

MIOSS : 3次元空間を重畳する鏡インタフェース MIOSS: Mirror Interface to Overlay Remote 3D Spaces

石井 亮[†] 小澤 史朗[†] 小島 明[†] 中野 有紀子^{††} 林 佑樹^{†††} 大塚 和弘[†]
Ryo Ishii Shiro Ozawa Akira Kojima Yukiko I. Nakano Yuki Hayashi Kazuhiro Otsuka

1. はじめに

我々は、遠隔地間において、お互いの実空間およびオブジェクトを共有しながら、部屋のレイアウトの相談や、空間中にある物の移動、運搬といった協調作業を円滑に行うことのできるメディアスペースの実現を目指している。そのようなシステムの実現するためには、2つの空間をどのように共有しユーザに映像として提示するかということが課題である。そこで考慮されるべきことは、互いの空間とユーザ同士の位置関係といった空間の整合性を保った表現を行うことである。

これに対して、我々は、対話相手の空間とユーザの空間が重なり、一つの空間となったかのような鏡映像を提示可能な鏡インタフェース MIOSS (Mirror Overlaid Shared Space) を提案する。MIOSS は、実際の鏡と同様に位置関係や大きさが再現されると共に、オブジェクトの遮蔽関係を正しく再現する映像をディスプレイに提示する。また、実際の鏡と同様にユーザの観察位置に応じた運動視差を再現する。このため、2つの空間のオブジェクトを共有しながらの協調作業において、互いの空間のオブジェクトの位置関係を正しく認知し、スムーズに作業を行えることが期待される。本稿では、提案する MIOSS および、MIOSS がオブジェクトの位置や大きさの認知が正しくできること、スムーズに作業が行えることを評価実験を行って確認したので報告をする。

2. 関連研究

従来から遠隔地にある実空間を共有し、遠隔指示や協調作業を支援する試みがなされている。例えば、SharedView[1]では、遠隔地のユーザの頭部に装着したカメラで撮影された視野映像をユーザ同士で共有し、ユーザがその映像をディスプレイで観察し手で指示を与えると、それを撮影した映像が遠隔地のユーザの頭部に設置されたディスプレイに表示される。これにより、ユーザは遠隔地のユーザの視野を共有し、指示を与えることができる。また、GestureMan[2] やテレグジスタンス[3]の研究では、遠隔地に代理ロボットを介してユーザが入り込みコミュニケーションをとることを提案している。これらのシステムは、どちらか一方がメインで対話相手に指示するようなシーンを想定している。また、ディスプレイを窓に見立てて、窓越しに2つの空間をシームレスに接合する窓越しインタフェース MoPaCo[4,5]では、映像中の人物から自然な視線や指差しといった非言語情報を円滑に伝えられることが示されている。2つの空間を窓越しに共有し、相互に指示を行うような作業が可能であるが、あくまでも窓越しでの指示

に留まり、一方の空間のオブジェクトを他方に移動することは想定されていない。

一方、2つの実空間を共有可能な通信システムの提案も行われている。Agora[6]では、互いの机上の共有を試みている。同室間通信 t-room[7]では、複数台のディスプレイに囲まれた同じ空間を各遠隔地に用意し、ディスプレイ上にカメラの遠隔地のユーザを表示することで、あたかも同じ部屋にいるかのような体験が可能である。ただし、あらかじめ決められた2次元的な共有面のみでの使用が前提となり、それ以外の領域での利用はできない。よって、空間全体のオブジェクトを共有することはできない。これに対して、HyperMirror (超鏡) [8]では、ディスプレイ面を鏡に見立てて、遠隔地の空間を合成し表示することが提案されている。このシステムでは、カメラで撮影された映像から、互いの空間の映像を重畳して表示することで、撮影された空間全体を共有可能である。また、映像提示ディスプレイとして、各地点で平面ディスプレイのみという小規模なシステム構成で実現できることに利点がある。しかしながら、一方の空間が後方、もう一方の空間が前方といったように、単純に画像を重ねて表示を行うため、奥行き位置に応じたオブジェクトの配置は考慮されていなかった。また、前後の動きによって被写体の大きさが変化するため、互いの空間の大きさや位置関係を正しく表現することはできなかった。

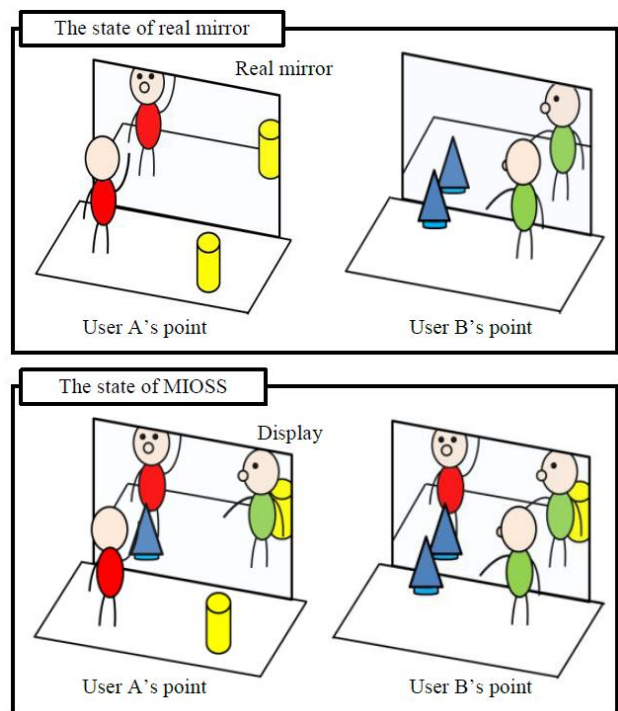


図1 鏡状況とMIOSS状況での映像表現イメージ

[†] 日本電信電話株式会社 NTT Corporation

^{††} 成蹊大学工学部 Seikei University

^{†††} 大阪府立大学 Osaka Prefecture University

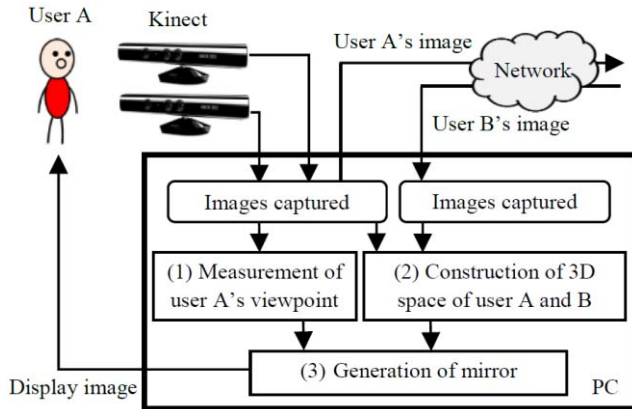


図 2 MIOSS のシステム構成

これに対して、本論文で提案する MIOSS は、HyperMirror を拡張し、実際の鏡越しに空間が重畳されたように、大きさや位置関係が考慮され、さらに観察位置に応じた運動視差を実現する。本論文では、MIOSS の提案に加えて、プロトタイプシステムを用いた評価実験を通じて、鏡メタファを用いた映像表現の、円滑な遠隔協調作業の実現への有用性を探る。

3. MIOSS の提案

3.1 システム概要

鏡状況と MIOSS が実現する映像表現のイメージを図 1 に示す。鏡状況では、ユーザ A とユーザ B のそれぞれの空間のオブジェクトが鏡に表示されている。MIOSS 状況では、ユーザ A の空間とユーザ B の空間が合成されて、鏡中で映像を観察している様子である。ユーザ A とユーザ B の空間が、あたかも重なって同じ空間になったかのように、鏡映像が提示される。このとき、鏡面に対して、ユーザやオブジェクトの大きさや位置関係を正しく考慮し表示される。たとえば、ユーザ A 側にある黄色い円柱形のオブジェクトはユーザ B の後方に表示され、ユーザ B 側の青い家の形をしたオブジェクトはユーザ A の前方に配置されている。

次に、MIOSS のシステム構成を図 2 に示す。本システムは、各地点において RGB 画像および距離画像の撮影機器として Kinect を 2 台、映像表示機器として 2D ディスプレイまたはプロジェクタを使用する。ユーザ A にユーザ B (対話相手) の窓越し映像を提示するためには、ユーザ A, B の RGB 画像、距離画像を用いて、大きく以下の 3 つの処理を行う。

- (1) ユーザの観察位置の 3 次元位置の測定
 - (2) ユーザ A, B の空間の 3 次元モデルの構築
 - (3) 観察位置に応じた鏡映像の生成
- 次に具体的な処理について述べる。

3.2 ユーザの視点位置の 3 次元位置の測定

本システムでは、ユーザの視点位置を両眼の中心位置と定義し、1 台の固定カメラの映像からカメラ座標系における 3 次元的なユーザの視点位置 (X, Y, Z) を既存の顔検出技術を利用して測定する。

まず顔の検出処理には、状態履歴を参照したパーティクルフィルタによる顔追尾手法[9]を用いた。本技術を用いる

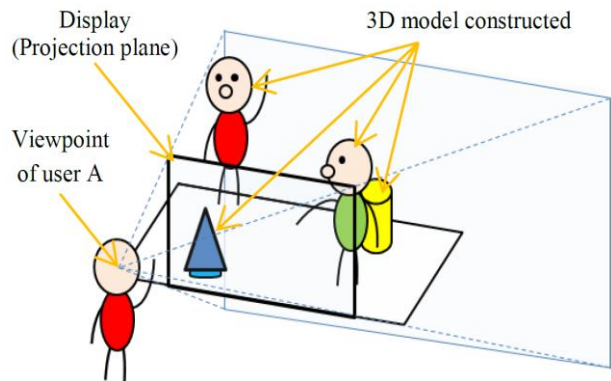


図 3 観察位置に応じた鏡映像の生成

ことで、画像上の両眼の中心位置 (E_x, E_y) がリアルタイムに測定される。

次に、カメラ座標系におけるユーザの視点位置 (X, Y, Z) を算出する。まず、ユーザの視点位置の奥行 Z の算出方法を示す。OpenNI ライブラリ[10]に用意されている、RGB 画像と距離画像の対応点を取得する機能を利用して、両眼の中心位置 (E_x, E_y) に対応する奥行き位置を取得することで Z が求まる。次に、 X と Y は、画像上の両眼の中心位置 (E_x, E_y) 、カメラの画角 θ_x , θ_y 、画像サイズ W , H 、そしてユーザの視点位置の奥行 Z により、次式 (1) で算出可能である。

$$\begin{aligned} X &= Z * \tan(\theta_x/2) * (W/2 - E_x), \\ Y &= Z * \tan(\theta_y/2) * (W/2 - E_y) \dots \quad (1) \end{aligned}$$

以上により、カメラ座標系における 3 次元的なユーザの視点位置 (X, Y, Z) を算出する。

3.3 ユーザ A, B の空間の 3 次元モデルの構築

ユーザ A, B の空間にそれぞれ設置された 2 台の Kinect から各ユーザの空間の 3 次元モデルを構築し、次にユーザ A, B の空間の 3 次元モデルを統合する。

まず、各ユーザの空間の 3 次元モデルの構築方法として、Point Cloud Library (PCL) [11]の機能を利用して、Kinect の RGB 画像と距離画像を入力として、点群データとして空間の 3 次元モデルを構築した。2 台の Kinect で生成された点群データは、事前に 2 台の Kinect の位置のキャリブレーション[12]を行っておき、その結果を基に統合する。

最後に、ユーザ A, B の空間の 3 次元モデルを合成する。このとき、互いの Kinect と映像を表示するディスプレイ面の位置を考慮し、ディスプレイ面を基点として位置関係が保持されるようにする。

3.4 観察位置に応じた鏡映像の生成

次に、図 3 のように、ディスプレイを投影面として、3.3 節で生成した 3 次元モデルを 3.2 節で取得したユーザの視点位置に応じて、透視投影変換を行うことで視点位置に応じた映像を生成する。このとき、Kinect カメラとディスプレイの位置関係を考慮して、座標系を統一する。

3.5 実装結果

上記の提案手法を用いて MIOSS のプロトタイプを実装した。実装環境および性能を表 1 に示す。表 1 中の“ユー

ザの動作からの遅延”は、ユーザ視点位置が動いてから映像に視差が反映されるまでの時間、“カメラ画像の遅延”はキャプチャ画像が提示されるまでの時間である。

このとき、映像コミュニケーションシステムでは、“カメラ画像の遅延”に通常、ネットワーク遅延が発生するが、地点間のデータはネットワーク経由ではなく直結をしたため、“カメラ画像の遅延”は映像のネットワーク伝送による遅延を含まない。

表 1 実装環境および性能

実装環境	CPU	Intel Core™ i7 -3960X
	メモリ	16GB
	グラフィックボード	NVIDIA GeForce GTX580
性能	Kinect の 入力画像サイズ	640×480 (RGB) 640×480 (デプス)
	描画速度	15fps
	ユーザ動作からの遅延	約 500ms
	カメラ映像の遅延	約 500ms

4. オブジェクトの位置の知覚に関する実験

4.1 実験内容

MIOSS が実現するような空間の 3 次元処理による奥行き位置と遮蔽関係の表現と、運動視差による観察位置による正しい見え方が、対話相手の空間のオブジェクトの位置の正しい認識に寄与するかを評価する実験を行った。実験条件として、以下のような 2D 条件、3D 条件、3DM 条件を設定した。

- **2D 条件**：Kinect で撮影した被験者側の RGB 画像から、人物とオブジェクトのみを抽出し、対話相手側の RGB 画像の上に合成されるシステムを実装し用いる。そのため、被験者が前後に移動すると、被験者の大きさは実際とは異なる大きさで表示される。また遮蔽関係も考慮されず、前後に移動しても常に対話相手の映像の前方に表示される。すなわち、HyperMirror[8] と同様の見え方が再現される。
- **3D 条件**：MIOSS から運動視差を除外したもの。すなわち、大きさ、位置関係が正しく考慮され、遮蔽関係が再現される。
- **3DM (MIOSS) 条件**：MIOSS を使用。大きさ、奥行き位置が正しく考慮され、遮蔽関係が再現される。また、ユーザの視点位置に応じた運動視差が付与される。

2D 条件と 3D 条件を比較することで、空間の 3 次元処理による奥行き位置と遮蔽関係の表現の効果を検証する。また、3D 条件と 3DM 条件を比較することで、前後の動きに伴う大きさの変化と遮蔽関係の表現に加えて、ユーザの視点位置に応じた運動視差が再現されることの効果を検証する。さらに、2D 条件と 3DM 条件を比較することで、奥行き位置と遮蔽関係の表現とユーザの視点位置に応じた運動視差の複合的な効果を検証する。

実験環境を図 4 に示す。400×700cm の部屋の真ん中をカーテンで仕切り、400×350cm の 2 つの部屋に分けた。それぞれの部屋のカーテンの前に、100 インチ (縦 125cm × 横 221cm) のスクリーンを床面から 60cm の高さに設置した。スクリーンにはプロジェクタ (NP-U310WJD) から 80 インチ (縦 108cm × 横 172cm) の映像 (解像度

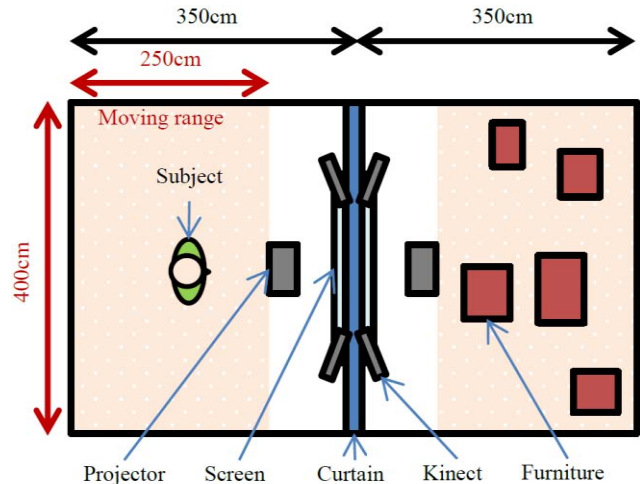


図 4 実験環境

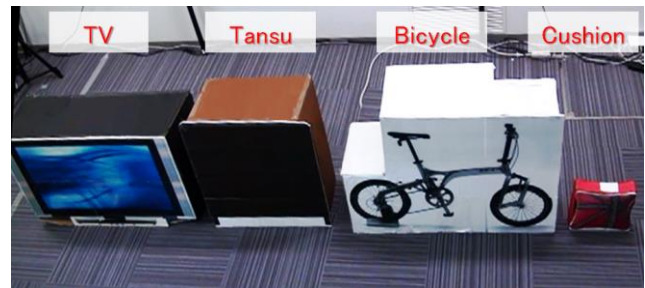


図 5 使用したオブジェクトの例

1280×800) が投射された。Kinect は各部屋のスクリーン上部に 2 台ずつ設置した。被験者のいない対地側には、5 つのオブジェクトが配置された。オブジェクトには、実寸大の家具をモチーフにして、家具の写真が貼られた立方体の紙箱を用いた。図 5 に、用いたオブジェクトの一例を示す。家具は異なる 5 個の家具の組み合わせを 3 グループ用意した。配置位置は、家具の組み合わせごとに全体的に同程度に分散するように設定した。なお、Kinect センサの画面の都合上、被験者が移動可能な領域を横 400cm、スクリーンから 100cm 離れた箇所から奥行き 250cm の範囲とした。

実験手順として、開始の合図とともにスクリーンに映像が映し出された後、被験者は 3 分以内に遠隔地の部屋に置かれたオブジェクトの位置と大きさを回答用紙に記入した。回答用紙には、上方から見た部屋の俯瞰図が図示されており、その図中の床面には 5cm 間隔の方眼が示されていた。被験者は、その俯瞰図内に配置されたオブジェクトの位置と大きさを記入した。

試行順序は、順序効果を考慮して実験条件および家具の組み合わせ条件をランダムに選択し、各条件それぞれ 1 試行ずつ行った。

4.2 実験結果

被験者 16 人 (20 代の男性 12 人、女性 4 人) で実験を施行した。まず、各実験条件において被験者が回答したオブジェクトの中心位置と、実際の配置位置の中心との誤差を横方向 (スクリーンに平行方向) と奥行き方向 (スクリ

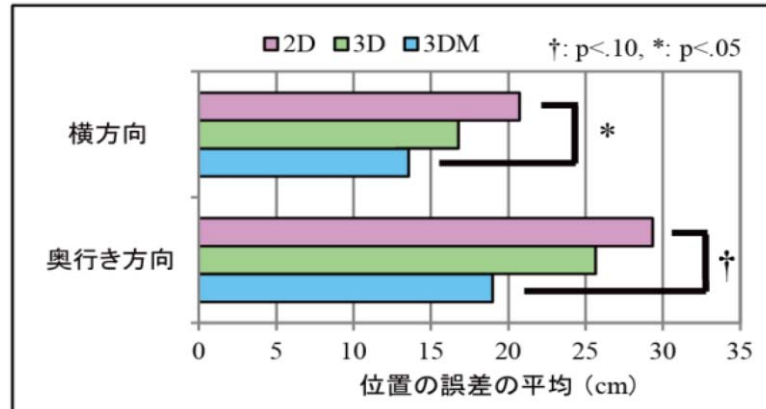
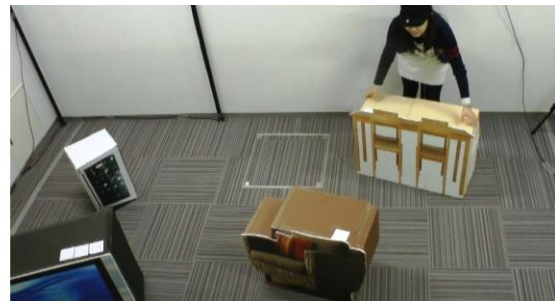


図6 被験者の知覚位置と実際の位置の誤差

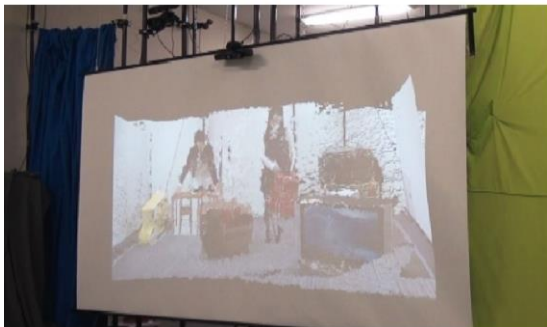
(a) 被験者 A の部屋の様子



(b) 被験者 B の部屋の様子



(c) 被験者 A のスクリーンの様子



(d) 被験者 B のスクリーンの様子



図7 協調作業シーンでの評価実験の様子

ーンに垂直方向)それぞれで集計し、誤差の平均と標準偏差を求めた。その結果を図6に示す。また、実験条件の因子が、誤差に影響を与えるかを検証するために、3つの実験条件ごとに、繰り返しの一元配置の分散分析を行った。その結果、実験条件の要因は有意であった(横方向: $F(2,45)=3.55, p<.05$, 縦方向: $F(2,45)=2.51, p<.10$)¹。

次に、実験条件の要因に効果が確認されたため、3つの実験条件について、Tukey-Kramer法による多重比較を行った。その結果、2D条件と3DM条件の間で有意差または有意傾向(横方向で $p<.05$, 奥行き方向で $p<.10$)が認められた。よって、遠隔地の空間にあるオブジェクトの知覚位置と実際の位置との誤差は、2D条件に比べて3DM条件で小さくなること示唆された。

¹ 分散分析の結果は、「 F (グループ間の自由度,グループ内の自由度) = F 値」の書式で示した。

5. 協調作業シーンでの評価実験

5.1 実験内容

次に、MIOSSが実現するような2つの空間と物を共有して行う協調作業時に、MIOSSがどのような影響を与えるかを検証するために、被験者2人による実験を実施した。実験条件は、4章の実験と同様に2D条件、3D条件、3DM条件を設定した。

実験環境は図4と同様であり、2つの部屋にそれぞれ被験者を一人ずつ配置した。それぞれの部屋には、先の実験と同様な4つのオブジェクトを被験者が移動可能な領域の中央に横並びに配置した。家具は異なる8個の家具の組み合わせを3グループ用意した。

実験手順として、開始の合図とともにスクリーンに映像が映し出された後、被験者は10分以内に2つの部屋にあ

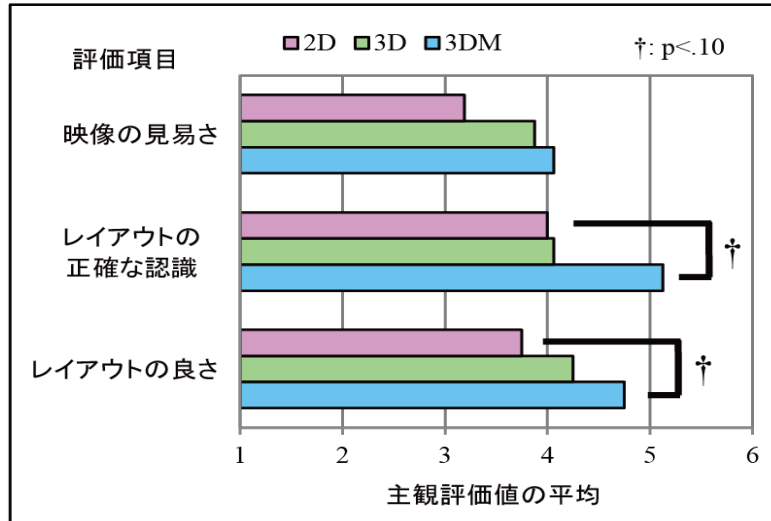


図 8 主観評価の結果

る 8 つの家具を自由に移動させて、良いレイアウトを協力して作成するよう指示が出された。制限時間は 10 分間であった。実験終了後、実験操作者が、被験者によって配置された 2 つの部屋の家具を、実際の配置位置どおりに片方の部屋に配置し、共同作業によって作られたレイアウトを実際に再現した。これを、被験者にはよく観察してもらった。これは、実験施行後に行うアンケート内容に関連するためである。レイアウトを観察した後、映像や実験中の作業、作成したレイアウトに関する印象を問う主観評価（リッカート法による 1～6 点の 6 段階評価）を実施した。具体的な 3 つの質問項目を以下に示す。

- **映像の見易さ**：映像は見易かったですか？（映像変化による酔いや煩わしさ、画質に問題はなかったか？）
- **レイアウトの正確な認識**：（実際に再現されたレイアウトを見てみて、）実験中に映像を見ながら作業をしていた時に認識していたレイアウトと、実際の家具のレイアウトは一致していましたか？
- **レイアウトの良さ**：（実際に再現されたレイアウトを見てみて、）完成したレイアウトは良いものができましたか？

試行順序は、順序効果を考慮して実験条件および家具の組み合わせ条件をランダムに選択し、各条件それぞれ 1 試行ずつ行った。実験の様子を例を図 7 に示す。図 7 は、3DM 条件下で実験を行っている際の、スクリーン上部から部屋を撮影したもの（(a)、(b)）と、スクリーンの斜め前方からスクリーンを撮影したもの（(c)、(d)）である。

5.2 主観評価の結果

4 章の実験と同じ被験者 16 人を 2 人 1 組のペアとして、8 組の被験者の組ごとに実験を実施した。その主観評価値の平均を図 8 に示す。まず、実験条件の因子が、主観評価値に影響を与えるかを検証するために、3 つの評価項目ごとに、繰り返しのある一元配置の分散分析を行った。その結果、実験条件の要因は、レイアウトの正確な認識と、レイアウトの良さの 2 つの項目においてのみ有意傾向が認められた（レイアウトの正確な理解： $F(2,45)=2.52, p<.10$,

レイアウトの良さ： $F(2,45)=2.58, p<.10$ ）。映像の見やすさの項目では、有意差は見られなかった。

次に、実験条件の要因に効果が確認されたため、3 つの実験条件について、Tukey-Kramer 法による多重比較を行った。その結果、2D 条件と 3DM 条件の間で有意傾向が認められた（レイアウトの正確な理解： $p<.10$ 、レイアウトの良さ： $p<.10$ ）。よって、遠隔地の空間にあるオブジェクトの知覚位置と実際の位置との誤差は、2D 条件に比べて 3DM 条件で小さくなることが示唆された。

6. 考察

対話相手の空間のオブジェクトの位置の正しい認識に寄与するかを評価する実験において、2D 条件に比べて 3DM 条件で物体の位置の誤差が小さかったことから、MIOSS が実現するような空間の 3 次元処理による奥行き位置と遮蔽関係の表現と、運動視差による観察位置による正しい見え方の両方を備えることで、対話相手の空間のオブジェクトの位置や大きさをより正しく知覚させることができると考えられる。3DM 条件での平均誤差は、横方向で約 13.5cm、縦方向で約 19cm であり、人手による書き取りの際の誤差も考えると、かなり正確な位置をユーザは知覚できると考えられる。

次に、協調作業実験において、まず映像の見易さに関する主観評価結果は、どの条件間でも有意な差は見られなかった。特に 3DM 条件では、運動視差によって映像が絶えず変化するため、酔いや煩わしさなどの負の影響を与える可能性も懸念されるが、それらの影響は見られなかったと考えられる。これは、MIOSS と同様に運動視差を付与する映像表現[4]によっても、同様の結果が示されている。ユーザの視点位置に合わせて、映像に運動視差を与えることは人間にとって自然な視覚効果であると考えられる。レイアウトの正確な認識、レイアウトの良さに関する主観評価結果は、2D 条件に比べて 3DM 条件で向上した。レイアウトの正確な認識については、オブジェクトの位置の正しい認識に寄与するかを評価実験の結果で示された通りの結果が、実際の作業時にも見られたと考えられる。また、3DM 条件の平均評価値は約 5.15 と非常に高く、MIOSS によりレ

レイアウトを正確に把握することが可能であることが示された。位置の知覚の正確さや、容易さの向上が、協調作業によって完成したレイアウトの良さの評価値の向上に貢献したのではないかと考えられる。以上より、MIOSS を利用することで、2つの空間を共有し、あたかも一つの空間になったかのように、円滑に協調作業を行うことの一手段として有効である可能性が示唆された。

今後は、

7. まとめ

本稿では、空間を3次元的に処理して合成することで、奥行きや位置関係を考慮して、2つの空間が合わさって一つの空間に重畳されたかのような鏡映像を表示する鏡インタフェース MIOSS を提案した。評価実験から、MIOSS が実現する奥行きや位置関係および運動視差を考慮する鏡映像によって、オブジェクトの位置の認識の正確性が向上することが確認された。また MIOSS では、遠隔地間で部屋のレイアウトを相談する協調作業実験において、映像から実際のレイアウトを正しく認識でき、作業後のレイアウトがより良いものができる可能性が示唆された。

参考文献

- [1] H. Kuzuoka, "Spatial workspace collaboration: a SharedView video support system for remote collaboration capability," in Proc. CHI '92, pp.533-540, 1992.
- [2] H. Kuzuoka, J. Kosaka, K. Yamazaki, Y. Suga, A. Yamazaki, P. Luff, and C. Heath, "Mediating Dual. Ecologies," in Proc. CSCW2004, pp.477-486, 2004.
- [3] M. Minsky, "Telepresence," OMNI magazine, June, 1980.
- [4] 石井亮, 小澤史朗, 川村春美, 小島明, 中野有紀子, "映像コミュニケーションにおける窓越しインタフェースM o P a C o によるテレプレゼンスの増強," 電子情報通信学会論文誌, Vol.J96-D, No.1, pp.110-119, 2013.
- [5] 石井亮, 小澤史朗, 川村春美, 小島明, 中野有紀子, "窓越しインタフェースM o P a C o による指示作業への効果検証," 電子情報通信学会論文誌, Vol.J96-D, No.12, pp. 3044-3054, 2013.
- [6] H. Kuzuoka, J. Yamashita, K. Yamazaki, and A. Yamazaki, "Agora: A Remote Collaboration System that Enables Mutual Monitoring," in CHI'99 Extended Abstracts, pp.190-191, 1999.
- [7] K. Hirata, K. Kaji, Y. Harada, N. Yamashita, and S. Aoyagi, "t-Room: Remote Collaboration Apparatus Enhancing Spatio-Temporal Experiences", Adjunct Electronic Proceedings of Computer Supported Collaborative Work (CSCW'08), 2008.
- [8] O. Morikawa and T. Maesako, "HyperMirror: Toward Pleasant-to-use Video Mediated Communication System," Proceedings of Computer Supported Collaborative Work (CSCW'98) pp.149-158, 1998.
- [9] 三上 弾, 大塚和弘, 大和淳司, "メモリベースパーティクルフィルタ: 状態履歴に基づく事前分布予測を用いた頑健な対象追跡," 電子情報通信学会論文誌, Vol.J93-D, No.8, 2010.
- [10] Open NI (The standard framework of 3D sensing), <http://www.openni.org/>
- [11] PCL (The point cloud library), <http://pointclouds.org/>
- [12] Zhengyou Zhang, "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.