

# 可動式カメラを用いたメディアスペースの開発 Development of a Media Space with Movable Cameras

村上 友樹  
Yuki Murakami

中西 英之  
Hideyuki Nakanishi

加藤 慶  
Kato Kei

大阪大学大学院 工学研究科 知能・機能創成工学専攻  
Department of Adaptive Machine Systems, Graduate School of Engineering, Osaka University

Media spaces are expected to enable social interaction between distributed sites much like face-to-face interactions do. Media spaces, however, still have some problems. We developed a media space using a movable camera which moves forward at a remote site when a local user approached the display. The camera's movement generates motion parallax. We conducted an experiment and observed that a movable camera enhanced telepresence, which is the feeling of facing a remote person in the same room. A zoom-in camera is also able to enlarge the remote person's image as the movable camera does. However, the zoom-in camera did not enhance telepresence, most likely because of a lack of motion parallax. If a remote person initiated the camera's movement, it was found that the movable camera did not enhance telepresence.

## 1. はじめに

大型ディスプレイやネットワークカメラの普及によって、メディアスペースの実装は容易になってきた。メディアスペースとは映像と音声を常時接続することで遠隔地間を接続するシステムである[Bly 93]。メディアスペースによって、対面環境のようなインフォーマルな社会的インタラクションが、遠隔地間において促進されることが期待されている。しかし、それは完全には実現されていない[Jancke 01]。対面環境で実際に人と出会った場合、メディアスペースを通して出会った場合より頻繁に会話が行われるという報告がある[Fish 90]。メディアスペースによって2つのオフィスを接続した場合においても、別のオフィスにいる人との会話より、同じオフィスにいる人との会話の方が多いうことが報告されている[Mantei 91]。また、メディアスペースにおいては、映像の中の人間の体の動きは、実空間での体の動きよりも目立たないので、遠隔地の人の行動を把握するのは難しい [Heath 92]。以上のことは、メディアスペースは対面環境より人と対面している感覚が小さいということを示している。

メディアスペースの古典的な研究における主題は、離れたオフィス間のアウェアネス支援だった。近年においては、アウェアネス支援に加えて、遠隔地間インタラクションの社会的な側面に着目した研究がなされている。例えば、離れたパーティ会場間において陽気な雰囲気共有するために、卓上ディスプレイを用いたという研究がある[Roussel 02]。ユーザがテーブルの周りに集まるという行動は、体と体が接近するという感覚を生じさせる。また、ユーザが画面に近づくとユーザの映像を鮮明にするという機構によって、体と体が接近するという感覚を増幅するという研究もある[Roussel 04]。ユーザの体の動きを利用してユーザをより社会的にすることも提案されている[Karahalios 04]。ユーザが体を動かした際に、ビデオ映像を用いたインタラクションがより社会的になったということも観測されている[Mueller 03]。これらの研究は、ディスプレイの周りにおけるユーザの自発的な動きを利用することによって、ビデオ映像を用いたインタラクションを改善することを目的としていた。本研究においては、そのようなユーザの動きを用いて、ユーザが遠隔地の対話者と対面

している感覚を強める方法を示す。

リアルタイムのビデオ映像がテレプレゼンスを強化するということが知られている。本論文で着目しているテレプレゼンスとは遠隔地の人間と同じ空間で対面しているような感覚、あるいは単純に対面環境にどれほど近いかの度合いである[de Greef 01]。テレプレゼンスはアイコンタクト[Bondareva 04]、等身大映像、立体視映像[Prussog 94]によって強化される。等身大の映像を用いると、4分の1の映像を用いる場合より会話が促進されたという報告がある[Mantei 91]。しかし、アイコンタクトや等身大映像に比べて、立体視映像を用いたメディアスペースの実装例は少ない。これは、特に公共の場にメディアスペースを設置する場合などにおいて、利用する際に眼鏡をかけなければならないことが不便であるためだと考えられる。大型かつ眼鏡不要の立体視映像を用いることでこの問題は解決できるが、そのような立体視映像を実現する装置は依然として高価であるため、実装するのは難しい。

奥行き情報は立体視映像以外に、運動視差によっても伝達することができる。ラリーカーの操縦席に設置したカメラの映像を見た被験者は、立体視映像を見る場合よりも大きなテレプレゼンスを感じたという報告がある[Ljsselsteijn 01]。我々の先行研究においても、カメラを搭載したロボットの移動が引き起こす運動視差がテレプレゼンスを強化することが示されている[Nakanishi 08]。この先行研究のもうひとつの重要な点は、運動視差によるテレプレゼンスの強化は、ロボットが自動的に移動した場合には見られなかったという点である。テレプレゼンスには3つの独立した要素がある。それは知覚情報、センサのコントロール、環境変更能力である[Sheridan 95]。恐らく、ロボットが自動的に移動した場合、センサのコントロール、環境変更能力の2つの要素が欠けることになるため、テレプレゼンスが強化されなかったのであろうと考えられる。また、神経学的な既存研究によると、人が自己発生的動作と外部発生的動作を区別することができる理由は、予測フィードフォワードモデルによって説明することができることとされている[Vogeley 03]。このことから、ロボットが自動的に移動する場合は、ユーザはその移動が予測できないので、移動は外部発生的動作であると認識されるため、テレプレゼンスが強化されないのではないかと考えられる。

本研究ではメディアスペースにおいて、移動可能なカメラを用いることでテレプレゼンスを強化することを目的とする。本研究で想定しているメディアスペースとは、ディスプレイに近づくだ

連絡先: 中西 英之, 大阪大学大学院 工学研究科 知能・機能創成工学専攻, nakanishi@ams.eng.osaka-u.ac.jp

けで使用可能で、ビデオ会議を始める際の呼び出し操作などは必要としないものである。我々はディスプレイに近づくだけで使用可能であるという利点を損なわずにテレプレゼンスを強化することを試みた。テレプレゼンスを強化するには、カメラの移動はユーザに予測可能でなければならない。しかし、カメラは操作されるべきではない。我々はユーザのディスプレイへの接近動作を利用することでこの問題を解決した。

## 2. システム設計

我々は以下の 3 つの条件に従ってメディアスペースを設計した。1)ユーザの動きを取得してカメラを移動させ、それによって生じる運動視差によりテレプレゼンスを強化する。2)ユーザがビデオ映像の変化を予想でき、ユーザ自身が映像の変化を引き起こしたと感ずることができるように、カメラの移動はユーザに同期する。3)キーボードやマウスなどを用いた操作は付け加えず、通常のメディアスペースと同様に使用できる。

デザインを図1に示す。カメラはもう一方の部屋のユーザとディスプレイとの間の距離に従って移動する。ユーザとディスプレイとの距離が小さくなるとカメラは前方に移動し、大きくなるとカメラは後方に移動する。図1は、ユーザ A がディスプレイに近づき、ユーザ B の部屋のカメラが前方に移動している状態を示している。ディスプレイの中のユーザの映像は両方の部屋で大きくなるが、カメラが動くのはユーザ B 側の部屋のみであるので、運動視差が生じるのはユーザ A 側のディスプレイの中のユーザ B の映像のみである。図中の箱の絵が運動視差を表している。ユーザ B 側のディスプレイにおけるユーザ A の映像の拡大は、ユーザ A の接近動作によるものであり、運動視差を伴わない。したがって、ユーザ A の感じるユーザ B の存在感のみが運動視差によって強化される。以下、設計の方針を示す。

### 2.1 ユーザの動作の増幅

メディアスペースを使用する際、ユーザはディスプレイへと接近する。我々は、メディアスペースを使用する際に自然と行われるこの動作を増幅することにした。カメラの移動はユーザの動きに比例するように設定した。これにより、ユーザの目に映るもう一方のユーザの映像が拡大され、ユーザの接近動作が増幅される。

逆に、ユーザが近づいたときにカメラが後方に移動するようにした場合、ユーザの接近動作は増幅されるのではなく打ち消される。しかし、この場合も運動視差は発生する。我々は予備的な実験を行ってこれを試したが、不自然であり、テレプレゼンスが強化されるようには思えなかった。また、カメラを前方に移動させつつズームアウトを行うと、ディスプレイ上のユーザの映像の

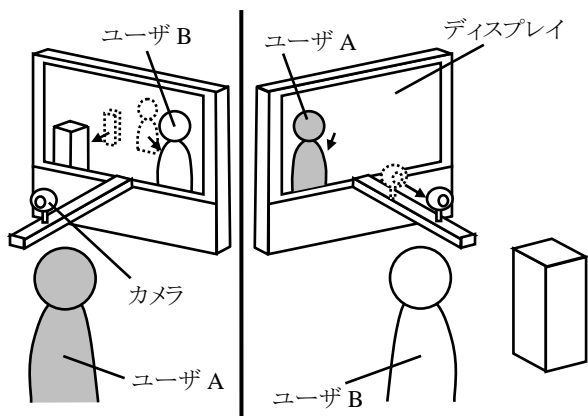


図 1. 可動式カメラを用いたメディアスペース

大きさを変化させずに運動視差を発生させることができるが、この場合も不自然に思えた。

### 2.2 知覚インタフェース

ビデオ会議において、視点を動かす方法が 3 つある。1)ユーザが直接カメラの位置を操作する[Kuzuoka 04]。2)ユーザの行動をウェアラブルなセンサによって取得し、その値を用いる[Morita 07]。3)ユーザの行動を環境中のセンサによって取得し、その値を用いる[Gaver 95]。我々は通常のメディアスペースと同様に使用できるように、3)の方法を選択した。

ユーザの頭の三次元的な位置を、ユーザに何もデバイスを取り付けずに取得するのは依然として困難である[Gaver 95]。しかし、ユーザとディスプレイとの間の距離ならば、ユーザにデバイスを取り付けなくても、レーザレンジファインダなどのセンサを用いて比較的正確に取得することができる。

### 2.3 カメラの可動範囲の最小化

本論文で提案するメディアスペースを容易に導入できるように、我々はカメラの可動範囲を最小化した。まず、カメラの移動は前後移動のみの一次元にし、ユーザの動きのディスプレイに垂直方向成分に同期させた。ユーザがディスプレイの前を動き回る際、ユーザは垂直方向だけではなく水平方向にも動くので、それを利用してカメラを左右に動かすこともできる[Gaver 95]。二次元の動きはより大きな運動視差を発生させることができるが、可動範囲が大きくなってしまいうので、水平方向の動きは用いないことにした。

次に我々はカメラの移動距離を最小化した。我々の先行研究においてはカメラを 1m 移動させていたので、我々は 1m 未満の距離にすることを試みた。予備実験において、テレプレゼンスの強化は移動距離が 750mm の際に見られた。移動距離は長い方がテレプレゼンスを強化する効果は大きくなるであろうと考えられるが、我々はこの距離を用いることにした。

### 2.4 ユーザとの同期

我々はユーザのディスプレイへの接近動作に同期させてカメラを移動させた。これはユーザにカメラの移動操作をさせずに、相手の映像が拡大されることを予測させるためである。ユーザに同期する場合とユーザが操作する場合の違いで最も重要なことはユーザの意図があるかどうかである。ユーザがカメラを操作する場合、人や物を見ようというユーザの意図が存在するが、カメラがユーザの動作に従って自動的に移動する場合、ユーザは単に自分の動きとカメラの移動が関係していることに気付くだけであり、カメラが移動する目的は知らない。本論文で提案するメ



図 2. 実験環境

ディアスペースにおいては、ユーザが可動式カメラのメカニズムを知る必要は無い。後に述べる実験においては、我々は実験が終了するまでは被験者に対してカメラが移動することは伝えなかった。

### 3. 実験

学部生の被験者に、研究室とは別の部屋から、メディアスペースを用いて説明者から 3 台のロボットの説明を聞くというタスクを行ってもらった。

図 2 に実験の様子を示す。被験者側の部屋には通常のネットワークカメラを設置し、説明者側の部屋には可動式カメラを設置した。可動式カメラはネットワークカメラを直動位置決めテーブルに装着したものである。直動位置決めテーブルは、被験者側の部屋の机の上に設置したレーザレンジファインダの値に従って動作する。机の約 1.2m 後ろには 65 インチのプラズマディスプレイを設置した。ディスプレイには可動式カメラからのリアルタイム映像を表示した。映像の解像度は 704x480 ピクセルであり、フォーマットは Motion JPEG である。各画像は約 17 キロバイトでフレームレートは 26fps、水平画角は 52 度であった。

被験者は実験開始時には机から 3m 以上離れた位置にある椅子に座っていた。実験開始後、説明者はロボットを説明する前に、被験者にディスプレイに近づくように言った。ロボットの説明を受けている最中は、被験者は机の前から 750mm の範囲内に立っていた。机の前から 750mm の位置から 2750mm の位置の間の 2m の範囲において、被験者の位置をレーザレンジファインダで取得した。説明者はロボットの説明が終わる度に、被験者にもう一度椅子に座るように言った。説明者は 3 台のロボットを説明したので、被験者は 3 回ディスプレイへの接近動作を行った。

#### 3.1 仮説と実験条件

実験では 4 つの条件を比較した。各条件において実験環境は同様であり、被験者のディスプレイへの接近動作とカメラの移動との関係のみが各条件で違った。

本論文で提案した設計をそのまま反映した条件として、同期条件を設定した。この条件においては、カメラは被験者のディスプレイへの接近動作に比例して 750mm 移動する。カメラが説明者から最も遠い状態と最も近い状態を比較すると、被験者が見る説明者の映像は約 2 倍になっていた。

我々は 3 つの仮説を立てた。1 つ目の仮説はカメラが被験者のディスプレイへの接近動作に合わせて前後移動すると、被験者の感じるテレプレゼンス、すなわち説明者と同じ部屋で対面し

ている感覚が強化されるというものである。この仮説を示すために、我々は固定条件を設定し、同期条件と比較した。固定条件においては、被験者の位置に関わらずカメラは常に説明者に最も接近した状態であった。

2 つ目の仮説は、被験者がカメラの移動が自己発生的でないと認識すると仮説 1 の効果は消えるというものである。カメラの移動が外部発生的である場合、被験者は運動視差の効果を予測不可能であり、テレプレゼンスは強化されないと考えられる。この仮説を示すために、我々は非同期条件を追加した。非同期条件では、カメラは被験者の接近動作と同期せず、被験者には説明者がカメラを移動させたことと認識させた。具体的には、被験者がディスプレイへの接近を完了した後、説明者はカメラに向けてリモコンのボタンを押した。カメラはその説明者の動作のあと前方に移動した。

3 つ目の仮説はカメラの前後移動ではなくズームイン、ズームアウトで説明者の映像の拡大縮小を行った場合、仮説 1 の効果は無くなるというものである。ズームイン、ズームアウトは説明者の映像を大きくする効果はあるが、運動視差は発生しない。従ってテレプレゼンスは強化されないと考えられる。この仮説を確かめるためにズーム条件を追加した。ズーム条件ではカメラは被験者のディスプレイへの接近動作に比例してズームイン、ズームアウトを行った。最もズームインした際には被験者から見た説明者の映像は同期条件と同様に約 2 倍になるように設定した。同期条件とズーム条件における、被験者が見る説明者の映像の違いは図 3 のようになった。

#### 3.2 実験結果

固定条件、ズーム条件、非同期条件、同期条件を比較した結果を示す。各条件 7 人ずつの被験者にメディアスペースを使用してもらい、アンケートによる主観評価を用いて比較した。アンケートは各項目に対して 9 段階のリッカート尺度を用いた。4 つの条件は被験者間要因の一元配置の分散分析と Tukey の多重比較を用いて比較した。

同じ部屋の中で実際に会話している感じがしたという項目に有意差が見られた( $F(3,24)=8.980, p<0.001$ )。多重比較によると、この感覚は同期条件の場合は固定条件より有意に強く( $p<0.001$ )、ズーム条件よりも有意に強かった( $p<0.01$ )。また、非同期条件よりも有意に強かった( $p<0.01$ )。固定条件とズーム条件と非同期条件の間には有意差は見られなかった。同じ部屋の中で実際に説明者を眺めている感じがしたという項目に有意差が見られた( $F(3,24)=5.398, p<0.01$ )。多重比較によると、この感覚は同期条件の場合は固定条件より有意に強く( $p<0.05$ )、非同期条件よりも有意に強かった( $p<0.01$ )。固定条件とズーム条件と非同期条件の間には有意差は見られなかった。

以下の項目に関しては有意差が見られなかった。映像は十分綺麗だと感じた( $F(3,24)=0.417$ )。音声は会話を行う上で問題が無かった( $F(3,24)=0.085$ )。説明者の振る舞いが良く分かった( $F(3,24)=0.321$ )。ロボットの説明は良く分かった( $F(3,24)=1.917$ )。このことから、条件間の説明の質の違いによってテレプレゼンスの違いが発生しているわけではないことが分かった。

加えて、説明者側のカメラが動いていたことに気がついたかどうかをアンケートで質問した。その結果、固定条件以外のすべての被験者は、映像の変化には気付いたが、カメラが前後に移動していたのではなくズームしていたと思ったと答えた。

カメラの動きはあった方が良いという項目に有意傾向が見られた( $F(2,18)=2.630, p<0.1$ )。多重比較によると、同期条件の場合の方がズーム条件の場合より映像の変化が好意的に受け取られる傾向があった( $p=0.103$ )。

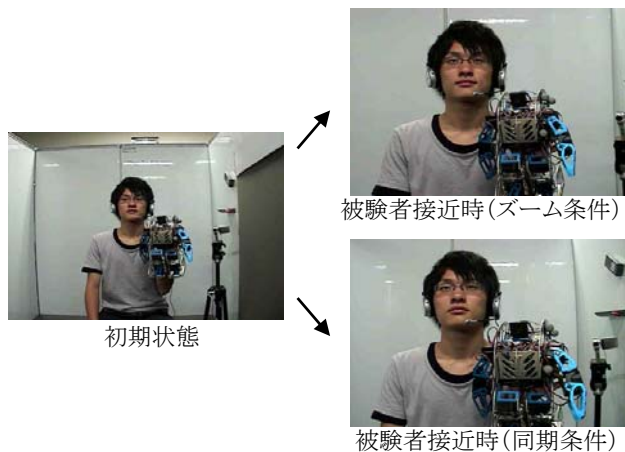


図 3. 同期条件とズーム条件の違い

## 4. 考察

実験結果は仮説を支持するものだった。同期条件や非同期条件の被験者は、実際にはカメラは前後移動しているにも関わらずズームしていると感じた。被験者は無意識的にはズームと前後移動の違いを認識しているが、意識的にはズームと前後移動の違いを認識していないと考えられる。しかし、この結果は単に現在普及しているカメラには、ズーム機能はあっても前後移動する機能が無いためであるとも考えられる。

同期条件の方がズーム条件の場合より映像の変化が好意的に受け取られる傾向があった。これは本実験では被験者のディスプレイへの接近動作を増幅したからであると考えられる。被験者の焦点の移動に反応する場合はズーム条件の方が好意的に受け取られる可能性がある。

## 5. おわりに

本論文では、メディアスペースにおいて、遠隔地にあるカメラをユーザのディスプレイへの接近動作に合わせて前後移動させることにより、ユーザが感じるテレプレゼンスを強化できることを示した。本論文で着目しているテレプレゼンスとは、遠隔地の人間と同じ空間で対面しているような感覚のことである。映像の拡大縮小は通常ズーム機能によって行われるが、ズーム機能では運動視差が発生しないため、少なくとも我々が行った実験においてはテレプレゼンスの強化は確認されなかった。

遠隔地の人間がカメラを動かしていると被験者が認識した場合は、テレプレゼンスは強化されなかった。これは我々の先行研究に即している [Nakanishi 08]。先行研究においては被験者がカメラの前後移動を操作しなければテレプレゼンスは強化されなかったが、本研究では被験者がカメラを操作しなくても、被験者がカメラは自分の動きに反応していると認識すればテレプレゼンスが強化されることが分かった。

## 謝辞

可動式カメラの構築に協力して頂いた松村礼央氏に感謝する。本研究の一部は科学研究費補助金基盤研究(S)「遠隔操作アンドロイドによる存在感の研究」(代表研究者:石黒浩)及び科学技術振興調整費「先端融合領域イノベーション創出拠点の形成:ゆらぎプロジェクト」によるものである。

## 参考文献

- [Bly 93] Bly, S.A., Harrison, S.R. and Irwin, S. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment. *Communications of the ACM*, 36(1), pp.28-46, 1993.
- [Bondareva 04] Bondareva, Y. and Bouwhuis, D. Determinants of Social Presence in Videoconferencing. *Proc. AVI2004 Workshop on Environments for Personalized Information Access*, pp.1-9, 2004.
- [de Greef 01] de Greef, P. and Ijsselstein, W. Social Presence in a Home Tele-Application. *CyberPsychology & Behavior*, 4(2), pp.307-315, 2001.
- [Fish 90] Fish, R.S., Kraut, R.E. and Chalfonte, B.L. The VideoWindow System in Informal Communication. *Proc. CSCW90*, pp.1-11, 1990.
- [Gaver 95] Gaver, W.W., Smets, G. and Overbeeke, K. A Virtual Window on Media Space. *Proc. CHI95*, pp.257-264, 1995.
- [Heath 92] Heath, C. and Luff, P. Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction. *Human-Computer Interaction*, 7(3), pp.315-346, 1992.
- [Ijsselstein 01] Ijsselstein, W., de Ridder, H., Freeman, J., Avons, S.E. and Bouwhuis, D. Effects of Stereoscopic Presentation, Image Motion, and Screen Size on Subjective and Objective Corroborative Measures of Presence. *Presence: Teleoperators & Virtual Environments*, 10(3), pp.298-311, 2001.
- [Jancke 01] Jancke, G., Venolia, G.D., Grudin, J., Cadiz, J.J. and Gupta, A. Linking Public Spaces: Technical and Social Issues. *Proc. CHI2001*, pp.530-537, 2001.
- [Karahalios 04] Karahalios, K. and Donath, J. Telemurals: Linking Remote Spaces with Social Catalysts. *Proc. CHI2004*, pp.615-622, 2004.
- [Kuzuoka 04] Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y. and Heath, C. Dual Ecologies of Robot as Communication Media: Thoughts on Coordinating Orientations and Projectability. *Proc. CHI2004*, pp.183-190, 2004.
- [Mantei 91] Mantei, M.M., Baecker, R.M., Sellen, A.J., Buxton, W.A.S., Milligan, T. and Wellman, B. Experiences in the Use of a Media Space. *Proc. CHI91*, pp.203-208, 1991.
- [Morita 07] Morita, T., Mase, K., Hirano, Y. and Kajita, S. Reciprocal Attentive Communication in Remote Meeting with a Humanoid Robot. *Proc. ICM2007*, pp.228-235, 2007.
- [Mueller 03] Mueller, F., Agamanolis, S. and Picard, R. Exertion Interfaces: Sports over a Distance for Social Bonding and Fun. *Proc. CHI2003*, pp.561-568, 2003.
- [Nakanishi 08] Nakanishi, H., Murakami, Y., Nogami, D. and Ishiguro, H. Minimum Movement Matters: Impact of Robot-Mounted Cameras on Social Telepresence. *Proc. CSCW2008*, pp.303-312, 2008.
- [Prussog 94] Prussog, A., Muhlbach, L. and Bocker, M. Telepresence in Videocommunications. *Proc. Annual Meeting of Human Factors and Ergonomics Society*, pp.25-38, 1994.
- [Roussel 02] Roussel, H. Experiences in the Design of the Well, a Group Communication Device for Teleconviviality. *Proc. Multimedia2002*, pp.146-152, 2002.
- [Roussel 04] Roussel, N., Evans, H. and Hansen, H. Proximity as an Interface for Video Communication. *IEEE Multimedia*, 11(3), pp.12-16, 2004.
- [Sheridan 95] Sheridan, T.B. Teleoperation, Telerobotics and Telepresence: A Progress Report. *Control Engineering Practice*, 3(2), pp.205-214, 1995.
- [Vogeley 03] Vogeley, K. and Fink, G.R. Neural Correlates of The First-Person-Perspective. *Trends in Cognitive Sciences*, 7(1), pp.38-42, 2003.