌跡（以後 motion flow と呼ぶ）を表示する．

ただし本論文では Remote Lag の効果を確認することを目的とするため，Remote Lag におけるこれらの設定は暫定的なものであり，より効果的な設定については今後の課題とする．

4. 評価実験

遠隔協調作業における Remote Lag の効果を評価するために，t-Room を用いて被験者実験を実施した．本実験では特に，実物体が介在する遠隔指示に焦点を当て，Remote Lag が遠隔地間での実物体の扱いや情報伝達にどのような影響を及ぼすかを調べる．具体的に，実験を通して解明したい項目は以下のとおりである．

- Remote Lag の表示によって遠隔ジェスチャの不可視性の問題は緩和されるか．
- Remote Lag の表示によって遠隔指示における発話効率が向上するか．たとえば，作業者による質問や確認および指示者による冗長な指示が減少するか．
- Remote Lag の表示によってユーザにどの程度の認知負荷がかかるか．

4.1 ビデオシステム：t-Room

テーブルを共有した対面型協調作業の特徴として，テーブルの周囲に複数の参加者が自由に身体を位置取ることができることや，テーブルへの参加の仕方が時間とともに動的に変化することが指摘されている[17,18]．t-Room は，こうした特徴を遠隔環境でも実現することを目指して構築されたビデオシステムである[9]．t-Room ではテーブル面と上半身投影ディスプレイをうまく配置することによって遠隔ユーザのテーブル上（作業領域）の作業を周辺環境（テーブル周りの身体行為）と関連付けて理解することを可能にしている[26]．また，Agora[14] などの従来の遠隔テーブル共有システムが協調作業中のユーザの立ち位置や

© 2011 Information Processing Society of Japan

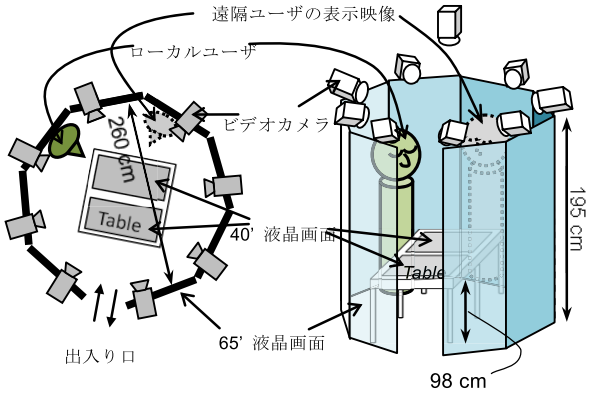

図 4　t-Room のハードウェア構成
Fig. 4　Hardware configuration of t-Room.

テーブル周りの移動を制限するのに対し，t-Room はユーザの立ち位置に柔軟性を持たせ協調作業中の移動できるようにしている．さらに，t-Room は 1 地点につき複数ユーザが参加可能であり，ユーザが t-Room 内のどこにいても遠隔ユーザの立ち位置をほぼ正確に伝達できる．

t-Room は，ユーザのいる空間を囲むように配置された複数のカメラ，縦置きディスプレイ，2 枚のディスプレイで構成された中央テーブル，そしてテーブル面を撮影するビデオカメラから構成される（**図 4**）．t-Room にはマイクとスピーカも設置されているが図には示していない．複数地点に同一構成の t-Room を設置し，それらの間で映像・音声を通信し合うことによって，遠隔地にいる t-Room 内のユーザ同士が同じ作業空間を共有することができる．たとえば t-Room の 1 番目の縦置きディスプレイの前に立っているユーザは遠隔 t-Room の 1 番目の縦置きディスプレイに投影される．テーブル上の物体は遠隔テーブル上の同じ位置に 2D 映像として再現される．こうして，同じ t-Room 内の複数ユーザ同士だけでなく遠隔のユーザ同士でもテーブルを囲み自由な立ち位置をとりながら協調作業を行うことができる．

4.2　実 験 手 法

4.2.1　実 験 環 境

我々は神奈川県と京都府にまったく同じ構造の t-Room を設置し，これら 2 台の t-Room をギガビットネットワークで接続した．映像と音声の遅延は，各々約 0.3〜0.4 秒と 0.2 秒〜0.3 秒であり，映像と音声の同期はとらなかった．この遅延の差が Remote Lag の効果に影響する可能性もあり，その場合は Remote Lag の遅延時間を調整する必要が生ずるかもしれないが，これについては今後の課題とする．t-Room 内の音声は各ユーザが装着するワイヤレスマイクにより収音され，1 ch に集約され（モノラル），遠隔の t-Room の壁面下方に設置された 2 台のスピーカから流された．これは，遠隔ユーザの映像位置と音源位置の相違による影響を除く，つまり遠隔ユーザの位置が音源位置から推測できず映像からのみ推測できるようにするためである．

4.2.2　被　験　者

本実験には，互いに面識のない被験者 56 名（神奈川県に 28 名，京都府に 28 名）が有償で参加した．どの被験者もビデオ会議システムやテーブル共有システムの経験がないものとした．また，被験者の協調作業（特にテーブル周りの移動）において性別による影響を排除するために，すべての被験者は女性とした．

4.2.3　実 験 設 計

神奈川県と京都府から各 2 名ずつの合計 4 名で 1 グループとし，合計 14 グループを構成した．各グループは，2 つのメディア条件で 1 回ずつ（計 2 回）t-Room 内のテーブル上で遠隔作業指示に関するグループ課題に取り組んだ．2 つのメディア条件とは「**メディア条件 1**：Remote Lag を表示しない場合」と，「**メディア条件 2**：Remote Lag を表示した場合」である．これらのメディア条件を比較することによって，他の要因（たとえば音声と映像の時間差や被験者の個性による影響）を排除し，Remote Lag の影響のみを取り出して調べることができる．なお，両条件とも，ユーザ自身の動作が遠隔ユーザにどのように見えるかを確認できるようにするためにリアルタイムのフィードバック映像（エコー映像）を表示した．遅延映像は遠隔映像に対してのみ表示した．

実験の設計では，実験計画法に則り，メディア条件の順序と課題の順序の釣り合いをとった．

4.2.4　グループ課題

本実験の課題には，テーブルを用いた遠隔協調作業でよく題材として用いられる「実物体を用いた遠隔作業指示（mentoring collaborative physical task）[12]」を採用した．この課題は，専門家（指示者）が遠隔地にいる初心者（作業者）に指示を出すことによって目的を達成するというものである．本実験では，異なる専門知識を持ち指示を出す 2 名の専門家が神奈川県に，作業する初心者 2 名が京都府にいる状況を想定し，合計 4 名で遠隔協調作業を実施する．

© 2011 Information Processing Society of Japan

図 5　課題における完成図の一例
Fig. 5　A completion map used in this study.

具体的な協調作業内容は「レゴブロック*1の組み立ておよび配置」とした．レゴブロックの組み立て作業は，主に作業者が指示者の指示に従ってレゴブロックを操作（回転や接合）するものであり，実物体を扱う協調作業の実験で採用されることが多い[4),11)]．そして，配置作業は，組み立てたレゴブロックをテーブル上の定められた位置に置くものであり，都市計画や協調デザインと類似して高い精密性が求められる[23)]．

実験では，指示者はレゴブロックの完成図（**図 5**）を手渡され，これをもとに作業者に指示を行うこととした．完成図は t-Room 越しの作業者から視認できない大きさ（手のひらサイズ）にプリントアウトした．

部品は作業者のいる t-Room 内のテーブル上に過不足なく置いてあり，作業者だけが操作することができる．作業者は指示者の指示に従ってテーブル上の部品を組み立て，テーブル上に表示された白い円画像との相対的な位置関係や向きが完成図と一致するように部品を配置しなければならない．課題を遂行するにあたり，指示者と作業者 1 名ずつでペアを組み，直方体ブロックの組み立てのみを扱う直方体担当ペアと，曲面入りの部品を含んだ組み立てを扱う曲面部品担当ペアに役割分担した．複数のペアが同時並行的に協調作業を行う設定にした理由は，遠隔テーブル共有に関する最近の研究で同時並行的に複数の会話や作業が行われることが報告されており[20)]，そのような状況を確実に再現するためである．

*1 実験で用いたレゴブロックは 1.5 歳〜5 歳くらいの幼児を対象にして作られた「レゴデュプロ」と呼ばれる製品で，通常のレゴブロックの倍くらいの大きさを持っている．

被験者らはメディア条件を変えて 2 回課題に取り組んだ．ただし，課題に対する学習効果を軽減するために，完成図を 2 つ用意し，メディア条件と課題の種類はカウンターバランスをとった．

4.2.5　実験手順

手順 1. 実験開始前，各被験者は実験前に知己になることを防ぐため別々の部屋で待機した．各被験者は待機部屋で実験に参加するための同意書に署名し，その後 t-Room に案内された．4 名の被験者は t-Room 内で t-Room システムに関する簡単な説明を受けた後，自己紹介を行った．次に，神奈川県の指示者と京都府の作業者 1 名ずつをランダムにペアとし，役割分担（1 ペアは直方体担当，もう 1 ペアは曲面部品担当）を決めた．t-Room の環境やテーブルの使い方に慣れるために，各組は実際にテーブル上のレゴブロックの組み立てと配置作業の練習を 5 分程度行った．

手順 2. 被験者らは課題の流れについて説明や注意事項を受けた．注意事項は以下のとおりである：原則として，手順 1 で定めた役割分担に従って指示や作業を行うこととし，自分と関係のない部分については指示や作業を行わないものとした．ただし，誤りの発見など，課題の目的を達成するうえで重大な問題が生じた場合には分担を超えたアドバイスや手伝いをしてもよいことにした．また，組み立てる順番や t-Room の利用方法は自由とし，相談や話し合いは 4 名で自由に行ってよいこととした．なお，可能な限り早く作業を完成させるように伝えた．

手順 3. レゴブロックの組み立ておよび配置作業を，Remote Lag を表示した状態と表示しない状態で課題の完成図を変えて合計 2 回実施した．それぞれの課題終了後は，直前に行った作業に対する認知負荷に関するアンケートに回答した．

手順 4. 最後に，京都府で実験に参加した被験者は，Remote Lag の有無による違いや指示内容の理解のしやすさなどに関するインタビュを受けた．

5. 評価結果

我々はメディア条件の違いが協調作業に及ぼす影響を詳しく調べるために，両 t-Room の出入り口の頭上付近および出入り口と反対の位置にビデオカメラ（合計 4 台）を設置し，実験の様子を撮影した．以下に，これらのビデオ映像を分析した結果を述べる．まず，Remote Lag が実際にどのような場面でどのように有効活用されているかを理解するために，観察的な分析を行った．5.1 節にその分析結果の例を示す．さらに，Remote Lag の効果を定量的に評価した結果を 5.2 節以降に示す．

© 2011 Information Processing Society of Japan

5.1　不可視性の問題

5.1.1　注意の離脱

課題の進め方は，いずれのメディア条件の場合も，直方体担当と曲面部品担当が同時並行で作業するというものであった．どのペアもおおむね円滑に作業を進めることができたが，時折，作業者が予期せぬトラブル（指示者の手を見誤ったり，ブロックをうまく接合できないなど）や自分と異なるペアのやりとりに気をとられて指示内容を見逃す場面が観測された．

図 6 は，指示者が作業者に対して，目の前にある 2 つの赤いブロックのうちの 1 つを持ちライオンのブロックと接合するよう指示し，作業者がこれに従う一連の場面をとらえたものである．作業者が手元の作業に気をとられて指示者の次の指示を見逃してしまう場面である．右側の作業者が遠隔 t-Room の同じ位置に立つ指示者の指示に従ってブロックを組み合わせている．ただし，指示者の映像は t-Room の構造上，作業者の後ろのスクリーンに表示されている．図 6-a のシーンの直前に，指示者はそれら 2 つの赤いブロックを続けて指し示し，これらのブロックが同一のものかどうかを質問した．作業者はすぐに「はい」と答えたものの，それらのブロックが実際に同じものかどうか，手にとって確認した（図 6-a）．しかし指示者はこの作業者の行為に気づかず，続けてテーブルの奥側に腕を伸ばしてライオンのブロックを指し示しながら，「どっちかにこのライオン挿してください」と指示した（図 6-a）．指示者が「このライオン」と言ったとき，作業者は指示に追いつこうと顔をあげたが，このときリアルタイムの指差しはすでにライオンのブロックに到達していた．作業者は続けて現れた遅延映像の動きを追うことでライオンのブロックを発見することができた（図 6-b）．最終的に作業者はリアルタイムの指示に追いつくことができ（図 6-c），協調作業の進行に大きなロスは発生しなかった．

対面的な状況であれば，作業者が何に注目しているかということに指示者がすぐに気がつくため，指示者は作業者の視線を獲得してから指示動作を開始することが多い．また，作業者は指示者の指示動作が目の前で示されれば，それに容易に気づくことができる．しかし遠隔的な状況では，3 次元的空間が 2 次元に縮退するために，これに気づきにくくなってしまうのである．しかしこの例に示したように，そうした見逃しが Remote Lag によって補われる例が複数見られた．

5.1.2　オクルージョン

本実験では，オクルージョンが発生した場合に，指示者と作業者同士が冗長なインタラクションを行う傾向にあった．特に，指示者が言葉による十分な説明を行わずにジェスチャを用いて指示を行った場合（たとえば，「こんなふうにしてください」，「あそこに置いてください」など），そのジェスチャが実物体に隠されてしまうと作業者はその指示に従うことができなくなる．このようなシーンは，言葉だけでうまく表現することが難しく複雑なジェスチャを必要とする場合や，部品の配置の微調整を行う場合に多く見られた．隠された遠隔ジェスチャの視認性を Remote Lag が向上させた例として，ある部品の下に隠されたジェ

図 6-a

6-a 中, motion flow の実映像は薄く写真上では認識が難しかったため，白線を付加して強調してある

図 6-b

図 6-c

トランスクリプト 1.

```
指示者：  どっちかに ::  (1.3)  このライオンくんを:: (1.1)  挿してください
                <RT>       <LT> <RP-lion>
                          *図 6-a  *図 6-b      *図 6-c
作業者：            はい          <X>                 はい                 はい
```

<RT>:	手の移動を開始（ライオンを指差すために）	
<RP-lion>:	ライオンを指差す	“:”　0.1 秒の間
<LT>:	遅延映像の手の移動が開始	(1.3)　1.3 秒の間
<X>:	作業者が遅延映像を見た瞬間	“*”　シーン撮影の瞬間

図 6　Remote Lag が注意の離脱による不可視状態からの復帰に役立った例

Fig. 6　An example of a worker making use of the lagged gestures (when her attention was directed to something else).

© 2011 Information Processing Society of Japan

トランスクリプト 2.

```
01　指示者：　この四角の　あ　緑のかどが::
　　作業者：　　はい
02　指示者：　この辺りまで下げてもらっていいですか？
　　作業者：
03　指示者：　(3.0)　はい　そう::で::す
　　　　*図 7
　　作業者：
```

図 7　Remote Lag がオクルージョンによる不可視状態からの復帰に役立った例
Fig. 7　An example of a worker making use of the lagged gestures (when the remote gesture was occluded by a real object).

スチャを，作業者がその部品を持ち上げて確認したシーンを**図 7** に示す．

このシーンでは，指示者と作業者のペアが緑色の四角いブロックの配置を微調整している．まず白いジャケットを着た作業者が，そのブロックを大体の位置に置き，次に指示者が，指し示している箇所までブロックの角をもう少し下げるように指示している．しかしこの指差しのジェスチャは緑色のブロックの下に隠されてしまい，作業者は指示を見ることができなかった．そこで作業者はその指示を見るために緑色のブロックを持ち上げた．指示者は，作業者が行ったブロックを持ち上げるという動作を，指示を了解した合図であると解釈し，ただちに指差しをやめて手を引っ込めた．その結果，作業者が指し示された箇所を確認したときにはすでにリアルタイムのジェスチャ映像は消えてしまっていた．ここで，作業者は遅延映像を見ることで指示されていた箇所を認識し（図 7 の瞬間），適切な位置にブロックを

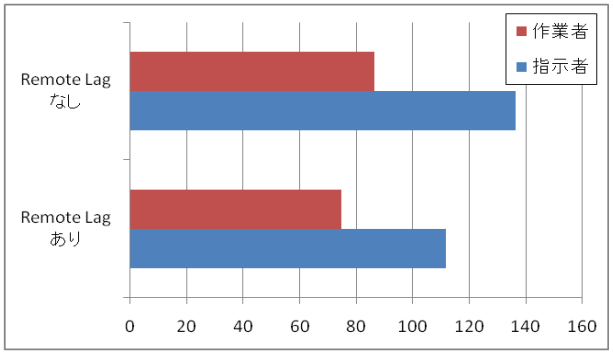

図 8　課題あたりの平均発話数
Fig. 8　Mean number of utterances per task by experimental role and condition.

移動することができた．

ここで興味深いのは，作業者が指示された対象物に触るという行為が，指示者から見ると指示の了解と解釈されてしまう点である．このような場合には指示者はすぐに指さしを終了してしまう傾向が見られた．このような場合に，Remote Lag によって不可視性の引き起こす効率低下が軽減される例が複数見られた．

5.2　発話効率

Remote Lag によって遠隔作業指示に関する冗長な発話がどの程度減少するかを検証するために，ビデオ映像から会話データを書き起こしたうえで，タスクごとの発話回数の平均を実験の役割と条件に分けて算出した（**図 8**）[4]．ここで発話回数とは，発話ターンの獲得回数のことである．対応のある両側 t 検定を用いて作業者と指示者それぞれについて発話回数を比較した．その結果，指示者について，Remote Lag ありの条件の方が Remote Lag なしの条件よりも発話回数が有意に少ないことが分かった（$t[27] = 2.74$，$p < .05$）．一方，作業者の発話回数については 2 つの条件間で有意な差が認められなかった．

Remote Lag が質問や確認の発話の減少に貢献するか否かを調査するため，役割と条件ごとに質問や確認に関する発話回数を数えた．ここで質問や確認に関する発話とは，タスクを遂行するにあたり不十分・不明瞭な情報を質問したり確認したりする発話を指す．たとえば作業者の場合，ブロックの配置場所を尋ねる発話（「ここ？」）や自分の作業を確認する発話（「これでいいですか？」）がこれに該当する．指示者の場合は，自分の指示した内容が伝わったかどうか確認する発話（「分かります？」）などがこれに該当する．質問と確認に関

© 2011 Information Processing Society of Japan

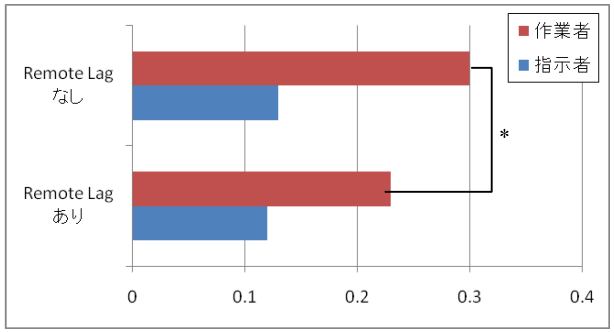

図 9　1 発話あたりに質問や確認を行う割合の平均
Fig. 9　Mean proportion of utterances on questions and confirmations per task by experimental role and condition.

する発話がタスク全体の発話回数に占める割合を**図 9** に示す．この結果から，Remote Lag ありの条件における作業者の質問や確認の発話割合が 23%（SD = 13%）であり，Remote Lag なしの条件の 30%（SD = 16%）よりも少ないことが示された．対応のある両側 t 検定を用いて，作業者と指示者がそれぞれ質問や確認に関して発話する割合を比較した．その結果，作業者について，Remote Lag ありの条件の方が Remote Lag なしの条件よりも質問や確認を行う発話割合が有意に少ないことが分かった（$t[27] = 2.71$，$p < .05$）．一方，指示者の質問や確認に関する発話割合は，メディア条件間で有意な差は見られなかった．

以上の結果から，Remote Lag の表示によって指示者が行うジェスチャの視認性が向上し，作業者が指示者の指示を理解しやすくなったため，質問や確認の頻度が減少したと考えられる．Remote Lag の条件において指示者の発話が少なくなったのは，遅延映像によって遠隔ジェスチャの不可視の問題が緩和されたことにより，指示者が作業者の質問や確認に回答しなければならない場面が減少したからであると考えられる．

5.3　認知的な負荷

5.2 節において，Remote Lag が遠隔作業指示を円滑にすることが示されたが，協調している作業者にどの程度の作業負荷がかかっているかを検証することは重要である．たとえば Remote Lag を導入したことによる問題点として，共有作業領域の表示が煩雑になることがあげられ，これは作業者に対してより集中や努力を強いることにつながる可能性がある．そこで，Remote Lag の表示による主観的な作業負荷を測定するために，NASA Task Load Index（TLX）をタスク後アンケートとして実施した．

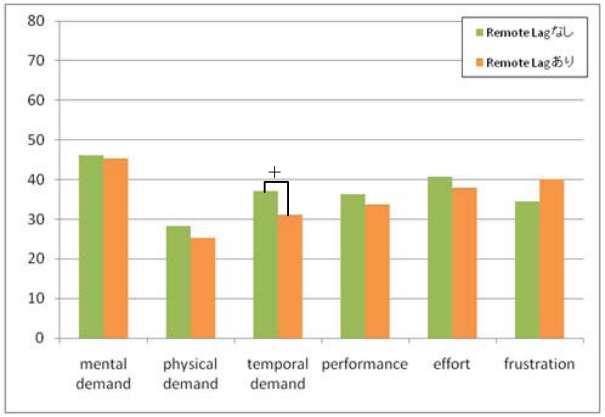

図 10　作業者の項目ごとの TLX 平均値
Fig. 10　Mean TLX ratings of the workers.

NASA TLX では全体的な作業負荷を 6 つの指標（知覚的要求（mental demand），身体的要求（physical demand），タイムプレッシャ（temporal demand），作業成績（performance），努力（effort），フラストレーション（frustration））によって評価する．全体的な作業負荷は各指標の重み付き平均に基づいて 0 から 100 までの範囲（スコアは低いほど良い）で算出される．この手法は，人々の主観的な作業負荷を評価する手法として，様々な人間とデジタル機器間のインタフェースの評価に広く用いられている．

図 10，**図 11** は，それぞれ作業者と指示者についての各指標のスコアの平均である．全体的な傾向として，指示者が作業者よりも高い作業負荷を感じていたことが見て取れる．次に，対応のある t 検定を用いて，指示者と作業者それぞれについて項目ごとに作業負荷の違いを比較した．その結果，作業者が感じるタイムプレッシャは，Remote Lag ありの条件の方が Remote Lag なしの条件よりも有意に低いことが示された（$t[27] = 1.72$，$p = .097$）（図 10）．その他の指標のスコアに関しても，Remote Lag ありの条件の方が Remote Lag なしの条件よりもおおむね負荷が低かったが，有意差は認められなかった．なお，フラストレーションの指標に関しては，Remote Lag の条件の方が Remote Lag なしの条件よりも高くなったが，これも有意差は認められなかった．

指示者の各指標についても同様に 2 つの条件間で t 検定を行った．その結果，被験者は Remote Lag ありの条件の方が Remote Lag なしの条件よりもタイムプレッシャによる作

© 2011 Information Processing Society of Japan

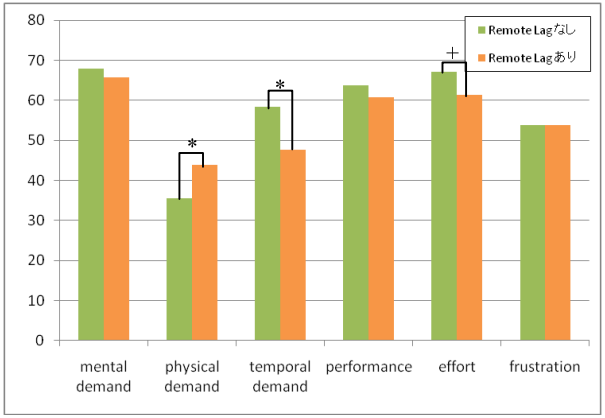

図 11　指示者の項目ごとの TLX 平均値
Fig. 11　Mean TLX ratings of the helpers.

業負荷を有意に低く（$t[27] = 2.38$，$p < .05$），少ない努力で作業を終了できると感じる傾向にあることが分かった（$t[27] = 1.83$，$p = .079$）．しかしその一方で，指示者の感じる身体的負荷は Remote Lag ありの条件の方が Remote Lag なしの条件よりも有意に高かった（$t[27] = 2.38$，$p < .05$）．

以上の結果から，作業負荷に関して Remote Lag は以下の利点を持つといえる．まず，指示者と作業者は，タスク中に感じるタイムプレッシャが低くなる．これは Remote Lag ありの条件における指示者と作業者は遠隔ジェスチャを見逃したとしても，遅延映像を追うことによってリアルタイムのジェスチャに追いつくことができるため，リアルタイムの遠隔ジェスチャを追う余分な労力を必要としないからであると推測できる．次に，指示者が作業者に指示を行う際に必要な努力が低い理由としては，作業者のジェスチャ見逃しやオクルージョンが生じたときに，指示の繰返しや補足の必要性が減少したためと推測される．一方，Remote Lag の問題点として，タスク中に指示者が感じる身体的負荷が増加することが分かった．これは，統計的な有意差こそ検出されなかったものの Remote Lag ありの条件では Remote Lag なしの条件よりも作業時間が短い傾向にあったため（Remote Lag なしの場合，平均 699 秒（SD = 278）; Remote Lag ありの場合，平均 606 秒（SD = 158）），指示者が次々と指示を行う必要があり忙しかったためと考えられるが，今後インタビュなどを通してさらに深く原因を調査する必要がある．

6. 結論と今後の展望

本論文では，遠隔ジェスチャを不可視にする 2 つの要因をあげ，この問題を緩和するために Remote Lag という手法を提案した．4 名の参加者による遠隔作業指示の実験において，Remote Lag ありの条件と Remote Lag なしの条件を比較した結果，Remote Lag は不可視性の問題の緩和に効果があることが示された．

一方で，実験後のインタビュから Remote Lag の表示方法に関していくつか課題が明らかになった．

- スクリーンの煩雑さ：何人かの作業者は，繰返しをともなう細かなジェスチャを指示者から提示された際，遅延映像や motion flow によって画面が煩雑になり，ジェスチャ認識が困難になることがあったという感想を述べた．この問題を解決する 1 つの方法として，Remote Lag の表示を作業に必要な箇所のみに制限するため，映像中から手や指の領域だけを抽出し，動きが大きい部分（特徴点の差分が閾値以上の部分）のみ Remote Lag を表示することが考えられる．
- 遅延映像の区別：今回の実験では遅延映像をグレースケールにすることでリアルタイムの映像と遅延映像を瞬時に判別することが可能になったが，一方，何人かの作業者は 2 名いる遠隔指示者のどちらの遅延映像なのかという区別がつきにくいという不満を訴えた．この問題を解決する 1 つの方法として，たとえば遅延映像全体をユーザの袖の色に合わせて表示することが考えられる．
- 遅延映像の枚数と遅延時間：遅延映像の区別やスクリーンの煩雑さを考慮すると，遅延映像を 2 枚以上表示する方式は現実的ではないだろう．本実験では遅延映像とリアルタイム映像との時間幅を 1.1 秒で固定したが，1 枚の遅延映像を効果的に使うために，動きの速さに合わせて時間幅を変化させることが考えられる．
- 他の遠隔ビデオシステムへの応用：遠隔地の映像が実寸より小さく表示されるようなシステムでは，スクリーンに表示する Remote Lag 映像の量を本実験の設定より減らさなければならない．たとえば，motion flow の本数を減らしたり，不可視性の問題が生じる場所や時間に限定して Remote Lag を表示したりするなどの工夫が必要になろう．また上面投影方式のシステムでは，実物体によって映像が隠される不可視性は起きにくいものの，上方から投影される映像の一部が実物体に被るような場合，実物体の色や形による映像の変色や変形が生じて物体の認識がうまくできない場合が起きる．したがって，このような上面投影方式のシステムでも Remote Lag は効果があるのではないかと期待される．

© 2011 Information Processing Society of Japan

今後の課題として，スクリーンが煩雑にならず遅延映像を区別しやすい Remote Lag 表示方法の検討が必要である．また，単純に音声と映像の時間差を広げた場合に Remote Lag と同様な効果が見い出せるか否かを調べることによって，Remote Lag がうまく機能するための条件について検討する必要がある．最後に，t-Room とは異なる遠隔ビデオコミュニケーションシステムでの効果を調べることを考えている．

参 考 文 献

1) Buxton, W.: Integrating the Periphery and Context: A New Model of Telematics, *Proc. GI'95*, pp.239–246, ACM Press (1995).
2) Dietz, P.H. and Leigh, D.L.: DiamondTouch: A Multi-User Touch Technology, *Proc. UIST'01*, pp.219–226, ACM Press (2001).
3) Ishii, H. and Kobayashi, M.: ClearBoard: A seamless medium for shared drawing and conversation with eye contact, *Proc. CHI'92*, pp.525–532, ACM Press (1992).
4) Fussell, S.R., Setlock, L.D., Yang, J., Ou, J., Mauer, E. and Kramer, A.: Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks, *Human-Computer Interaction*, Vol.19, pp.273–309 (2004).
5) Gaver, W.: The affordances of media spaces for collaboration, *Proc. CSCW'92*, pp.17–24, ACM Press (1992).
6) Gutwin, C.: Improving interpretation of remote gestures with telepointer traces, *Proc. CSCW'02*, pp.49–57, ACM Press (2002).
7) Gutwin, C.: The Effects of Local Lag on Tightly-Coupled Interaction in Distributed Groupware, *Proc. CSCW'08*, pp.447–456, ACM Press (2008).
8) Han, J.Y.: Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection, *Proc. UIST'05*, ACM Press (2005).
9) Hirata, K., Harada, Y., Takada, T., Yamashita, N., Aoyagi, S., Shirai, Y., Kaji, K., Yamato, J. and Nakazawa, K.: Basic Design of Video Communication System Enabling Users to Move Around in Shared Space, *IEICE Trans.*, Vol.E92-C, No.11, pp.1387–1395 (2009).
10) Kaji, K., Yamashita, N. and Hirata, K.: A Method for Supporting Recognition of Remote Gestures in a Video Communication System, *Proc. ASIAGRAPH'09*, pp.84–89 (2009).
11) Kirk, D., Crabtree, A. and Rodden, T.: Ways of the hands, *Proc. ECSCW'05*, pp.1–21, Springer (2005).
12) Kraut, R.E., Fussell, S.R. and Siegel, J.: Visual information as a conversational resource in collaborative physical tasks, *Human Computer Interaction*, Vol.18, No.1, pp.13–49 (2003).
13) Kraut, R.E., Gergle, D. and Fussell, S.R.: The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence, *Proc. CSCW'02*, pp.31–40, ACM Press (2002).
14) Kuzuoka, H., Yamashita, J., Yamazaki, K. and Yamazaki, A.: Agora: A Remote Collaboration System that Enables Mutual Monitoring, *Extended Abstracts of CHI'99*, pp.190–191, ACM Press (1999).
15) Kuzuoka, H., Kosaka, J., Yamazaki, K., Suga, S., Yamazaki, A., Luff, P. and Heath, C.: Mediating dual ecologies, *Proc. CSCW'04*, pp.477–486, ACM Press (2004).
16) Luff, P., Heath, C., Kuzuoka, H., Yamazaki, K. and Yamashita, J.: Handling documents and discrimination objects in hybrid spaces, *Proc. CHI'06*, pp.561–570, ACM Press (2006).
17) Minatani, S., Kitahara, I., Kameda, Y. and Ohta, Y.: Face-to-Face Tabletop Remote Collaboration in Mixed Reality, *Proc. ISMAR'07*, pp.43–46, IEEE Computer Society (2007).
18) Scott, S.D., Grant, K.D. and Mandryk, R.L.: System guidelines for co-located collaborative work on a tabletop display, *Proc. ECSCW'03*, pp.159–178, Springer (2003).
19) Tang, A., Neustaedter, C. and Greenberg, S.: VideoArms: Embodiments for mixed presence groupware, *Proc. British-HCI'06*, pp.85–102, Springer (2006).
20) Tang, A., Pahud, M., Inkepen, K., Benko, H., Tang, J. and Buxton, B.: Three's Company: Understanding Communication Channels in Three-way Distributed Collaboration, *Proc. CSCW'10*, pp.271–280, ACM Press (2010).
21) Tang, J. and Minneman, S.: VideoDraw: A video interface for collaborative drawing, *ACM Trans. Information Systems*, Vol.9, No.2, pp.453–469 (1991).
22) Tang, J.: Findings from observational studies of collaborative work, *Man-Machine Studies*, Vol.34, No.2, pp.143–160 (1991).
23) Underkoffler, J. and Ishii, H.: Urp: A luminous-tangible workbench for urban planning and design, *Proc. CHI'99*, pp.386–393, ACM Press (1999).
24) Wilson, A.D.: PlayAnywhere: A Compact Interactive Tabletop Projection-Vision System, *Proc. UIST'05*, pp.83–92, ACM Press (2005).
25) Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H. and Harada, Y.: Impact of Seating Positions on Group Video Communication, *Proc. CSCW'08*, pp.177–186, ACM Press (2008).
26) 山下直美，葛岡英明，平田圭二，青柳滋己，白井良成，梶　克彦，原田康則：身体の動きを伴う遠隔協調作業における上半身映像の効果，情報処理学会論文誌，Vol.51, No.4 (2010).

(平成 22 年 4 月 19 日受付)
(平成 22 年 10 月 4 日採録)

© 2011 Information Processing Society of Japan

山下　直美（正会員）

1999 年京都大学工学部情報工学科卒業．2001 年京都大学大学院情報学研究科数理工学専攻修士課程修了．同年日本電信電話（株）コミュニケーション科学基礎研究所入所．博士（情報学）．CSCW，HCI の研究に従事．

梶　　克彦（正会員）

2007 年名古屋大学大学院情報科学研究科博士課程修了．博士（情報科学）．同年 NTT コミュニケーション科学基礎研究所リサーチアソシエイトを経て，2010 年名古屋大学大学院工学研究科助教．人工知能学会，日本ソフトウェア科学会，ヒューマンインタフェース学会各会員．Web コンテンツ処理，音楽情報処理，実世界コミュニケーションに関する研究に従事．

葛岡　英明（正会員）

1992 年東京大学大学院工学系研究科情報工学専攻博士課程修了．博士（工学）．同年筑波大学構造工学系講師．現在，筑波大学大学院システム情報工学研究科教授．CSCW，HRI，HCI の研究に従事．

平田　圭二（正会員）

1987 年東京大学大学院工学系研究科情報工学専門課程博士課程修了．工学博士．同年 NTT 基礎研究所入所．1990〜1993 年（財）新世代コンピュータ技術開発機構（ICOT）．1999 年より NTT コミュニケーション科学基礎研究所．2001 年度論文賞，2003 年度山下記念研究賞，2005〜2007 年本会理事．現在，日本ソフトウェア科学会理事，t-Room プロジェクトリーダ．音楽情報処理に興味を持つ．

青柳　滋己（正会員）

1965 年生．1988 年東京工業大学理学部情報科学科卒業．1990 年 3 月同大学大学院理工学研究科情報科学専攻修士課程修了．同年日本電信電話株式会社入社．現在，同社コミュニケーション科学基礎研究所に所属．複合メディア情報処理の研究に従事．電子情報通信学会，日本ソフトウェア科学会各会員．

© 2011 Information Processing Society of Japan