

推薦論文

MM-Space：頭部運動の物理的補強表現 に基づく会話場再構成

大塚 和弘^{1,a)} 熊野 史朗^{1,b)} 松田 昌史^{1,c)} 大和 淳司^{1,d)}

受付日 2012年6月20日, 採録日 2013年1月11日

概要: 複数人物による対面会話場面を実世界に再構成するためのシステムを提案する。本研究では、時間・空間を隔てた会話の状況を、あたかもその場にいるかのように観測できるシステムの実現を目指し、会話者の顔画像をスクリーンに投影しつつ、その人物の頭部運動をスクリーンの物理的な運動として補強して提示するという表現方法を提案する。本システムでは、会話者の配置に合わせて、複数のプロジェクタ、および、透過型スクリーンが配置され、各スクリーンに等身大の人物像が投影される。スクリーンにはアクチュエータが接続され、計測された会話者の頭部姿勢と同期して、スクリーンの姿勢が動的に制御される。人物の頭部運動には、視線方向の遷移にともなう首振りや傾きが含まれ、このスクリーンの物理的な運動と画像運動との相乗効果により、会話者の動作がより明確に観測者により知覚されることが期待される。さらに枠なし透明スクリーンへの背景なし人物像の投影により、高い存在感の醸成も期待できる。実験により、このスクリーンの動きの効果として、会話者の視線方向がより分かりやすく、かつ、話し手の話しかける相手がより正確に理解されるという仮説の妥当性が示唆された。

キーワード: テレプレゼンス, 対面会話, マルチモーダルインタラクション, 視線, 視覚的注意, プロジェクションマッピング, 会話シーン分析

MM-Space: Re-creating Multiparty Conversation Space Based on Physically Augmented Head Motion

KAZUHIRO OTSUKA^{1,a)} SHIRO KUMANO^{1,b)} MASAFUMI MATSUDA^{1,c)} JUNJI YAMATO^{1,d)}

Received: June 20, 2012, Accepted: January 11, 2013

Abstract: A novel system is presented for reconstructing multiparty face-to-face conversation scenes in the real world. We aim to develop a system that can display and playback recorded conversations as if the remote people were taking in front of the viewers, and propose a novel representation scheme that physically augments human head motions with facial images of each conversation participant. This system consists of multiple projectors and transparent screens attached to actuators. Each screen displays the life-size face of a different meeting participant, and are spatially arranged to recreate the actual scene. The main feature of this system is dynamic projection; screen pose is dynamically controlled in synchronization with actual head motions of the participants to emulate their head motions, including head turning, which is typically accompanied with shifts in visual attention, and head nodding. It is expected that physically augmented screen movement with image motion can boost viewers' understanding of people's behaviors. In addition, we expect that the human images without background, which are projected on to the transparent screens, can enhance the presence of the remote people. Experiments suggest that viewers can more clearly discern the visual focus of attention of meeting participants, and more accurately identify the addressees.

Keywords: telepresence, face-to-face conversation, multimodal interactions, gaze, visual attention, projection mapping, conversation scene analysis

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 Labs., NTT Corporation, Atsugi, Kanagawa 243-0198, Japan

a) otsuka.kazuhiro@lab.ntt.co.jp

b) kumano.shiro@lab.ntt.co.jp

c) matsuda.masafumi@lab.ntt.co.jp

d) yamato.junji@lab.ntt.co.jp

1. はじめに

人と人が対面で行う会話は、最も基本的なコミュニ

本論文の内容は2012年3月のインタラクション2012で報告され、同プログラム委員長により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

ケーションの形態であり、日々、情報の伝達・共有、他者の感情や意図の理解、グループでの意志決定が会話を通じて行われている。昨今、空間を隔てた会話を実現するため、映像通信の活用が期待されているが、いまだ対面の会話ほど自然なコミュニケーションを実現するに至っていない。著者らは、より良い遠隔コミュニケーション環境を実現するには、まず、人の会話のメカニズムについて理解を深める必要があるという考えのもと、会話シーン分析の研究を進めている [1]。その一例として、頭部動作や発話区間などの非言語行動の計測、人物間のインタラクション (e.g. 「誰が誰に回答しているか」) や会話の構造 (e.g. モノログ、ダイアログ), 対人感情 (e.g. 「誰が誰に共感しているか」) の自動推定などの課題への取り組みがあげられる。今後、このような会話シーン分析の成果を実際のコミュニケーション・システムへと活用すべく、本論文ではその部分課題として、複数人物の対面会話の場を事後に再現するという問題に着目し、その1つの解として新しい表現法、および、システムの提案を行う。

対面会話においては、視線や頭部ジェスチャ、顔表情、韻律など様々な非言語情報が交わされている [2]。我々は、遠隔コミュニケーション環境では、これらの非言語情報が十分に伝達されないことが、その不自然さの一因であると考える。これを何らかの手段で補い提示する必要があると考えている。様々な非言語情報の中でも、本研究では、頭部の姿勢や運動として表出される情報に着目している。その1つが「誰が誰を見ているか」という視線方向 (視覚的注意の焦点とも呼ばれる) である。視線には、他者のモニタリング、自身の態度・意図の表出、会話の流れの調整といった機能がある [3]。人は、興味の対象を視野の中心にとらえることから、対象との位置関係に応じた頭部姿勢 (顔向き) として視線方向が表出される。さらに、頷き、首振り、傾げといった頭部ジェスチャも重要な非言語情報である。話し手の頭部ジェスチャは、話しかけ・問いかけや発話の強調のサインであり、また、聞き手の頭部ジェスチャは、傾聴や承認、同意・不同意、理解の程度を表すサインとして機能することが知られている [4]。

本研究では、このような非言語情報を含めた会話の場を実空間上に再構成し、時間・空間を隔てた会話の状況を、あたかもその場にいるかのように観測可能なシステムの実現を目指している。特に視線方向や頭部ジェスチャの表出を担う頭部運動に着目し、人物像を投影するスクリーンの物理的な運動として頭部運動を補強して表現するという方法を提案する。本システムは、会話者の人物配置に合わせて、複数のプロジェクタおよび透明スクリーンを配置し、等身大の人物像を投影する。スクリーンにはアクチュエータが接続され、会話者の頭部運動と同期してスクリーンの姿勢が制御される。このスクリーンの物理的な運動と投影画像上の動きとの相乗効果により、会話者の頭部動作が

より明瞭に知覚でき、その結果、視線方向や話し手の話しかける相手 (受け手と呼ぶ)、および、頭部ジェスチャが、より明瞭に観測者によって理解されることを期待する。また、透明スクリーンへの背景除去された人物像の投影により、遠隔人物の顔が実際の部屋の背景に重畳されて表示され、高い存在感の実現も期待される。なお、本論文のシステムは、オフライン動作を対象とするが、双方向の遠隔会話を実現するシステムの構成要素として位置づけられる。

特に本論文では、頭部運動の補強表現の効果として、会話者の視線方向の理解が促進され、その結果、受け手人物がより正確に同定できるという仮説を立て、被験者実験によりその妥当性を示唆する。なお、会話場の再現には、視線方向以外の非言語情報の表現も重要であると考えられるが、本論文では、遠隔映像通信システムの典型的課題である「視線不一致」の現象と関連して、視線方向の正確な伝達を特に先決な課題として取り上げる。

頭部運動の物理的補強表現は、「バイオロジカルモーション」[5], [6], 「心的帰属」[7] と呼ばれる人間の知覚の性質に関連する。これら理論によると、人間には、光点や図形の動きを人や生物の動きに「見立て」、その原因を対象の心的状態や社会的文脈に帰属させて解釈する傾向があるとされる。それにより正方形スクリーンという単純形状であっても、人間の頭部運動に近い動きで動かせば、観測者はスクリーン平面を顔面と見立て、その動きから会話者の動作や会話者間の相互作用をある程度、読み取ることができると期待される。

また、人間の視覚は、周辺視野においても動きに対する感度が高いことが示唆されている [8]。この性質により、たとえば、視聴者が正面に位置する人物のスクリーンに着目している場合においても、周囲に位置する人物の動きを感じることができ、複数の会話者がおりなすその場全体の様子を感じ取ることができると期待される。

以上のように本研究は、会話参加者の運動を補強して提示することが、コミュニケーション行動の理解に有用であるとの仮説のもと、スクリーン姿勢の変化として、人物の頭部運動を補強表現する方式を提案するものである。以下、本論文は次のように構成される。まず、2章において関連研究に対する本研究の位置づけを述べる。3章では提案システムの構成を記し、4章では実験結果を示す。最後に5章において結びと今後の課題を述べる。

2. 関連研究

従来のテレビ会議システムやテレプレゼンスシステムにおいて、円滑な会話を妨げる原因の1つとして、「誰が誰を見ているか」、「自分に視線が向けられているか」が分かりにくいという視線不一致の問題が知られている。この問題の緩和には、地点間において人物の空間配置に整合性を持たせることが有効とされる [9], [10]。たとえば、各人物に

個別のディスプレイを割り当て、それを人物の配置に合わせて設置する Hydra システム [9] や、大型ディスプレイで囲まれた空間内で、ある地点の人物の画像を別地点の同一相対位置に表示する t-Room [10] が提案されている。このような空間的整合性によって、顔向きの手がかりが適切に伝達され、視線方向の理解が向上すると考えられている。しかし、会話者の画像が空間に分散されて表示されることにより、全人物を1つの視野内に収めることができず、ときに人物間の相互作用が分かりにくいという側面もある。本論文の提案システムでは、この空間的整合性の考え方を取り入れつつ、スクリーンの物理運動による補強表現により、広く会話の場に含まれる相互作用の知覚を増強する狙いを持つ。

また、視線不一致問題の解消やより高い存在感を狙ったアプローチとして、人物の3次元再形状データをセンサなどにより取得し、それを元に3次元の人物像をCG合成して3次元ディスプレイなどに表示する方式が有望視されている。しかし、立体視のため眼鏡型デバイスの装着が必要なシステム（たとえば文献 [11], [12]）では、ユーザ間のアイコンタクトが妨げられるなどの問題があり、また、裸眼視が可能な3次元ディスプレイを用いるシステム（たとえば文献 [13], [14]）においても、正しく視聴できる視点位置が限定されるなど、ユーザの自然なコミュニケーション行動が制約される。本研究では、これら人の3次元形状の正確な提示を目指すアプローチとは異なり、2次元画像から他者の表情を読み解く人の知覚能力の高さを活用し、平面スクリーンに2次元画像を表示しつつ、人物の頭部運動を物理的なスクリーンの運動として補強的に表現することで高い存在感の醸成や、自然なコミュニケーション行動の伝達を狙うという新しい表現法を探る立場をとる。

さらに近年ではテレプレゼンスロボットと呼ばれる、ユーザの遠隔操作により会話への参加を可能とするロボットが注目を集めている [15], [16]。見る者に高い存在感を与える点が利点とされるが、一方、機械としての存在感が突出し、ユーザの個性が表現されにくいという問題がある。また、究極の個人化として、アンドロイドの利用が検討されているが [17]、いわゆる「不気味の谷」[18]の克服が長期的課題とされる。また、ロボットには、人の非言語行動を表出する手段としての可能性もある。たとえば、頷きの表現や、ロボットアームによるジェスチャ表現 [19]、ユーザとの対人距離が可変なディスプレイによる近接学（プロセミクス）の再現 [20] などが試みられている。本研究では、ロボティクスによる動き表現の可能性に着目し、複数人会話における対人視線方向、および、話しかけの方向性の提示のため、スクリーン姿勢の動的制御による頭部運動の補強表現を提案する点において新規性を有すると考える。また、提案システムでは、機械としての存在感を極力消し、また、合成画像の不気味さ・不自然さを排除し、さらにそ

の人自身の存在感を高めるため人物の2次元画像を（最小限の画像処理のみで）透明スクリーンへ投影するという方式をとる。

また、遠隔人物のアバタとして、平面以外の形状を持つスクリーンを用いるシステムとして、球面スクリーンへ顔画像を投影する SphereAvatar [21]、円筒形スクリーンへ全身像を投影する TeleHuman [22]、顔面形状を持つスクリーンへ顔画像を投影する TalkingHeads [23]、LiveMask [24] がそれぞれ提案されている。SphereAvatar と TeleHuman において、スクリーンは卓上または床面に固定され、人物の頭部運動は、投影画像上の運動として表現される。これらは注視方向の正確な伝達を志向したシステムであるが、人の3次元形状とスクリーンの形状との差異が目立つため、その表現には不自然さがともなう。また、TalkingHeads と LiveMask では、人物の顔形状を正確に反映したスクリーンを用いることにより、投影像とスクリーン形状との乖離が小さく、注視方向の正確な伝達や個人性の表現が可能とされる。しかし、個人特化したスクリーンの製作が煩雑である点や、顔面以外の身体部位が具現化されていないため、見る者に不自然さや不気味さを感じさせる点が課題である。

また、本論文の提案法もスクリーンを物理的に動かすという点において、上記、顔形状スクリーンを用いる方法との類似性を有するが、以下の点において異なる表現方式である。顔型スクリーンを用いる場合、スクリーンの物理的な運動が人の頭部運動を表現する唯一のモダリティであるのに対して、提案法では、スクリーンの動きは頭部運動を表現する補足的なモダリティとして機能する。顔形状スクリーンの場合、その機械的運動と実際の頭部運動に差異がある場合、それが直接的に見る者に伝わり、不自然さや機械らしさを感じさせる要因となる。一方、提案法では、スクリーン上の投影画像として、すでに人物の運動が表現されており、スクリーンの物理運動は補足的に動き知覚のトリガとして働くことが期待される。そのため限定された頭部運動の表現であっても見る者に不自然さを与えにくい。また、通常のディスプレイ同様の平面上の表示により、円筒形や球面状のスクリーンの場合とは異なり、スクリーンの形状を特段意識することなく、画像より人物の外見や動作を自然に読み取ることができる。なお、注視先の正確な空間定位という点では、顔形状スクリーンを用いる方式が優位であると考えられる。しかし、本論文で対象とする対人視線方向の伝達という目的に際しては、提案法でも効果を発揮することが後述の実験により示されている。

さらに上述の従来のスクリーンアバタの研究では、遠隔人物に対応した1体のアバタに対して1人もしくは複数の人が対峙する状況を主たる対象としており、本論文で対象とする複数人物の会話の「場」を再現するという課題設定にはほとんど言及がなされていない。その点、本論文は新しい問題を取り上げるものと位置づけられる。



図 1 提案システム (MM-Space) の外観. (a) 背景除去あり, (b) 背景除去なし, (c) 元になった会話シーン

Fig. 1 Overview of proposed system (MM-Space). (a) background removal, (b) without background removal, (c) original meeting scene.

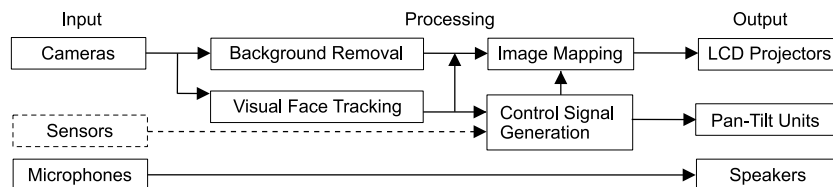


図 2 ブロック図

Fig. 2 Block diagram of system.

3. 提案するシステム：MM-Space

図 1 に提案システム (MM-Space と呼ぶ^{*1}) の外観を示す。また、図 2 には、本システムのブロック図、図 3 には空間配置をそれぞれ示す^{*2}。実際に会話を行う空間 (図 1(c)) では、各会話人物の顔画像や音声を取得する (図 2 の入力部、図 3(a) の配置)^{*3}。また、会話を再現する空間 (図 1(a), (b)) では、会話者の配置にあわせて、複数のプロジェクタ、スクリーン、アクチュエータ、スピーカが配置され (図 3(b))、各人の顔画像が透過型平面スクリーンへ投射され、そのスクリーンの姿勢が人物の頭部運動と同期して制御される。また、各人物の音声は、各々のスクリーンの前方に設置されたスピーカより再生され、観測者は、聴覚的にも話者の位置を特定することが可能である。なお、本論文の提案システムは、頭部運動の物理的補強表示の効果を検証するため、オフラインの再生に特化した初期システムであり、今後、実時間・双方向システムへ

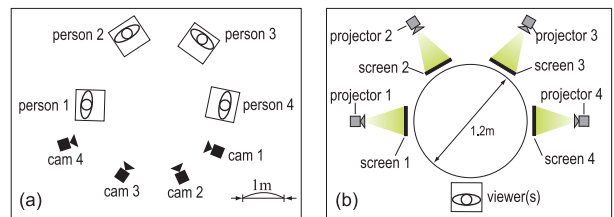


図 3 空間配置, (a) 実際の会話空間, (b) 再現される空間

Fig. 3 Spatial configuration of people and devices, (a) actual meeting space, (b) reconstructed space.

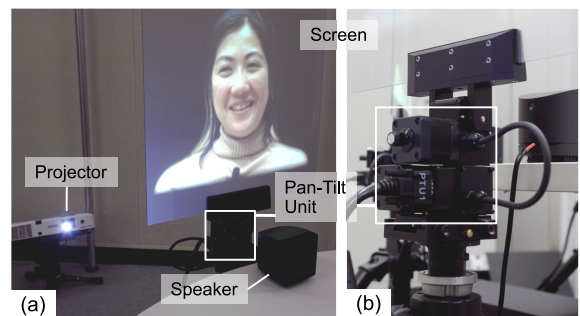


図 4 動的ディスプレイ装置, (a) スクリーン, プロジェクタ, および, スピーカ, (b) パンチルトユニット (側面)

Fig. 4 Dynamic display devices, (a) screen, projector, and speaker, (b) pan-tilt unit from in side view.

の拡張が予定されている。

図 4 に、スクリーン、プロジェクタ、および、アクチュエータの外観を示す。スクリーンは、アクリル板に光線

*1 MM-Space の MM は、Multimodal, Multiparty, Meeting, Motion, Minimalismなどを象徴する。

*2 このような人物配置を採用した1つの理由としては、遠隔システムへの拡張を念頭にし、遠隔の参加者(本論文でいう観測者)を配置する空間を設けるという意図がある。また、画像上の顔方向のみからは対人視線方向が判別しにくい状況を作り、提案方式の効果をより明瞭に浮き上がらせる意図もある。

*3 なお、各人物ごとに個別のカメラ、マイクを用いる代わりに、文献 [25] のような全方位センサも利用可能である。

拡散物質が配合されたものを用いる。また、その背後には、背面投射用の LCD プロジェクタが設置される。各スクリーンは、その下部よりアクチュエータにより支持される。このアクチュエータは、鉛直軸周りの回転運動（パンと呼ぶ）、および、水平軸周りの回転運動（チルトと呼ぶ）の運動を生成する 2 つのユニットから構成される。以後、これを PTU（パンチルトユニット）と呼ぶ。この PTU によりスクリーンの姿勢が動的に制御される。

図 2 の処理部では、顔姿勢追跡、背景除去、制御信号生成、画像写像生成が実行される。顔姿勢追跡部では、撮影された画像から各人物の顔の位置と姿勢が推定される。なお、頭部の姿勢計測のためモーションキャプチャなどのセンサを利用することもできる。背景除去部では、入力画像から背景領域を除去し、人物領域のみの画像が生成される。制御信号生成部では、計測された人物の頭部姿勢に基づいて、スクリーンの姿勢を制御するための制御信号を生成する。画像写像生成部では、スクリーンの姿勢と連動して、歪みのない投影画像を得るため、入力画像（または背景除去後の画像）に写像を施す。以下、各部の概要を説明する。

3.1 顔の位置と姿勢の計測

図 5 (a) のような撮影画像から人物の画像上での顔の位置、および、顔姿勢（頭部姿勢）の計測が行われる。その方法として、STCTracker（疎テンプレートコンデンセーション追跡法）[26] と呼ばれる顔姿勢追跡法を用いる。STCTracker では、顔の 2 次元位置、および、3 軸周りの回転角、スケーリング係数、照明変動係数からなる 7 次元の状態変数が、GPU 実装されたパーティクルフィルタにより逐次推定される。図 5 (b) の例では、顔姿勢の推定結果がメッシュとして重畳表示されている。また、顔姿勢の計測には磁気式のセンサを用いることもできる。

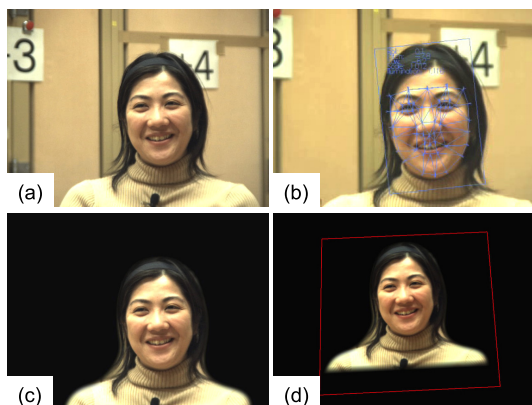


図 5 画像処理の様子、(a) 撮影画像、(b) 顔姿勢追跡、(c) 背景除去、(d) 投影画像（赤枠は説明用）

Fig. 5 Image processing, (a) original images, (b) faces under pose tracking, (c) background-removed images, (d) projection images (red frame is just for illustration).

3.2 背景除去

入力画像に対して背景除去の処理が施される。図 5 (c) に結果例を示す。本システムでは、混合正規分布に基づく背景モデル化 [27]、および、動的輪郭法による前景抽出 [28] を組み合わせた方法が使用される。

3.3 アクチュエータの制御信号の生成

PTU の機械的特性を加味しつつ、できる限り実際に近い頭部運動を再現することを狙い、計測された頭部姿勢の時系列データから、PTU に入力される制御信号を生成する。なお、頭部姿勢は画像としてすでに表現されている点、および、スクリーンの姿勢を、観測者からみてスクリーン像が良好に視認できる範囲に保つという観点より、PTU の制御姿勢は、頭部姿勢角度に一定のスケーリングが施されたものとする。なお、このスケーリングの係数の設定については 4.1 節で後述する。

図 6 には、頭部姿勢角（パン方向）の計測値（黒線）、および、生成された制御信号の姿勢角（緑線）、および、その速度成分（青線）の時系列をそれぞれ例示する。計測信号に含まれる微小ノイズや微小運動を除外しつつ、大きな姿勢変化のみを再現していることが分かる。この処理は、1) ダウンサンプリング、振幅のスケーリング、ゼロ値のシフト、時間差分からなる前処理、2) 速度のピーク区間の検出と選別、3) 最低・最高速度の制約下での波形の整形、からなる。また、チルト運動成分についても、パン運動と同様の手順を踏むが、特に人物の顔に焦点を当て、連続した運動の後、直立姿勢に回帰するよう制御信号が生成される。こうして得られた制御信号が、映像の再生と同期し

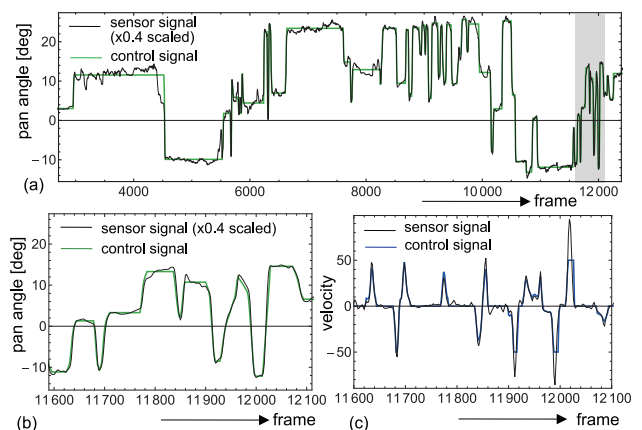


図 6 頭部姿勢時系列（パン角度）の例（ $\times 0.4$ スケーリング）、黒線：計測信号、緑線：PTU 制御信号（角度）、青線：PTU 制御信号の速度成分、(a) 会話セッション全体（ ~ 5.2 [min]）、(b) 部分拡大（ ~ 17 [sec]）、(c) (b) の区間の速度成分

Fig. 6 Example time series of pan angle and its velocity, black line is original signal, green line is generated PTU control signal, and blue line (a) whole meeting duration, (b) close-up over grayed interval in (a), (c) velocity of pan angle during same period as in (b).

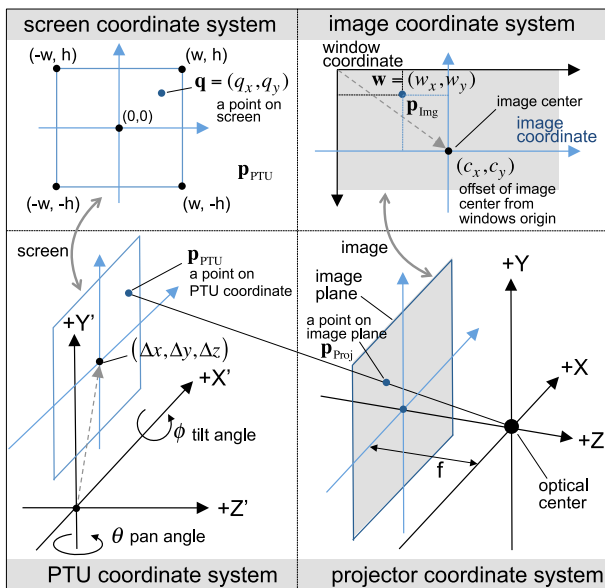


図 7 投影のための座標系
Fig. 7 Projector-screen coordinate systems.

て逐次 PTU に入力される。

3.4 投影画像の生成

スクリーン上に歪みのない画像を投影するため、撮影された画像から人物の顔を中心とする部分画像が切り出され、スクリーンとプロジェクタの相対位置関係、および、スクリーンの姿勢に応じた写像（マッピング）が行われる（図 5(d) は結果の一例）。この変換は、いくつかの座標系間の写像の系列に分解して、記述することができる。図 7 に、これら座標系間の関係を図示する。この写像の系列は、スクリーン上の 1 点、 q 、を始点とし、PTU 座標系上の点、 p_{PTU} 、プロジェクタ座標系上の点、 p_{Proj} 、画像座標系上の点、 p_{Img} 、ウィンドウ座標系上の点、 w 、へと至る； $q \rightarrow p_{PTU} \rightarrow p_{Proj} \rightarrow p_{Img} \rightarrow w$ 。スクリーン座標系と PTU 座標系間の変換は、並進、および、スクリーンの回転（パンとチルト）として記述される。PTU 座標系とプロジェクタ座標系間の変換は、並進と回転、また、プロジェクタ座標系と画像座標系間には透視投影を仮定する。最後に画像座標系からウィンドウ座標系への変換は、LCD プロジェクタのレンズシフトを反映した原点シフトが行われる。

この変換において、機器の設置後には定数となるパラメータの値は、事前のキャリブレーションにより推定される。まず、既知の異なる PTU の姿勢について、スクリーンの 4 つの角に対応するウィンドウ座標の値の実測値が取得される。その後、この実測値と変換式により計算される理論値の間の二乗距離の合計の最小化により、未知パラメータが計算される。

4. 実験

本実験では、動的スクリーンによる頭部運動の物理的補強表現の効果として、観測者にとって会話者の視線方向、および、話し手の話しかける相手（受け手）の理解が促進されるという仮説に着目し、会話者の視線方向、および、受け手の同定の正確さと分かりやすさの観点からスクリーンの動きあり（動的条件と呼ぶ）と、動きなし（静的条件と呼ぶ）の 2 つの表示条件を対比した被験者実験を行った。

4.1 使用機器、および、動作設定

カメラには、Point Grey Research Dragonfly™ を使用し、30 [frame/sec] で VGA 画像を取得した。PTU には、Directed Perception PTU-D46-17 を用いた。スクリーン（G-Screen Through 型、透過率 97%）のサイズは 45 cm × 45 cm、重量は約 570 g であった。プロジェクタには EPSON 1925W（4000 [lm] XGA 解像度）を用いた。なお、本実験では、対人視線方向の表現に焦点を当てるため、パン運動のみが再現され、チルト姿勢は直立に固定した。また、PTU のパン動作のスケーリング係数として、0.4 という値を経験的に設定した。この設定に際しては、会話者の実際の頭部の回転角^{*4}を参考にして、観測者からみてスクリーン像がおおむね良好に視認できる範囲にスクリーンの姿勢を保ち、かつ、PTU の機械的な特性の範囲内^{*5}であり、かつ、視線方向の推移にともなう頭部運動が明確にパン運動として表現できるような値を選択した。また、画像処理による微小な画像欠損などによる観測者の注意の散逸を防ぐため、投影画像は背景除去なしの原画像とし、同様の理由により磁気式センサ（Polhemus FASTRAK™）による頭部姿勢計測値を用いた^{*6,*7}。

4.2 実験 1

動的条件、静的条件における観測者の定性的な振舞いの違い、および、観測者の受けた印象を分析した。

4.2.1 データ

4 人の女性（20 代後半）グループ 2 組による 3 つの会話セッションを用いた。会話参加者は、与えられた議題について 5 分以内にグループとしての結論を出すよう指示された。グループ 1 による会話セッションでは、C1：「恋愛と結婚は一緒か別か」、C2：「専業主婦に税制優遇措置を与え

^{*4} 4 章の実験で用いたデータにおいて、各会話参加者のパン方向の頭部姿勢角の最大値（絶対値）は約 36~71 [deg]（ただし、各人の正面を 0 [deg] とする）であり、角速度の最大値（絶対値）は、約 81~201 [deg/sec] であった。

^{*5} 今回の実験で使用した PTU の仕様上の最大速度は、300 [deg/sec] であった。

^{*6} ただし、顔領域抽出は顔姿勢追跡による。センサは会話者の頭部にヘアバンドで不自然さなく固定された（図 1、図 5 参照）。

^{*7} 事後の検証により、顔追跡結果を用いた場合もセンサの場合と遜色ない動作を確認している。



図 8 実験中の観測者の様子
Fig. 8 A viewer participating in experiments.

るべきか否か], グループ 2 による会話 C3 では「安楽死は合法化すべきか否か」が議題として, 実験者により選択された. 各会話 C1, C2, C3 の長さは, それぞれ 5.1 分, 5.5 分, 5.2 分であった.

4.2.2 実験参加者

会話に参加していない 8 人の女性 (20 代~30 代) が本実験に参加した*8. 本論文では実験参加者のことを観測者と呼ぶ.

4.2.3 手続き

図 8*9のように, 観測者は 1 人ずつ別個に会話セッションをその開始から終了まで連続して視聴した. 各観測者は, 動的条件, 静的条件, それぞれ 1 つ会話セッションを視聴した. 会話, および, 表示条件, その提示順番は, 観測者ごとに混合された.

4.2.4 質問紙

すべての実験 (後述の実験 2 も含む) が終了した後, 観測者が質問紙への回答を行った. ここでは会話の流れ, 表情, 視線, 感情, 雰囲気について各項目の分かりやすさ, および, 身近さ, のめり込み, 楽しさの各項目の印象の強さについて, 動的条件と静的条件を比較し, 順位をつけた. また, 自由記述形式で実験の感想を記入させた.

4.2.5 定性的分析

視聴中に撮影された観測者の正面ビデオ画像に基づいて, 観測者の視線方向 (いつ誰を見ていたか) のアノテーションを手動で作成し, 視線行動の分析を行った. このアノテーションは, 会話, および, 本実験に参加していない 1 人の女性により行われた. 図 9 にはスクリーンの物理的運動が, 観測者の視線行動に影響を与えたと解釈しうる一例を示す. この場面では, 話し手である P2 が, 聞き手の 1 人である P1 の反応を確認しようと P1 の方に視線を向ける場面である. 動的条件下で視聴した 3 人中 2 人が, P2 が視線を P1 に向けた直後に, 視線を P2 から P1 へと変化させ, その結果, P1 が P2 に向かって頷く場面を確認している. 一方, 静的条件下で視聴した 2 人の観測者は, このタイミングで視線を P2 から P1 へと変化させることがなく,

*8 会話者, および, 会話内容への親和性が高いと思われる同性, および, 近い年代を実験参加者とした.

*9 本図の撮影に用いたカメラの画角上, 図 1(a), (b), 図 3(b) の配置と若干異なって見えるが, 実際は同一である.

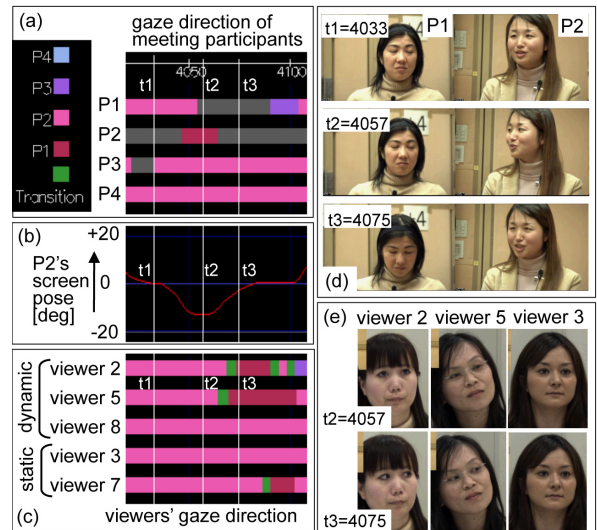


図 9 観測者の振舞い, (a) 会話者の視線方向の時系列, (b) 人物 P1 のスクリーンのパン角度, (c) 観測者の視線方向の時系列, (d) 会話者 P1 と P2 の様子, (e) 観測者の様子

Fig. 9 Behavior analysis of viewers. (a) time series of gaze directions of meeting participants, (b) time series of P1's pan angle for PTU control, (c) time series of viewers' gaze directions, (d) snapshots of meeting participants P2 and P1, (e) snapshot of viewers.

表 1 実験 1 の設問と回答. *は統計的有意差があったことを示す (5%水準)

Table 1 Question items for Experiment 1 and the count of answers from viewers (8 in total). * indicates statistical significance $p < 0.05$.

items	dynamic > static	static > dynamic
conversation flow	6	2
facial expression	5	3
gaze	7*	1
emotion	4	4
mood	5	3
proximity	6	2
involvement	1	7*
enjoyable	7*	1

P1 の頷きを見逃している. この例では, 動的スクリーンによって, 表示された人物の視線変化に誘導されて観測者の視線が変化したとの解釈が成り立つ. 今後, より多くの事例を含めたより精緻な分析が待たれる.

4.2.6 結果と議論

表 1 に, 質問紙への回答をまとめた結果を示す. 今回の実験では実験参加者数は 8 人と少数であったものの, 提案システムの特徴を示唆する興味深い結果が得られた. 質問紙への回答を分析した結果, 特に「視線」と「楽しさ」の項目について, 8 人中 7 人が動的条件をより優位と記入し, 片側二項検定によって統計的有意差 ($p < 0.05$) が確認された. また, 「会話の流れ」と「身近さ」については, 8 人中 6 人が動的条件を優位と回答した. また, 自由記述欄か

らも、8人中5人より、「動くスクリーンによって、視線の方向が分かりやすかった」という旨の回答を得た。これらより限定的ではあるものの、動的なスクリーンにより視線方向が分かりやすいという仮説の妥当性を示唆する結果が得られた。また、4人よりスクリーンの動くノイズ音を指摘された。

4.3 実験2

頭部運動の物理的補強表現の効果を定量的に検証することを狙いとして、話し手、および、その受け手の同定の正確さと明瞭さの観点から評価を行った。

4.3.1 データ

実験1とは異なり、文脈情報を除くため、観測者は、会話全体から切り出された数秒程度の短い映像（クリップと呼ぶ）を視聴した。会話セッション C1, C2, および、C3 の各々について、それぞれ18個のクリップを手動で抽出した。各クリップには、1人の話者が他者（1人または複数）に対して話しかける/問いかける場面を含む。なお、その相手として複数の可能性がある場面のみを対象とした。たとえば、図3(a)において、人物2が人物1に話しかける場合など、人物の顔向きから一意にその相手が特定できる場面は除外した。また、視聴するクリップの時間順序は、文脈手がかりを排除するためランダムとした。クリップの長さは最小1.9秒、最大14.3秒、平均5.4秒であった。

4.3.2 実験参加者

実験1と同じとした。

4.3.3 質問紙

話し手と受け手の同定の正確さと明瞭さを評価するため、表2の設問を用いた。Q1-1とQ2-1では人物の番号を1つ回答させた。Q1-2とQ2-2では、7ポイントのスケール（-3（非常に分かりにくい）から+3（非常に分かりやすい）より1つを選択させた。

4.3.4 正解データ

設問 Q1-1 については、発話内容の書き起こしより話者が特定され、設問 Q2-1 については、話し手の視線方向より受け手が特定された。なお、受け手については、クリップの時間内で話者が複数の人物に視線を向けていた場合、それらの人物を正解とした。

4.3.5 手続き

図8のように、8人の観測者は各々別個に実験に参加し

表2 実験2の設問

Table 2 Questionnaire for Experiment 2.

	Questions
Q1-1	話し手は誰だったと思いますか？
Q1-2	話し手は誰だったかよく分かりましたか？
Q2-1	話し手が話しかけていた相手は誰だったと思いますか？
Q2-2	その相手は誰だったかよく分かりましたか？

た。各観測者は、2つの条件（動的条件、静的条件）について、各1つの会話セッションから抽出された18個のクリップを視聴した。各クリップの再生の時間順序はランダムとした。各クリップの視聴後、20秒間で質問紙への記入を行った。なお、視聴条件と会話セッションの組合せは、観測者ごとに混合された。

4.3.6 結果と議論

Q1-1とQ2-1については各条件、各クリップごとに平均の正答率（正解した観測者数/そのクリップを視聴した観測者数）が計算された。なお、受け手の同定（Q2-1）については、質問紙に回答された受け手が、正解の中に含まれていた場合には正解と判定した。明瞭性の設問（Q1-2とQ2-2）については、各条件、各クリップごとに平均の評定値が計算された。図10には、動的条件、静的条件のそれぞれについて平均正答率、および、平均の明瞭性評定値のそれぞれ全体平均を記す。図10(a)より、話し手の同定（Q1-1）については両条件とも高い正答率が得られたことが分かる。図10(b)より、話し手同定の明瞭性（Q1-2）についても、両条件で高い評価が得られた（評定値の+2は、「分かりやすい」に対応）。正答率、明瞭性ともに、両条件間に統計的有意差はなかった（Mann-WhitneyのU検定の5%水準^{*10}）。なお、話し手の同定は、次の受け手の設問の前提となるものであり、両条件間で差があることはシステムの設計上、想定していない。また、各人物の音声がスク

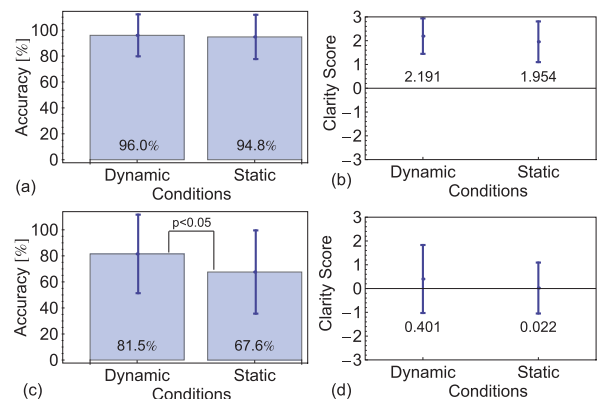


図10 実験2の評価結果。(a) Q1-1:話し手の同定の平均正答率、(b) Q1-2:話し手同定の明瞭性の平均評定値、(c) Q2-1:受け手の同定の平均正答率、(d) Q2-2:受け手同定の明瞭性の平均評定値。なお、Q1-1、Q1-2、Q2-1については、データが正規性の条件を満たさなかったためエラーバーは参考までに付記した

Fig. 10 Evaluation results. (a) Q1-1: accuracy of speaker identification, (b) Q1-2: clarity of speaker identification, (c) Q2-1: accuracy of addressee identification, (d) Q2-2: clarity of addressee identification. Note error bars in the graphs of Q1-1, Q1-2, and Q2-1 are just for reference, because data did not satisfy normality test.

*10 本論文では、データが正規性の条件を満たす場合にはT検定、そうでない場合にはMann-WhitneyのU検定を用いた。

表 3 受け手同定の設問 (Q2-1) における視線方向の過大評価と過小評価の観点からの誤答の傾向

Table 3 Tendency of wrong answer on addressee identification question (Q2-1), in terms of overestimate/underestimate of gaze direction.

conditions	overestimate	underestimate
dynamic condition	42.9%	57.1%
static condition	37.9%	62.1%

リーン前方の個別スピーカから再生されたため、音の手がかりも有効に作用したと考えられる。

次に、図 10(c) には、設問 Q2-1 の受け手の同定の正答率を示す。動的条件では 81.5%、静的条件では 67.6%と、動的条件がより高い正答率を示し、Mann-Whitney の U 検定の結果、統計的有意差が認められた ($p = 0.019 < 0.05$)。図 10(d) には、設問 Q2-2 の受け手の同定の明瞭性に関する結果を示す。動的条件の評定値がわずかに静的条件を上回るが、T 検定の結果、統計的有意差は示唆されなかった ($p = 0.12 > 0.05$)。以上の結果より、スクリーンの物理的な運動による補強表現が、受け手の同定の正確さに寄与していることが示唆されたといえる。

4.3.7 受け手の同定における回答の傾向

受け手の同定の設問 (Q2-1) における正答、および、誤答の傾向をより詳細に分析する。

まず、提案法により過度に視線方向が強調され、受け手の同定を誤ることがないかという観点より、表 3 のように、観測者が受け手を誤って回答したケースについて、視線の方向を過大評価した割合、および、視線の方向を過小評価した割合を算出した。この集計では 1 人の話者が 1 人の他者に対して話しかける/問いかける場面のクリップを対象とした*11。ここで視線方向を過大評価するとは、たとえば、図 3(a) の配置において、人物 1 (以後、人物 i のことを P_i と記す) が P_3 に話しかけている場合に、観測者が P_2 を受け手として回答した場合を指し、また、過小評価するとはこの場合、 P_4 を受け手として回答することを指す。表 3 より、動的条件、静的条件ともに視線の過大評価、および、過小評価の両方のケースがあるが、これらの割合には極端な差はなく、動的条件の方が若干、両者の差が小さい。これにより提案法によって視線方向がつねに過度に強調され、受け手を誤るといった画一的な傾向はないことが読み取れる。

さらに表 4 には、話し手と受け手の位置関係に応じた観測者の回答の傾向を集計した結果を示す。表 4 より、両端に位置する話者 (P_1 と P_4) が 1 つ隣、または、3 つ隣の人物 (話者からみて左右の端にいる他者) に話しかける場合、動的条件では 8 割超の正答率を示した。それに対して、静

*11 全クリップの約 6 割に相当する。また、前述のとおり、頭部の左右の向きより、受け手が一意に定まる状況は、最初からクリップには含まれていない。

的条件の正答率は 25%~44%と低く、話者からみて 2 つ隣の中央の人物を受け手として偏って判断した傾向が読み取れる。特に話者の隣にいる受け手に話しかける場合、静的条件の正答率は 25%と低い。このことから視線を左右に大きく振り向ける場合、静的条件では、画像のみから視線方向を正しく読み取ることが難しいのに対して、動的条件では大幅に正答率が向上していることから、この点において提案法の効果が明確に確認できる。

また、両端に位置する話者 (P_1 と P_4) が、2 つ隣の受け手 (話者からみて中央の人物) に話しかける場合には、動的条件と静的条件は類似した傾向を示し、本来の受け手の左右に位置する人物を、受け手として誤答する傾向がみられた。

さらに話者が中央に位置する場合 (P_2 と P_3)、話者が両端にいるケースと比較して、受け手の正答率は、動的条件で 9 割超、静的条件で 7~8 割程度とより高い値を示した。これは受け手の候補が少ないほど、チャンスレベルが向上することに加えて、各々の受け手に対する話し手の頭の向きに明確な差ができ、観測者はより正確に受け手を同定できるようになるという一般的な推測と合致する。

以上のように受け手同定のタスクにおける提案手法の効果、ならびに、静的条件との差異が明らかになった。なお、提案法による視線方向の伝達効果は、人数や人物配置、パン動作のスケーリング係数などと密接に関連すると考えられるが、上記の実験結果は、3 章や 4.1 節で述べたこれらの条件下における結果であることに留意されたい。

5. 結びと今後の課題

本論文では、複数人物の会話場面を表示するための新しいシステムを提案した。本システムは、人物像を投影するスクリーンの物理的な運動として、人物の頭部運動を補強的に提示するという表現法の提案と実装を特徴とする。この物理的なスクリーンの動きと画像上の動きとの相乗効果により、会話者の視線方向、および、話し手の話しかける相手 (受け手) の理解が促進されるという仮説を立て、観測実験により、視線方向の分かりやすさ、および、受け手の同定の正確さの点において、仮説の妥当性を示唆した。

本論文の実験で確認した提案法の効果は、PTU のパン動作のスケーリング係数 0.4 の条件下におけるものである。したがって、異なるスケーリング係数を用いた場合、今回の実験とは異なる結果が得られる可能性もある。よって様々なスケーリング係数を用いて提案法の効果を比較することや、最適な値を決定する方法を検討することが今後の課題とされる。

また、本論文の実験では、対人視線方向の表現を検証するためパン運動に絞った評価を行ったが、今後、チルト運動の効果も検証する必要がある。チルト運動によって、会話人物の顔がより強調されて表現され、観測者の注意を

表 4 受け手同定の設問 (Q2-1) における話者と受け手の位置関係に応じた回答の割合 (下線が正答を示す)

Table 4 Rate of answer on addressee identification question (Q2-1), depending on the relative position of speaker and addressee; underline indicates right answer.

Speaker	Side speaker (P1 and P4)								
	1st next			2nd next			3rd next		
Addressee position									
Answered addressee	<u>1st next</u>	2nd next	3rd next	1st next	<u>2nd next</u>	3rd next	1st next	2nd next	<u>3rd next</u>
Dynamic condition	83.3%	16.7%	0.0 %	19.2%	<u>61.5%</u>	19.2%	0.0%	14.3%	<u>85.7%</u>
Static condition	<u>25.0%</u>	75.0%	0.0 %	14.8%	<u>63.0%</u>	22.2%	11.1%	44.4%	<u>44.4%</u>

Speaker	Middle speaker (P2 and P3)			
	1st next		2nd next	
Addressee position				
Answered addressee	<u>1st next</u>	2nd next	1st next	<u>2nd next</u>
Dynamic condition	<u>93.8%</u>	6.3%	0.0%	<u>100.0%</u>
Static condition	<u>70.0%</u>	30.0%	15.4%	<u>84.6%</u>

引きつける効果のほか、顔きの機能である態度表出や発話強調をとともなう社会的相互作用の理解も促進されることが期待できる。また、提案法では、枠なしの透明なスクリーンに背景なしの人物像を投影することにより、さらなる存在感の増強、たとえば、従来法のディスプレイの中に人が映っているという印象から、目の前に人が存在しているという印象への変化を期待しているが、その効果についても今後、評価が必要である。さらに、会話の内容や会話人物の感情・性格などより高次の情報の表現・伝達についても今後、長期的に検証していく必要がある。加えて、他の再生環境との比較も課題である。

また、現状の提案システムでは、人物の肩より上部のみをスクリーン上に表示している。そのため、手や身体の動作やジェスチャなどが十分に表現されないという問題もある。今後、スクリーンのサイズや投影倍率などの比較検討が要される。さらに、会話の場の表現には、音場の正確な再現も重要な要因である。本論文の提案システムでは、固定されたスピーカを用いたが、今後、頭部方向と連動して、音源の位置や指向性を変化させるような音声再生方式も検討する価値がある。

これまで一般的に遠隔人物の存在感を高める方策として人物の3次元形状の正確な再現が本命の課題とされ、様々な3次元表示方式の研究が行われてきた。それに対して本研究は、2次元表示に若干の3次元物理運動を加えるだけでも存在感が増強されるという「動きがリアリティを作る」という立場から一種の代替方式を提案するものであり、現状、入手可能なデバイスやソフトウェアの組合せによって、複数人物の会話の場を再現する現実解を提供するものと位置づけられる。この方式には、動きに対する知覚の鋭敏性という人の知覚特性が関わっており、スクリーン上の画像情報と周辺視野も含めて知覚されるスクリーンの物理的運動とが結びつき、会話の場全体が高い存在感をもって知覚されうるとの仮説に基づいている。今後、画像上の動きと

物理的な動きの作用の違いなど、人の知覚メカニズムの側面から提案システムをより綿密に検証することも興味深い課題である。

さらに本システムは、今後、多地点間の複数人物による実時間コミュニケーションシステムへの発展が期待される。その場合、視線不一致問題と関連して、本論文提案の物理的補強表現が、観測者からみて「話しかけられる」ことの知覚に寄与しうるものかの検証が必要である。また、カメラの最適配置や、物理系を駆動する際の遅延の問題もあわせて検討が求められる。

参考文献

- [1] Otsuka, K.: Conversation Scene Analysis, *IEEE Signal Processing Magazine*, Vol.28, No.4, pp.127–131 (2011).
- [2] Argyle, M.: *Bodily Communication - 2nd ed.*, Routledge, London and New York (1988).
- [3] Kendon, A.: Some Functions of Gaze-Direction in Social Interaction, *Acta Psychologica*, Vol.26, pp.22–63 (1967).
- [4] Maynard, S.K.: Interactional Functions of a Nonverbal Sign: Head Movement in Japanese Dyadic Casual Conversation, *J. Pragmatics*, Vol.11, pp.589–606 (1987).
- [5] Johansson, G.: Visual Perception of Biological Motion and a Model for its Analysis, *PERCEPTION & PSYCHOPHYSICS*, Vol.14, No.2, pp.201–211 (1973).
- [6] Neri, P., Luu, J. and Levi, D.: Meaningful Interactions Can Enhance Visual Discrimination of Human Agents, *Nature Neuroscience*, Vol.9, pp.1186–1192 (2006).
- [7] Heider, F. and Simmel, M.: An Experimental Study of Apparent Behavior, *American Journal of Psychology*, Vol.57, No.2, pp.243–259 (1944).
- [8] Shapiro, K.L. and Johnson, T.L.: Effects of Arousal on Attention to Central and Peripheral Visual Stimuli, *Acta Psychologica*, Vol.66, No.2, pp.157–172 (1987).
- [9] Sellen, A.J.: Speech Patterns in Video-Mediated Conversations, *Proc. CHI*, pp.49–59 (1992).
- [10] Hirata, K., Kaji, K., Harada, Y., Yamashita, N. and Aoyagi, S.: t-Room: Remote Collaboration Apparatus Enhancing Spatio-Temporal Experiences, *Proc. CSCW'08* (2008).
- [11] Chen, W.-C., Towles, H., Nyland, L., Welch, G. and

- Fuchs, H.: Toward a Compelling Sensation of Telepresence: Demonstrating a Portal to a Distant (Static) Office, *Proc. Visualization '00*, pp.327-333 (2000).
- [12] Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A.V. and Staadt, O.: blue-c: A Spatially Immersive Display and 3D Video Portal for Telepresence, *ACM Trans. Graph.*, Vol.22, No.3, pp.819-827 (2003).
- [13] Jones, A., Lang, M., Fyffe, G., Yu, X., Busch, J., McDowall, I., Bolas, M. and Debevec, P.: Achieving Eye Contact in a One-to-many 3D Video Teleconferencing System, *ACM Trans. Graph.*, Vol.28, pp.64:1-64:8 (2009).
- [14] Matusik, W. and Pfister, H.: 3D TV: A Scalable System for Real-time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes, *ACM Trans. Graph.*, Vol.23, No.3, pp.814-824 (2004).
- [15] Guizzo, E.: When My Avatar Went to Work, *IEEE Spectrum*, Vol.47, pp.26-50 (2010).
- [16] Tsui, K.M., Desai, M., Yanco, H.A. and Uhlik, C.: Exploring Use Cases for Telepresence Robots, *Proc. ACM/IEEE HRI '11*, pp.11-18 (2011).
- [17] Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N.: Android as a Telecommunication Medium with a Human-like Presence, *Proc. ACM/IEEE HRI '07*, pp.193-200 (2007).
- [18] Mori, M.: The Uncanny Valley, *Energy*, Vol.7, No.4, pp.33-35 (1970).
- [19] Adalgeirsson, S.O. and Breazeal, C.: MeBot: A Robotic Platform for Socially Embodied Presence, *Proc. ACM/IEEE HRI '10*, pp.15-22 (2010).
- [20] Nakanishi, H., Kato, K. and Ishiguro, H.: Zoom Cameras and Movable Displays Enhance Social Telepresence, *Proc. CHI '11*, pp.63-72 (2011).
- [21] Oyekoya, O., Steptoe, W. and Steed, A.: SphereAvatar: A Situated Display to Represent a Remote Collaborator, *Proc. CHI '12*, pp.2551-2560 (2012).
- [22] Kim, K., Bolton, J., Girouard, A., Cooperstock, J. and Vertegaal, R.: TeleHuman: Effects of 3D Perspective on Gaze and Pose Estimation with a Life-size Cylindrical Telepresence Pod, *Proc. CHI '12*, pp.2531-2540 (2012).
- [23] Naimark, M.: Talking Head Projection, available from <http://www.naimark.net/projects/head.html>.
- [24] 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).
- [25] Otsuka, K., Araki, S., Ishizuka, K., Fujimoto, M., Heinrich, M. and Yamato, J.: A Realtime Multimodal System for Analyzing Group Meetings by Combining Face Pose Tracking and Speaker Diarization, *Proc. ACM ICMI '08*, pp.257-264 (2008).
- [26] Lozano, O.M. and Otsuka, K.: Real-time Visual Tracker by Stream Processing, *Journal of Signal Processing Systems*, Vol.57, No.2, pp.285-295 (2008).
- [27] Pham, V., Vo, P., Hung, V.T. and Bac, L.H.: GPU implementation of Extended Gaussian mixture model for Background Subtraction, *Proc. IEEE RIVF*, pp.1-4 (2010).
- [28] Bojsen-Hansen, M.: Active Contours without Edges on the GPU, Technical Report 13, Project Paper for the Course in Parallel Computing for Medical Imaging and Simulation (2010).

推薦文

インタラクシオン 2012 では、87 人から構成されるプログラム委員会によって投稿数 43 件の中から優秀な論文 18 件を一般講演発表として採択し、インタラクティブ発表は 149 件の投稿から 19 件をファイナリストとして選出いたしました。本論文は、これらの 37 件からさらにプログラム委員会による投票によって、論文誌に推薦すべき論文であるとの評価を得たものであり、論文誌編集委員長としてもぜひ推薦したいと考えました。

(インタラクシオン 2012 プログラム委員長 宮下芳明)



大塚 和弘 (正会員)

1993 年横浜国立大学工学部電子情報工学科卒業。1995 年同大学大学院工学研究科博士課程前期修了。2007 年名古屋大学大学院情報科学研究科博士後期課程修了。1995 年日本電信電話(株)入社。以後、NTT ヒューマンインタフェース研究所、NTT サイバースペース研究所を経て、2001 年より NTT コミュニケーション科学基礎研究所。2010 年スイス Idiap Research Institute 特別招聘研究員。現在、NTT コミュニケーション科学基礎研究所メディア情報研究部主幹研究員(特別研究員)。コンピュータビジョン、コミュニケーション科学に興味を持つ。IAPR 10th Int. Conf. Image Analysis and Processing, Best Paper Award, ACM 9th and 14th Int. Conf. Multimodal Interfaces (Interaction) (ICMI2007, ICMI2012) Outstanding Paper Award, 電子情報通信学会 MVE 賞 (2008 年), MIRU2009 優秀論文賞, 電子情報通信学会 H22 年度論文賞, および, 同喜安善市賞, 経産省 Innovative Technologies, および, 同特別賞 (2012 年), ほか受賞, 電子情報通信学会, IEEE 各会員。博士(情報科学)。



熊野 史朗 (正会員)

2009 年東京大学大学院情報理工学系研究科博士課程修了。現在、NTT コミュニケーション科学基礎研究所研究員。表情認識や感情推定等コンピュータビジョンや感性情報処理に関する研究に従事。8th Asian Conf. Computer Vision (ACCV2007) Honorable Mention, MIRU2011 インタラクティブセッション賞ほか受賞。電子情報通信学会, IEEE 各会員。博士(情報理工学)。



松田 昌史

2003年北海道大学大学院文学研究科人間システム科学専攻博士後期課程単位取得退学。同年日本電信電話(株)入社。現在、NTTコミュニケーション科学基礎研究所勤務。対人コミュニケーション、対人信頼関係の形成過程、行動科学、実験ゲーム研究等の研究に興味を持つ。博士(文学)。日本心理学会優秀論文賞(2002年)、電子情報通信学会ヒューマンコミュニケーション賞(2007年、2010年)各受賞。日本心理学会、日本社会心理学会、日本グループダイナミクス学会各会員。



大和 淳司

1988年東京大学工学部精密機械工学科卒業、1990年同大学大学院工学系研究科精密機械工学専攻修士課程修了。1998年マサチューセッツ工科大学電気工学およびコンピュータサイエンス学科修了。1990年日本電信電話(株)入社。画像認識等の研究に従事。NTTヒューマンインターフェース研究所、MIT人工知能研究所、NTT第三部門を経て、現在、NTTコミュニケーション科学基礎研究所メディア情報研究部部長、IEEE、ACM各会員。博士(工学)。