

360度テーブルトップ型3Dスクリーンシステムにおける 視点検出による垂直視差再現に関する基礎検討

巻口 誉宗^{1,2} 高田 英明¹ 本田 健悟² 坂本 大介² 小野 哲雄²

概要：被写体をテーブル上に表示し、全周囲から立体的に視聴可能な映像表示技術は、エンターテインメント分野や産業分野での幅広い応用が考えられる。我々はこれまで、複数のユーザがテーブルトップ型ディスプレイの周囲360度好きな方向から、その角度に応じた3D映像を同時に視聴できるスクリーンシステムを提案した [1], [2]。このシステムでは、水平方向に対してなめらかな運動視差を提示できる一方、ユーザの身長差への対応や、視点の上下移動といった垂直方向の運動視差の提示は困難であった。そこで本稿では、スクリーン中心に設置した360度カメラの映像から画像認識によってテーブル周囲のユーザの視点位置を検出し、ユーザに提示する視点画像の仮想カメラを上下させることで垂直視差を再現する手法を提案する。この手法により、身長差のある複数のユーザの視聴や上下方向に視点移動を行った場合でもそれぞれのユーザに正確な3D映像を提示できる。我々は提案手法を実装し、テーブル周囲のユーザの視点位置に合わせて対応する視点画像をリアルタイムに変更できることを確認し、垂直視差再現への提案手法の有効性を示した。

Vertical Motion Parallax by Viewpoint Detection for 360-degree Table-top type 3D Screen System

Motohiro MAKIGUCHI^{1,2} Hideaki TAKADA¹ Kengo HONDA² Daisuke SAKAMOTO² Tetsuo ONO²

1. はじめに

被写体があたかもテーブル上に実在するかのように再現する360度テーブルトップ型3D映像表示技術は、複数のユーザが自分の応援するチームサイドからスポーツ競技を観戦できるライブビューイングや、デザイナーとエンジニアが製品の3Dデータを直感的に共有して協調作業を行える製品設計など、幅広い応用が考えられる。我々はこれまで、テーブル上に設置したスクリーンに対して複数のプロジェクタから異なる視点画像を同時に投影することで、テーブル周囲360度好きな方向からその角度に応じた3D映像を3Dメガネ無しで提示できるシステムを提案した [1]。このシステムはテーブルを取り囲む複数のユーザが同時に視聴でき、水平方向の視点移動に対してなめらかな運動視差を提示できる一方、ユーザの身長差への対応や、上下の視点移動といった垂直方向の運動視差の提示は困難であった。

そこで我々は、一台で360度全周囲の映像を取得可能な360度カメラによってユーザの視点位置を取得し、視点位置に合わせて投影する視点画像をリアルタイムに変更することで垂直視差を再現する手法を提案する。本手法はスクリーン中心の上部に設置した360度カメラの映像から画像認識によってリアルタイムにスクリーン周囲のユーザのボーンを推定し、視点位置の方位角と仰角を求める。この方位角を元にユーザに視点画像を提示しているプロジェクタを特定し、仰角に基づいて視点画像の仮想カメラの高さを変更する。我々は提案手法を360度テーブルトップ型システムに実装し、本手法が垂直視差再現に効果があることを確認した。

2. 関連研究

従来の360度3Dディスプレイにおける視点位置検出手法として、Jonesら [3] はユーザのこめかみに装着したセンサを電磁追跡システム (Electromagnetic tracking system)

¹ 日本電信電話株式会社 NTT サービスエボリューション研究所

² 北海道大学 大学院情報科学研究科

によってトラッキングして視点位置を求め、垂直視差を再現する手法を提案している。また、Seifert ら [4] は光学モーションキャプチャシステムを使用してユーザに装着したマーカを元に視点位置を求め、360 度スクリーンのパララックスバリアの方向を制御する手法を提案している。こうした装着物を必要とする手法は、事前にすべてのユーザへの装置装着が必要となることや、装着物の形状や大きさによってはユーザの表情を隠し、コミュニケーションの妨げとなるため、我々が目指す複数ユーザ向けの裸眼 3D のシステムに対しては望ましくない。

装着物が不要な手法として、Takeuchi ら [5] や Benko ら [6] は、深度カメラを用いてユーザの視点位置をトラッキングする手法を提案している。しかし、深度カメラ 1 台だけでは視野角が狭く、360 度の視域をカバーすることが困難な点に加え、観察者が複数いる場合には観察者同士の遮蔽が生じやすく、視点検出の精度低下が懸念される。こうした問題を解決するため、Wilson らの LightSpace [8] や Sousa らの Creepy Tracker Toolkit [7] では、複数のカメラやセンサ情報を統合することで部屋全体をセンシングし、ユーザの位置や行動をインタラクションシステムに反映する手法を提案している。これらは非常にパワフルな仕組みであるが、複数カメラの設置位置調整や座標のキャリブレーション、カメラ同士の時刻同期は簡単ではない。また、我々が必要な視点位置検出のみの用途ではオーバスペックである。

1 台のカメラでユーザを検出する手法として、Callemein ら [9] は会議室の天井に設置した 360 度カメラで会議室内の人を認識し、会議室の使用率を数値化するシステムを提案している。このシステムは機械学習ベースの画像認識技術である Darknet と OpenPose [11] を組み合わせて人の認識を行うことで、画像中のユーザのプライバシーを確保するために低解像度化処理を行った画像であっても認識できることを示している。このシステムは人数把握が目的であるため、視点位置検出は行われていない。Komiya ら [10] は円形テーブル上に置いた 360 度カメラの映像に対して OpenPose による画像認識を行い、テーブル周囲に座ったユーザの頭部姿勢と視線方向の推定によってユーザ同士のコミュニケーションにおける注意レベルを推定する手法を提案している。この手法ではテーブル上でタスクを行うことは考慮されておらず、カメラはテーブル上のユーザの視点よりも低い位置に設置されているため、顔のパーツ位置も取得することができる。一方で、我々のシステムではテーブル上に 3D 映像を表示するため、ユーザの視点位置よりも低い位置にカメラを設置することはコンテンツ視聴の妨げとなり、望ましくない。そこで我々は、ユーザの視点位置よりも高い位置に設置した 360 度カメラ映像に対して OpenPose による画像認識を行うことで観察者の視点位置を取得し、3D 映像における垂直視差を再現する手法を

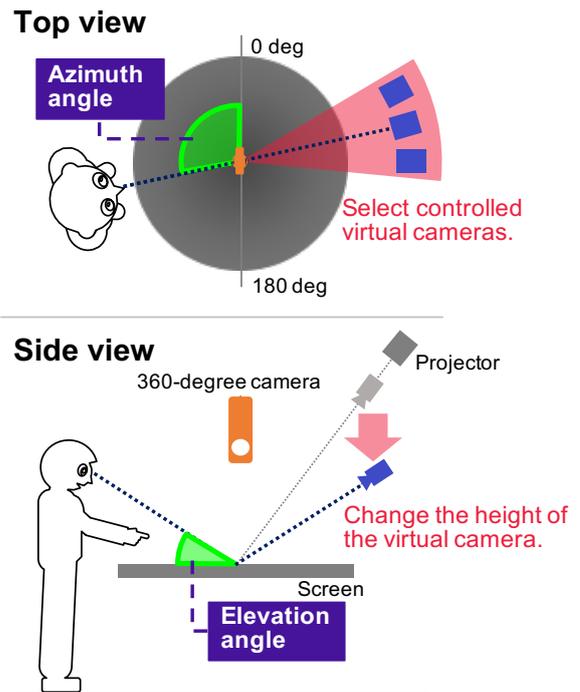


図 1 垂直視差再現の原理

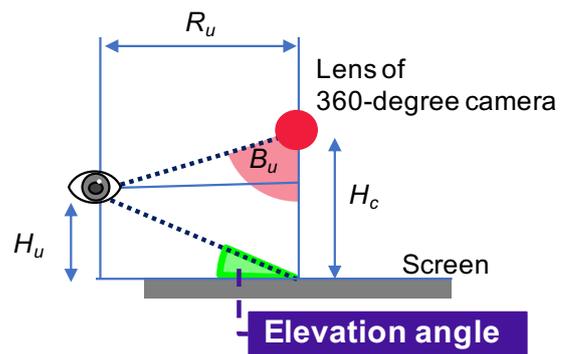


図 2 仰角算出に関連するパラメータ。

検討した。

3. 提案手法

図 1 に垂直視差再現の原理を示す。360 度カメラはスクリーン中央の、ユーザの視点位置より高い位置に設置し、スクリーン周囲に立つユーザの視点位置の方位角・仰角を算出する。方位角を元に、ユーザが視認するプロジェクタを特定し、そのプロジェクタの視点画像の視点の高さを仰角に基づいてリアルタイムに変更する。

OpenPose は画像中の人の全身 17 点のボーンを検出し、画像中の座標を出力する。本稿ではユーザの鼻のボーン位置を視点位置とし、画像中の xy 座標から方位角と仰角を算出する。方位角 A_z は、画像の横方向の解像度を W_i 、検出された視点位置の x 座標を X とすると以下の式で与えられる。

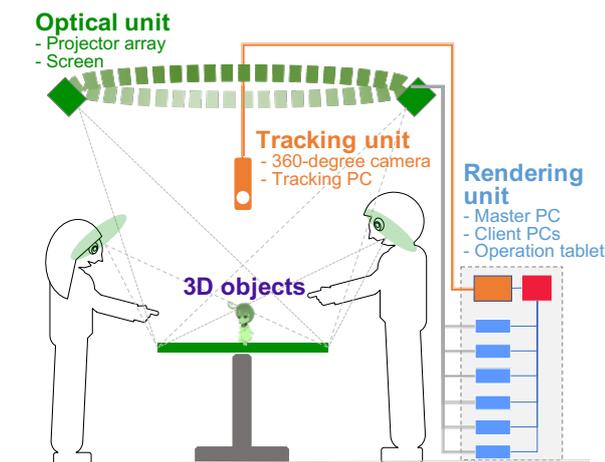


図3 360度テーブルトップ型3Dスクリーンシステムのシステム構成

$$Az = \frac{360 \times X}{W_i}$$

仰角 El は、画像の縦方向の解像度を H_i 、検出された視点位置の y 座標を Y 、スクリーンに対するカメラの設置高さを H_c 、カメラからユーザまでの水平距離を R_u とすると以下の式で与えられる。

$$El = \arctan\left(\frac{H_u}{R_u}\right)$$

$$H_u = H_c - \frac{R_u}{\tan B_u}$$

$$B_u = 180 - \frac{180 \times Y}{H_i}$$

本稿では、ユーザはスクリーン近傍でコンテンツを観察するという前提を置き、カメラからユーザまでの水平距離 R_u を一定と仮定することで奥行方向の距離算出を省略し、1台の360度カメラのみで仰角を算出した。

4. システム構成

図3にシステム構成を示す。システムは6度間隔で円形に配置された60台のプロジェクタと特殊な光学スクリーンから構成される光学部、プロジェクタと1対1で対応した60台のクライアントPCと、1台のマスタPCから構成されるレンダリング部、360度カメラと画像認識用のトラッキングPCから構成されるトラッキング部からなる。視点映像を投影するプロジェクタは6度間隔で設置され、60台で360度の視域をカバーする。各プロジェクタにはそれぞれクライアントPCがHDMIケーブルで接続されている。すべてのクライアントPCはUnity3Dで作られたアプリケーションによってCG空間を共有しており、視点画像はリアルタイムにレンダリングされる。

図4にレンダリング部の構成を示す。60台のクライアントPCは1台のマスタPCとローカルネットワークで接続されている。視点画像をすべてのクライアントPCで同

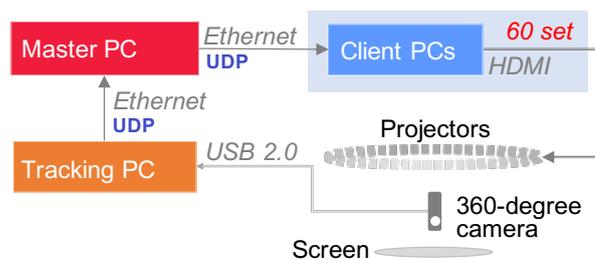


図4 レンダリング部の構成

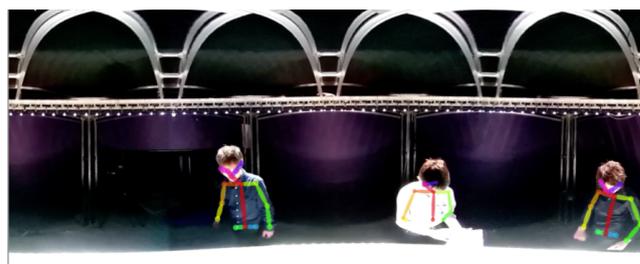


図5 3名のユーザの観察時におけるボーン検出結果。

期するため、マスタPCはCG空間中のオブジェクトのアニメーション・移動・回転の情報を毎フレームごとにUDPで送信する。360度カメラにはRICOHO社のTHETA Vを用いる。THETA VはUSBケーブルでトラッキングPCに接続される。THETA Vは、内部処理によって魚眼レンズによる画像歪みを打ち消し、横1920px縦960pxの全周囲画像を出力する。トラッキングPCではこの画像にOpenPoseによる画像認識を行い、視点位置の方位角と仰角を算出する。トラッキングPCはマスタPCとローカルネットワークで接続されており、方位角と仰角を送り続ける。マスタPCは上記のオブジェクト制御の情報に加え、各クライアントPCの仮想カメラの仰角を毎フレームごとにUDPで送信する。

上記システムを実装し、3名のユーザでの動作検証を行った。本稿ではカメラからユーザまでの水平距離 R_u をスクリーン半径600mmとし、スクリーンからカメラまでの高さ H_c は950mmとした。

5. 結果

図5に3名のユーザがスクリーン周囲から3Dコンテンツを視聴している際の360度カメラの映像と、その映像中のOpenPoseによるボーン検出の結果を示す。この図から、カメラ設置位置がユーザの視点位置より高い今回の設定においてもユーザの鼻も含めたボーン検出が行えることがわかる。さらに、スクリーンによってユーザの下半身が隠れているにもかかわらず上半身のボーンが検出されており、機械学習ベースのOpenPoseを用いるアプローチが視点位置検出に有効であることも確認できる。

図6にトラッキングがある場合とない場合において、視

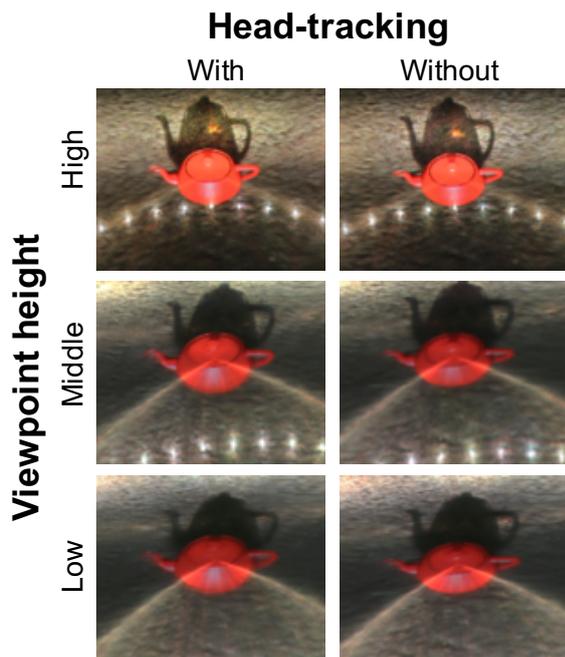


図6 視点高さを変化させた際のトラッキング有無による観察画像の比較

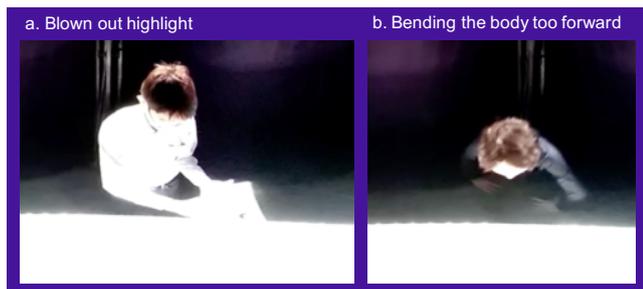


図7 視点検出エラーの例

点の高さを3段階に変化させた際の観察画像の変化を示す。この図から、トラッキングがない場合には視点の高さ変更によってオブジェクトのパスが崩れ、正確な3D映像を提示できない一方で、トラッキングがある場合にはパスが保たれ、視点の高さが変わっても本来の形状を提示できることがわかる。このことから、提案手法による垂直視差再現の有効性が確認できる。

一方で提案手法では、スクリーンの反射光によって白飛びが生じた場合(図7左)や、スクリーン上の被写体を覗き込むように覆いかぶさった場合(図7右)は、OpenPoseによるボーン検出の失敗が見られた。カメラ設置や画像認識の学習データの検討が今後の課題である。

6. まとめ

360度テーブルトップ型3Dスクリーンシステムに対し、1台の360度カメラを用いてユーザの視点位置を検出し、視点映像を制御することで垂直視差を提示する手法を提案

した。提案システムは機械学習を用いた画像認識によって360度カメラの映像中からユーザの視点位置を検出する。この視点位置に応じて仮想カメラの高さを変更することで身長差やユーザの視点の上下移動に対応した3D映像を提示する。本手法を、360度テーブルトップ型3Dスクリーンシステムに実装し、提案手法によって垂直方向の運動視差を提示可能なことを示した。提案手法は1台のカメラで全周囲をカバーできることで、複数のカメラを用いた手法と比較してカメラの位置合わせや時刻合わせ、内部パラメータのキャリブレーションが不要なことから、よりシンプルな機材構成、ソフトウェアで実装できる。今後はスクリーンの反射への対策やユーザの姿勢へのロバスト性を高め、視点位置の認識精度向上を検討する。

参考文献

- [1] 巻口誉宗, 高田英明, 川上徹, 篠井むつみ: "光学リニアレンディングを用いた360度テーブル型裸眼3Dスクリーンシステムの提案と基礎検討", 2018年映像情報メディア学会年次大会, 12A-5(2018)
- [2] M.Makiguchi and H.Takada: "360-Degree Tabletop Type 3D Screen System", Proceedings of LDC'19, LDC-2-01(2019)
- [3] A.Jones and I.McDowall, H.Yamada, M.Bolas and P.Debevec: "Rendering for an Interactive 360 Light Field Display", ACM Transactions on Graphics (TOG)", Vol.26, No.3(2007)
- [4] H.Seifert and Q.Smithwick: "360-degree Multi-Viewer Autostereoscopic Tabletop Display with Omnidirectional Dynamic Parallax Barrier and Novel Time-Multiplexed Directional Backlight", Electronic Imaging: Stereoscopic Displays and Applications, vol.XXVII, pp.1-8(2016)
- [5] T.Yusuke and M.Sugimoto: "An Immersive Surface for 3D Interactions", Proceedings of ITS'12, pp.359-362(2012)
- [6] H.Benko, R.Jota and A.D.Wilson: "MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop", Proceedings of SIGCHI'12", pp. 199-208(2012)
- [7] M.Sousa, D.Mendes, R.K.D.Anjos, Rafael, D.Medeiros, A.Ferreira, A.Raposo, J.M.Pereira and J.Jorge: "Creepy Tracker Toolkit for Context-aware Interface", Proceedings of ISS'17, pp.191-200(2017)
- [8] A.D.Wilson and H.Benko: "Combining Multiple Depth Cameras and Projectors for Interactions On Above and Between Surfaces", Proceedings of UIST'10, pp.273-282(2010)
- [9] T.Callemein, K. V. Beeck and T.Goedeme: "How Low Can You Go? Privacy-preserving People Detection with an Omni-directional Camera", Proceedings of VISIGRAPP'19, VISAPP19-RP-238(2019)
- [10] R.Komiya, T.Saitoh and K.Shimada: "Image-based Attention Level Estimation of Interaction Scene by Head Pose and Gaze Information", Proceedings of ICIS'18, pp.497-501(2018)
- [11] Z.Cao, T.Simon, S.E.Wei and Y. Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", Proceedings of CVPR'17, pp.7291-7299(2017)