

MoPaCo: 単眼カメラを用いた運動視差による 3次元映像コミュニケーションシステム

石井 亮[†]

小澤 史朗[†]

向内 隆文[†]

松浦 宣彦[†]

本研究では、単眼カメラと2次元ディスプレイという一般的な映像コミュニケーションにおける装置構成で、ユーザの視点位置に合わせて運動視差を付与し、奥行きを知覚可能な立体映像を視聴できる映像コミュニケーションシステム MoPaCo を提案する。システムを実装し、主観評価実験を行った結果、本運動視差によって映像に違和感なく立体知覚が可能であり、対話相手が実際に在りし対面している感覚を増強すること、さらに対話相手の空間の位置関係の把握が容易となることが確認された。

MoPaCo: 3D Video Communication System using Motion Parallax with Monocular Camera

RYO ISHII[†]

SHIRO OZAWA[†]

TAKAFUMI MUKOUCHI[†]

NORHIKO MATSUURA[†]

In this study, we propose a pseudo 3D video communication system that imparts motion parallax which adjusts to the viewpoint position of the user and enables the user to view video pictures in which depth can be perceived, with an ordinary equipment setup, namely a monocular camera and a 2D display. We have implemented a system and the results of subjective evaluation experiments showed that motion parallax gives the user stereoscopy without a uncomfortable feeling and a feeling as if the conversational partner is actually present. In addition, motion parallax makes it easier to comprehend the positional relationship of the conversational partner in space.

1. はじめに

我々は一般的な会話はもちろんのこと、お互いの空間や物を共有しながら行う協調作業を、実際に対面したときと同等に自然に行える臨場感の高い映像コミュニケーションシステムの実現を目指している。そのためには、対話相手があたかも目の前に実際にいるように感じることや、お互いの空間の繋がりを表現し、対話相手との距離感（対人距離）、視線や指差しといった非言語情報を自然に伝達可能にすることが重要であると考えている。一般的な、単眼カメラで撮影された映像を2次元ディスプレイに表示する映像コミュニケーションシステムでは、対話相手の空間の奥行き情報が大きく欠落する。また、ユーザが移動してもその視点に合わせて、対話相手や背景の見え方は変化しない。そのため、お互いの空間の繋がりは表現されず地点間に空間的な隔たりを感じる[1,2]、対話相手の存在感が感じられない[3]、さらには奥行きを伴う非言語情報（対人距離、指差しジェスチャなど）が正しく伝達できない等の大きな問題点がある[4-6]。これらの情報を伝達するためには、互いの空間の繋がりを表現すると共に、映像中に奥行情報を与えることが重要となる。

人間が奥行きを感じる手がかりとして、近距離では運動視差と両眼視差が大きく寄与していることが知られている[7]。このうち、両眼視差は比較的容易に利用が可能なため、両眼視差を利用して対話相手を立体表示することで奥行きを表現する手法が多く提案されている[8]。しかしながら、両眼視差を与える従来の方式ではピント位置の生理的な矛盾による目の疲労や、メガネの装着による視線伝達の妨げといった課題がある。これに対して、運動視差を利用して奥行きのある映像を表現する手法として、2次元ディスプレイに、ユーザの視点位置に応じて映像を投影する手法がある。この手法において、ユーザの視点位置の検出を画像ベースで行えば、ユーザに特殊な機器を装着させる必要がないため、ユーザ負荷の観点からみても映像コミュニケーションでの利用は有効的であると考えられる。この運動視差を、実写のオンライン映像に付与する試みでは[9]、高精度にユーザ視点位置を検出するためにステレオ視を行い、さらに視点位置に合わせてカメラを物理的に制御していた。そのため、大がかりなシステム構成となっており、映像コミュニケーションに利用できなかった。

そこで本研究では、運動視差を伴うリアルタイム映像コミュニケーションシステム MoPaCo (Motion Parallax Communication video system)を提案する。本システム

[†] 日本電信電話株式会社 NTT サイバースペース研究所
NTT Cyber Space Laboratories, NTT Corporation

は、既存の映像コミュニケーションシステムと同様な、単眼カメラおよび2次元ディスプレイでの構成で実現される。さらに、MoPaCoを用いて評価実験を行った結果、本運動視差映像が、違和感なく立体知覚が可能であり、対話相手が実際に実在し対面している感覚を増強すること、さらにお互いの空間が窓越しに繋がっている感覚を与え、対話相手の空間の位置関係の把握が容易となることが確認されたので報告する。

2. 運動視差映像コミュニケーションシステム MoPaCo の提案

運動視差映像コミュニケーションシステム MoPaCo における、映像表現のイメージを図1に示す。ユーザ自身のディスプレイと遠隔側の対話相手のディスプレイがあたかも一つの窓として繋がったかのように、ユーザの視点位置に応じた運動視差映像をディスプレイに表示する。ここで、ユーザに対話相手の運動視差映像を2次元ディスプレイに提示するためには、

- (1) ユーザおよび対話相手の視点位置の測定処理
 - (2) 対話相手の空間の3次元情報（テクスチャ、実寸の大きさ、位置情報を持つ）の構成処理
 - (3) ユーザの視点位置（(1)で取得）に応じた、3次元空間情報（(2)で構成）の描画処理
- を行う必要がある。MoPaCoでは上記(1)、(2)の処理を単眼カメラのみを用いて実現する。次に、(1)～(3)における具体的な処理について説明する。

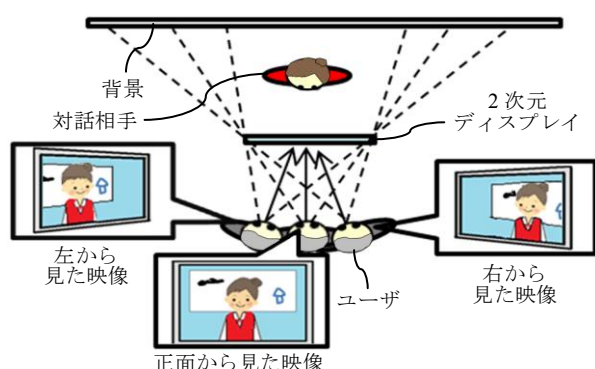


図1 MoPaCoによって実現される映像表現のイメージ

2.1 ユーザの視点位置の測定手法

本研究では、単眼カメラのキャプチャ画像中の顔のパーツの位置情報（2次元座標）から、ユーザの視点位置（両眼の中心）の実空間上の3次元位置を取得する。より厳密な3次元位置を取得するため、事前にユーザの眼間距離を計測し、それをリファレンスとする。まず前処理として、カメラのオートフォーカス（AF）機能実行時に算出される被写体（ユーザ）ま

での距離をカメラから直接取得する。これと同時に画像をキャプチャし、顔認識処理を行うことでユーザの両眼の位置（画像中の2次元座標）および頭部の姿勢を取得する。これらの情報に加えて、カメラの画素ピッチから、ユーザの眼間の実寸距離を算出する。

以上の前処理を終えた後、リアルタイムキャプチャを開始すると、先と同様にキャプチャ画像からユーザの両眼の位置および頭部の姿勢を顔認識処理により取得する。この取得情報と事前に算出した眼間の実寸距離情報を基に、ユーザの視点の奥行き位置 z を算出する。なお、縦、横(x, y)の座標は、画像中の2次元座標と画像の画素ピッチから算出する。

2.2 3次元空間情報構成部

実写ベースの3次元空間を再現する技術として、IBRやMBRなど様々な手法が提案されているが、複数カメラが必要かつ処理時間が膨大といった問題があるため、映像コミュニケーションへの適用は困難である。本研究では、1対1の映像コミュニケーションにおける2つの制約条件に着目した。

- 人が会話中に大きくディスプレイの前から動かないため正面からの映像で十分である。
- カメラは事前に設置されるため、背景画像の取得が可能である。

そこで、単眼カメラのキャプチャ画像に対して、事前に取得された背景画像を用いて背景差分処理を行うことで、人物画像と背景画像を分離し、それらをレイヤとして奥行き位置に合わせて多層化することで3次元情報を構成する手法を提案する。

人物画像および背景画像の生成手法の具体的な処理フローを図2に示す。背景差分処理により生成されたマスク画像（図2(b)参照）に膨張収縮処理、ラベリング処理を行い、孤立した小領域などの細かいノイズ除去した後人物領域を抽出する（図2(c)参照）。この

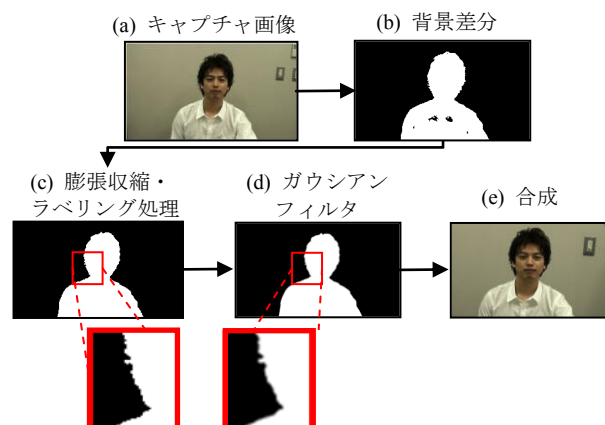


図2 人物画像と背景画像の生成処理フロー

マスク領域を用いて人物画像と背景画像を生成する。また、描画時に背景画像と人物画像の境界面を滑らかにブレンディングできるように、人物のマスク領域に対してガウシアンフィルタ処理を行った結果を人物画像の α 値とする（図2(d),(e)参照）。背景画像は背景領域を過去の背景画像に逐次、更新することで生成する。次に、人物画像と背景画像を用いて、実寸サイズとカメラからの距離関係を持つ多層レイヤを生成する。背景レイヤの距離情報には、背景差分画像取得前に、カメラのAF機能実行時に算出した値を利用する。人物レイヤについては、対話相手の位置情報（2.1節にて取得方法を説明）を用いる。これらの距離情報 d_i 、カメラ画角（横 θ_w 、縦 θ_h ）から、各レイヤ i の実寸サイズ（横 $w_i \times$ 縦 h_i ）は、式(1)より算出する。以上より、テクスチャ、実寸のサイズおよび位置情報を持つ3次元空間が構成される。

$$w_i = 2 * d_i * \tan(\theta_w / 2), h_i = 2 * d_i * \tan(\theta_h / 2) \quad \dots \dots \dots (1)$$

2.3 運動視差映像の描画処理部

描画処理では、2次元ディスプレイをディスプレイ面として、3次元空間情報構成部で生成した人物レイヤと背景レイヤを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（図3参照）。このとき、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。

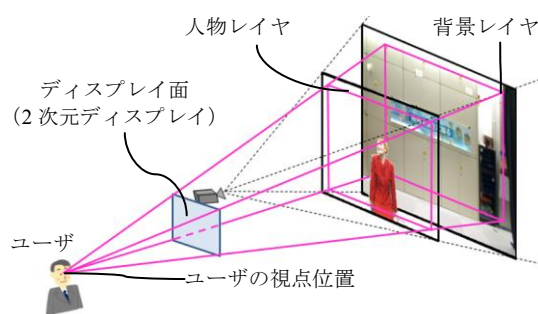


図3 透視投影変換による映像の描画

2.4 実装結果

以上の提案手法を用いて MoPaCo システムを実装した。高速化処理のため、3次元空間情報生成、運動視差映像の描画処理は GPU 処理を行った。実装環境は、CPU が Intel Core i7 Extreme 980X、メモリ 12GB、グラフィックボード NVIDIA GeForce GTX480 であり、動作結果を表1に示す。表1中の“ユーザ動作からの遅延”は、ユーザの視点位置が実際に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”

延”はキャプチャ画像が提示されるまでの時間である。

表1 実装結果

入力画像サイズ	1920×1080 (HD)
描画速度	30fps
ユーザ動作からの遅延	200ms 程度
カメラ画像の遅延	200ms 程度

3. 評価実験

3.1 実験概要

提案する運動視差映像の効果を検証するために、MoPaCo による運動視差映像（MoPaCo 条件）と、従来の 2D 映像（2D 条件）の 2 つの条件下において、対話相手の映像を被験者に観察させ、印象評価を実施した。評価方法として、映像や対話相手に関する印象を、質問紙を用いたリッカート法による 6 段階（1～6 点）評価で問いた。評価項目は、表2に示す7項目である。なお、視覚刺激を統一するために対話相手を実際の人ではなく、実物大の人型マネキンに代替した。

実験環境を図4に示す。2D ディスプレイを設置したパーテーションから被験者は 150cm の位置に座席した。また、対話相手の空間はディスプレイの奥 150cm に対話相手（マネキン）、340cm に壁を配置した。なお 2D 条件では、ユーザが椅子に座った位置にユーザ視点があるときの映像を提示した。

次に、実験手順について説明する。被験者は、まず実験の前に実物のマネキンの大きさを確認し、その後、座席に着席してから顔を下げて待機した。映像が提示された後、開始の合図と同時に顔を上げ、椅子に座っ

表2 実験における主観評価項目

- **映像の見易さ**：映像は見やすいか（違和感はないか）？
- **立体感**：相手と相手の空間を立体的に感じたか？
- **距離感のつかみ易さ**：相手との距離感がつかみ易い？
- **対話相手の実在感**：相手が実在しているように感じたか？
- **対面感**：実際に相手と会っているように感じるか？
- **空間把握の容易さ**：相手の空間の位置関係が分かるか？
- **窓越し感**：窓越しに相手と会っているように感じたか？

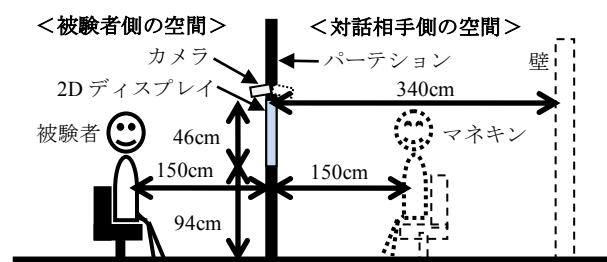


図4 実験における実験環境

たままの状態自由に頭部を動かシマネキンを観察した。時間制限は無制限であったが、どの被験者もおおよそ 10 秒以内で観察を止め評価を行った。試行順序は、順序効果を考慮して 2 条件をランダムに実施した。被験者ごとの試行回数は各条件 1 施行である。

3.2 実験結果・考察

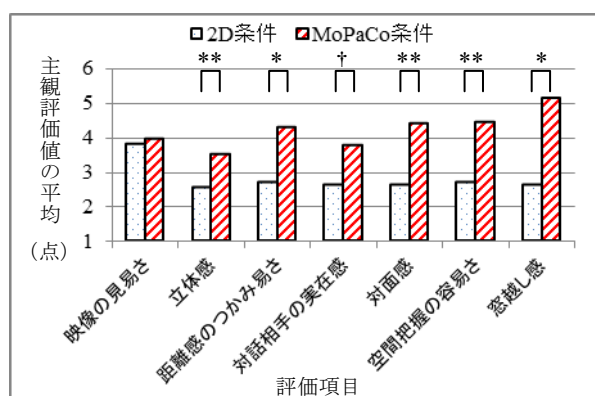


図5 主観評価結果

被験者 10 人 (20～50 代の男性 9 人, 女性 1 人) の主観評価結果の平均値を図 5 に示す。図 5 を見ると、映像の見易さの項目では、2D 条件と MoPaCo 条件で差は見られなかった。すなわち、本運動視差映像は、酔いや映像処理によるノイズなどによる映像の見易さへの悪影響を生じないことが示された。また、映像の見易さ以外のすべての評価項目で MoPaCo 条件が 2D 条件よりも評価値が高く、対応のある t 検定の結果、有意差が見られた (立体感: $t(9)=2.03$, $p<.01$, 距離感のつかみ易さ: $t(9)=4.00$, $p<.05$, 実在感: $t(9)=2.68$, $.10<p<.05$, 対面感: $t(9)=5.51$, $p<.01$, 空間把握の容易さ: $t(9)=3.21$, $p<.01$, 窓越し感: $t(9)=7.39$, $p<.05$)。これにより、本運動視差映像は 2D 映像に比べて立体的な知覚を可能にし、対話相手との距離感のつかみ易さを可能にすることが示唆された。また、対面感および対話相手の実在感の項目において、運動視差映像が 2D 映像に比べて向上が見られことは興味深い。これは、より実際の対面状況に近い視覚効果である運動視差を与えることで実際に対面相手と会っている感覚が増強される可能性があると考えられる。空間把握の容易さの項目について MoPaCo 条件で向上したことから、運動視差映像は対話相手の空間を観察する必要のある場面に有効であると考えられる。最後に、窓越し感について、MoPaCo 条件では評価値の平均が 5 点と高得点であり、実際に窓を覗いているのに近い感覚をユーザに与えられたと考えられる。これにより、ユーザがお互いの空間の接合面を自然に感じることや、

対話相手が自分の空間のどの範囲を観察しているかを直感的に理解できる可能性が考えられる。

4. まとめ

本稿では、対話相手の映像に運動視差を与える新しい映像表現を利用したコミュニケーションシステム MoPaCo を提案した。実現手法として、単眼カメラを用いてユーザの視点位置を検出する手法と、対話相手の 3 次元空間情報を生成する手法を提案し、システムの実装を行った。システムを用いて主観評価実験を行った結果、本運動視差映像は違和感なく立体知覚が可能であり、対話相手が実際に実在し会っている感覚を増強すること、ならびに対話相手の空間の位置関係の把握がし易くなることが確認された。今後は、実装時の各種処理の高速化、および運動視差映像による会話への効果を実験的に検証する予定である。

参 考 文 献

- 1) Strickland, L. H., Guild, P. D., Barefoot, J. and Paterson, S. A. Teleconferencing and leadership emergence. *Human Relations*, 31 (1978), 583-596.
- 2) Williams, E. Coalition formation over telecommunications media. *European Journal of Social Psychology*, 5 (1975), 503-507.
- 3) Heath C. et al.: "Disembodied conduct: communication through video in a multimedia environment", In *Proc. CHI 91*, ACM Press (1991), 99-103.
- 4) de Greef, P. and Ijsselstein, W. Social Presence in a Home Tele-Application. *CyberPsychology & Behavior*, 4(2) (2001), 307- 315.
- 5) 橋本雅行, 譚玉昆, 森本一成, 黒川隆夫, "ビデオ・コミュニケーションにおける画像表示条件と距離感の関係," 電子情報通信学会技術研究報告, HC, Vol.1994, No.89, pp.39-44, (1994).
- 6) 宮里勉, 岸野文郎, "ポインティング知覚における指差し情報と視線情報による正答率, The Journal of the Institute of Television Engineers of Japan 50(1) pp.137-139, (1996).
- 7) Cutting, J. and Vishton, P. Perceiving Layout and Knowing Distances. *Perception of Space and Motion*, W. Epstein and S. Rogers, eds., Academic Press (1995), 69-117.
- 8) Towles, H., Chen, W.C., Yang, R., Kum, S.U., Fuchs, H., Kelshikar, N., Mulligan, J., Daniilidis, K., Holden, L., Zeleznik, B., Sadagic, A., and Lanier, J. 3D Tele-Immersion Over Internet2. *IITP2002*, (2002).
- 9) 末永剛, 松本吉央, 小笠原司, "非拘束な運動視差提示 3 次元ディスプレイの提案と評価," ヒューマンインタフェース学会論文誌, Vol.9, No.1, pp.49-56, (2007).