

遠隔空間内の物に対する回転と 設置が表現可能なビデオチャット

濱上 宏樹¹ 吉野 孝²

概要：我々はこれまでに可動枠を用いた部分重畳表示型ビデオチャットシステム「ドアコム Z」を開発してきた。ドアコム Z は、現実中存在する枠を介して遠隔地間を仮想的につなぎ、相手の空間と繋がっているような表示を行うビデオチャットシステムである。深度情報を用いることで、遠隔空間内における三次元的な移動と遠隔空間内の物に対する接触を表現した。遠隔空間内における三次元的な移動の実現により、「遠隔空間内に隠れる」「遠隔空間内を指差す」という表現が可能になった。また、遠隔空間内の物に対する接触の実現により、「遠隔空間内の物を掴む」「遠隔空間内の物を動かす」という表現が可能になった。しかし、「遠隔空間内の物を回転させる」「遠隔空間内に物を置く」という表現は実現できていない。そこで本研究では、骨格情報と深度情報を用いた、遠隔空間内の物に対する接触 後 のインタラクションを提案する。

Video Chat Which can Express Rotation and Placement to a Thing in Remote Space

HIROKI HAMAUE¹ TAKASHI YOSHINO²

1. はじめに

近年、Skype などの無料でビデオチャットを行えるツールが普及してきている [1]。スマートフォンの普及 [2] により、容易にビデオチャットを行える環境が整ってきた。そのため、今後は会議などのフォーマルな場面での利用だけでなく、ビデオチャットの家族や友人とのインフォーマルなコミュニケーションでの利用が増えてくると考えられる。しかし、従来のビデオチャットには、対面環境に比べて対話相手との距離感を感じてしまう問題点がある。この問題点を改善するため、従来、メディアスペースによって遠隔地間を対面環境に近づける試みが多くなされてきた。大画面のディスプレイで遠隔地の相手を等身大に表示すること [3] や、ハーフミラーを用いたシステムでアイコンタクトを可能にすること [4] によって、遠隔地にいる相手の存在感が増すことがわかっている。また、遠隔地間を対面環境に近づける以外の試みも多くなされてきた。相手と自分が鏡の中で同じ場所にいるような表示を行うこと [5] や、聴

覚や触覚に情報提示を行い相手が隣にいるような感覚を与える方法 [6] によって対話相手との同室感が増すことがわかっている。

しかし、これらの方法を用いても対面環境との差は依然大きい。そこで我々は、遠隔空間内における三次元的な移動と遠隔空間内の物に対する接触が表現可能なビデオチャットシステム「ドアコム Z」を開発してきた。本システムは、対話相手と自身の映像の重畳表示を行う部分重畳表示型ビデオチャットである。ドア型の専用インタフェースを用いている。部分重畳表示型ビデオチャットとは、遠隔地間の一方または両方のカメラの映像の一部を、他方の映像や別の映像に重畳し、互いが同じ映像を見て会話を行うシステムである。本システムは、従来のビデオチャットでは不可能であった遠隔空間における接触の表現を実現した。しかし、「遠隔空間内の物を回転させる」「遠隔空間内に物を置く」という表現は実現できていない。そこで本研究では、骨格情報と深度情報を用いた、遠隔空間内の物に対する接触 後 のインタラクションを提案する。

¹ 和歌山大学院システム工学研究科

² 和歌山大学システム工学部

2. 関連研究

遠隔重畳型ビデオチャットとして森川らの超鏡 [5] がある。超鏡では、相手と自分が鏡の中で同じ場所にいるような表示を行うことにより、存在感を実現するシステムである。この研究は、ビデオ対話に適した映像であれば、現実にはありえない状況でも存在感を実現することが可能であることを示した。超鏡と本システムの類似点としては、重畳型ビデオチャットである点や、対話者同士が同じ映像を見ているため WISI-WYS(What I See Is What You See) を満たしている点などがある。相違点としては、超鏡が鏡をメタファとして、遠隔対話者同士が同じ部屋にいるような映像を作り出しているのに対し、本システムは遠隔地の空間がドアによって繋がっている点がある。また、それを実現するために実際のオブジェクトをインタフェースとしている点も異なる。

ビデオチャットにおいてはユーザの位置関係も重要な要素である。実際の部屋と同じような空間を作り出すシステムとして、平田らの t-Room[7] がある。t-Room ではユーザが同じ部屋にいるような感覚として同室感を定義している。ユーザの位置関係を考慮して、実対面しているような協調コミュニケーション環境を実現し、同室感を高めることを目標にしている。

ビデオチャットにおいて存在感を向上させるための、腕を用いた様々な手法が提案されている。Tang らの VideoArms[8] では共有している作業領域上にある腕の映像を対話相手に表示することにより、対話相手との共有空間を指示することが可能である。大西らの PopArm [9] では、対話相手の映像から境界面を超えて実体化したかのように見せることにより、存在感を出すシステムである。実体化させた腕を用いて対話相手側の空間を指示することが可能である。HANDY System[10] では利用者の手の映像を対話相手の対話相手の顔の映像と組み合わせることで、存在感の向上を行った。これらの研究により、ビデオチャットにおいて腕を対話相手の空間に侵入させる手法は存在感が向上することが分かっている。しかし、これら手法では遠隔地の物に対する指示は可能であるが、遠隔地の物に対する接触や移動に関しては表現できていない。

3. ドアコム Z

3.1 システム構成

図 1 にドアコム Z のシステム構成を示す。本システムはドアを持たないドア無し側 (図 1(a)) とドア型インタフェースを操作するドア操作側 (図 1(b)) とで通信を行う。図 2 にドアコム Z の表示映像を示す。図 2 内でドアを持ちドアから手を伸ばしているユーザがドア操作側、地図を持っているユーザがドア無し側である。ドア操作側ユーザは、ド

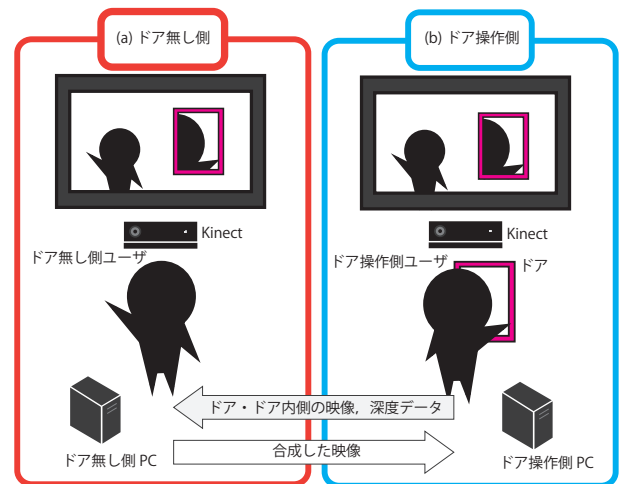


図 1 システム構成図



図 2 ドアコム表示映像

アを顔の前に掲げてドアを開き、その中からディスプレイを覗くようにして会話をを行う。各ユーザのディスプレイには、ドア無し側の部屋の映像にドア操作側のドアとドアの内側が重畳表示される。これにより、ドアの操作側とドア無し側の空間がドアによって繋がっているような映像となる。Kinect によりカラー情報、深度情報、骨格情報を取得し通信を行う。

(A) ドア無し側

ドア無し側の構成は図 1(a) である。ドア無し側の使用機器はドア無し側 Kinect, ドア無し側 PC である。ドア無し側では、ドア無し側 Kinect でユーザを撮影し、ドア操作側に送信する。ドアの操作側 PC から送られてきたドアのフレームとその内側の映像を、ドア無し側 PC でドア無し側の映像とオーバーラップさせる。また、ドア操作側ユーザの動作に応じた処理を行う。

(B) ドア操作側

ドア操作側の構成は図 1(b) である。ドア操作側の使用機器はドア, ドア操作側 Kinect, ドア操作側 PC で

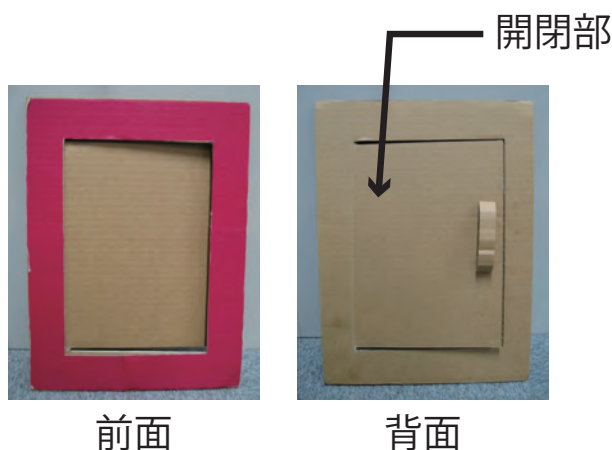


図3 ドアコムのインタフェース

ある。ドア操作側 Kinect でドア操作側ユーザを撮影し、ドア操作側 PC でドアのフレームを認識してドアのフレームとその内側の映像、枠外に出た手の映像と深度情報、右手の状態、骨格情報をドア無し側に送信する。

3.2 ドア型インタフェース

図3に、ドアコムのインタフェースを示す。ドアコムは、他人の部屋に入るために使用する「ドア」をメタファとした。本システムのインタフェースはユーザの負担を軽減するため、ダンボールを材料としている。カメラと向かい合わせとなる方を前面、ユーザ側に向く方と背面とする。前面には色を着ける。

ドア前面の着色は以下の条件を満たす必要がある。

(1) 色

色は後述するドアの認識に用いられる。ドアの色がユーザの服の色と近似色である場合、後述する理由によりドアの認識精度が低下する。また、ドアの枠の色は遠隔空間との境界を示す役割を持つため、彩度の高い色が望ましい。

(2) 材質

前面の材質は光を反射しない物が良い。本システムで用いている深度センサーは金属などの光を反射する材質に対して、深度を計測することができない。そのため、光沢のある材質を用いるとドアを認識できない。

本研究では、赤の色画用紙を用いてドアの前面を着色した。

3.3 ドアの認識

ドア操作側をドア無し側に重畳表示するために、ドアとドアの内部を検出する必要がある。Kinect から得られる情報を基にドアの認識を行う方法を検討した。図4にドア認識方法の図を示す。図4(a)はKinect から得られるカラー画像と深度画像の座標を対応させ表示したものである。

まずは、色情報を用いてドア領域の検出を行う方法を開発した。ドアコムZには赤いドアをインタフェースとして使用する。この赤色をカラー画像より抽出してドアの領域と認識する。画面の赤がドア枠のみの場合、この方法によりドアの領域を認識可能であった。しかし、画面内にドア枠と同系色の物が存在した場合、ドアの領域を正しく認識することができなかった。例として図5(a)のようにドア操作側ユーザがドア枠と同系色の服を着用している場合である。この時のドア認識結果が図4(b)である。服の一部がドアの領域と認識されている。

次に、深度情報を用いてドア領域の検出を行う方法を開発した。インタフェースとして用いられるドア枠は平面である。そのため、Kinect に対してドアを平面方向に持った場合、ドア領域の深度は近い値となる。人体は曲面で構成されているため、平面であるドアの領域の深度情報の特徴とは異なり深度値に幅がある。この深度情報の特徴を用いてドアの領域の検出を行った。また、検出を行う深度の範囲を設定することで、背景や深度が測定できていない画素の除去が可能である。この方法を用いて図4(a)からのドア認識を行った結果が図4(c)である。色情報を用いた検出と異なり、服の一部をドアと誤って認識することはない。深度情報を用いたドア領域の検出の問題点は、ドアがKinect に対して平行である必要がある点である。ドアの面とKinect との角度が増大するとドアの認識精度は低下する。これは、深度情報に平面である特徴が弱くなるためである。

以上のことから、色情報もしくは深度情報の一方を用いたドア領域の検出ではそれぞれ問題点があった。そこで、本システムでは色情報と深度情報の両方を用いたドア領域の検出を行う。また、本システムでは遠隔空間内におけるインタラクションを行うため、ドア枠から伸ばした腕の領域の検出を行う。腕の検出にはKinect から得られる深度情報と身体情報を用いる。ドアと腕の領域検出の流れは以下の通りである。図5に重畳画像作成手順を示す。

(1) 取得画像

Kinect から取得したカラー画像と深度画像の座標を一致させる。

(2) 色情報による検出

赤色をカラー画像より抽出し、ドアの領域の検出を行う。

(3) 深度情報の処理

背景やノイズを除去するため、深度値の範囲設定を行う。平面の深度情報の特徴からドアの領域を検出する。ドアの領域の深度値からドアの奥行方向の位置を求める。

(4) ドアと腕の検出

カラー画像よりドア領域であると検出された画素の深

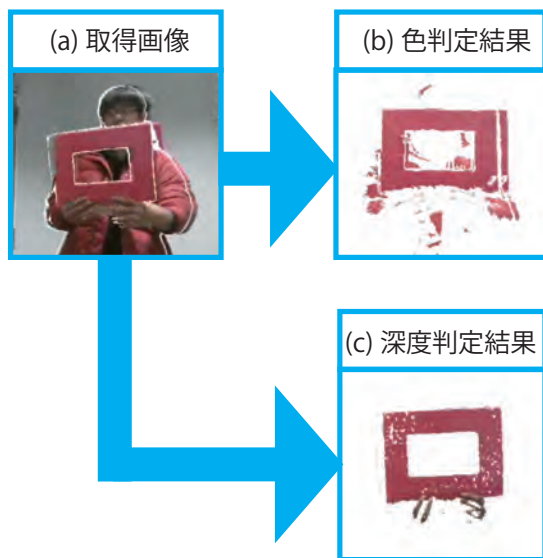


図 4 ドア認識方法

度情報とドアの奥行方向の位置を比較する。ドアの領域であると色情報から判定されているが、ドアの位置から奥行方向に離れている画素はドア領域から除外する。次に腕の領域の判定をおこなう。ドアの位置より前面にあり、身体情報から体と判定されている画素を腕と認識する。

(5) 重畳画像

(4) までの手順によって作成された、ドアと腕の画像をドア無し側の映像に重畳表示する。

4. 提案手法

4.1 従来のドアコム Z における課題

これまでに開発してきたドアコム Z においては、前述したように「ドア」を用いて遠隔地間を仮想的に接続し、ドアを用いた、ドア枠から手を出すなどのジェスチャが可能となっている。また、Kinect により取得可能な深度情報を用いて、ユーザや周辺物体の前後関係や前後への移動を認識し、前後関係を考慮した重畳表示を行うことにより、遠隔地における三次元的な移動の表現が可能となっている。このシステムにより、「相手の空間に侵入している感覚」や、「自身の空間へ侵入されている感覚」はある程度確認できているものの、得られている効果は未だ十分とは言えない。ビデオチャットにおいて存在感を向上させるための、手を用いた様々な手法が提案されている。これらの研究により、ビデオチャットにおいて手を対話相手の空間に侵入させる手法は存在感が向上することが分かっている。従来のドアコム Z では、ドア枠から手を出すことで遠隔空間内の物に対する指示や、遠隔地の物に対する接触を表現可能であった。しかし、「遠隔空間内の物を回転させる」「遠隔空間内に物を置く」という遠隔空間内の物に対する接触後

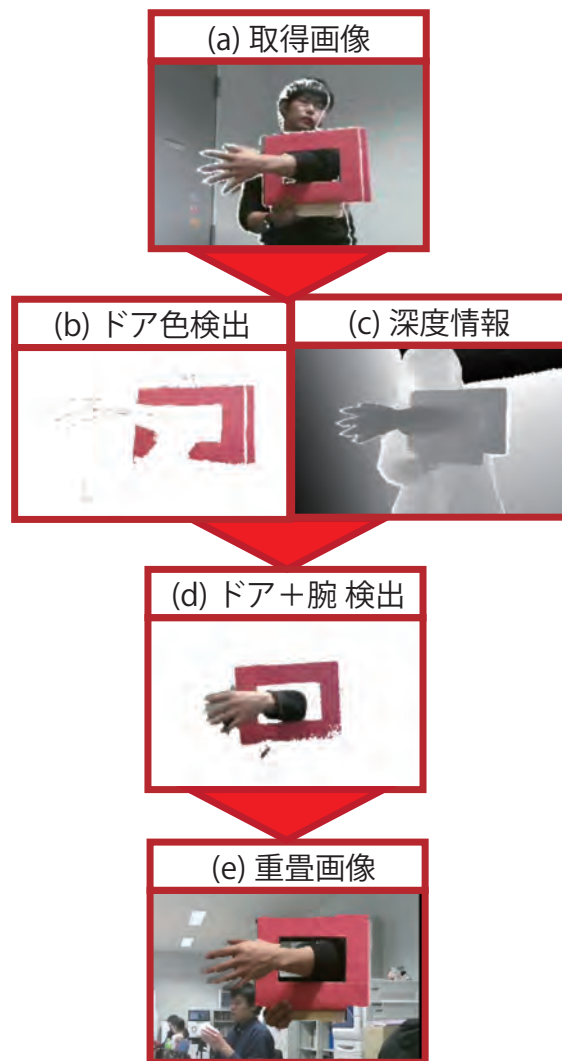


図 5 重畳画像作成手順

のインタラクションは表現できていない。

4.2 遠隔空間内における接触の表現

遠隔空間内における三次元移動の表現とドア枠から手を出す表現により、遠隔空間内の指差しが可能である。しかし、実際に空間に侵入しているのであれば、指差しだけでなく、その空間にあるモノに対する接触と移動を行うことが可能である。ドアコム Z では、Kinect により取得可能な深度情報と骨格情報を用いることにより、従来のドアコムでは実現できていなかった、「遠隔空間内の物を掴む」「遠隔空間内の物を動かす」という表現を実現する。

まず、遠隔空間内の物に対する接触の表現方法について述べる。

図 6 に提案手法による接触判定の仕組みを示す。図 6 の $d1$ は Kinect と侵入側利用者の右手との距離、 $y1$ は右手の位置を表している。Kinect から侵入側ユーザの骨格情報を取得し、右手の二次元座標を得る。右手の二次元座標を基準に二次元的な接触判定領域を定める。Kinect から侵入側

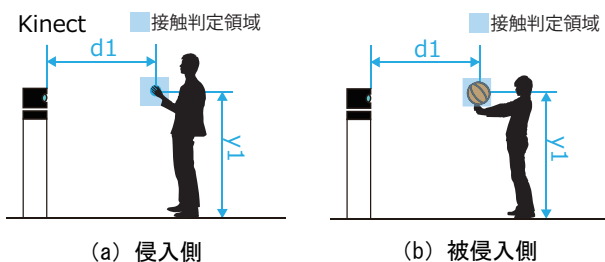


図 6 接触判定



図 7 接触表現

ユーザの右手の深度情報を取得する。右手の深度情報と二次元的な接触判定領域を用いて、三次元的な接触判定領域を作成する(図 6(a))。三次元的な接触判定領域を被侵入側に送信する。被侵入側にて、三次元的な接触判定領域にある映像を接触したと判定する(図 6(b))。図 6(b) では被侵入側ユーザが持っているボールが三次元的な接触判定領域内にあり、接触していると判定される。深度情報を用いて接触判定を行うことで、遠隔空間内のモノを映像から背景と区別して判定することができる。

次に、遠隔空間内にあるモノを動かす表現方法について述べる。遠隔空間内のモノを動かす表現には、Kinect により取得可能な骨格情報と深度情報、カラー画像を用いる。

(1) 掴み表現

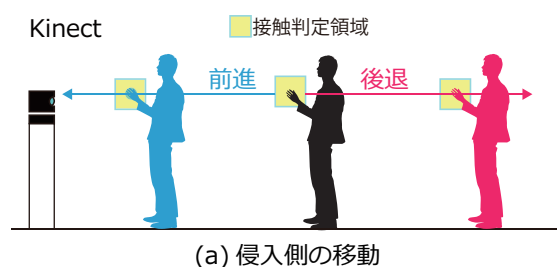
遠隔空間内のモノが接触している状態が図 7 である。接触している部分は青く表示されている。Kinect からユーザの骨格情報を取得し、右手の状態を判定する。右手が閉じている状態のときは接触状態にあるモノを掴んでいると判定する。掴んでいる状態では、掴んだ時のモノの映像を複製し保持する(図 8)。遠隔地内のモノが移動し接触状態では無くなっても、掴んでいる間は複製された映像は保持されたままである。右手が開いている状態のときは接触していても掴んでいないと判定する。掴んでいる状態から右手を開くと、保持していた映像は消える。

(2) 上下左右移動表現

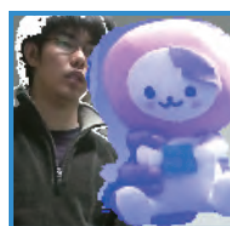
Kinect からユーザの骨格情報を取得し、右手の座標を取得する。掴んだ時の位置から右手が移動した分、



図 8 掴み表現



(a) 侵入側の移動



(b) 前進時の表示映像



(c) 後退時の表示映像

図 9 前後移動表現

保持されている映像を平面的に移動させる。これにより、遠隔空間内の掴んだモノの上下左右移動表現が可能である。

(3) 前後移動表現

図 9 に提案手法による前後移動表現の仕組みを示す。Kinect からユーザの骨格情報と深度情報を取得し、右手の深度情報を取得する。モノを掴んだ状態で Kinect に対して前後に移動したとき、変化した右手の深度情報により保持した映像の拡大縮小を行う。拡大縮小の倍率は深度情報の変化量により求める。Kinect に対して前進すると Kinect からの距離が近くなる(図 9(a))。前進したことにより深度情報が変化し、保持された映像は拡大される(9(b))。Kinect に対して後退すると Kinect からの距離が遠くなる(図 9(a))。後退したことにより深度情報が変化し、保持された映像は縮小される(9(c))。これにより、遠隔空間内の掴んだモノの前後移動表現が可能である。

(4) 回転表現

Kinect から被侵入側の深度情報、カラー画像と侵入側ユーザの骨格情報を取得する。被侵入側の掴みにより



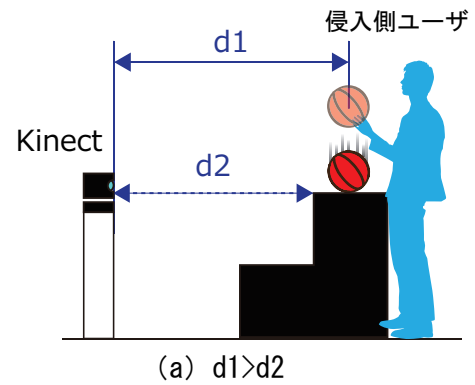
図 10 回転表現

保持されている映像のカラー画像と深度情報を組み合わせて三次元画像情報とする。侵入側ユーザの骨格情報から右手首の角度を取得する。右手首の角度に応じて三次元画像を回転させる。回転させた三次元画像を Kinect を正面とした二次元画像に変換し、画面に表示する。回転前の映像が図 10(a)、回転後の映像が図 10(b) である。これにより、遠隔空間内の掴んだモノの回転表現が可能である。

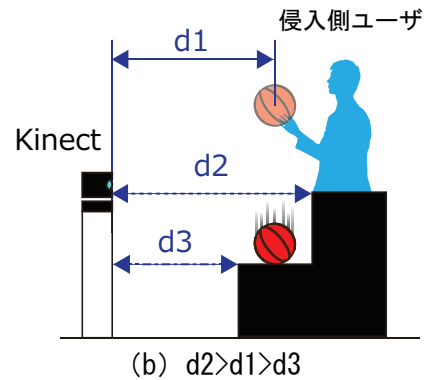
(5) 設置表現

Kinect から被侵入側にある物の深度情報、侵入側ユーザの骨格情報と深度情報を取得する。侵入側ユーザの右手の平面座標を骨格情報から取得し、その座標を用いて右手の深度値を取得する。侵入側ユーザの右手の平面座標と深度値と、被侵入側の右手と同一座標の深度値を比較して設置判定を行う。図 11 に提案手法による設置判定の仕組みを示す。図 11 の d_1 は Kinect から侵入側ユーザの右手の距離、 d_2 は被侵入側にある段の上段までの距離、 d_3 は下段までの距離を表している。侵入側ユーザによって掴まれていた被侵入側の物が放されたとき、設置判定が行われるまで下方向に落下する。 $d_1 > d_2$ の場合 (図 11(a))、落下している物は段の上段にて停止する。次に、 $d_2 > d_1 > d_3$ (図 11(b)) の場合、落下している物は段の上段の前を通り、下段にて停止する。最後に、 $d_3 > d_1$ (図 11(c)) の場合、落下している物は段の上段と下段の前を通り床にて停止する。この判定により、遠隔空間にある物を掴み、遠隔空間内に設置することが可能である。図 12 に設置表現の映像を示す。侵入側ユーザは遠隔空間にある物を掴む (図 12(a))。次に、掴んだ物を移動させる (図 12(b))。移動させた座標で、物を離すと物は落下し、設置判定された座標で停止する (図 12(c))。

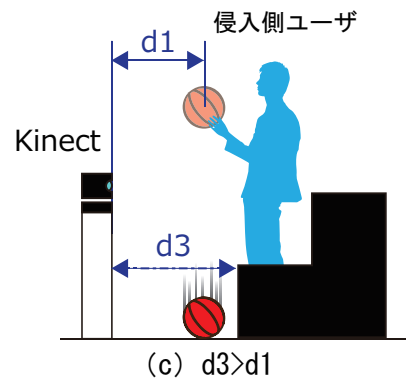
従来のビデオチャットでは不可能だった「遠隔空間内の物を掴む」「遠隔空間内の物を動かす」「遠隔空間内の物を回転させる」「遠隔空間内に物を設置する」といった表現が、本提案手法により可能となる。



(a) $d_1 > d_2$



(b) $d_2 > d_1 > d_3$



(c) $d_3 > d_1$

図 11 設置判定

5. おわりに

本研究では「自身の空間に侵入されている感覚」や「相手の空間へ侵入している感覚」を与えることが可能なビデオチャットの実現を目指し、遠隔空間内の物に対する接触後のインタラクションを提案した。提案手法により、「遠隔空間内の物を回転させる」「遠隔空間内に物を置く」という表現が可能となった。今後、遠隔空間内の物に対する接触後のインタラクションを用いた遠隔コミュニケーションについて評価実験を行う。

謝辞

本研究は JSPS 科研費 15K12085 の助成を受けたものです。



図 12 設置表現

参考文献

- [1] Skype Blogs: Skype の同時接続ユーザー数、4000 万人を達成, <http://blogs.skype.com/2012/04/13/skype4000/> (参照 2015.1.9).
- [2] 総務省:平成 24 年度版 情報通信白書, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc122110.html> (参照 2015.1.9).
- [3] Marilyn M. Mantel, Ronald M. Baecker, Abigail J. Sellen, et al. : Experiences in the Use of a Media Space, CHI '99, pp.203-208 (1991).
- [4] Yevgenia Bondareva and Don Bouwhuis: Determinants of Social Presence in Videoconferencing, AVI2004 Work-

- shop on Environments for Personalized Information Access, pp.1-9 (2004).
- [5] 森川 治: 「超鏡」: 魅力あるビデオ対話方式をめざして, 情報処理学会論文誌, Vol.41, No.3, pp.815-822 (2012).
- [6] 熊谷 真吾, 横山 牧, 佐藤 未知, 福嶋 政期, 梶本 裕之: ソファを介した「隣り合う」遠隔コミュニケーション, 情報処理学会, インタラクション 2012, pp.965-969(2012).
- [7] Kenji Hirata, Yasunori Harada,Takehiro Ohno, et al. : t-Room: Telecollaborative Room for Everyday Interaction, 情報処理学会第 66 回全国大会, 4B-3, pp.4.97-4.98 (2004).
- [8] Anthony Tang, Carman Neustaedter, Saul Greenberg: VideoArms: Embodiments for Mixed Presence Groupware, Proceedings of BCS HCI 2006, pp.85-102(2006).
- [9] 大西 裕也, 田中 一品, 中西 英之: PopArm: 身体映像の部分実体化によるソーシャルテレプレゼンスの強化, 情報処理学会, インタラクション 2015 論文集, pp.38-46(2015).
- [10] Igor de Souza Almeida, Marina Atsumi Oikawa, Jordi Polo Carres, Jun Miyazaki, Hirokazu Kato : AR-based video-mediated communication : A social presence enhancing experience, 14th Symposium on Virtual and Augmented Reality, pp.125-130(2012).