

TEllipsoid:遠隔会議のための視線方向を正しく伝える 顔画像提示用楕円球型ディスプレイ

一居 太朗

概要：遠隔会議システムにおいては、非言語情報が伝達しづらくコミュニケーションが円滑に行なわれない問題がある。これは平面ディスプレイに顔画像を投影するとモノリザ効果が発生し、視線方向が適切に伝達されないことが一因と考えられる。そこで我々は、実用時に重要な利便性および実物と提示像の同一性を確保しながら適切な視線方向の伝達が行なえる映像提示手法として、楕円球型のスクリーンに顔画像を投影するシステム（TEllipsoid）を提案する。このシステムは楕円球型スクリーン、小型プロジェクタ、凸面鏡によって構成され、底部の小型プロジェクタから投影された映像を凸面鏡で反射させ、楕円球型スクリーンに投影することで顔画像を表示する。顔画像は人物を 360 度全周撮影し合成した画像を基に作成し、実物との同一性を高めるために一部を加工した。また視線提示機能を、眼球の 3D モデルを楕円球型スクリーンに投影することで実装した。視線伝達能を評価した結果、会議で隣り合った参加者の角距離が 38.5° 以上離れた場合であれば、適切な視線方向の伝達が可能であることが示された。

1. 背景

今日では、電話をはじめメールや SNS、ビデオ会議など様々な情報通信技術が発達・普及しているが、未だに対面会話が最も基本的なコミュニケーション手段である。特に重要な意思決定などは依然対面会話で行われることが多く、空間的に離れた人々が対面しているかのように会話できる技術の実現は、人件費・交通費の削減など大きなメリットをもたらす。しかしながら、現状の遠隔会議システムで多く用いられる平面ディスプレイでは、視線や頭の動きなど非言語情報の適切な伝達が難しく、コミュニケーションが円滑に行なわれない問題がある。コミュニケーションにおいて非言語情報は 93% を占めるともいわれ大きな役割を果たしている [1]。特に 3 人以上の多人数会話においては、視線や身振りによって話者の交替が調整されることが示されている [2][3]。

中でも視線は、モノリザ効果 [4] が発生してしまうために適切な視線の提示が困難となる。モノリザ効果とは、平面ディスプレイ内の顔画像が正面を見ている場合、見る方向に関わらず自分と目が合っているように感じ、正面でない場合は見る方向に関わらず目が合っていないように感じる現象であり、これにより遠隔参加者が誰に話しかけているか判断することが困難になる。モノリザ効果を解消するため、多視点 3D ディスプレイ [5] や立体顔形状ディスプレイ [6] などが提案されてきたが、これらはアクチュエー



図 1 想定される使用法

タを必要とするため騒音や保守運用などに問題がある。よりリーズナブルな手法として、曲面スクリーンに顔画像を投影することでモノリザ効果が解消されることが報告されている [7]。

また実用的な遠隔会議システムには、視線方向の適切な伝達に加え、システムの利便性および実物と提示像の同一性も必要であると考えられる。

実用時には不特定多数の利用者が使用する可能性があるため、利用者固有の設備が不要な汎用性の高いシステムが求められる。また、製造コストや故障リスクの低さも重要であり、これらを実現するためには駆動機構を最小限にする必要がある。テーブルを囲む会議では、顔を左右 90° まで向ける可能性があるため通常は駆動機構が必要となってしまうが、ディスプレイが鉛直軸対象な回転体形状であれ

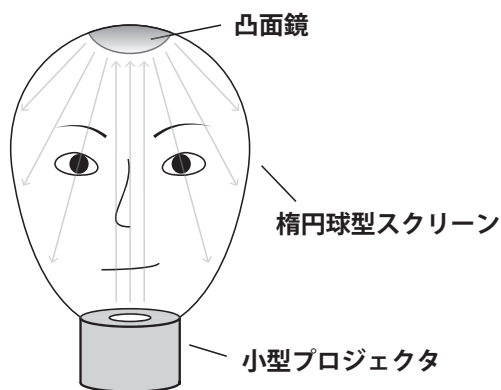


図 2 TELLipsoid の構成

ば左右方向の回転を映像のみで表現でき駆動機構を削減できる。

また現状では、遠隔会議システムで対話する相手は以前に対面会話したことがある顔見知り、あるいは採用応募者など以後対面会話する可能性がある人物である場合が殆どである。遠隔会議システムで表示される顔画像が実際の外見と大きく異なった場合、受話者が顔画像を対話者として認識しづらくなる問題がある [8]。対話者が誰であるかを同定する顔認識の過程において、我々は主に顔パーツおよび輪郭の全体的な配置情報で識別するといわれている [9][10]。そのため実用的な遠隔会議システムには、提示像の顔パーツ・輪郭の配置が実際の外見と同様に見えるディスプレイが必要となる。特にテーブルを囲んだ会議では正面側 180° の領域から観察され得るため、顔パーツ配置の同一性が保たれる視野角は広い必要がある。

そこで我々は、実用時に重要な利便性および実物と提示像の同一性を確保しながら適切な視線方向の伝達が行なえる映像提示手法として、鉛直軸対称な曲面を持ち人間の顔形状に近い楕円球型のスクリーンに顔画像を投影し、表示するシステム (TELLipsoid) を提案した。このシステムは特に図 1 に示すような、参加者が 3 人から 15 人の少人数会議を想定して設計された。

本論文ではまず提案するシステム (TELLipsoid) の構成や設計について詳述し、さらに視線方向の伝達能および印象について評価を行なう。本論文の一部は、筆者の過去の報告 [11][12] を基に作成している。

2. 提案手法

図 2 に、提案するシステム TELLipsoid (Tele+Ellipsoid) の構成を示す。このシステムは楕円球型スクリーン、小型プロジェクタおよび凸面鏡によって構成され、底部の小型プロジェクタから投影された映像を凸面鏡で反射させ、楕円球型スクリーンに投影することで顔画像を表示する。スクリーン形状は日本人青年男子の平均顔寸法 [13] を基に設計しており、表示する顔画像には事前生成した静止画像を

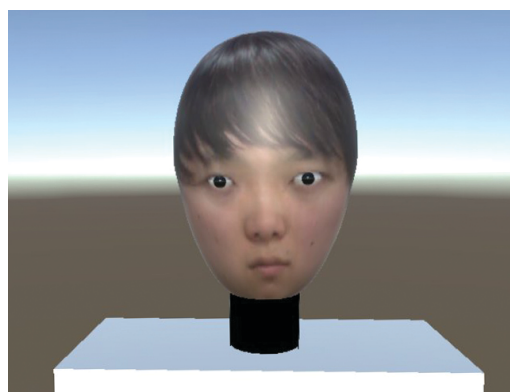
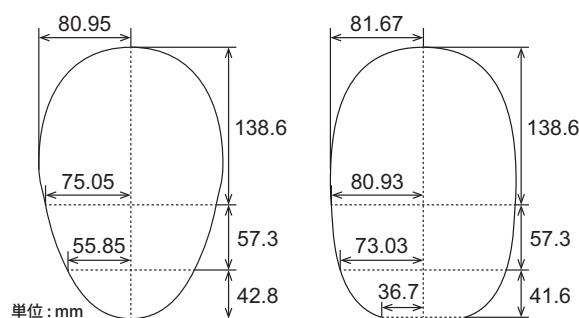


図 3 VR 空間内で製作されたプロトタイプ



(a) プロトタイプ1:VR (b) プロトタイプ2:実デバイス

図 4 平均顔寸法を基に設計されたスクリーン形状

用いている。また眼球の 3D モデルを楕円球型スクリーンに投影することで、視線提示機能も実装している。

2.1 プロトタイプ1: VR 空間内

スクリーン形状や投影する顔画像、視線提示機能の検討・設計を効率的に行なうため、実デバイスを製作する前にバーチャルリアリティ (VR) 空間内でプロトタイプを製作した (図 3)。VR 空間は Unity (2019.1.2f1) を用いて製作し、ヘッドマウントディスプレイ (Oculus Quest) を用いて観察を行なった。

スクリーン形状は、日本人青年男子の平均顔寸法 [13] を基に設計された図 4 (a) の楕円形状を鉛直軸に対して回転させた回転体形状とした。

表示する顔画像は、人物を 360 度全周撮影した 425 枚の画像から、図 5 のように顔の中心部を切り出し水平方向に合成した画像 (図 6 (a)) を生成し、楕円球型スクリーンの 3D モデルにテクスチャとして逆円筒状投影したものとした。また、後述する視線提示機能のため眼球部分を透過処理した。

2.2 視線提示機能

実際の会議で参加者が他の参加者を見る際、頭部と共に視線方向も動かす場合があるため、TELLipsoid にも視線提示機能を実装した。視線の提示および投影手法を図 7 に示す。まず Unity 内で楕円球型スクリーンの目に相当する位

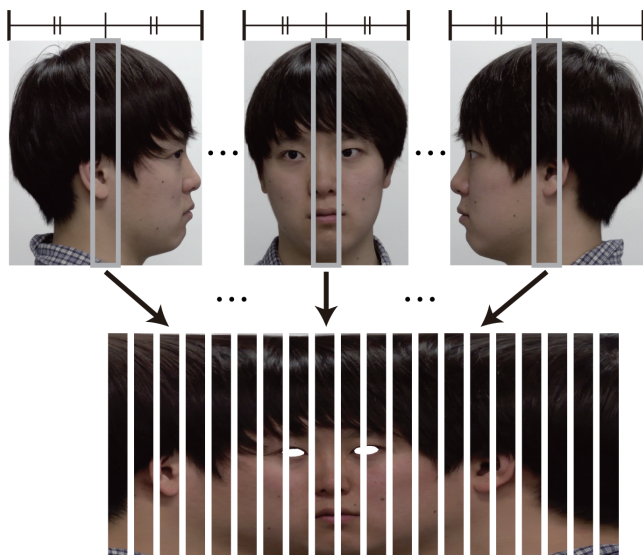


図 5 表示する顔画像の生成手法



(a) プロトタイプ 1: VR



(b) プロトタイプ 2: 実デバイス

図 6 表示する顔画像

置に眼球の 3D モデルを配置し、瞳孔を目標点に向けることで視線を提示した。眼球モデルの上下/前後位置および大きさは実際の顔と一致するように、左右位置および回転中心は眼窩の中心に揃えて調整した。次にこの眼球モデルを視線方向からスクリーン表面に平行投影し、実デバイスに用いる画像を得た。視線方向からの投影は、目標点から観察した際、原理的に視線一致が知覚される投影手法であり、球面スクリーンに線画の顔画像を投影した Kamin-FA1[14]でも用いられている。

2.3 プロトタイプ 2: 実デバイス

製作したプロトタイプ 1 とその評価を基に、実デバイス

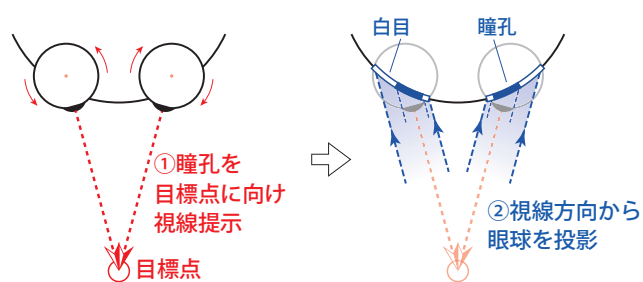


図 7 視線の提示および投影手法



図 8 実デバイスのプロトタイプ (TEllipsoid)

のプロトタイプ (図 8) を製作した。楕円球型スクリーンは厚さを 0.4 mm とし、PETG 製フィラメントを用いて 3D プリンタで出力した。プロジェクタは、スクリーン上の位置によって投影距離が大きく変わるため、小型レーザプロジェクタである Cremotech 社製の Laser Beam Pro C200 を用いた。凸面鏡には直径 110 mm、曲率半径 102.5 mm のプラスチック製球面鏡を用いた。

スクリーン形状はプロトタイプ 1 の評価・フィードバックを踏まえ、鼻や口の配置および横顔の自然さを改善するため、下側を膨大させる変更を加えた (図 4 (b))。スクリーン色は、コントラストを高めるために灰色とした。

表示する顔画像についても、一部を黒く塗り潰す加工を行なった (図 6(b))。この加工は顎の輪郭を表現し、提示像を実際の顔に近づける意図で行なわれ、塗り潰し領域は正面から見た顔が豊満にならず、かつ輪郭は自然になるよう手作業で調整された。この加工により、顔の輪郭が視線方向に影響を与える Wollaston effect (図 9) の誘起も期待される [15]。

プロジェクタから投影する顔画像は、Unity で製作した実デバイスと等価な光学系を用いて生成した。実デバイスにおけるプロジェクタの位置には仮想カメラを置き、画角をプロジェクタのスローレシオと合うように調整した。楕円球型スクリーンの 3D モデルに図 6 (b) の顔画像をテクスチャとして貼り付け、これを仮想カメラで撮影することで、投影する顔画像を得た (図 10)。最後にこの画像を実デバイスのプロジェクタから投影することで、図 8 の立体像を得ることができた。

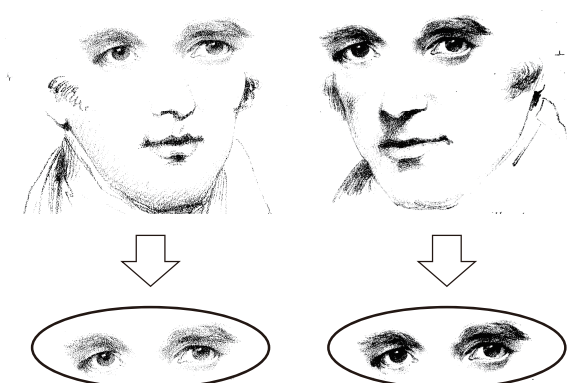


図 9 Wollaston effect (in the public domain)



図 10 プロジェクタから投影する顔画像

3. 実験

視線方向の伝達がどれだけ正しく行なえるか評価するため、3種類の実験を行なった。各実験の目的は以下の通り。

- 実験 1：特定方向からのみ視線が合ったと知覚できるか評価する。また、視線角度と視線一致を知覚する観察角度の関係を調査する。
- 実験 2.1：各観察角度について、参加者が視線一致を最も強く知覚する厳密な視線角度を調査する。
- 実験 2.2：実験 2.1 で得られた各視線角度について、参加者が視線一致を知覚する観察角度の範囲を調査する。

上述した視線角度 (θ) と観察角度 (ϕ) の定義は図 11 に示す通りである。視線角度は TELLipsoid を中心とした、正面から目標点までの方位角として定義した。また観察角度についても同様に、正面から観察位置までの方位角として定義した。

これら 3 種類の実験に加え、提示像が遠隔参加者であると認識できるか、および提示像にどのような印象を抱くか評価するため、実験 3 を行なった。

3.1 実験 1

この実験はプロトタイプ 1 の評価として、VR 空間内で行なわれた。実験参加者は研究室内の 22 歳 ~ 68 歳の成人男女 11 名であった。実験参加者はディスプレイから 1 m 離れた位置に座り、目の高さがディスプレイと一致するよ

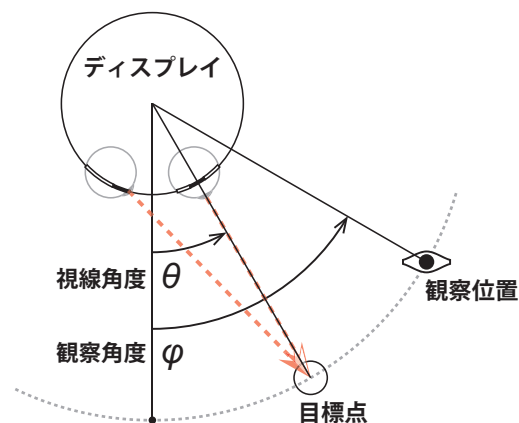


図 11 視線角度と観察角度

う高さ調整を行なった。視線提示の目標点はディスプレイから 1 m 離れた位置に設置した。

視線角度 θ は、瞳孔が自然に表示できる範囲である -40° から $+40^\circ$ までを 8° 刻みで変化させた。また観察角度 ϕ は、テーブルを囲んだ会議で観察され得る範囲として、 -90° から $+90^\circ$ までとし 10° 刻みで変化させた。

これら視線角度・観察角度の全組み合わせである 77 通りの刺激をランダムに実験参加者に提示し、各刺激について視線一致を知覚したかどうか、知覚した場合のみコントロールのボタンを押すことで回答してもらった。より直感的な判断結果を得るため、各刺激は 5 秒ずつの提示とした。

3.2 実験 2

3.2.1 実験 2.1

この実験はプロトタイプ 2 の評価として現実空間で行なわれ、実験参加者は研究室内の 22 歳 ~ 38 歳の成人男女 11 名であった。ディスプレイの視認性を高めるため、実験は暗室で行なわれた。実験参加者は椅子に座り、目の高さがディスプレイと一致するよう高さ調整を行なった。少人数会議における参加者同士の社会的距離 (Social Distance) は 1.2 m ~ 2.1 m であること [16] を考慮して、ディスプレイと実験参加者の距離は 1.5 m とした。

観察角度 ϕ は、実験 1 と同様に -90° から $+90^\circ$ まで変化させ、間隔は隣り合った会議参加者の最小角距離に近い 30° 刻みで変化させた。実験参加者の移動の手間を省くため、観察角度の変化は、位置関係が同値となるように顔画像を鉛直軸周りに回転させることで代替した。

実験参加者はランダムに提示される各観察角度に対してボタンを押すことで視線角度を調整し、視線一致を最も強く知覚する視線角度を回答した。ボタンが一度押される毎に視線角度は 2° あるいは 0.5° ずつ離散的に変化し、また順応による効果を軽減するため、ディスプレイは 0.5 秒間消灯した後再点灯するようにした。

3.2.2 実験 2.2

この実験は実験 2.1 の終了後に続けて行なわれた。実験

参加者は実験 2.1 で回答した視線角度と観察角度の組み合わせに対して、観察角度をずらしていき、視線一致を全く知覚しなくなった角度を回答した。

観察角度の変化は、顔画像を鉛直軸回転させることで代替した。実験参加者はボタンを押して顔画像を一方方向に回転させ、初めて視線一致を全く知覚しなくなったときに決定ボタンを押すことで、目的の観察角度を回答した。ボタンを押した際の挙動は実験 2.1 と同様であった。この操作は各観察角度および左右方向の全 14 通りについて、ランダムな順序で行なわれた。

3.3 実験 3

この実験では、アンケート用紙に印刷されたビジュアルアナログスケール (VAS) を用いて、提示像がモデルの人物であると認識できるか、また提示像にどのような印象を抱くか定量評価を行なった。VAS とは、2 つの単語に挟まれた直線上の任意の位置に実験参加者が印をつけることで、対象に対する印象がどちらに近いか評価する手法である。本実験は英語で行なわれ、実験参加者は以下の質問および単語対について、VAS を用いて印象を評価した。また実験参加者は全員、モデルの人物と面識があった。

- “How do you feel about the face?”
 - “Beautiful” – “Ugly”
 - “Normal” – “Creepy”
 - “Friendly” – “Unfriendly”
 - “Humanlike” – “Robot-like”
- “Can you recognize who is that person?”
 - “Easy to recognize” – “Hard to recognize”

この実験は同様に暗室で行なわれ、実験参加者は 22 歳 ~ 28 歳の成人 11 名であった。実験参加者は評価中、ボタンを押して視線角度および観察角度を調整することが認められていた。

4. 結果

4.1 実験 1

各刺激に対して視線一致を知覚した人数を図 12 に示す。左右両方向について、視線角度が大きくなるほど、視線一致を知覚する観察角度が大きくなる傾向が見られたが、視線角度と観察角度は必ずしも一致していなかった。特に視線角度 $\theta = \pm 8^\circ, \pm 16^\circ, \pm 24^\circ$ のとき、視線一致を知覚した人数の多い観察角度は複数存在していた。

また実験参加者から、ディスプレイの見た目について「シャイな印象を受けた」や「怪訝そうな印象を感じる場合がある」とのコメントがあった。

4.2 実験 2

4.2.1 実験 2.1

各観察角度に対する視線一致を最も強く知覚する視線角

視線角度 $\theta / ^\circ$	観察角度 $\phi / ^\circ$							
	-90	-60	-30	0	30	60	90	
-40	2	4	0	0	1	0	0	
-32	0	5	1	0	0	0	0	
-24	0	10	7	0	0	0	0	
-16	0	8	9	0	0	2	0	
-8	0	3	7	2	0	0	0	
0	0	1	1	11	1	1	1	
8	0	0	0	1	7	5	0	
16	0	0	0	0	9	9	0	
24	1	0	0	0	3	10	0	
32	0	0	0	0	0	6	0	
40	0	0	0	0	1	2	2	

図 12 各刺激に対して視線一致を知覚した人数

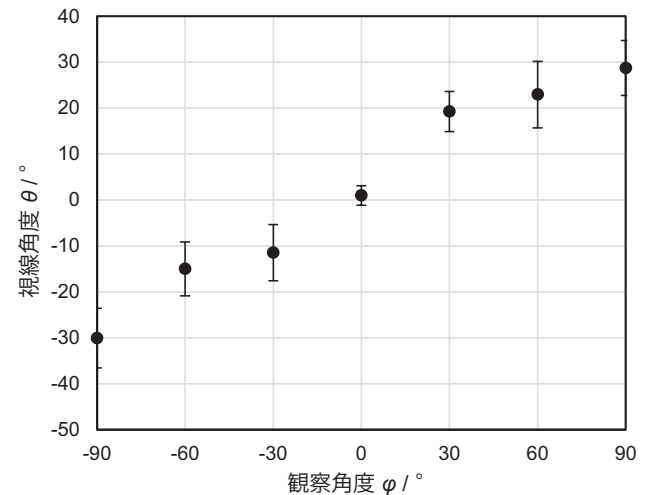


図 13 各観察角度に対する視線一致を最も強く知覚する視線角度

度を図 13 に示す。プロトタイプ 1 の結果と同様に、観察角度が大きくなるほど視線一致を知覚する視線角度が大きくなる傾向が左右両方向について見られたが、視線角度と観察角度は必ずしも一致しなかった。

標準偏差は比較的大きく、最大で観察角度 $\phi = +60^\circ$ のとき 7.25° であった。また最小でも観察角度 $\phi = 0^\circ$ のとき 2.12° であり、これは実際の顔における結果より数倍大きかった [17]。

4.2.2 実験 2.2

実験 2.1 で回答した視線一致を最も強く知覚する視線角度を提示した際に、視線一致を知覚する観察角度の範囲を図 14 に示す。また詳細な値を表 1 に示す。観察角度 $\phi = +90^\circ (-90^\circ)$ の右端 (左端) の値は、実際の会議でこの位置に参加者が座ることがないため無視する。

視線一致を最も強く知覚する観察角度から範囲の端までの角度幅は、会議参加者同士の最小角距離に近い 30° より

も概ね小さく、最大でも 38.5° であった。これは 1.5 m 離れた位置では 99 cm に相当する。また角度幅は $\phi = \pm 30^\circ$ のとき右に大きく、 $\phi = \pm 60^\circ, \pm 90^\circ$ のとき左に大きかった。 $\phi = 0^\circ$ のとき角度幅は比較的小さく、偏りも少なかった。

4.3 実験 3

実験参加者がビジュアルアナログスケール (VAS) に記録した印は、線分の左端から印までの距離として数値化し、さらに線分全体の長さに対する比に変換した。各質問および単語対についての、比の平均値・標準偏差・中央値を表 2 に示す。

質問 “How do you feel about the face?” について、比の平均値は全ての単語対で二単語の間である 0.5 に近かった。一方で質問 “Can you recognize who is that person?” については、平均値は “Easy to recognize” に比較的近い 0.357 であり、中央値は更に近い 0.271 であった。

5. 考察

5.1 モナリザ効果の解消

実験 1 の結果より、顔画像が正面を見ているときに視線一致を知覚するのは概ね正面のみであり、また顔画像が正面を見ていないときも、特定の観察範囲で視線一致を知覚しやすい傾向があった。さらに実験 2 でも同様の傾向が示されており、どの方向から観察しても、ある狭い範囲に視線が向いたときにのみ視線一致を知覚する傾向があった。

したがって今回の提案手法では、モナリザ効果はある程度解消されたといえる。

5.2 視線角度と観察角度の関係

実験 1, 2 いずれにおいても、視線角度と視線一致を知覚する観察角度は実際と異なり必ずしも一致しておらず、観察角度が視線角度より大きくなる傾向にあった。この傾向の一因として、提示像の鼻が平面的であったことが考えられる。

実際の顔における視線方向の知覚について、立体的な鼻の向きが影響を及ぼすこと [18] が報告されている。しかし今回の提示像は実際の顔を楕円球形状に投影したものであり、鼻は平面的になっていた。そのため立体的な鼻の向きの変化が提示できず、これが視線方向の知覚に影響を与えて視線角度と観察角度の不一致をもたらしたと考えられる。また同様に、眼球部の形状が異なっていたことも不一致の一因として考えられる。この不一致は眼球オブジェクトの前後位置および曲率を最適化することで改善される可能性があると考えられる。

いずれにせよ、視線角度と視線一致を知覚する観察角度の関係をより詳細に評価し定式化することができれば、視線角度と観察角度を一致させなくとも、任意の観察角度で視線一致を知覚させることができると考えられる。

5.3 観察角度についての Cone of Gaze

実験 1 の結果から、視線角度が $\pm 8^\circ, \pm 16^\circ, \pm 24^\circ$ のとき、テーブルを囲んだ会議において同時に複数の参加者が視線一致を知覚するおそれがあった。そこで視線一致を知覚しうる観察範囲を実験 2.2 で詳細に調べた結果、視線一致を最も強く知覚する観察角度から観察範囲の端までの最大幅は 38.5° であった。したがって、会議で隣り合った参加者の角距離が 38.5° 以上あれば、TEllipsoid は参加者一人ずつに視線提示することが可能であると考えられる。

この角距離は、視線提示機能における眼球位置の調整や、瞳孔のハイライトの実装によってさらに小さくできると考えられる。

知覚心理学において、ある観察角度から視線一致を知覚する視線角度の範囲に明確な境界はなく、視線角度に対する視線一致確率の正規分布 (Cone of Gaze) として存在することが報告されている [19]。本実験で求めた、ある視線角度に対して視線一致を知覚する観察角度の範囲も同様に境界ではなく分布を持つことが考えられ、この分布は Cone of Gaze の畳み込みとなることから同様に正規分布となることが推測される。

実験 2.2 において実験参加者は視線一致を全く知覚しなくなる境界を回答したが、これは分布の端に相当すると考えられ、心理学的に有意な視線一致を知覚する観察角度の範囲はさらに狭い可能性がある。この分布および絶対閾は、階段法や PEST [20] などの心理物理学測定法を用いて求めることができる。

5.4 実物と提示像の同一性

実験 3 の結果より、提示像に対する印象は 4 つの単語対についてほぼ中立であり、提示像が悪印象を与えてインタラクションを阻害することは起こりにくいと考えられる。しかしながら、単語対 “Normal” – “Creepy” についてネガティブな印象を感じた参加者もあり、不気味の谷現象 [21] が起こっている可能性がある。要因の一つとして瞳孔運動が連続的であったことが考えられ、アイトラッキング機能を実装し実際の目のサッケード運動を反映させることでこの印象は改善されることが考えられる。また人間に酷似したアンドロイドではまばたきが好印象形成に強く寄与すること [22] が示されており、TEllipsoid でもまばたきの表出機能を実装することで印象を改善できる可能性がある。

質問 “Can you recognize who is that person?” における単語対 “Easy to recognize” – “Hard to recognize” の中央値は 0.271 であり、半分以上の実験参加者は提示像が遠隔参加者であると簡単に認識できたことが示された。したがって今回の提案手法では、実物と提示像の同一性は高いといえる。

一方で認識しやすさは個人によるばらつきも比較的大きく、要因の一つとして参加者がこれまでに見てきた顔の種

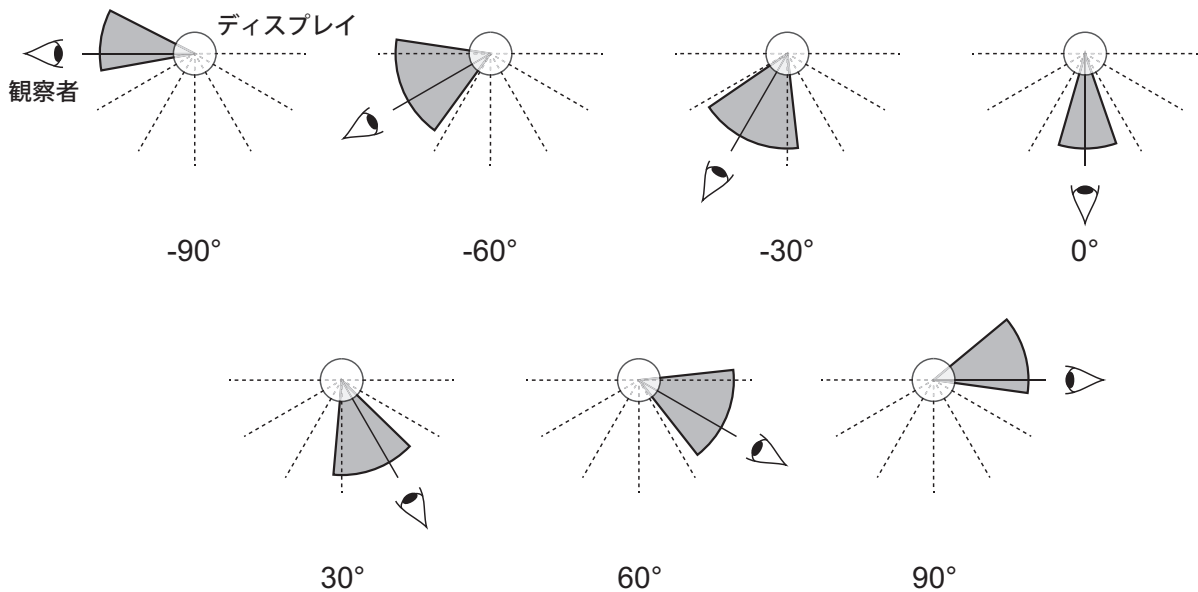


図 14 実験 2.1 で回答した視線角度を提示した際に視線一致を知覚しうる観察角度の範囲

表 1 実験 2.2 の結果 (単位: °)

	観察角度	-90	-60	-30	0	30	60	90
左端	平均	(-117.0)	-98.5	-55.6	-16.1	-5.0	38.1	82.1
	標準偏差	(15.5)	14.5	13.5	5.5	12.7	10.0	3.5
	角度幅	(27.0)	38.5	25.6	16.1	35.0	21.9	7.9
右端	平均	-79.8	-36.1	6.4	19.1	45.8	96.2	(129.6)
	標準偏差	3.9	8.3	17.3	8.9	6.9	14.1	(21.0)
	角度幅	10.2	23.9	36.4	19.1	15.8	36.2	(39.6)

類の違いがあると考えられる。人間はこれまでに会った顔を基に平均的な顔（プロトタイプ）を抽出し、相手を認識する際にはこのプロトタイプからの逸脱情報を参照しているというプロトタイプ仮説 [23] が提唱されており、プロトタイプに近い顔ほど認識しやすいといわれている。このプロトタイプが個人によって異なっているために、認識しやすさにばらつきがあったと考えられる。

6. 結論

本論文では、実用時に重要な利便性および実物と提示像の同一性を確保しながら、適切な視線方向の伝達が行なえる映像提示手法として、楕円球型のスクリーンに顔画像を投影するシステム（TEllipsoid）を提案した。このシステムは、底部の小型プロジェクタから投影された映像を凸面鏡で反射させ、楕円球型スクリーンに投影することで顔画像を表示するものである。

本研究で我々はまず VR 空間内でプロトタイプを製作して仕様を検討し、その後実デバイスの製作を行なった。提示する顔画像は人物を 360° 全周撮影し合成した画像を基に作成し、実物との同一性を高めるために一部を加工した。また視線提示機能も、眼球の 3D モデルを楕円球型スクリーンに投影することで実装した。

TEllipsoid の視線伝達能を評価したところモノリザ効果は解消されており、テーブルを囲んだ会議で隣り合った参加者の角距離が 38.5° 以上ある場合には、適切な視線方向の伝達が可能であることが示された。さらに印象評価を行なったところ、提示像に対する印象は中立であり、また半分以上の実験参加者が提示像が誰であるかを簡単に認識できたことが示された。このことから、本研究の提示像は実際の外見との同一性が高いといえる。

今後は遠隔会議システムとして完成させることを目指しており、以下の機能を実装することを考えている。

- 表情・まばたき・口唇動作の表出機能
- 頭部運動・表情のキャプチャおよび同期機能
- 遠隔参加者への会議室の 180° パノラマ映像伝送機能

それに加え、別の顔画像を用いた場合の視線提示能・印象の評価やデバイスと視線一致しない場合の視線方向の伝達能評価、視線伝達能の上下方向・前後方向依存性の評価など、さらなる評価も今後必要であると考えている。

本研究では顔パーツ配置以外の情報、すなわちパーツ要素の形や色などが同一性に影響する可能性を現状排除できないため、実写の顔画像を用いた。しかしながら、顔パーツ配置が実物と同様に見える線画でも同一性の高さを担保できる可能性は否定できず、特に類似度と好感度を両立で

表 2 実験 3 の結果

	平均	標準偏差	中央値
質問: “How do you feel about the face?”			
Beautiful(0) – Ugly(1)	0.499	0.123	0.504
Normal(0) – Creepy(1)	0.453	0.175	0.421
Friendly(0) – Unfriendly(1)	0.481	0.219	0.504
Humanlike(0) – Robot-like(1)	0.405	0.201	0.329
質問: “Can you recognize who is that person?”			
Easy to recognize(0) – Hard to recognize(1)	0.357	0.262	0.271

きる普遍的なデフォルメ則 [8] が報告されていることから、提示像を実写顔画像に適切なデフォルメを加えた線画顔画像とすることで、いまの特性を保ちながらさらに好感度を高められる可能性も示唆される。

また遠隔医療用アバタについて、利用者と同体型のアバタよりも太っているアバタの方が信頼されやすいこと [24] が報告されており、本研究においても実物との同一性を保ちつつ、より太って見える顔画像を提示することで対話者に与える印象を改善できる可能性がある。

デバイス構成について、現状ではレーザプロジェクトの出力が低く一般的な会議室の明るさでは顔画像が不明瞭になる問題がある。今後の技術革新によって高出力化されることを期待したい。

7. 関連研究

7.1 話者交替

人間が他者に対して好意 (liking) を伝達する際に、言葉・声調・表情が占める割合は、

$$\text{Total Liking} = 7\% \text{ verbal liking} + 38\% \text{ vocal liking} + 55\% \text{ facial liking}$$

であり、非言語情報が 93% を占めるといわれている [1]。

対面会話において会話の参加者は、承認された参加者 (Ratified-participant) と立ち聞き者 (Overhearer) に分けられ、前者はさらに話し手 (Speaker)・聞き手 (Addressee)・傍参加者 (Side-participant) に分けられる [25][26]。話し手のターンが終了し次話者が話し始める、すなわち参加者の役割が交替することを話者交替といい、この話者交替は視線や身振りによって調整されることが示されている [2][3]。特に視線については、話し手が聞き手を凝視することが他者にターンを譲る合図となり、目の合った聞き手が視線を逸らすことがターンを受け入れる合図となることが知られている。遠隔会議で平面ディスプレイを用いた場合、モナリザ効果が発生することで単一の聞き手への凝視が難しくなり、話者交替が円滑に行なわれなくなる問題がある。

この話者交替を人間とロボットエージェントとの間で行なう際、両者が同時に発話して発話衝突が起きてしまい、円滑なコミュニケーションが阻害される問題があった。これに対し、エージェントが明滅光源を用いて内部状態を表出し衝突を抑制する手法 [27] が提案された。エージェン

トが人間の発話を認識し応答するまでの間に光源を明滅させたところ、発話衝突をもたらす人間の不要な発話の言い直しが抑制されたほか、望ましい言い直しは促進され、またロボットに対する印象も改善されることが示唆された。TEllipsoid についても、顔画像を明滅させて遠隔参加者の状態を補助的に表現することで、コミュニケーションをさらに円滑化できる可能性がある。

7.2 顔パーツ配置の重要性

顔認識におけるプロトタイプ仮説の適応可能範囲を調べた研究 [9] では、顔画像の目・鼻・口の位置を変えたものおよびモーフィングにより目・鼻・口の形を変えたものについて、顔画像の変化の認識性が評価された。その結果、目・鼻・口の位置の変化がより認識されやすく、顔パーツ配置が顔認識でより重要であることが示された。本研究ではこれに基づき、顔パーツ配置が実際の外見と同様に見える楕円球型ディスプレイを提案した。

顔形態印象の客観的評価手法の研究 [10] において、著者らは類似性を基準とした顔画像の分類課題から、目の大きさ印象・目より下の顔輪郭印象が顔の認識において重要であると仮定した。さらに著者らは、この 2 印象を基準とした顔画像の評定課題から印象評定値を推定し、また 2 印象を軸とした顔形態マップを作成した。このマップに分類課題の顔画像をプロットしたところ、分類課題と同様の分類ができ、仮定の妥当性が確認された。本研究ではこれに基づいて、顔の輪郭が実際の外見と同様に見える楕円球型ディスプレイを提案した。

7.3 非言語情報の伝達

水平視差型裸眼立体ディスプレイ [5] は、モナリザ効果を解消可能な一手法である。このディスプレイは、回転するアルミニウム板および高速プロジェクトを用いて遠隔参加者の顔画像を立体的に表示できる。著者らが視線伝達能を評価したところ、提示する視線方向と認識される視線方向の誤差は $3^{\circ} \sim 5^{\circ}$ の範囲に収まっていた。しかしながら、駆動機構を必要とするため騒音や保守運用の問題がある。

人の顔型を模したスクリーンに顔画像を投影する遠隔会議用立体顔形状ディスプレイ [6][28] でも、モナリザ効果を解消できることが報告されている。しかしながら頭部の

回転には駆動機構が必要で騒音などの問題があり、特にスクリーン形状を利用者の顔から型取りするもの [6] は、利用者それぞれに固有の設備が必要となるため汎用性は高くない。

円筒型多視点ディスプレイを用いた遠隔会議システム [7] は、観察方向に応じて異なる顔画像を表示することで、適切な視線方向の伝達を行なうシステムとして提案された。著者らが視線伝達能を評価したところ、このシステムの伝達能は平面ディスプレイに比べて高かった。また著者らは、円筒に単一の顔画像を投影した円筒型一視点ディスプレイについても対照評価を行ない、こちらも視線伝達能は平面ディスプレイより高く、モノリザ効果が解消できることを報告した。特にこの円筒型一視点ディスプレイはシステムが簡易で利便性も高いが、顔パーツ配置が実際の外見と同様に見える視野角は限られている。

空間的に正確な多対多遠隔会議システム [29] は、参加者ごとに用意されたプロジェクタとカメラ、および再帰性反射スクリーンを用いて、参加者各々に空間的に正しい別の映像を提示し、対称な配置のグループ同士での遠隔会議を行なうシステムである。実験の結果、適切な視線方向の伝達はある程度行なえていたが、視線一致を知覚しづらい問題があった。また参加者それぞれに設備が必要であり、3対3の会議では片側の表示部だけで\$3,700と、コストもTEllipsoidの数倍高い。

移動・回転する平面ディスプレイを用いた対称型多対多遠隔会議システム [30] は、顔画像が投影された平面パネルを4軸のアクチュエータで移動・回転させることで身体動作を表現し、また参加者ごとに配置されたカメラを用い擬似的なアイコンタクトを可能にする遠隔会議システムである。著者らは印象評価および非言語行動分析を行ない、機械的運動によって非言語情報がより適切に伝達されることを示した。しかしながらシステムが複雑で汎用性が低く、駆動機構を用いるため保守などの問題もあり、またモノリザ効果自体は解消されていなかった。

DFD現象を利用した2層式ディスプレイ [31] は、手前の層に顔画像の中心部、奥の層にそれ以外を表示した2枚の平面ディスプレイを用い、観察位置で変化する2画像の相対位置によって視線方向の提示を行なうデバイスである。正面から見て2画像が整合している場合に視線一致を知覚する観察角度の幅は約 17° であり、整合していない場合も特定の狭い範囲でのみ視線一致を知覚する傾向があったことから、適切に視線方向が伝達できる可能性が示唆された。しかしながら、斜め方向から見たときに2画像が分離して見えやすい問題があり、広い範囲に視線方向を提示するには駆動機構が必要になる問題もあった。

対話者の視線を提示する眼球型ディスプレイ [32] は、平面ディスプレイと半球型に加工されたテレビ石を用いて眼球を立体的に表示し視線方向の提示を行なうデバイスであ

り、遠隔会議システムのアドオンとしての利用が想定されている。著者らは視線伝達能の評価を行ない、眼球型ディスプレイの視線伝達能が平面ディスプレイより高いことを示した。しかしながら眼球型ディスプレイと平面ディスプレイを組み合わせた想定用途での評価は行なわれておらず、眼球型ディスプレイに注意が向けられないおそれもあることから視線伝達能が改善されるかは不明瞭である。

半球型スクリーンに線画の顔画像を投影した顔ロボット、Kamin-FA1 [14] の視線提示手法は本研究の手法と類似しており、眼球オブジェクトの前後位置のみ異なる。本研究では実際の顔を基にした顔画像を用いているため、眼球オブジェクトは実際の顔パーツ配置に合わせ眼窩から突出する位置にあるが、Kamin-FA1では半球型スクリーンの内側に配置されている。

デフォルメされた顔形状スクリーンにアニメ調の顔画像を背面投影した顔形状ディスプレイに関する研究 [33] では、顔形状ディスプレイ・半球型スクリーンにアニメ調の顔画像を投影した半球型ディスプレイ・実際の顔・実際の顔を表示した平面ディスプレイ、の4種類について視線伝達能の評価が行なわれた。半球型ディスプレイの視線提示手法は、Kamin-FA1と同じ手法が用いられた。その結果、半球型スクリーンの視線伝達能が最も低かったことが報告されている。半球型スクリーンでは顔パーツや輪郭を人間の顔と同様に表示することが難しく、顎の向きなど視線方向の知覚にかかわる顔パーツ等の情報 [18] が欠如したため、視線伝達能が低くなったと考えられる。

7.4 視線方向の知覚

Cone of Gazeに観察距離や頭の向きがどう影響するか調べた研究 [17] では、3Dキャラクタだけでなく実際の顔のCone of Gazeも測定された。実験参加者はアシスタントの視線方向を操作し、正面顔を1mの距離から観察したときに視線一致を知覚する視線方向および視線一致を知覚する視線方向の範囲(Cone of Gaze)を回答した。視線方向の操作は、実験参加者がボタン操作で機械的にターゲット位置を調整し、観察対象のアシスタントが追従してターゲットを見ることで行なった。その結果、視線一致を知覚する視線方向はほぼ正面であり、標準偏差は 0.5° 以内であることが示された。またCone of Gazeの大きさが 8.12° であることも示された。本研究でも視線一致を知覚する視線方向について同様の評価が行なわれ、類似した結果が得られたものの標準偏差は 2.12° と実際の顔に比べて大きかった。

鼻の向きが視線方向の知覚に与える影響を調べた研究 [18] では、正面を見ている/見ていない顔画像の弁別性を、鼻の向きが視線方向と一致した顔画像(A)・鼻の向きが視線方向と一致しない顔画像(B)・目の部分を切り出した画像(C)について調査した。その結果、弁別性は大きい順

に A,C,B であり, 鼻の向きが視線方向の知覚に影響を与えることが示された. 本研究では, 提示像の鼻が平面的であることが視線方向の知覚に影響を与え, 視線角度と観察角度の不一致をもたらしたと考えられる.

7.5 デバイスを与える印象

実物と外見を一致させた線画アバタは不気味になりやすく, これを改善するために類似度と好感度を両立する個人化アバタのデフォルメ法 [8] が提案された. 著者らはマンガにおける人物画が親しみを持たれていることに注目して, 実在モデルのマンガ風デフォルメ顔画像から顔パーツの拡大率と下方移動度をパラメータとしたデフォルメ則を抽出した. その結果, 個人に依存しない共通したデフォルメ傾向が得られた. また類似度と好感度が評価され, 実物と外見を一致させた線画アバタよりも類似度・好感度共に高いデフォルメアバタが生成できたことが示された. 本研究においても, 提示像を実写顔画像に適切なデフォルメを加えた線画顔画像とすることで, さらに好感度を高められる可能性がある.

見かけも振る舞いも人間に酷似したアンドロイド Repliee[22] は, ロボットの見かけと振る舞いを区別してコミュニケーションに与える影響を調べるために開発された. 人間の型を取った子供型アンドロイド Repliee R1 をはじめ, 世界の女性の平均顔を参考に造形した大人型アンドロイド Repliee Q1, Repliee Q1 の頭部の人間らしさを改良した Repliee Q2 が開発された. Repliee Q2 について無自覚な動きによる印象を評価した結果, 視線を動かす動作・まばたき・頭の向きを変える動作・呼吸による肩の上下運動のうち, まばたきが最も好印象を与える動作であることが示された. 本研究においてもまばたきの表出機能を実装することで印象を改善できる可能性がある.

謝辞 本研究は, JSPS 科研費 19K12284 の助成を一部受けたものです. 本研究を遂行するにあたり熱心にご指導してくださった長谷川晶一准教授, 些細な相談にも乗ってくださり沢山のアドバイスをくださった三武裕玄助教, また日々の生活を楽しいものにして頂き, 実験にもご協力頂いた長谷川研究室の皆様へ深く感謝申し上げます. また技術的なアドバイスを下さった山口雅浩研究室の五十嵐俊亮様にもこの場を借りて感謝申し上げます. Finally, I deeply appreciate Mr. KITMUTI Tanapon for all your kindness and help in my research and life.

参考文献

- [1] A. Mehrabian. Silent messages. Wadsworth Pub. Co., 1971
- [2] S.Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292, Aug. 1972. doi: 10.1037/h0033031
- [3] A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26(1):22–63, Jan. 1967. doi: 10.1016/0001-6918(67)90005-4
- [4] S.Rogers, M.Lunsford, L.Strother, and M.Kubovy. *The Mona Lisa effect: Perception of gaze direction in real and pictured faces*, pp. 19–24. Lawrence Erlbaum, Jan. 2003.
- [5] A.Jones, M.Lang, G.Fyffe, X.Yu, J.Busch, I.Mcdowall, M.Bolas, and P. Debevec. Achieving eye contact in a one-to-many 3d video teleconferencing system. *ACM Transactions on Graphics*, 28(3):64:1–8, Aug. 2009. doi: 10.1145/1531326.1531370
- [6] 三澤 加奈, 石黒 祥生, 暦本 純一, 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, インタラクシオン 2012, pp.41-48, 2012
- [7] Y.Pan and A.Steed. A gaze-preserving situated multi-view telepresence system. In *Proc. CHI*, pp. 2173–2176. ACM, New York, NY, USA, 2014.
- [8] 平家 雅之, 川崎 裕達, 田中 貴紘, 藤田 欣也, 類似度と好感度を両立する個人化アバタのデフォルメ法の検討, ヒューマンインタフェース学会論文誌, Vol.13, No.3, p. 243-254, 2011.
- [9] R. Cabeza, V. Bruce, T. Kato, and M. Oda. The prototype effect in face recognition: Extension and limits. *Memory & Cognition*, 27(1):139–151, Jan 1999. doi: 10.3758/BF03201220
- [10] 井上 さくら, 山本 美恵子, 山崎 和広, 顔形態印象の客観的評価技術の開発, 日本化粧品技術者会誌, Vol.34, No.3, pp. 249-254, 2000.
- [11] 一居 太朗, 三武 裕玄, 長谷川 晶一, 楕円球面スクリーンに投影された顔画像による視線方向の伝達, 第 24 回日本バーチャルリアリティ学会大会論文集, 1C-04, 2019.
- [12] T.Ichii, H.Mitake, and S.Hasegawa. Tellipsoid:Ellipsoidal display for videoconference system transmitting accurate gaze direction. In *Proc. IEEE VR*.
- [13] 河内まき子・持丸正明, 2005AIST 人体寸法データベース, 産業技術総合研究所 H16PRO 287
- [14] M. Hashimoto, H. Kondo, and Y. Tamatsu. Effect of emotional expression to gaze guidance using a face robot. In *Proc. RO-MAN*, pp. 95–100. IEEE, Munich, Germany, Aug 2008. doi: 10.1109/ROMAN.2008.4600649
- [15] W.H.Wollaston. Xiii. on the apparent direction of eyes in a portrait. *Philosophical Transactions of the Royal Society of London*, 114(1):247–256, Jan. 1824. doi: 10.1098/rstl.1824.0016
- [16] E.T.Hall. *The hidden dimension*. Anchor Books, Garden City, N.Y., United States, 1969.
- [17] M.Gamer and H.Hecht. Are you looking at me? measuring the cone of gaze. *Journal of experimental psychology. Human perception and performance*, 33(3):705–715, Jun. 2007. doi: 10.1037/0096-1523.33.3.705
- [18] S. Langton, H. Honeyman, and E. Tessler. The influence of head con-tour and nose angle on the perception of eye-gaze direction.*Perception & psychophysics*, 66:752–71, 08 2004. doi: 10.3758/BF03194970
- [19] J. Hakala, J.Kätsyri, T. Takala, and J.Häkkinen. Perception of stereo-scopic direct gaze: The effects of interaxial distance and emotional facial expressions.*Journal of vision*, 16:5, 2016. doi: 10.1167/16.9.5
- [20] M. Taylor and D. Creelman. PEST: Efficient estimates on probability functions.*The Journal of the Acoustical Society of America*, 41:782–787, 04 1967. doi: 10.1121/1.1910407
- [21] M.Mori, K.F.MacDorman, and N.Kageki. The un-

- canny valley [from the field]. *IEEE Robotics Automation Magazine*, 19(2):98–100, Jun. 2012. doi: 10.1109/MRA.2012.2192811
- [22] 港 隆史, 石黒浩, アンドロイドの顔における人間らしさ (顔を見る, 顔を思う, 顔をつくる-顔知覚研究の行方-,2005 年度 第 2 回フォーラム), 基礎心理学研究, Vol.25, No.1, p. 96-102, 2006.
 - [23] 九島 紀子, 齊藤 勇, 顔パーツ配置の差異による顔印象の検討, 立正大学心理学研究年報, No.6, pp.35-52, 2015.
 - [24] H. Vugt, E. Konijn, J. Hoorn, and J. Veldhuis. Why fat interface characters are better e-health advisors. vol. 4133, pp. 1–13, 08 2006. doi: 10.1007/11821830_1
 - [25] E. Goffman. *Replies and responses. Language in Society*, 5(3):257–313, 1976.
 - [26] H. H. Clark and T. B. Carlson. Hearers and speech acts. *Language*, 58(2):332–373, 1982
 - [27] 船越 孝太郎, 小林 一樹, 中野 幹生, 山田 誠二, 北村 泰彦, 辻野 広司, Artificial Subtle Expression としての明滅光源による音声対話の円滑化, 電子情報通信学会論文誌, Vol.J92-A, No.11, pp.818-827, 2009.
 - [28] S. Al Moubayed, J. Edlund, and J. Beskow, Taming Mona Lisa: communicating gaze faithfully in 2D and 3D facial projections, *ACM Transactions on Interactive Intelligent Systems*, vol. 1, no. 2, p. 25, 2012.
 - [29] D. T. Nguyen and J. Canny. Multiview: Improving trust in group video conferencing through spatial faithfulness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI' 07, p. 1465–1474. Association for Computing Machinery, New York, NY, USA, 2007. doi:10.1145/1240624.1240846
 - [30] 大塚和弘, 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, 電子情報通信学会 信学技報 メディアエクスペリエンス・バーチャル環境基礎研究会 MVE2018-2, pp.27-32, 2018.
 - [31] K. Iso, S. Ozawa, M. Date, H. Takada, Y. Andoh, and N. Matsuura. Video conference 3D display that fuses images to replicate gaze direction. *Journal of Display Technology*, 8(9):511–520, Sep. 2012. doi: 10.1109/JDT.2012.2193662
 - [32] 大槻 麻衣, 丸山 啓太, 葛岡 英明, 鈴木 雄介, ThirdEye: 遠隔対話システムにおける対話者の視線を提示する眼球ディスプレイの開発と評価, ヒューマンインタフェース学会論文誌, Vol. 19, No.2, p. 211-218, 2017.
 - [33] F.Delaunay, J.deGreeff, and T.Belpaeme. A study of a retro-projected robotic face and its effectiveness for gaze reading by humans. In *Proc. HRI*, pp. 39–44. IEEE, Osaka, Japan, March 2010. doi: 10.1109/HRI.2010.5453271