

LiveMask : 立体顔形状ディスプレイを用いた テレプレゼンスシステムにおけるコミュニケーションの評価

三 澤 加 奈[†] 石 黒 祥 生^{†,††} 暦 本 純 一^{†,†††}

我々は、立体的な顔スクリーンを持つテレプレゼンスシステムを提案する。このシステムは、遠隔話者の顔をトラッキングして頭部運動と顔領域を抽出する。頭部運動に応じて三軸で動く機構を持ち、遠隔話者本人と完全に同じ形状の顔スクリーンに本人の顔映像を投影する。スクリーンの立体形状を利用し、ユーザの非言語情報、特に 2D スクリーンでは伝達できない視線方向の伝達を目指している (“Mona Lisa effect”)。本システムがどのようにコミュニケーションに効果があるか調べるため、3 種類の実験を行った。1 つは、顔スクリーンと 2D スクリーンを比較し、投影された人物の判別ができる死角を調べ、その結果を元に表情の読み取りやすさを比較した。2 つめは平面ディスプレイと LiveMask システムを比較し、遠隔ユーザの指し示す方向を判別する評価を行った。3 つめは、スクリーンの違いによる視線方向の評価を行った。その結果、顔スクリーンに投影すると広範囲な角度から人物の認識ができ、LiveMask の首の動作は、平面ディスプレイより遠隔ユーザの指示方向が明確に伝わった。そして最も重要なことに、顔スクリーンは正確に視線方向を伝えることができ、Mona Lisa effect を解消することが分かった。

LiveMask: A Telepresence Surrogate System with a Face-Shaped Screen for Supporting Nonverbal Communication

KANA MISAWA,[†] YOSHIO ISHIGURO^{†,††} and JUN REKIMOTO^{†,†††}

We propose a telepresence system with a real human face-shaped screen. This system tracks the remote user's face and extracts the head motion and the face image. The face-shaped screen moves along three degree-of-freedom (DOF) by reflecting the user's head gestures. We expect this system can accurately convey the user's non-verbal communication, in particular the user's gaze direction in the 3D space that is not correctly transmitted by using a 2D screen ("The Mona Lisa effect"). To evaluate how this system can contribute to the communication, we conducted three experiments. The first one is to examine the blind angle of a face-shaped screen and a flat screen, and evaluated the ease of distinguishing facial expressions. The second experiment is to evaluate how the face directions of the remote user can be correctly transmitted. The third experiment is to evaluate how the gaze direction can be correctly transmitted. We found that the recognizable angles of the face-shaped screen were bigger, and the recognition of the head directions was better than those of the flat 2D screen. More importantly, we also found the face-shaped screen could accurately convey the gaze directions and it solves the Mona Lisa effect problem.

1. はじめに

テレプレゼンスとは、遠隔地の人や場所をネットワークを通してつなぎ、臨場感や存在感のあるコミュニケーションを実現する技術のことである。システムの特徴により、臨場感か存在感を重視したものがあり、大きく 3 つに分類できると考えている。第一の方式は、HP Halo¹⁾ や Cisco TelePresence²⁾ といったテレビ



図 1 LiveMask を通したコミュニケーション
Fig. 1 Communication applied on LiveMask

[†] 東京大学 The University of Tokyo

^{††} 日本学術振興会 JSPS Research Fellow

^{†††} 株式会社ソニーコンピュータサイエンス研究所

Sony Computer Science Laboratories, Inc.

会議型システムであり、それらの感覚を同程度伝える。第二の方式は、臨場感の伝達を重視しており、利用者が HMD などの機器を装着し、高臨場感や没入感を得る装着型システムである³⁾。第三の方式は、Texai⁵⁾や QB⁴⁾ などアバター操作型システムである。ユーザはプロジェクタやマイクを搭載した自身のアバターを遠隔操作し存在感を提示する。我々のシステムは、第三の方式に関係している。

豊かなコミュニケーションを実現するため、臨場感や存在感を高める研究が行われているが、直接会って会話をするに勝る方法は確立されていない。それを改善する一つの方法として、非言語情報の伝達が考えられる。非言語情報は、言葉以外の表情、視線、人の仕草などを指し、コミュニケーションにおいて必要な役割を果たすと言われている。従来のテレビ会議システムでは、ユーザの外見的特徴や音声を伝達することができた。しかし、空間的な位置把握を必要とする頭部の動作や視線方向の読み取りは困難であった。

首の動作は、うなずきとしての相槌や相手の申し出に対する意思表示：賛成、反対、難色などが表現される。また、視線は会話において重要な役割を果たしていることが知られており、会話の順序や流れを円滑にする機能がある¹⁴⁾。

そこで、我々は実際の会話でしているような非言語情報である、首の動作と視線方向を伝達できるテレプレゼンスシステムを提案する。本システムは、遠隔話者の顔をリアルタイムでトラッキングし、頭部運動と顔領域を抽出する。そして、仕草を表現するために頭部運動に応じて 3DOF (Degree of Freedom) で動く機構を搭載し、遠隔話者本人と完全に同じ形状の顔型立体スクリーン (以降、顔スクリーン) に本人の顔映像を投影することで、正確な視線方向の伝達を目指す。

顔スクリーンに話者の顔映像を投影するプロジェクトに Talking Heads がある⁶⁾。しかし、テレプレゼンスシステムとしてディスプレイが平面ではなく立体である時に顔の印象が及ぼす効果や、優位性は明らかにされていない。また、本研究では顔や頭部のトラッキングにカメラを用いているため、何も機材を装着する必要がなく、自然な状態で会話ができる。

本論文では、開発した LiveMask システム (図 1) の説明をし、従来の平面ディスプレイと比較したときの、表情の見え方、首の動作による指示、そして視線方向の伝わりやすさを論じる。

2. 関連研究

テレプレゼンスでは、同じ空間にいない遠隔話者を

どのように表現するかが一つの課題である。近年では、臨場感や存在感だけではなく、対話をしている時に実際に感じられる親近感や信頼感、人を惹き付ける魅力といった心理的要素も伝えようとしている。我々は、会話において重要な非言語情報に注目し、遠隔地に頭部の動きと視線方向が伝達できる自身の“代理人”を開発した。

Kidd と Breazeal は、ロボットはコンピュータグラフィックスでつくられたエージェントに比較して、より魅力的であり、人に信頼されやすいことを示した⁷⁾。人間型ロボットをアンドロイドというが、Geminoid¹¹⁾は遠隔コミュニケーションを実現するための遠隔操作型アンドロイド・ロボットシステムである。非常に人間らしい外見を持っているため存在感がある一方で、不気味と評価された。石黒らは「人らしい外見を持ったメディアのほうが強い存在感をしめす」可能性があると感じている。顔は対人インタラクションで重要な要素でありながら、現在の研究では顔そのものの形状に注目して開発されたテレプレゼンスシステムは少なく、多くのシステムは二次元のディスプレイでつくられている。そこで、我々は特に人間の顔の形状に着目し、人の顔から型取りをした立体的なスクリーンを搭載したテレプレゼンスシステムを開発した。

また遠隔話者の動作を模倣する研究も行われている。中西らはディスプレイを前後に移動することによって人物を強調させると、社会的存在感が強化されることを証明した⁸⁾。また Master-slave 型のロボットをテレロボットというが、遠隔操作によって社会的な表現ができる MeBot⁹⁾ は人を惹き付け、好意的に受け取られることが示されている。そこで、我々は人の仕草、特に頭部の動作を伝達するために、3DOF で動作する機構を実装した。

Hydra¹⁰⁾ は遠隔にいる 4 人でビデオ会議をする卓上型のテレプレゼンスシステムである。各話者それぞれの小型ユニットには、対話者の代理 (surrogate) になる顔を映し出すモニター、眼となるカメラ、口としてのスピーカーが搭載され、“代理人”として卓上でコミュニケーションを行う。我々は、顔の代理になるスクリーンとして、その人本人の顔形状から形成されたスクリーンを使用すべきだと考えている。他人の顔から成形されたスクリーンより自身のスクリーンを使用すると“代理人”としての意味合いが増し、信頼性が上がると考えているからだ。

視線は、人と人の会話において古くから重要な役割を果たしていることが知られている。視線は会話の順序や流れを円滑にする機能や¹⁴⁾、相手が自分に対して

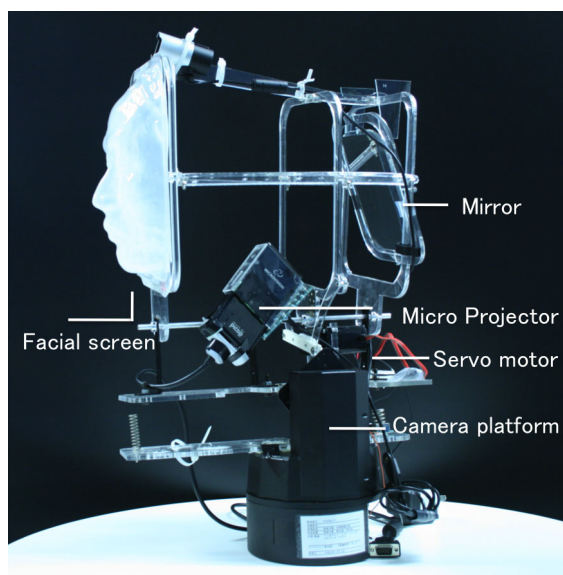


図 2 LiveMask：全体を小型化するために、プロジェクタの画角を調べ表面鏡を用いて映像を反射させている。黒い土台がカメラ雲台である。

Fig.2 Face Display current model

感じる印象を操作し、適切な親密さのレベルを維持する機能¹⁵⁾などが知られている。現在多くのテレプレゼンスシステムでは、平面のスクリーンに人物の顔を表示し使用している。しかし、平面のスクリーンでは The Mona Lisa effect¹⁷⁾ と呼ばれる、どの方向から見ても視線が合うように感じる現象が発生する。Jonas らは、2次元のディスプレイと3次元のマネキンに顔画像を投影すると、視線の方向が異なることに注目している¹⁶⁾。Humanoidの研究分野ではロボットに表情をつける研究が行われており、Delauny らは幼児顔型のスクリーンに Computer Graphics の画像を投影している。また、スクリーンの形状による視線方向の違いについて論じており、半球型の結果が悪く、顔型と平面にさほど違いがないと述べている¹³⁾。しかし、テレプレゼンス場面で使用するには高臨場感のシステムを実現するために人に置き換えた実証が必要になる。CG 投影時と同等の結果が得られるのか知見がないので我々は、実際の人の顔型のスクリーンに映像を投影し、視線方向の違いを調べることにした。

遠隔話者の顔を投影する既存研究はあるが、詳細な効果は明らかにされていない⁶⁾¹²⁾。我々はシステムの提案と開発した LiveMask と従来の 2D ディスプレイの比較を行うことを行った。

3. LiveMask: 立体顔形状ディスプレイを用いたテレプレゼンスシステム

遠隔にいる自身の代理人を再現するため、LiveMask



図 3 首の三軸の動作：パン-左右に首を振る動作，チルト-上下に傾く動作，ロール-左右に首を傾げる動作に対応する。

Fig.3 3DOF: Pan, Tilt, Roll

(図 2) の基本構成は、実寸大サイズの顔スクリーン、顔の映像を投影するための小型プロジェクタ、頭部の運動を伝える駆動部、小型の USB カメラを搭載している。頭部運動の自由度を三軸で設計しており、図 3 のように、チルト角：上下に傾く動作，パン角：左右に首を振る動作，ロール角：首を傾げる仕草が遠隔話者の動きに合わせて駆動する。

3.1 顔スクリーン

顔スクリーンは実際の人間の顔から型取りを行って成形している。必要な材料は、型とり材、石膏、熱可塑性のスクリーンシートである。水分を加えて攪拌して液体状になった型取り材に顔を挿入し固形化するまで待った後、外すと雌型ができる。そして雌型に石膏を流し込み、乾燥させると雄型ができる。この方法は手作業で行っていたが、3D スキャナーを用いると、顔の 3 次元データが取得できるため、そのデータを補正し 3D プリンタへデータを入力すると、同様の顔型をつくるのが可能である。雄型から、雌型のスクリーンは熱可塑性の素材を使用しバキューム装置を使用することにより、実際の人型の顔スクリーンを作成した。

開発初期の段階で、顔映像を投影する立体スクリーンの形状にいくつか候補があり、楕円型、本人ではない顔型、曖昧な輪郭の型などへの投影を試みた。その結果、楕円の形状や曖昧な輪郭の形状でも人の判別はできるが、横顔の印象が人の形状でなかったり、本人と異なることで違和感が生じ、立体であることの利点が損なわれた。

3.2 顔映像の画像処理

投影する映像をつくるために、遠隔話者の話している様子をカメラで撮影し、顔画像の抽出や、補正といった画像処理を行っている。立体的なスクリーンにリアルタイムで顔画像を投影をするため、サイズの不整合や画像のずれを改善させる必要がある。SeeingMachine 社の faceAPI¹⁸⁾ は精度が高いフェイストラッキングができるライブラリである。眼鏡を着用している人物でも、眼や鼻といった部位の追跡ができたり、人物の素早い動作でも安定してトラッキングができる点で使用に適していた。投影する映像は、カメラに向かって正面を向いている画像が必要である。faceAPI が自動処



図 4 顔画像の処理：顔認識が行われている時のスクリーンショットである。左図のように首を傾げた状態から画像補正後、右図に示す顔テクスチャが生成される。画像は投影用に反転させる。

Fig. 4 Image processing

理で正面を向いた顔になるよう画像補正を行っており、これをもとにスクリーンに合うよう画像調節を行って投影させている (図 4)。

3.3 駆動構造

頭部の運動を伝えるため本システムはパン、チルト、ロール角の 3DOF を搭載した。頭部の三次元位置座標は、faceAPI で取得できるヘッドトラッキングデータを元になっている。カメラで顔を認識し人が正面を向いた状態を 0° とした時、左右や上下に首を動作させて検出される座標値の幅は $\pm 45^\circ$ 以内に計測される。それ以上の大きな値ではカメラから顔を背けた状態になっており、顔認識がされない。以上を踏まえ、駆動の自由度は $\pm 45^\circ$ に設定した。

3.4 プロジェクタ

システムに搭載したのは Microvision 社の Pico Projector である。レーザータイプのプロジェクタのため、ピント調節の必要がなく、不定形のディスプレイに投影するのに優れている。提案システムのように顔スクリーンをバックプロジェクションする機構で回転させるには、プロジェクタの重さが問題になるが軽量で小型であったため、優れていた。

4. 実験

実験 1：LiveMask の死角、表情判別の比較

実際の会議では、何人かでテーブルを囲んで会話をするため、常にテレプレゼンスシステムの正面に相手がいるわけではない。遠隔参加するユーザ毎に人数分のディスプレイをテーブルに配置する会議環境を想定する。もしスクリーンが平面であった時、横隣に座った人物は遠隔話者を覗き込むようにして見る必要があり、表情を読み取るのが困難である。一方、LiveMask は凹凸感があるため覗かなくても、“横顔”が見れる。ディスプレイに凹凸感があると、実際の人のように横隣や上から見下した角度からも遠隔話者の表情を読み取ることができると予想される。そこで 2 種類の実験

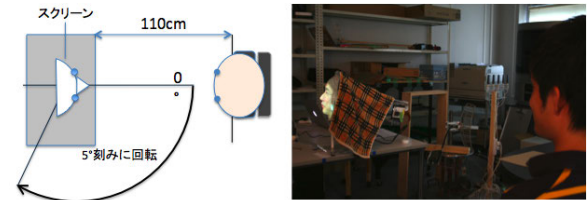


図 5 実験 1 の様子：左図は上から見た様子を示している。右図のように被験者は離れた位置からスクリーンを眺め解答をした。

Fig. 5 Experiment1 setup



図 6 スクリーンの比較：2 種類のスクリーンに同じ画像を投影している。左図が平面スクリーン、右図が顔スクリーンである。

Fig. 6 2 kinds of screens

を行った。被験者の数は、実験 1 から 3 まで 8 人ずつ実施した。

4.1 実験 1.1 予備実験：死角を調べる

実験目的：LiveMask がどれぐらいの角度から死角となるのか調べる。死角とは、投影された人物が誰であるか判断できない角度と定義する。死角の判断基準は、スクリーンに 3 人の人物をランダムに表示し、それが誰であるか認識できるかを調べる。

実験方法：システムは机の上に固定されて配置されており、被験者の座席は図 5 のようにシステムと向き合うように固定されている。実験用に使用する平面スクリーンは、型取りに使用するシートをディスプレイの枠の大きさに合わせて加工する。LiveMask のスクリーンを顔形状と、平面形状で交互に切り替えて実験を行う (図 6)。被験者を椅子に座らせ、システムが見えるように座席の高さを調節させる。システムにつけられた平面のスクリーンと正面に向かい合った角度を 0° とし実験を開始する。

実験手順：(1) 3 人の顔写真をランダムに表示し、それが誰であるか記号で解答させる。(2) スクリーンの角度は 5 度刻みで時計回りに回転させ、解答ができるまで続ける。(3) 顔スクリーンに替え、同様の手順で実験を行う。

実験結果：平面スクリーンと顔スクリーンによる人の判別できる角度を調べた結果を図 7 に示す。実験では時計回りにスクリーンを回転させたが、半時計回りに

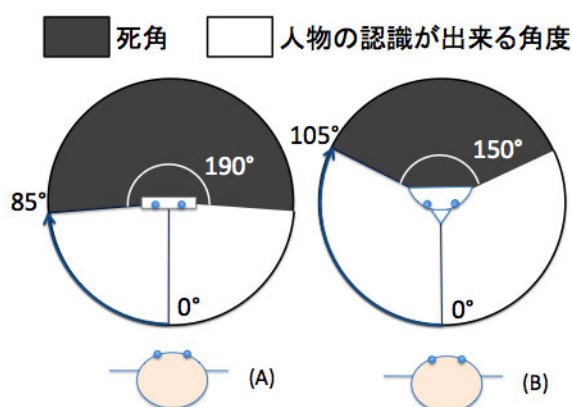


図 7 2つのスクリーンの死角：(A) は平面スクリーンであり、正面から 85 度以上横を向くと死角となる。(B) は顔スクリーンであり、凹凸があるため視認される角度が広く、105 度以上で死角となる。

Fig. 7 Comparison of two screens



図 8 実験 1.2 に使用した 5 種類の表情の例：ある人物の喜び、悲しみ、怒り、驚き、恐れ表情を示す。実験ではそれぞれの表情に 4 枚ずつバリエーションがあり、計 20 枚の画像を見せた。

Fig. 8 Five kinds of emotion used for experiment 1.2

回転させた時に左右反転した結果が得られるとすると図 7 で色付けした範囲が認識できる角度となる。

平面スクリーン：被験者に対し、正面を向かい合った状態から 90° 時計回りに回転した状態であると、誰も人物を判断することができなかったが、85° では 87.5 % の被験者が人物の識別をすることができた。

顔スクリーン：一方、平面スクリーンでは人物が識別されなかった 90° でも全員の被験者が投影された人物を判別することができた。被験者の解答から、識別できる角度は 105° ～110° の間であることが分かった。

4.2 実験 1.2 表情の判別をする実験

実験目的：顔スクリーンと平面スクリーンに顔画像が投影された時、人の判断ができるだけでなく、表情の読み取りやすさに違いがあるか比較実験を行う。

実験方法：事前に 1 人の人物の 5 種類の表情（喜び、悲しみ、怒り、驚き、恐れ）(図 8) の写真を計 20 枚用意した。実験 1.1 と同様の実験環境で平面スクリーンと顔スクリーンに投影し、どんな表情をしているのか解答させる。手元の配布資料として 5 種類の表情それぞれの写真と表情名を明記したものを用意した。実

験 1.1 より平面スクリーンで認識できる角度が明らかにされているので、その角度で比較を行う。

実験手順：(1) 実験開始前に、投影される人物のどのような表情がどの感情に相当するか被験者間で認識に差が出ないように、顔画像をそれぞれのスクリーンに投影し、確認を行う。(2) 実験 1.1 で得られた平面スクリーンで人物が認識できる限界の角度 (85°) に設定し、20 枚の顔写真をランダムに見せ解答用紙に記入させる。(3) 顔スクリーンに取り替え、上記と同様の内容を繰り返し行う。(4) 最後に被験者毎に得られた顔スクリーンの人物が認識できる限界の角度で、どのくらい表情の判別ができるか調査する。

実験結果：2つのスクリーンによる表情の読み取りやすさを比較した結果を表 1 にまとめた。平面スクリーンの方が顔スクリーンより正答率は若干高いものの、ほぼ同等の結果を得られた。また、顔スクリーンの人物が特定できる角度の 105° では 85° に比べ、表情の読み取り正答率は格段に下がり、人物の特定はできるが表情の読み取りは難しいことが分かった。実験終了後、ヒアリングを行い、表情を判断する時に何を基準にしたか、顔スクリーンと平面スクリーンではどちらかが判断しやすかったかを尋ねた。

表情判断の基準：喜びや、驚きの表情が分かりやすい一方で、怒りや不安な表情の判別が難しいというコメントを得た。これら 2つの表情は、スクリーンが立体か平面のどちらでも正答率が低かった。表情の判断には、目や口元で判断する人が多く、例えば口角が上向きか下向きかを手がかりにすると表情の判断がしやすいとコメントを得た。

スクリーンによる比較：平面と立体による表情の判別は、正答率がそれぞれ 71.9 %、70.0 % とほぼ同等の結果を得た。被験者にどちらが判断しやすいか尋ねたところ、6 人は立体、2 人は平面スクリーンが分かりやすいと答えた。顔スクリーンは「立体形状により顔の表面積が広いので、判断基準が多くて見やすい。表情と骨格が合っているため目元や口元の印象が分かりやすい」一方、「顔の凹凸があるため顔面の平面しか見えず、判断しにくい」という声も得た。

表 1 5 種類の表情：認識の精度

Table 1 Five kinds of emotional state : The accuracy of recognition

	平均	標準偏差
i) Flat screen 85°	71.9 %	0.113
i) Facial screen 85°	70.0 %	0.963
ii) Facial screen 105°-110°	25.0 %	0.134

(i)：平面スクリーンの死角で比較

(ii)：顔スクリーンの死角で比較

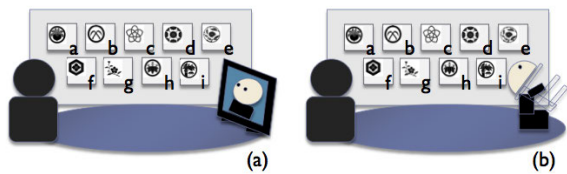


図 9 Experiment 2 Condition: (a) is a image for study setting of 2D display. (b) is a image for study setting of LiveMask system.

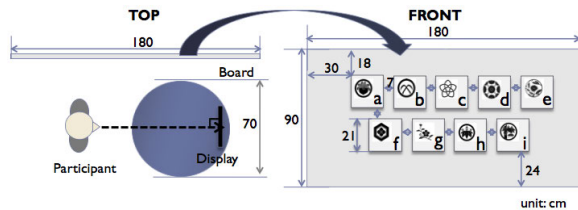


図 10 Experiment 2 Study Setting

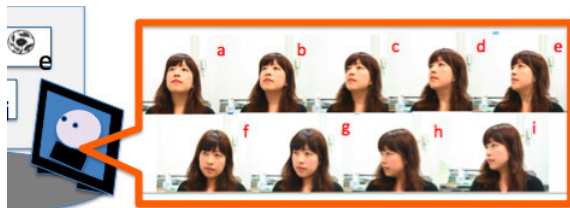


図 11 2D display condition: Each 9 images is projected on a flat display.

4.3 実験 2: LiveMask の首の動作による方向指示と平面ディスプレイに表示された画像による方向指示の比較

次に LiveMask に実装している仕組みである、頭部の方向指示に関する実験を行った。実際の会議では、話している人の方向を向いて議論をしたり、壁面やホワイトボード上に記述された内容の方向を向いて説明をする時がある。遠隔地にいる話者がどこを指し示しているのか、平面のディスプレイを使用した時と頭部の動作を持つテレプレゼンスを使用した時のどちらがわかりやすいかを調べた (図 9, 図 10)。

実験目的: LiveMask の首の動作による方向指示と平面ディスプレイに表示された画像による方向指示を比較する。

実験方法: 壁に (a) から (i) までアルファベットが振られた 9 種類の記号が貼られている。遠隔話者がこの 9 種類のいずれかを指し示す。それぞれの指示の出し方は以下のような事前設定を行う。

首のジェスチャー: LiveMask に 9 つの記号それぞれを指している首のポジションを事前登録した。この時、顔の表情、視線の位置による影響を受けないよう、表情は同じものを使用した。



図 12 Experiment 2 Participant try to distinguish the sign LiveMask indicates.

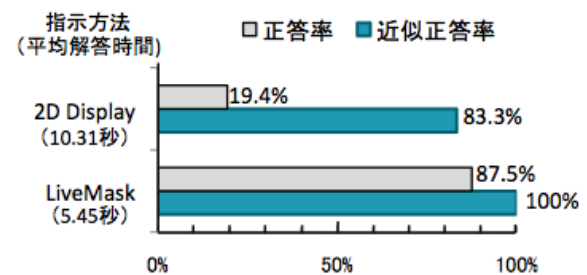


図 13 実験 2: 方向の判別のしやすさ

Fig. 13 Experiment 2: Direction

平面のディスプレイ: あらかじめ指示を伝える遠隔話者が実験環境の椅子に座り、各記号の首の向きや視線を変えて指示している様子を投影用に撮影する (図 11)。**実験手順:** (1) LiveMask が壁に貼ってる 9 種類の記号のいずれかを指し示すので被験者はその記号を答える。被験者は図 12 のように左側に着席し、オペレータが、LiveMask を 9 種類の方向のいずれかにセットする間、目隠しの仕切りを持ち待機する。(2) 開始の合図で被験者は目隠しを外し、オペレータは開始の合図から被験者が解答するまでの時間を計測する。(3) 9 種類の方向全てをランダムにテストした後、平面のディスプレイに切り替え同様に実験を行う。それぞれ 1 セット 9 回、2 種類ずつ計 18 回を行う。

実験結果: LiveMask による方向指示は多くの被験者が正解し、正答率は 87.5 % であった。一方、平面ディスプレイに画像を表示した場合は、正解率が 19.4 % であり (図 13)、解答時間は、判断するまでに約 2 倍長く時間がかかっている。しかし、正解に隣接する記号も正解とカウントすると正答率 (近似正答率と定義する) は 83.3 % にまで上がった。これは、被験者が答えた記号が、正解の両隣のいずれか、もしくは上下で間違えたことを示している。

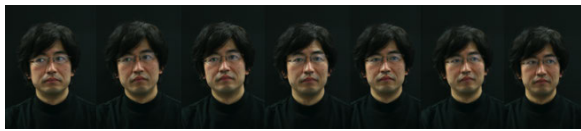


図 14 実験 3 視線の画像：顔の位置は固定で視線のみ変えて撮影した。左から 30°、20° と 10° ずつ視線の位置が変化している。

Fig. 14 Experiment 3: Eye sight

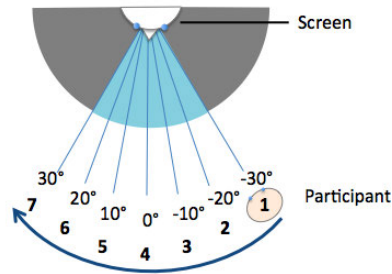


図 15 視線方向：-30° から 30° まで 10° 刻みに被験者は移動し、スクリーンに投影された画像と視線が合うか解答する。

Fig. 15 Experiment 3: Eye direction

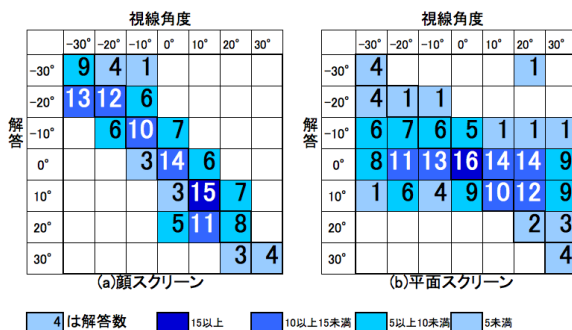


図 16 実験 3 視線の認識：横軸に画像中の視線角度、縦軸に、被験者が視線が合ったと答えた画像中の視線角度をとっている。顔スクリーンでは結果が対角線上に分布しているのに対し、平面スクリーンは 0° 周辺に帯状に分布している。

Fig. 16 Experiment 3: Test Environment

4.4 実験 3：視線の認識

実験目的：平面スクリーンと顔スクリーンで視線方向に違いが生じるか調べる。

実験方法：予め、-30°～30° まで 10° 刻みに視線を移動させた 7 枚の写真を用意し（図 14）、実験 1 と同様の環境で LiveMask のスクリーンに写真を投影する。1 カ所あたり 7 枚の写真投影を計 7 カ所で行い（1

表 2 実験 3：視線の認識
Table 2 The eye sights

	正答率	標準偏差
Facial screen	75.9 %	0.166
Flat screen	36.6 %	0.30

正答率：全問題（14）中の正解数の割合

セット）、各スクリーンで 2 セットずつ実施する。

実験手順：（1）被験者は図 15 のように LiveMask から -30° の位置で目線の高さが合うように椅子を調節する。（2）被験者は 7 枚の写真を順次見て視線が合うか答え、終了後時計回りに 10° ずれた位置で解答する（図 15）。（3）上記手順を 30° まで繰り返す、もう 1 セット行った後、オペレータは平面スクリーンに替え、同様にテストを行う。

実験結果：結果を表 2 に示した。顔スクリーンを使用したときの正答率は 75.9 % に対し、平面スクリーンは 36.6 % と低かった。また、各視線画像のうち、図 16 に示すように顔スクリーンでは解答結果が対角線上に分布しているのに対し、平面スクリーンは 0° 付近に帯状に分布している。これは、顔スクリーンでは視線方向が被験者に伝わっていることに対し、平面スクリーンは正面を向いた画像（0°）がどの角度から見ても視線が合いやすいことを示している。

5. 考 察

実験 1.1 より、顔スクリーンは平面スクリーンと比較すると広範囲な角度から人物が認識可能であることが分かった。死角が少ないことから、テーブルを囲んで会議をする時だけでなく、部屋にランダムに人物がいる状況下でも顔を認識することができ、存在感を出すために有用可能性がある。**実験 1.2** では、顔の表情の読み取りやすさにスクリーンによる違いは見られなかったが、判断のしやすさに対する被験者の主観的意見は分かれた。平面スクリーンの方が分かりやすいと答えた 2 人は、顔の凹凸、つまり鼻の高さや頬の膨らみによって片側の顔が見えないことによる分かりにくさを感じていた。これら 2 人は、表情の判断を顔全面を見て判断しようとしていた。今回は、顔が動かない環境で実験を行っていたが、実際の利用場面では頭部の動作を使用しユーザ側に向くことができるためこのような問題は解決できる。初めて LiveMask を目にした被験者からは「存在感があり、話してみたい」という声を得た。**実験 2** の方向指示を読み取る実験では結果に大きな差が出た。画像で判断した時の正解率は特に低かった。しかし、近似正答率で評価するとそれぞれの被験者が 8 割近い正解率を記録している。画像で判断した時の正答率が低い理由はいくつか考えられる。「人の視界は広いのでディスプレイに表示された人物が 3～4 つの記号を見ているように感じる」とのコメントを得た。2D 画像は、見ている大体の方向を伝えることができるが、正確な位置や方向を伝えることが、現状は難しい。それは人の視野は広いので、画

像中の人物がどこを見ているのか局所的に判断することができないからだ。実験に使用した画像では、首の方向と視線を変えているが、実際には首を動かして特定の方向を向かなくても、人は視線だけでその方向を指し示すことができる。そこで**実験 3**ではスクリーンの違いにより、視線方向の伝わり方が変わるか調べた。その結果、顔スクリーンではそれぞれの視線方向が伝わりやすかったが、平面スクリーンでは正面を向いている顔がどの角度から見ても視線が合うことが起こった。従来の平面的なディスプレイでは、厳密な視線による方向指示やアイコンタクトをとることが困難であり、それを解決する方法として顔スクリーンが使用できることが分かった。

6. 結 論

我々は、実際の顔から型取りした顔形状のスクリーンを持ち、遠隔話者の動きに合わせて駆動する、自身の“代理人”としてのテレプレゼンスシステムを提案した。本論文では、立体的なディスプレイの効果を調べるため、平面スクリーンと顔スクリーンを比較し、人物の判別ができる角度や、表情の読み取りやすさを調べた。その結果、立体的なディスプレイは広範囲な角度で見えるため、正面を向かい合ったコミュニケーションだけではなく、さまざまなインタラクションで応用できる可能性があることが分かった。その一方で、平面スクリーンと顔スクリーンの表情の認識のしやすさに関しては、差異が見られなかった。しかしながら、多くの被験者は「顔スクリーンの形状は、顔の表情が分かりやすく、通常の平面的なディスプレイと比較すると人らしい形状を持っているため、親しみやすい」と述べており、表情以外の非言語情報を受け取った被験者もいた。また遠隔話者による方向指示は、平面ディスプレイに表示された画像より LiveMask による首の動作の方が判断しやすいことが分かった。そして、平面ディスプレイでは Mona lisa effect の発生により遠隔ユーザとアイコンタクトを取ることは困難であったが、LiveMask の立体形状は、視線の方向を正確に伝達できることが分かった。

本研究では、LiveMask の基本的な特徴を評価した。今後はテレプレゼンスシステムを使用して、共同作業をした時の遠隔存在感や、顔スクリーンを使用した時に話しをしている人が、どのような感情を抱くかを明らかにしたい。

謝辞 ハードウェアの設計において、多大な助言をして頂いたソニーコンピュータサイエンス研究所の田島茂氏に感謝の意を表します。

参 考 文 献

- 1) HP Halo <http://hphalo.org/>
- 2) Cisco TelePresence <http://www.cisco.com/en/US/products/ps7060/index.html>
- 3) H. Nagahara, Y. Yagi, M. Yachida. Wide Field of View Head Mounted Display for Telepresence with An Omnidirectional Image Sensor. *IEEE/CVPR'03*,88,2003.
- 4) AnyBots <https://www.anybots.com>.
- 5) Lee, Min Kyung and Takayama, Leila. "Now, i have a body": uses and social norms for mobile remote presence in the workplace. *Proc. CHI'11*,33-42,2011.
- 6) Talking Heads projection <http://www.naimark.net/projects/head.html>.
- 7) C.Kidd and C. Breazeal. Effect of a Robot on User Perceptions. *IEEE/RSJ IROS'04*,2004.
- 8) H. Nakanishi, K. Kato, H. Ishiguro. Zoom cameras and movable displays enhance social telepresence. *Proc. CHI'11*,63-72, 2011.
- 9) Adalgeirsson, C. Breazeal. MeBots A Robotics Platform for Socially Embodied Telepresence. *Proc. HRI'10*,15-22,2010.
- 10) Hydra <http://www.billbuxton.com/hydraNarrative.htm>
- 11) D. Sakamoto , T. Kanda , T. Ono , H. Ishiguro , N. Hagita. Android as a telecommunication medium with a human-like presence. *Proc. HRI'07*,10-12, 2007
- 12) R. Tadakuma, Y. Asahara, H. Kajimoto, N. Kawakami, and S. Tachi, Development of Anthropomorphic Multi-D.O.F. Master-Slave Arm for Mutual Telexistence, *IEEE Trans. on VCG, vol. 11* ,626-636,2005.
- 13) Frederic and De Greeff, Joachim and Belpaeme, Tony. A study of a retro-projected robotic face and its effectiveness for gaze reading by humans. *Proc. HRI'10*,39-44,2010
- 14) Kendon, A. Some functions of gaze-direction in social interaction. *Acta Psychologica*,22-63,1967
- 15) M.Argyle, L. Lefebvre and M. Cook. The meaning of five patterns of gaze *European Journal of Social Psychology*,4,2,125-136,1974
- 16) Beskow, Jonas and Al Moubayed, Samer Perception of gaze direction in 2D and 3D facial projections. *Proc. FAA '10*,24-24, 2010
- 17) Jens Edlund, Samer Al Moubayed and Jonas, Beskow The Mona Lisa Gaze Effect as an Objective Metric for Perceived Cospatality. *IVA'11*,439-440,2011
- 18) Seeing Machine FaceAPI <http://www.seeingmachines.com/product/faceapi/>