

2人用テーブルトップ型運動視差立体視CGシステムとインタラクション手法の提案

水野 慎士¹

概要：本稿では2人が同時使用可能なテーブルトップ型運動視差立体視CGシステムと、そのシステムでのインタラクションの実現手法を提案する。著者らが開発した従来のテーブルトップ型運動視差立体視CGシステムは、1人のユーザを検出してそのユーザに特化した映像を表示する必要があるため、2人で同時に使用することはできなかった。そこで、従来の手法と実装ハードウェアを拡張することで、2人のユーザが1つのテーブル上で同時に三次元CG物体を観察したり操作することができる運動視差立体視CGシステムを提案する。

Proposals of a tabletop stereoscopic 3DCG system with motion parallax for two users and a method to interact on the system

Shinji Mizuno¹

1. はじめに

3Dディスプレイやホログラムなどの立体視映像を取り扱うことが可能な映像提示装置が進歩、普及するのに伴って、立体視映像を用いたコンテンツも増えてきた。立体視映像では表示物体がまるでその場に存在するような臨場感を実現することができるため、エンタテインメント、コミュニケーション、医療など、様々な分野での活用が期待されている。

立体視映像の生成には3DディスプレイやHMDなど両眼視差に基づく手法が一般的だが、著者らは運動視差を用いた立体視CGシステムの開発を行ってきた[1]。運動視差は観察する位置の移動による立体物の見え方の変化で、両眼視差に匹敵する重要な立体知覚である。筆者らが開発してきた運動視差立体視CGシステムではユーザ視点を追跡しながら視点位置に応じたCG映像を逐次生成することで運動視差を再現する。それにより、一般的なプロジェクタやディスプレイを用いた立体視映像の生成を実現している。そして、運動視差立体視映像をテーブルトップに表示した場合には、テーブル周囲のあらゆる方向からCG物

体を立体視で観察することができ、CG物体を直接手で操作するインタラクションも可能である。そのため、実際にテーブル上に存在するような感覚でCG物体を観察することができるという特徴があり、科学館展示物やエンタテインメントコンテンツとして利用されている[2][3]。

著者らが開発してきたテーブルトップ型運動視差立体視CGシステムでは、1人のユーザの視点位置に特化した映像をテーブルトップ上に表示する必要があり、2人での同時には対応していなかった。2人以上のユーザに対応した運動視差立体視映像を実現する手法としては、視点位置別に用意した多数のプロジェクタを組み合わせさせた手法[4]や、特殊なレンズを用いた手法[5]などが提案されているが、これらの手法は特殊なハードウェアが必要な上、視点位置の制限、限られたインタラクションなどの問題がある。また、Oculus[6]やHoloLens[7]などのHMDを用いて観察者ごとに異なる映像を提示する方法もあるが、機器の重量や価格、視野角の狭さなど、一般的に普及するには解決すべき問題がある。

そこで、本研究では著者らが開発してきた手法を拡張して、2人のユーザが同時に使用することができる運動視差立体視CGシステムを提案する。提案システムでは従来システムと同様にテーブルトップ上に投影して、2人のユー

¹ 愛知工業大学 情報科学部
Faculty of Information Science, Aichi Institute of Technology

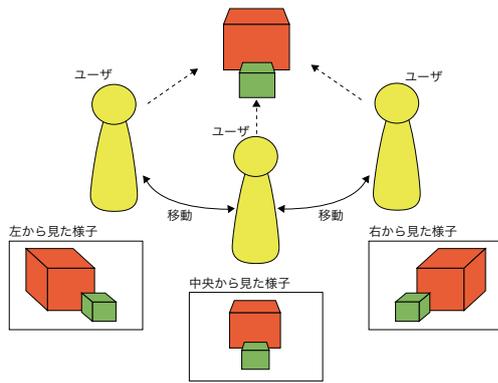


図 1 運動視差の概要

ずは同じテーブル上の映像を観察する。このとき、システムは各ユーザの視点三次元位置を追跡しながら、視点位置に応じた 2 枚の運動視差立体視映像を生成して、それらを 3D 対応ディスプレイを用いて同じ場所に重畳表示する。そして、各ユーザはアクティブシャッターグラスを用いて自分の視点に対応する映像だけを見る。これにより、同じテーブルトップ上に表示された運動視差立体視による三次元 CG 物体を 2 人のユーザで共有することが可能となる。さらに、2 人のユーザのそれぞれの手の座標を検出することで、各ユーザが立体視三次元 CG 物体とインタラクションを行うことも可能にする。

2. 運動視差立体視 CG システムの概要

2.1 運動視差

運動視差とは、立体物を観察するときに観察者または立体物が移動することで生じる見え方の変化である。図 1 に示すように、運動視差では観察者の移動に伴って今まで見えなかった部分が見えるようになったり、近くの物体が遠くの物体に比べて見え方が大きく変化する。Rogers らの研究では、運動視差のみで三次元形状と奥行きに関する十分な情報が得られることを示している [8]。

2.2 運動視差立体視 CG システム

運動視差立体視 CG システムは、ユーザの視点位置の移動に合わせて映像を変化させて運動視差を再現することで立体視を実現する CG 生成システムである。システムはテーブルトップディスプレイ、処理用 PC、RGB-D カメラで構成される。図 2 にシステム構成を示す。

ユーザの視点座標は深度画像に基づいて計算する。そのため、RGB-D カメラはテーブル上方に設置して、テーブル上および周囲の深度画像を取得する。ユーザがテーブルの周囲で映像を観察するとき、ユーザの頭部は RGB-D カメラから見て最も高い点となる。そこで、深度画像中で最も大きな高さ d_{max} を求めて、 d_{max} との差がしきい値 d_t 以下の高さを持つ領域をユーザの頭部領域として深度画像から抽出する (図 3)。そして頭部領域の重心点の三次元座



図 2 運動視差立体視 CG システムの構成



(a) 深度画像から最も高い点の検出 (b) しきい値処理による頭部領域抽出

図 3 深度画像に基づく頭部領域抽出



図 4 運動視差立体視 CG の表示例

標を深度情報を用いて求めて、その点から頭頂部と目の高さの差 w に相当する高さを下げた空間中の点をユーザの視点三次元座標 e とする。なお、 d_t の値および w の値は実験的にそれぞれ 150(mm)、50(mm) としている。

システムでは取得したユーザ視点座標に応じた三次元 CG 物体の映像を生成する。このとき、テーブルトップに三次元 CG 物体が存在するよう感じさせるため、実空間と CG 空間の座標系を一致させる。すなわち、CG 生成用視点座標をユーザ視点座標に一致させ、CG 生成用投影面をテーブルトップスクリーンに一致させる。この環境で生成された映像は、ユーザ視点位置以外から観察した場合は歪んで表示されるが、ユーザ視点からは CG 物体がテーブルトップに配置されているように観察される。そして、ユーザの位置の移動に応じて CG 映像を逐次更新することで運動視差が再現され、まるで三次元 CG 物体がテーブルトップに存在するように立体感を感じながら、様々な位置から観察することが可能となる (図 4)。

三次元 CG 物体とのインタラクションを実現するため、



(a) 深度画像からテーブル面上で最も低い点の検出 (b) しきい値処理による手領域抽出

図 5 深度画像に基づく手領域抽出

システムは深度画像に基づいてテーブル上のユーザの片手の位置も取得している。ユーザがテーブル上に自然に手を差し出した場合、手の位置は RGB-D カメラから見てテーブル面上で最も低い位置となる場合が多い。そこで、テーブル面の上方 d_{th} 以内の空間で最も低い位置 d_{min} を求めて、 d_{min} との差がしきい値 w 以下の高さを持つ領域をユーザの手領域として深度画像から抽出する (図 5)。そして、頭部の場合と同じ手法で手の三次元座標 h を求める。なお、 d_{th} の値は実験的に 300(mm) としている。

3. 2人ユーザに対応した運動視差立体視 CG システム

3.1 従来システムの問題点と改良

2章で述べた従来のテーブルトップ型運動視差立体視 CG システムでは、ユーザ視点座標を取得してから、その座標に特化した CG 映像をテーブルトップディスプレイに表示している。そのため、システムは 1人で使用することが前提となっており、協調作業のためのシステムや対戦ゲームなどへの応用はできなかった。

そこで、本研究では従来の運動視差立体視 CG システムを 2人で同時使用ができるような改良を行う。これを実現するには、テーブルの周囲に立つ最大 2名のユーザの視点座標をそれぞれ同時に取得するとともに、テーブルトップディスプレイに 2つの映像を同時に表示させながら、各ユーザに個別に提示する必要がある。また、各ユーザが立体視 CG 映像とのインタラクションを実現するため、各ユーザの手の座標も取得する必要がある。

3.2 2人ユーザの視点座標の同時取得

2人のユーザの視点座標をそれぞれ取得する手法は、従来システムの手法を拡張したもので、テーブル上部に設置した RGB-D カメラで取得した深度画像に基づいて行う。図 6 に深度画像に基づく 2人のユーザの視点座標取得の様子を示す。

1人目のユーザの視点座標の取得方法は従来システムとほぼ同様である。深度画像中で最も高い点を抽出して、しきい値処理を施すことで頭部領域を抽出する (図 6(b))。ただし、2人のユーザの頭部の高さがほぼ同じときには、深度画像から 2つの領域が抽出される場合がある。そのた

め、最大の面積を持つ領域を 1人目のユーザの頭部領域とする。抽出した頭部領域に対して従来手法と同様の処理を施して、1人目のユーザの視点三次元座標 e_1 を計算する。

1人目のユーザの視点座標を取得したら、1人目のユーザの頭部領域に対して領域拡張処理を施す。そして、拡張した領域の内部に対応する深度画像領域の深度値を削除する (図 6(c))。これにより、1人目のユーザの周辺は次の視点位置の候補から外れる。

次に、1人目のユーザの視点座標の高さの位置から d_t ずつしきい値を下げながら、深度画像から領域抽出を行う (図 6(d)(e)(f))。このとき、しきい値以上の面積を持つ領域が抽出された場合に、その領域を 2人目のユーザの頭部領域とする (図 6(f))。抽出した頭部領域に対して従来手法と同様の処理を施して、2人目のユーザの視点三次元座標 e_2 を計算する。

以上の処理により、深度画像から 2人のユーザの視点座標を計算することができる (図 6(g))。

3.3 2人ユーザへの個別映像提示

システムは 2人のユーザの視点座標をそれぞれ用いて、2つの視点からの CG 映像を同時に生成する。2人で同時に運動視差立体視を行うには、生成した 2つの映像を同じテーブルトップディスプレイに表示する必要があるが、通常のディスプレイではどちらか一方のユーザのための映像しか表示できない。

そこで、2つの CG 映像を同じテーブルトップディスプレイに表示させるために、本研究では 3D 対応プロジェクタを用いる。3D 対応プロジェクタでは両眼視差立体視映像を表示するために右目用と左目用の映像を重畳して表示している。提案手法ではそれを応用して、1人目のユーザ視点座標に対応した CG 映像を右目用映像、2人目のユーザ視点座標に対応した CG 映像を左目用映像として、3D 対応プロジェクタでテーブルトップ上に重畳表示する。

重畳された 2つの CG 映像は二重に表示されているため、そのままでは正しく映像を見ることができない。そこで、3D 対応プロジェクタ用のアクティブシャッターグラスを用いる。アクティブシャッターグラスの片側のレンズを塞いで用いることで、重畳された 2つの映像を分離することができる (図 7)。

1人目のユーザはアクティブシャッターグラスの左目側レンズを塞いで装着する。そのため、3D 対応プロジェクタの右目用映像だけ、すなわち 1人目のユーザ用の CG 映像だけを観察することになる。2人目のユーザはアクティブシャッターグラスの右目側レンズを塞いで装着することで、左目用映像として表示されている 2人目のユーザ用の CG 映像だけを観察する。以上により、2人のユーザはデスクトップに重畳表示された 2つの CG 映像のうち、それぞれの視点座標に特化した運動視差立体視 CG 映像を観察

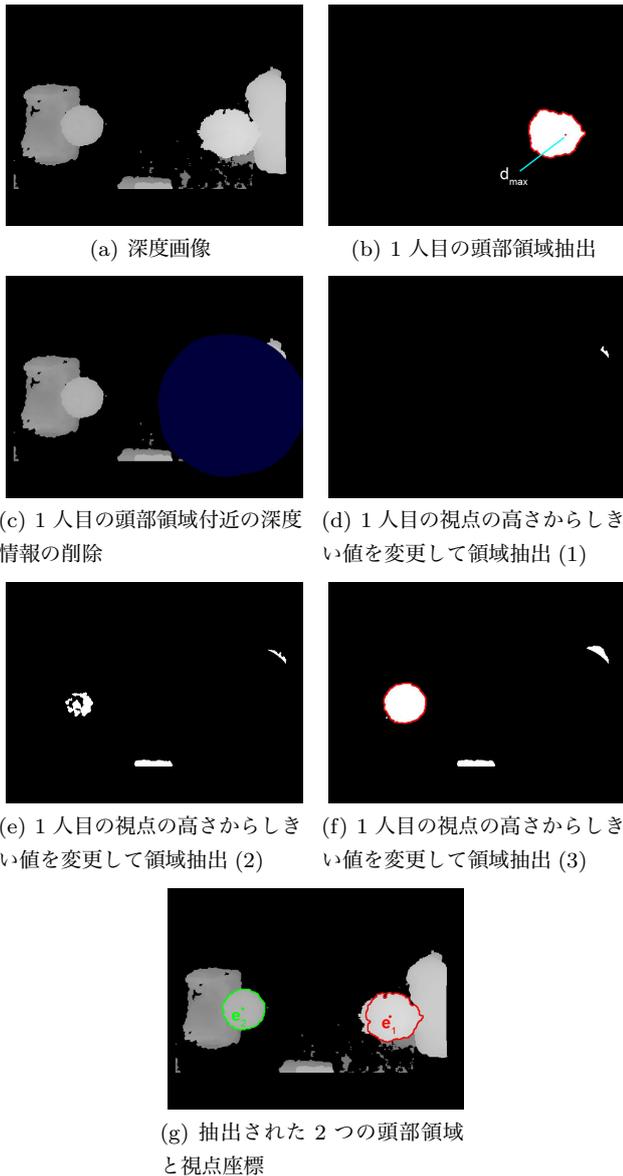


図 6 深度画像に 2 人のユーザの視点座標の取得の様子



図 7 片側のレンズを塞いだアクティブシャッターグラス

することが可能となる。

3.2 節で述べた 2 人のユーザの視点座標取得手法では、頭部位置の高い方が 1 人目のユーザの視点座標として取得されるため、ユーザの頭部位置が上下に移動した場合には 1 人目と 2 人目のユーザが切り替わってしまい、左右の映像が入れ替わる場合がある。そこで、前フレームの各ユーザの視点座標と現フレームで計算された 2 つの視点座標間の距離を計算して、現フレームの各ユーザの視点座標として現

在の距離が近い方の視点座標を採用する。これにより、頭部位置の上下移動による映像の入れ替わりを防ぎ、各ユーザに常に適切な映像を提示することを実現している。

なお、アクティブシャッターグラスは偏向グラスと異なり観察方向の影響を受けない。そのため、各ユーザは従来通りテーブル周囲のあらゆる方向から運動視差立体視映像を観察することができる。

3.4 2 ユーザによるインタラクション

2 人のユーザの手の座標をそれぞれ取得する手法も、テーブル上部に設置した RGB-D カメラで取得した深度画像に基づいている。ただし、従来手法のようにテーブル面上で最も低い点を抽出するのではなく、先に取得した各ユーザの視点座標を利用して各ユーザの手の座標を計算する。なお、2 人ユーザに拡張したシステムにおいても、現状では各ユーザの片手操作を実現する。

まず、3.2 節で求めた 1 人目のユーザの視点座標とテーブル面を高さを用いて深度画像にしきい値処理を施して、一定の面積以上の領域をすべて抽出する。このとき、2 人のユーザが接触していない限り、理論的にはユーザ数と同じ数の領域が抽出されて、各領域には各ユーザの視点座標 e_1 と e_2 に対応する点が 1 つずつ含まれる。

次に、各領域と e_1 および e_2 に対応する点との内包関係を調べる。 e_1 に対応する点を含む領域は 1 人目のユーザの領域、 e_2 に対応する点を含む領域が 2 人目のユーザの領域となる。

ユーザがテーブル上の物体に片手を伸ばして操作する場合、手の先の点はユーザの頭部から最も遠い点となることが多いと考えられる。そこで、各ユーザの領域の輪郭を追跡して、対応ユーザ視点座標から最も離れた場所にある輪郭点を求める。そして、深度画像に基づいて輪郭点の三次元座標を求めて、これをユーザの手の三次元座標とする。

処理に用いる領域は各ユーザの視点座標に基づいて識別されているため、取得した手の座標も各ユーザに識別することが可能である。そこで、1 人目のユーザと 2 人目のユーザの手の三次元座標をそれぞれ h_1 、 h_2 とする (図 8)。

各ユーザの手の座標を取得することで、運動視差立体視で表示された CG 物体との接触、移動、変形などのインタラクションを 2 人のユーザが同じテーブル上で同時に行うことが可能になる。

4. 実験

提案手法に基づいてプロトタイプシステムを実装して実験を行った。使用した PC は iMac(Core i7 3.4GHz, 32GB MM, NVIDIA GeForce GTX 780M)、プロジェクトは超短焦点 3D 対応の NEC NP-U321HJD、アクティブシャッターグラスは NEC NP02GL、RGB-D カメラは Kinect を使用している。

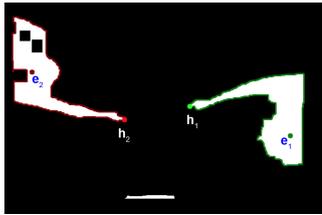
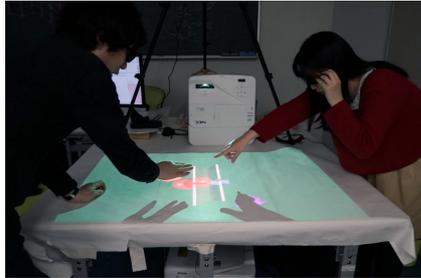
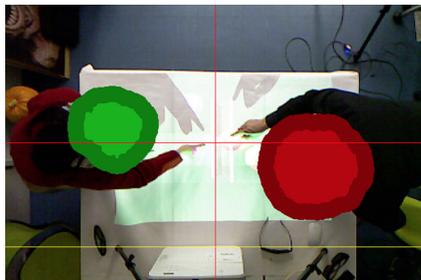


図 8 ユーザ領域の抽出と深度画像に基づく各ユーザの手の座標の取得



(a) アクティブシャッターグラスを装着した 2 人のユーザ



(b) 抽出された各ユーザの頭部領域

図 9 システム使用の様子

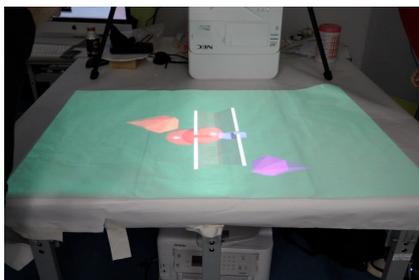
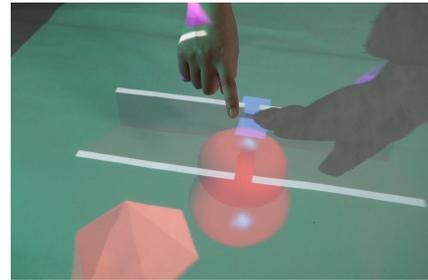


図 10 重畳表示された各ユーザ視点用映像

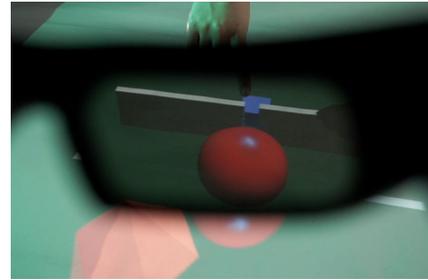
図 9 にシステム使用の様子を示す。2 人のユーザはそれぞれ片側のレンズが塞がれたアクティブシャッターグラスを装着して、テーブル周囲の自由な場所からテーブルトップの映像を観察する (図 9(a))。このとき、3.2 節で述べた手法により、2 人のユーザの頭部領域が正しく抽出されて、それぞれの視点座標が正しく計算された (図 9(b))。

システムは各ユーザの視点座標に応じた 2 つの映像から 1 つのトップアンドボトム映像を生成してプロジェクタの入力する。それにより、図 10 に示すようにテーブルトップ上には 2 枚の CG 映像が重畳表示された。

ここで、ユーザがテーブルトップ上の映像を観察すると、



(a) 重畳表示された CG 映像



(b) アクティブシャッターグラスを装着して観察した CG 映像

図 11 アクティブシャッターグラスを用いた重畳表示映像の分離

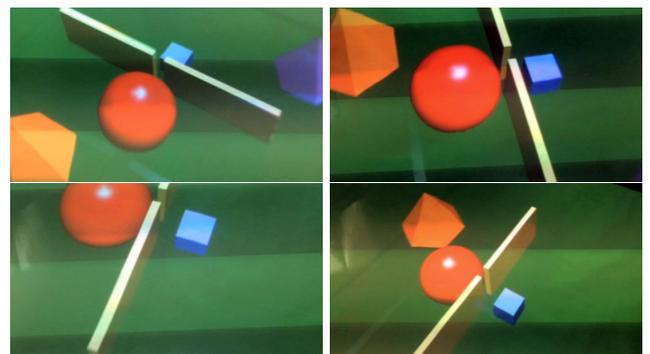
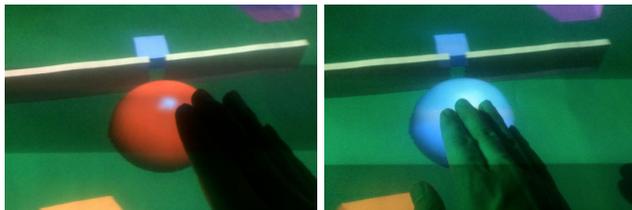


図 12 視点位置に応じた CG 映像生成による運動視差立体視映像

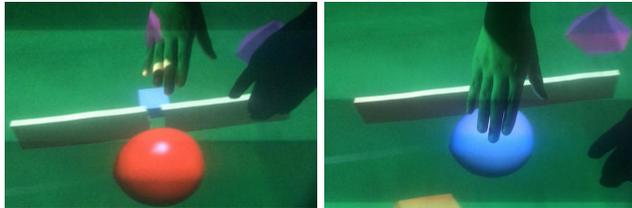
アクティブシャッターグラスによって映像が分離されて、図 11 に示すように各ユーザの視点座標に特化した映像だけが観察された。そして、図 12 に示すように、ユーザの移動に応じて逐次各ユーザの視点座標に応じた CG 映像が生成更新されて、運動視差を再現した。

図 13 に 2 人ユーザによる運動視差立体視 CG 映像とのインタラクションの様子を示す。ユーザが運動視差立体視で表示された CG 物体に手を差し出すと、システムはユーザの手の三次元座標を適切に取得して、ユーザ自身がボール表面に手が到達したと感じた場所でボールの色が変化した (図 13(a))。そして、他のユーザが手を伸ばした場合でも、システムは他のユーザの手の三次元座標を適切に取得して、ボール表面に手が到達したと感じた場所でボールの色が変化するのを確認した (図 13(b))。

以上により、提案手法を用いることで、2 人のユーザが同時に同じテーブル上に表示された運動視差立体視 CG 物体を共有して、テーブル周囲の様々な方向から観察したり触れたりすることが可能であることを確認した。



(a) 自身によるボールへのタッチ



(b) 他ユーザによるボールへのタッチ

図 13 2人ユーザによる運動視差立体視 CG 映像とのインタラクションの様子 (タッチによる色の变化)

5. まとめ

本研究では、著者らが開発してきたテーブルトップ型運動視差立体視 CG システムを 2 人で同時使用しながらインタラクションを可能にする拡張を行った。その実現ために、深度画像に基づく 2 人のユーザの視点座標と手の座標の取得方法の開発、3D 対応プロジェクタとアクティブシャッターグラスを用いた 2 つの映像の重畳表示と分離手法の提案を行った。実験では提案手法によって運動視差立体視 CG 空間を 2 人のユーザで共有可能であることを確認した。

今後の課題としては、各ユーザの両手による操作に対応させるとともに、物体の移動や変形など、より複雑なインタラクションを実現することが挙げられる。また、運動視差立体視 CG 映像の共有という提案システムの特徴を生かしたコンテンツの制作を行うつもりである。

謝辞 本研究の一部は科研費基盤研究 (C)(26330420) による。

参考文献

- [1] 塚田真未, 水野慎士: 運動視差立体視を用いた三次元 CG 天体ビューアの開発, 芸術科学会論文誌, Vol. 13, No. 3, pp. 134-143 (2014).
- [2] 上原悠永, 水野慎士: 擬似的三次元コピーの生成とインタラクションの実現方法, 情報処理学会論文誌・デジタルコンテンツ, Vol. 3, No. 2, pp. 22-31 (2015).
- [3] 水野慎士, 磯田麻梨乃, 伊藤玲, 岡本芽唯, 近藤桃子, 杉浦沙弥, 中谷有希, 廣瀬元美: インタラクティブコンテンツ「お絵描きダンスステージ」の開発, DICOMO2015 論文集, pp. 1841-1846 (2015).
- [4] S. Yoshida: fVisiOn: Interactive Glasses-free Tabletop 3D Images Floated by Conical Screen and Modular Projector Arrays, SIGGRAPH ASIA 2015 Emerging Technology (2015).
- [5] Microsoft Applied Science Group: The Wedge, [https://www.microsoft.com/appliedsciences/content/](https://www.microsoft.com/appliedsciences/content/projects/wedge.aspx)

- projects/wedge.aspx.
- [6] Oculus: Oculus Rift, <https://www.oculus.com/en-us>.
- [7] Microsoft: HoloLens, <https://www.microsoft.com/microsoft-hololens/en-us>.
- [8] Rogers, B. J., Graham, M.: Motion Parallax as an Independent Cue for Depth Perception, *Perception*, No. 8, pp. 125-134 (1979).