

特別研究報告書

バーチャルスタジオにおける  
仮想物体把持を支援する視聴覚情報の提示

指導教員 美濃 導彦 教授

京都大学工学部情報学科

高橋 康輔

平成 22 年 2 月 3 日

## バーチャルスタジオにおける 仮想物体把持を支援する視聴覚情報の提示

高橋 康輔

### 内容梗概

バーチャルスタジオでは、実写映像にあらかじめ作成されたCGを組み合わせて、演者があたかも仮想世界で演技を行っているかのような合成映像を作成することができる。その合成映像はスタジオ内に設置されたフロアモニタを通して演者に提示され、演者はその映像を参照することで自身がどのように映っているかを確認しながら演技を行う。

そのような合成映像において、演者がCGで作成された仮想物体を直接操作してその説明を行うような、仮想物体とのインタラクションを行うことができれば、より表現力の高い映像の作成が可能となり、有用であると考えられる。そこで、本研究はバーチャルスタジオにおいて演者が仮想物体を直接操作できるシステムの構築を目指す。

このようなインタラクションを行う際、実世界の演者の振る舞いを仮想世界に反映させる必要がある。その方法として尾原らは演者の手と同じ振る舞いを行う仮想手を導入し、演者による仮想物体の直接把持を実現した。このとき、演者の手に三次元センサとデータグローブを装着させることで手の三次元位置と姿勢を獲得し、その値を基に仮想手の仮想物体に対する把持を判定した。把持が成功した場合は演者の手の動きに合わせて仮想物体を移動させ、その様子はフロアモニタに表示した合成映像を通して演者に提示した。

しかし、演者は把持をしようと意図しても、演者が差し出している手の三次元位置と仮想物体の三次元位置が異なり、把持に失敗するということが少なくなかった。このような失敗は不自然であり、視聴者に違和感を与える可能性があるため映像コンテンツとして望ましくない。そこで、本研究は仮想物体の把持を行う際に演者が把持に失敗すると言った不自然な挙動を行うことなく把持を与えるように支援することを目的とする。

このような失敗は、演者は仮想物体を把持する際に二次元の合成映像を提示するフロアモニタを参照したとき、合成映像から仮想物体の正確な奥行きを獲得することが困難であることが原因と考えられる。つまり、演者はフロアモニタの提示する合成映像からは仮想物体を把持する上で十分な情報を獲得できていな

いと言える。このことから、フロアモニタの提示する視覚情報に加えて演者に対して何らかの手段で新たな情報提示を行うことが必要であると考えられる。本研究では、その手段としてフロアモニタによる別の視覚情報の提示と、イヤホンによる聴覚情報の提示を提案する。

仮想物体の把持の際に演者が獲得すべき情報は、我々が普段把持を行う際に利用している情報に照らして考えたときに、演者の視点にたったとき仮想物体がどこにあるかという演者に対する仮想物体の相対的な三次元位置であると言える。フロアモニタの提示する合成映像には演者と仮想物体の両方が映っているため、演者は合成映像から奥行きを除く仮想物体の相対的な位置を獲得することができると考えられる。そのため、本研究では演者に対する仮想物体の奥行き、そして奥行きを含む仮想物体の三次元位置そのものの提示を考える。

仮想物体の奥行きを提示するため、視覚では演者と仮想物体の相対的な位置関係を反映する別視点の映像の提示を考える。演者はこの映像から奥行きを知覚し、合成映像の提示する位置と合わせて仮想物体の三次元位置を獲得することができる。聴覚では音を用いて演者を適切な場所まで誘導することを考える。演者を仮想物体を把持しやすい奥行きまで誘導することで、演者は物体の奥行きを知覚できると考えられる。仮想物体の三次元位置の提示として、仮想物体と同じ位置に仮想音源を設置することを考える。我々は音が耳に到達する時に発生する差異から音源の方向や距離などの情報を取得している。このことから、仮想物体の位置に音源がある場合に演者の位置で聴取する音を再現することで、その音源の距離や方向を獲得できることが期待できる。

実験では、これらの視聴覚情報による支援の効果を検証するために、バーチャルスタジオにおいて仮想物体の把持を行う際に、合成映像に加えてこれらの情報を提示することで把持の支援を行った。被験者には提示された情報をもとに仮想物体が存在すると推測する位置で把持の姿勢をとるというタスクを課した。このとき、被験者が仮想物体を把持したと認識した時点での仮想物体と被験者の手の距離と奥行きのずれを計測した。結果からは、合成映像のみの場合と比べて、いずれの情報も提示した場合に仮想物体と被験者の手の距離と奥行きのずれが小さくなっていた。その中でも別の視点からの映像を提示する場合と、音によって演者を仮想物体を把持しやすい奥行きまで案内する場合の効果が高かった。このことから、上で述べた視聴覚情報が演者による仮想物体の把持において有用であることが確認できた。

## Display Visual and Auditory Information to Support Manipulation of Virtual Object in Virtual Studio

Kosuke TAKAHASHI

### Abstract

In the virtual studio, we can make synthetic videos by overlaying image sequences of actors captured by a studio camera on a background image created by CG. The actors perform as if they are in the virtual world. For helping such performance, the synthetic video is displayed on a floor monitor in the virtual studio.

We can enhance the quality of the synthetic videos by making it possible for the actor to interact with a virtual object. In this paper, we aim to construct the system that enables the actors to directly manipulate a virtual object in the virtual studio.

The actor is not able to directly see and grasp the virtual object with his/her hand without no support, because it does not exist in the real world. Therefore it is necessary to bridge the gap between the real world and the virtual world by introducing some support system into the virtual studio.

Obara, et al.[3] proposed a system to bridge the gap and to enable the actor to grasp the virtual object. The system captures the data such as the posture and the 3D positions of the actor's hand, and the system judges whether the hand touches the virtual object or not. When the system judges that the hand touches the virtual object, the system moves the virtual object as if the actor directly grasps and moves it, and the actor can recognize it through the synthetic video displayed on the floor monitor.

Although the system enables the actor to manipulate the virtual object, the actor sometimes fails to grasp the virtual object. The video with such failures is inappropriate as the contents because the fault might cause a sense of incongruity to viewers.

One of the reasons of the failure is that the actor cannot recognize the precise position of the virtual object from the synthetic video on the floor monitor. Therefore, we try to add the new way to display the information to the actor. In this paper, we introduce a wireless ear phone, to provide additional information

to the actor for manipulating virtual object.

When we grasp an object, we usually recognize the relative 3D positions of it before grasping. The actor can recognize relative 2D positions of the virtual object through the synthetic video that captures both the actor and the virtual object. Therefore, we attempt that the system shows the actor the depth of the virtual object.

In order to let the actor recognize the position of the virtual object with the information, we propose three methods of additional information to the actor. First method is to display a map from a different point of view with the studio camera to him/her (Visual Map). The actor can recognize the 3D positions of the virtual object by watching this video and the synthetic video. Second method is to guide the actor to a place which is the same depth of the virtual object by displaying auditory information based on the relative 3D positions (Auditory Guidance). Third method is to set a virtual sound source on the 3D positions of virtual object (Auditory Positioning). We can acquire sound source's positional information such as its direction and distance from the sound.

In experiment, in order to evaluate that these visual and auditory information support the actor, we displayed one of these information and the synthetic video to the actors and we made them to grasp a virtual object in virtual studio. We imposed the task on the actor that he/she poses as if he/she grasps the virtual object based on the information. When the actor grasps it, we measured the distance from his/her hand to the virtual object. As the result of the experiment, we confirmed that these information supported the actor. In particular, Visual Map and Auditory Guidance were more effective than the other methods.

# バーチャルスタジオにおける 仮想物体把持を支援する視聴覚情報の提示

## 目次

第1章	序論	1
第2章	バーチャルスタジオにおける仮想物体とのインタラクション	3
2.1	バーチャルスタジオ	3
2.2	バーチャルスタジオにおける仮想物体操作	4
第3章	インタラクティブバーチャルスタジオにおける視聴覚情報の提示	7
3.1	バーチャルスタジオへの演者と仮想物体のインタラクションの導入	7
3.2	インタラクティブバーチャルスタジオにおいて演者が獲得すべき情報	9
3.2.1	バーチャルスタジオにおいて演者が獲得すべき情報	9
3.2.2	仮想物体を把持する際に演者が獲得すべき情報	10
3.3	情報を伝達するメディアの特性	10
3.3.1	視覚メディア	10
3.3.2	聴覚メディア	11
3.4	視聴覚情報による情報の提示	12
3.5	視聴覚情報の具体的なデザイン	13
3.5.1	視覚情報による奥行き情報のデザイン	13
3.5.2	聴覚情報による奥行き情報のデザイン	14
3.5.3	聴覚情報による三次元位置情報のデザイン	16
第4章	実験	17
4.1	把持を行う手の奥行きのばらつきを検証	17
4.1.1	実験の目的	17
4.1.2	実験の方法	17
4.1.3	実験手法	18
4.1.4	結果と考察	19
4.2	視聴覚情報の提示による仮想物体把持への影響の検証	20
4.2.1	実験の目的	20

4.2.2	実験の方法 . . . . .	20
4.2.3	結果 . . . . .	22
4.2.4	考察 . . . . .	23
第 5 章	結論	24
	謝辞	25
	参考文献	25

## 第1章 序論

ニュースや気象予報、教育番組といったテレビ番組において、スタジオのセットにコンピュータグラフィックス (CG:Computer Graphics) を用いた映像コンテンツが頻繁に制作されている。CGを用いることで、スタジオのセットとして用意できない物体を映像に登場させることができ、またその色や形を変化させるといった特殊な変化も容易に実現できる。このような映像をリアルタイムで作成するためにバーチャルスタジオというシステムが利用されている。バーチャルスタジオでは、壁や床が特定色 (青色や緑色) で統一されたスタジオで演技を行う演者の実写映像から特定色の領域を背景とみなすことで、実写映像を演者領域と背景領域に切り分け、その背景領域にあらかじめ作成しておいたCGの仮想背景や仮想物体を組み合わせることで合成映像を作成する。

そのような映像において、演者がCGで作成された仮想物体を直接操作するような、仮想物体とのインタラクションを行うことで、より表現力の高い映像の作成が可能となり、有用である。例えば教育番組において、教師である演者が説明したい場所を視聴者に見えやすくするために、教師自身の手で仮想物体を直接操作することができれば、視聴者にとってわかりやすい教育映像を作成することができる。しかし、バーチャルスタジオでは実写の映像に仮想背景や仮想物体を合成することしか実現されておらず、演者による仮想物体の直接操作は実現されていなかった。

演者による仮想物体の直接操作に関する研究として尾原ら [3] の研究がある。尾原らは物体を手で把持して移動・回転するという直接操作を実現している。仮想物体は実世界には存在しないため、演者はその仮想物体を直接把持することはできない。そこで、尾原らは演者が仮想物体を把持している状況を、演者の手が仮想物体の表面付近で把持の姿勢を取った時に把持を行っているとして定義した。このとき、演者に三次元センサとデータグローブを演者に装着させることで手の三次元位置と姿勢を検出し、その値をもとに演者が仮想物体を把持している状況を判定した。把持の検出後には、演者の手の動きに合わせて仮想物体も動くといったような、仮想物体の把持による操作を実現した。

仮想物体を把持する際、演者は仮想物体を直接視認することができないため、演者は従来のバーチャルスタジオと同様にフロアモニタを参照することで仮想物体の位置を確認しながら演技を行う。フロアモニタに映し出される映像はカ



メラ視点の映像であるため、カメラの光軸方向に対する物体の位置、つまりフロアモニタに提示される映像に対して垂直な方向の物体の位置によって映像中の物体の大きさが変化する。以下ではこの位置のことを奥行きと呼ぶこととする。しかし、奥行きが大きい場合には奥行きの変化による映像上の物体の大きさの変化も微小であり、合成映像中の仮想物体の大きさからその奥行きを推測することが困難である。そのため、演者が合成映像から仮想物体の正確な三次元位置を獲得することは難しい。この結果、把持の際に仮想物体の存在する三次元位置とは異なる位置に手を差し出してしまい、把持に失敗するという例が度々見受けられる。

演者が仮想世界にいるような映像において、演者がその世界にある物体の把持に失敗するという挙動は映像として不自然であり、視聴者に違和感を与え得るために映像コンテンツとして望ましくない。そこで本研究は、仮想物体の把持を行う際に演者が把持に失敗するといった不自然な挙動を行うことなく把持を行えるように支援することを目的とする。

上で述べた例からわかるように、把持に失敗するなどの不自然な挙動の理由として、フロアモニタの提示する視覚情報のみでは演者は仮想物体の三次元位置を十分に獲得できていないことが挙げられる。ゆえに、フロアモニタの提示する視覚情報に加えて演者に対して何らかの手段で新たな情報提示を行うことが必要であると考えられる。本研究では演者に対して新たに提示する情報として視覚情報の他に聴覚情報の利用を検討する。聴覚情報を提示する手段としてイヤホンを使用する。イヤホンはテレビ番組において出演者も装着している場面も見受けられるように、合成映像に映りこんでも不自然ではないと言える。加えて、演者の三次元位置によって出力する聴覚情報を変化させることで、聴覚情報に位置情報を持たせることが可能である。本稿ではイヤホンをを用い、視覚からは得られない仮想物体の奥行き、もしくは仮想物体の三次元位置そのものを提示することを提案する。また、仮想物体把持を行う際にそれらの情報を提示し、把持をした時の手の奥行きと仮想物体の奥行きとの差、手と仮想物体の重心との距離を計測し、それらの情報を提示することの有用性を評価した。

本稿の章構成は以下のとおりである。まず、第2章ではバーチャルスタジオとインタラクティブバーチャルスタジオについて述べ、仮想物体把持に関する従来研究とその問題点について述べる。続く第3章ではインタラクティブバーチャルスタジオにおいて演者が獲得すべき情報について述べ、視覚と聴覚を利用し

た情報の提示方法について検討する。第4章では第3章で挙げた提示方法に関する実験を行い、その結果に対する評価を行うことで提示方法の有用性を検討する。最後に第5章では結論と今後の課題について述べる。

## 第2章 バーチャルスタジオにおける仮想物体とのインタラクション

### 2.1 バーチャルスタジオ

バーチャルスタジオとは、現実世界とCGであらわされる仮想世界との合成映像を作成するためのシステムである。合成映像がバーチャルスタジオによって作成される際、撮影は図1のように、ある特定色（通常は青色か緑色）で統一された壁や床で覆われたスタジオで行われる。このスタジオで演者を撮影した実写映像からクロマキー合成装置で特定色（スタジオの壁や床の単一色）の領域を取り除くことにより、演者領域のみを切り出した実写映像が生成される。一方、スタジオカメラにはセンサが内蔵されており、これによって得られるスタジオカメラの位置や方向、ズーム値等のカメラパラメータに基づいて、仮想背景描写用コンピュータでスタジオカメラと同じカメラパラメータで撮影された仮想背景の映像が生成される。このようにして、生成された演者領域の実写映像と仮想背景の映像を映像合成装置で合成することで、図2のように演者があたかも仮想世界で演技をしているような映像が作成される。作成された合成映像はリアルタイムに視聴者に配信されるだけでなく、スタジオ内のカメラに撮影されない場所に設置されたフロアモニタにも映し出され、演者はこのフロアモニタを通して自身がどのように映っているかを確認しながら演技を行う。



図1: バーチャルスタジオ



図2: 作成される映像

バーチャルスタジオでは作成される映像にCGで作成された仮想背景や仮想物体を組み合わせることができ、本来ならばコストや物理的な大きさが原因で撮影することができないような背景や物体を利用することができる。さらに、仮想背景や仮想物体はCGで作成されているので、その色や形を変化させるといった特殊効果を容易に付与することが可能であり、映像コンテンツとしての表現力が高まると考えられる。

## 2.2 バーチャルスタジオにおける仮想物体操作

バーチャルスタジオで作成される映像において、演者がCGで作成された仮想物体と直接操作のようなインタラクションを行うことで、より表現力の高い映像の作成が可能になると考えられる。例えば教育番組において、教師である演者が説明したい場所を視聴者に見えやすくするために、教師自身の手で仮想物体を直接操作することができれば、視聴者にとってわかりやすい教育映像を作成することができる。しかし、バーチャルスタジオでは実写の映像に仮想背景や仮想物体を合成することしか実現されておらず、演者による仮想物体の直接操作は実現されていなかった。

本研究では、教師である演者が三次元の仮想物体を演者自身の手で操作し、自由に動かすことのできるようなシステムの構築を考える。このようなシステムが構築できれば、演者と仮想物体がインタラクションを行っているような合成映像をバーチャルスタジオで作成することができる。このように、演者が仮想物体とインタラクションを行うようなシステムを導入したバーチャルスタジオのことをインタラクティブバーチャルスタジオと呼ぶ。

仮想物体は実世界には存在しないため、実世界に存在する演者が仮想世界に存在する仮想物体とインタラクションを行うためには、何らかの方法で実世界と仮想世界をつなぐ必要がある。そこで、バーチャルスタジオにおける演者による仮想物体の直接操作に関する従来研究を、実世界と仮想世界をつなぐ方法の違いで二つに分類することができる。

一つ目は、演者が操作する現実物体を仮想物体に置き換えるアプローチである(図3の(1))。つまり、現実物体と仮想物体が実世界と仮想世界をつなぐ媒介となる。演者の操作対象となる仮想物体と同じ形状の現実物体を用いて撮影を行い、得られた実写映像中の現実物体を仮想物体に置き換えるという方法である。このアプローチに関する従来研究として大島ら[2]の研究がある。大島らは

操作対象の現実物体として球を用いた。大島らは演者が操作する球にマーカを張り、そのマーカ位置に基づいて球の姿勢を検出することで、仮想物体に演者が操作する球と同じ動きをさせている。演者は仮想物体と同形状の現実物体である球を操作しているため、球の位置・姿勢を目視で確認することで仮想物体の位置・姿勢を把握することができる。しかし、この手段では仮想物体の動きは現実物体の動きに制限されるため、物体の色を変えることなどはできても、例えば物体が突然消滅したり急激に大きさが変わるなどの現実物体に起こり得ない挙動は仮想物体においても実現できない。

二つ目は、演者の手を仮想手に置き換えて仮想物体操作を行うアプローチである(図3の(2))。つまり、演者の手と仮想手が現実世界と仮想世界とつなぐ媒介となる。仮想物体を操作する際に演者の手の情報をデバイスを用いて取得し、その手の情報をもとに仮想手を作成して仮想物体に対して操作を実現するという手法である。このアプローチに関する従来研究として尾原ら[3]の研究がある。

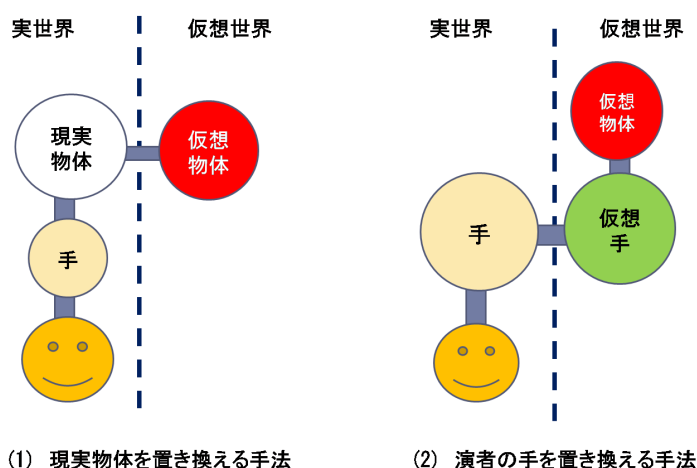


図3: 従来研究の比較

尾原らは演者の手に三次元センサとデータグローブを装着させることで手の三次元位置と姿勢を獲得し、その値を基に演者の手を仮想手に置き換えることで演者による仮想物体把持を実現した。仮想手が仮想物体を把持している時、つまり演者の手が仮想物体の表面付近で把持の姿勢を取った時に把持を行っているとした。このようにして、演者は仮想物体に対して把持を行うことが可能であり、把持された物体は演者の手の動きに合わせて移動をする。把持を行う際に演

者は仮想物体の表面の位置に手を差し出す必要があるが、仮想物体は実世界には存在しないため、演者はその位置を直接視認することはできない。そこで、演者は従来のバーチャルスタジオと同様にフロアモニタを参照することで仮想物体の位置を確認しながら演技を行う。

このときフロアモニタに映し出される映像はカメラ視点の映像であるため、物体の奥行きによって映像中の物体の大きさが変化する。このため、仮想物体の正確な大きさを把握していれば仮想物体の奥行きを推測することが可能である。しかし、奥行きが大きい場合には奥行きの変化による映像上の物体の大きさの変化も微少であり、合成映像中の仮想物体の大きさからその奥行きを推測することが困難である。そのため、演者が合成映像から仮想物体の正確な三次元位置を獲得することは難しい。この結果、把持の際に仮想物体の存在する三次元位置とは異なる位置に手を差し出してしまい、把持に失敗するという例が度々見受けられる。

このように、どちらの従来研究も十分に仮想物体の把持を実現しているとは言えない。バーチャルスタジオで仮想物体を使用することで、現実物体では用意できないような物体まで扱えることが利点であることを前節でも述べた。このことから、大島らのアプローチのように現実物体を仮想物体に置き換える場合、仮想物体の形状が現実物体の形状と同じものに制限されるため、物体を自由に變形できるというCGならではの表現力を活かしきれていないと言える。そのため、本研究では尾原らのアプローチのように、演者の手を仮想手に置き換えるアプローチに焦点を当てる。

このアプローチには上で述べたように把持の失敗が発生するという問題がある。合成映像上では演者はあたかも仮想世界にいるかのように映っているにもかかわらず、その世界にある物体の把持に失敗することは視聴者に違和感を抱かせる可能性があり、映像コンテンツとして望ましくない。そこで本研究は、仮想物体の把持の際に演者が把持に失敗するといった不自然な挙動を行うことなく把持を行えるように支援情報を提示する方法を提案する。

## 第3章 インタラクティブバーチャルスタジオにおける視聴覚情報の提示

### 3.1 バーチャルスタジオへの演者と仮想物体のインタラクションの導入

インタラクションとは、図4のように一方からのアクションに対してもう一方のリアクションが返ってくる、つまり双方向のアクションが成立することである。

インタラクティブバーチャルスタジオでは演者と仮想物体のインタラクションを実現する。つまり、図4で言うAを演者とし、Bを仮想物体として見る。仮想物体は実世界には存在せず、演者と仮想物体はお互いが直接アクションを行うことはできない。そのため、それらの双方向のアクションを成立させるため、従来のバーチャルスタジオに加えてお互いのアクションを他方へ伝達する機能を備える必要がある。ここで、図5のように演者から仮想物体に対するアクションを取得する機能を取得機能とし、逆に仮想物体から演者に対するアクションを提示する機能を提示機能とする。

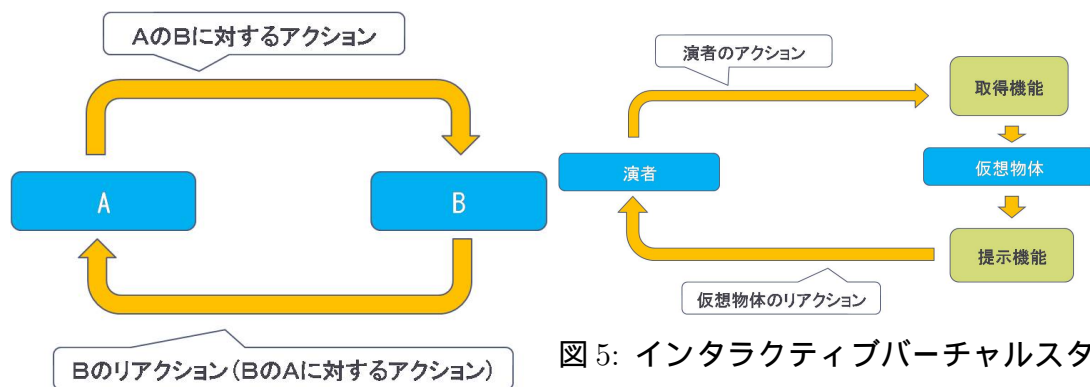


図4: インタラクション

図5: インタラクティブバーチャルスタジオ

このとき、従来のバーチャルスタジオは図6のように見ることができる。従来のバーチャルスタジオでは、実写の演者の映像がカメラを通して取得され、その映像は合成機能へ与えられる。合成機能はその映像とあらかじめ作成された仮想物体の映像を組み合わせることで合成映像を作成し、その映像をフロアモニタを通して演者に提示している。このことから、この合成機能は上で述べた提示

機能として見ることができる。このとき合成機能は仮想世界に対して変化を及ぼすものではないことから、合成機能は取得機能として働くことはない。そのため従来のバーチャルスタジオでは演者から仮想物体に対するアクションは伝達されず、演者と仮想物体はインタラクションを行っていないと考えられる。

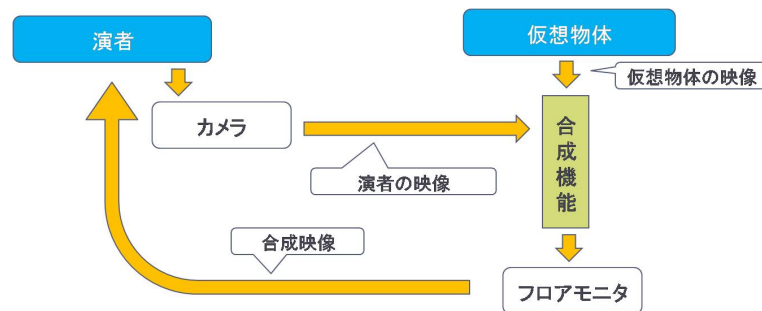


図 6: 従来のバーチャルスタジオ

ここで先ほど挙げた尾原ら [3] の研究を例として、具体的なインタラクティブバーチャルスタジオについて考える。尾原らの提案したシステムは図 7 のようにあらわされる。尾原らは従来のバーチャルスタジオに加えて、演者と仮想物体の間に把持検出機能を導入した。把持検出機能は演者の手が仮想物体の表面付近で把持を行う姿勢を取っているかどうかを判定する機能である。尾原らは演者の手に三次元センサとデータグローブを装着させることで、それらの値を入力として把持検出機能が仮想物体に対する把持を判定する。演者は把持検出機能を介することで仮想物体に対するアクションを行っていることから、把持検出機能を取得機能として見るることができる。演者の把持というアクションに対して、仮想物体が把持されていれば演者の手の移動に合わせて仮想物体も位置を変化させ、把持されていなければ位置を変化させないというリアクションを合成映像として演者に提示する。

演者がインタラクティブバーチャルスタジオで演技をすると、システムはそれに対して演者の装着している三次元センサやデータグローブから取得した情報を基に仮想物体に対して変化を加え、その仮想物体とカメラから取得した映像を合成してフロアモニタに提示している。演者はその合成映像を通して自身の仮想物体へのアクションの影響がどのように出ているかを確認し、その仮想世界のリアクションによって演者は次取るべきアクションを決定する。このように、尾原らのシステムでは取得機能として把持検出機能を、提示機能として

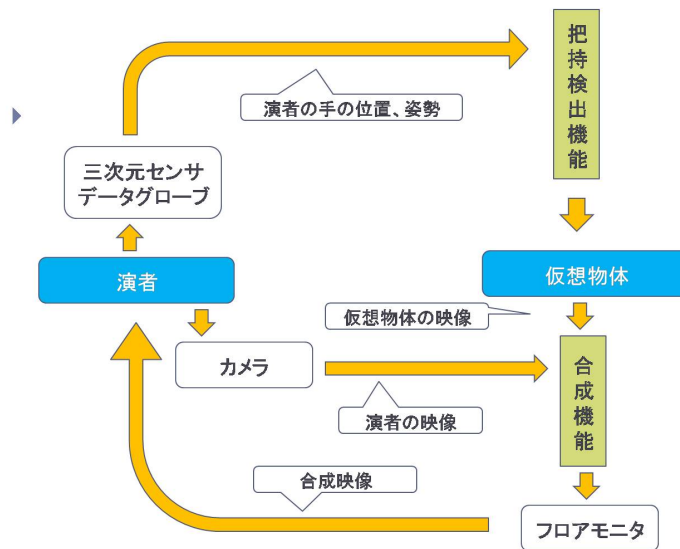


図 7: 尾原らの提案したインタラクティブバーチャルスタジオ

合成機能を利用することで演者は仮想物体とインタラクションを行っている。

しかし、一方で仮想物体を掴みに行った手が把持に失敗してしまうような問題も発生する。これは 2.2 でも述べたように、合成機能の提示する合成映像のみでは演者は仮想物体の三次元位置を十分に獲得できていないことが原因と考えられる。つまり、仮想世界からのリアクションを伝える提示機能が合成機能のみでは不十分であると考えられる。次節以降ではこの三次元位置をはじめとし、演者がバーチャルスタジオで仮想物体把持を行うために獲得すべき情報を整理し、その情報を提示する手段を提案する。

### 3.2 インタラクティブバーチャルスタジオにおいて演者が獲得すべき情報

仮想物体把持を実現するインタラクティブバーチャルスタジオは、従来のバーチャルスタジオのシステムに拡張として仮想物体の把持を実現するシステムを追加することで実現する。このため、演者は従来のバーチャルスタジオで演技をする時に獲得すべき情報と、仮想物体を把持する際に獲得すべき情報を獲得する必要がある。

#### 3.2.1 バーチャルスタジオにおいて演者が獲得すべき情報

バーチャルスタジオで合成映像を作成する際、演者は青色や緑色といった特定色で統一された床や壁に覆われたスタジオで演技を行う。通常のテレビ番組



のセットにおいて演技をする場合と異なり、演者は自身がどのような環境で演技をしているかを直接視認することはできない。そのため、自身が合成映像としてどのように映っているかが分からず、仮想背景にそぐわない不自然な演技を行う可能性がある。バーチャルスタジオの主目的は演者にあたかも仮想世界にいるような感覚で演技を行うことを支援することではなく、演者が仮想世界で演技をしているような合成映像を作成し、それを視聴者に配信することである。これらのことから、演者は自身が合成映像においてどのように映っているかといった、視聴者視点での情報を取得することが必要であると考えられる。

従来のバーチャルスタジオでは視聴者に配信する合成映像を演者に提示することで、これらの条件を満たす情報を提示できていると言える。したがって、本研究でもバーチャルスタジオで獲得すべき情報はフロアモニタを通して合成映像を提示することで実現する。

### 3.2.2 仮想物体を把持する際に演者が獲得すべき情報

物体の把持を行う場合、通常我々は物体の三次元位置を知覚することで、その位置に対して手を伸ばすことにより把持を行う。しかし、演者は仮想物体を直接視認することはできないため、何らかの方法で仮想物体の三次元位置を知覚する必要がある。例えば仮想物体の三次元位置情報を演者に提示する方法として、仮想物体の実空間中の三次元座標の値を演者のに伝える方法が考えられるが、この方法は普段空間の座標系を意識しない我々にとって直感的でなく望ましくない。我々が普段把持を行う際に利用しているのは仮想物体の三次元座標といった絶対的な三次元位置情報ではなく、演者の視点にたったときに仮想物体がどこにあるかという、演者に対する仮想物体の相対的な三次元位置情報である。ゆえに、演者が必要とする仮想物体の三次元位置情報は、演者視点の相対的なものとして提示されることが望ましい。

## 3.3 情報を伝達するメディアの特性

### 3.3.1 視覚メディア

日常的に人間が五感によってインプットする情報量の割合は視覚が83%を占め、五感の中でも圧倒的に取得する情報量が多い[1]。

従来のバーチャルスタジオでは視覚メディアの提示デバイスとしてフロアモニタが利用されている。フロアモニタはスタジオ内に設置されたモニタのことである。フロアモニタは通常カメラの横など撮影されない領域に設置されるた

め、合成映像に映りこむことはない。本研究では視覚情報を提示するメディアとして、このフロアモニタを使用する。

このフロアモニタを通して仮想世界に関する静止画像や映像、文字などを提示することが可能である。しかし、例えば仮想物体の三次元位置情報に関して、演者視点である仮想物体との距離を文字で提示しても直感的でない。このため、本研究ではフロアモニタを通して提示する視覚情報として映像を用いる。

また、一方でフロアモニタには演者が視線を向けなければ情報を取得できないという制約がある。さらに、演者が直接見ているスタジオは特定色（青色や緑色）で統一されたものであり、実際の仮想世界を視認しているわけではない。このとき演者が仮想世界を直接視認しているように、どのような方向を向いていても常に取得することができれば、より直感的であり自然な演技を行うことが可能であると考えられる。

このような視覚情報を提示できる視覚メディアのデバイスとしてヘッドマウントディスプレイがある。ヘッドマウントディスプレイとは、ユーザの頭部に装着する視覚提示装置である。ヘッドマウントディスプレイを利用することでユーザの見る現実世界の映像に仮想世界の映像を重畳して表示することができ、演者視点の仮想世界の情報が提示できると考えられる。しかし、演者がヘッドマウントディスプレイのような視聴者にとって見慣れない大規模なデバイスを装着していると、映像を見る視聴者に対して違和感を抱かせる可能性がある。そのため、ヘッドマウントディスプレイのような合成映像に映りこみ、かつ大規模なデバイスの使用は控えるべきである。

フロアモニタではそのような情報の提示方法は困難であることから、以下では視覚以外のメディアとして聴覚メディアの検討を行う。

### 3.3.2 聴覚メディア

視覚では演者視点での情報提示が困難であることから、視覚とは異なる情報提示の方法として、聴覚の利用を検討する。聴覚を通して得られる情報量は、日常的に人間が五感によってインプットする情報量の割合において11%を占め、五感の中では視覚に次いで二番目に情報量が多い[1]。

聴覚情報を提示する手段としてはスピーカーやイヤホンが考えられる。スピーカーなどはフロアモニタと同様にカメラの横などに設置すれば合成映像に映りこむこともないが、バーチャルスタジオにおいて発生している音は演者の声とともにスタジオのマイクによって録音されるので、その音は合成映像にも反映

されてしまう可能性がある。そのため、視聴者が聴取しても不自然でない程度の聴覚情報しか出力出来ず、提示できる聴覚情報の自由度が低いために支援を行う手法としては望ましくない。

それに対し、イヤホンはその聴覚情報を演者のみに提示できるので、聴覚情報が合成映像に反映される問題は発生しない。また、イヤホンは演者に装着するデバイスであるために合成映像に映りこむ。しかし、近年のテレビ番組において出演者たちがイヤホンを装着している場面が度々見受けられるため、合成映像に映りこんでも視聴者はさほど違和感を抱かないと考えられる。そのため、本研究では聴覚情報を提示するデバイスとしてイヤホンを使用する。

本研究ではイヤホンを通して提示する聴覚情報として、電話のベルや警告音に用いられているような信号音を利用する。和氣ら [4] は音を情報と言う観点から次のように分類している (表 1)

表 1: 日常音の分類

音	音の説明		
自然音	意図的に発せられるのではない音全般		
メッセージ音	音声	声	鳴き声, 笑い声, 動物の鳴き声など
		言語音声	言語を発話したもの 明確な意味を持つ
	非音声	信号音	電話のベル, 警告音など
		音楽	芸術, 音楽体系を有する

今回のように音で情報を表現する場合、意図的に発せられることのない自然音は適さない。さらに、メッセージ音において、例えば教育番組やニュースなどで演者は説明をしながら演技を行うために、音声などで情報を提示することは演技を阻害する恐れがある。そのため、本研究では演者に情報を提示する音として非音声の音を使用する。その中でも日常的に警告音や電話のベルなどで、“ある情報を聞き手に伝達する音”として利用されていることから信号音を利用することとする。

### 3.4 視聴覚情報による情報の提示

3.2.1 では演者がバーチャルスタジオで獲得すべき情報に対して、本研究はフロアモニタによって合成映像を提示することを述べた。以下では仮想物体の把

持の際に演者が獲得すべき情報について述べる。

3.2.2 では演者が仮想物体を把持する際に獲得すべき情報として仮想物体の三次元位置情報を挙げ、かつその位置情報は演者視点の相対的なものであることが望ましいと述べた。3.2.1 で述べたとおり、演者はフロアモニタを通して合成映像と言う視覚情報を既に得ているものとして考える。二次元の合成映像の提示する情報からは正確な奥行きを獲得することは困難であるが、合成映像には演者と仮想物体の両方が映っているため、奥行きを除く獲得できる二次元の位置情報は演者にとって相対的な位置情報となる。そのため、演者が仮想物体把持を行う際に不足している情報として、残りの一次元情報である仮想物体の奥行きの提示を考え、さらに奥行きを含む仮想物体の三次元位置そのものの提示を考える。

- 視覚情報による奥行き情報提示

フロアモニタの提示する合成映像に加えて新たに提示する視覚情報として、演者と仮想物体の奥行きの相対的な位置関係を反映する別視点の映像の提示を考える。演者はこの映像から奥行きを獲得し、合成映像の提示する位置情報と合わせて仮想物体の三次元位置を獲得することができると期待できる。

- 聴覚情報による奥行き情報提示

聴覚情報として、演者と仮想物体の奥行きに対して出力する音を変化させ、演者を仮想物体を把持しやすい奥行きまで誘導すること考える。演者は仮想物体を把持しやすい奥行きまで誘導されることで仮想物体の奥行きを獲得し、合成映像の提示する位置情報と合わせて仮想物体の三次元位置を獲得することができると期待できる。

- 聴覚情報による三次元位置情報提示

聴覚情報として、演者と仮想物体の三次元位置関係に対応して出力する音を変化させることを考える。演者はこの音から仮想物体の三次元位置を獲得することが期待できる。

### 3.5 視聴覚情報の具体的なデザイン

#### 3.5.1 視覚情報による奥行き情報のデザイン

3.4 で述べた視覚情報による奥行き情報提示の具体的なデザインに関して述べる。この映像のデザインを考えたとき、我々が普段から見慣れている別視点の映

像としてカーナビゲーションシステムのような上からの視点の映像が挙げられる。バーチャルスタジオでは演技をしながら仮想物体の位置を獲得することから、提示される視覚情報は演者にとって見慣れたものが望ましいと考えられる。そこで、本研究では図8のように演者と仮想物体を上から見た場合の位置関係をプロットした映像を提示する。この映像はフロアモニタを通して合成映像の横に提示されるものとし、この映像を作成する機能を奥行き可視化機能と呼ぶこととする。

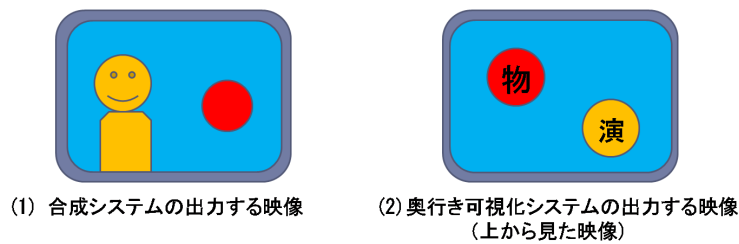


図8: 奥行き可視化機能の提示する映像

### 3.5.2 聴覚情報による奥行き情報のデザイン

3.4で述べた聴覚情報による一次元位置情報提示の具体的なデザインに関して述べる。3.3.2で述べたように、聴覚情報として信号音を利用する。信号音の表現方法は、表2のように、音量や高さをを用いた量的表現とリズムや音色を用いた記号的表現の2つに分けて考えることができる。

表2: 信号音の種類

音の種類	特徴	
量的表現	説明	音と情報をより直感的に結び付けて定めたもの。
	例	音量, 高さ。
	利点	直感性に優れ, 演技への影響が少ない。
	欠点	識別性, 記憶性に欠ける。表現できる情報の種類が少ない。
記号的表現	説明	音情報との対応関係を人為的なルールによって定めたもの。
	例	リズム, 音色。
	利点	識別性, 記憶性に優れる。
	欠点	音と情報の変換を経由し, 演技を阻害する可能性がある。

音量や高さをを用いた量的表現とは音とそのあらず情報をより直感的に結び

付けて定めたものである。音量の大小や音の高低が提示したいものの大小や高低に対応づけられるので、ある情報の変化を伝達させるのに有効である。

これに対してリズムや音色を用いた記号的表現とは、音とその音のあらゆる情報を人為的なルールによって定めたものである。特定のリズムや音色がそれぞれ異なる情報と対応づけられるので、それらの違いを利用することで異なる情報を伝達させるのに有効である。また、和氣ら [5] は、リズム、音色、高さの三要素に対してそれぞれの聞き分けやすさ、覚えやすさについて判定を行っており、聞き分けやすさについてはリズム、音色、高さの順で聞き分けやすく、覚えやすさについてもリズム、音色、高さの順で覚えやすいと報告している。このことから、リズムや音色といった記号的表現は識別性、記憶性に優れていると考えられる。

これらの表現方法を用いて、仮想物体の奥行きを表現する。

本研究では演者を仮想物体を把持しやすい奥行きまで誘導することで演者に対して仮想物体の奥行きを提示することを考える。このとき演者が必要とする情報は仮想物体が演者のどちら側(方向)に、どれだけ離れているか(距離)である。

これらの方向と距離に対して量的表現と記号的表現によるデザインを考える。まず量的表現について考える。量的表現は提示したいものの変化に対応づけられる。仮想物体が演者のどちら側かという位置関係の変化に対して音量や音の高さを変化させることで、奥行きに関して方向を量的に表現することが可能である。また、演者と仮想物体の距離は演者が移動することで連続的に変化する。この変化に対して音量や音の高さを変化させることで、距離を量的に表現することができる。このことから、方向情報、距離情報の提示に関して量的表現を用いることができる。

次に記号的表現について考える。記号的表現は特定のリズムや音色が、例えばこの音色はこの情報を意味するというように、提示したい情報と対応づけられる。そのため、仮想物体が演者のどちら側にかという方向に対してリズムや音色を定めることで、記号的に方向を表現することが可能である。一方、距離のように連続的に変化する情報を表現するためには膨大な数のルールが必要となり、そのルールを覚える演者に多大な負担を強いる可能性がある。そのため、方向の提示には記号的表現を用いるが、距離の提示に対しては用いない。

以上より、聴覚情報によって仮想物体の奥行きを提示するデザインとして、[手法 1] 方向:量的表現, 距離:量的表現と [手法 2] 方向:記号的表現, 距離:量的表現の

二通りが考えられる。

手法1について、奥行き方向に対する量的表現に関して、提示すべき情報を奥行き軸上の演者と仮想物体の位置関係の変化とした。本稿ではカメラに向かって仮想物体が演者より前方に存在する時は、前方から聞こえてくるように音を出力し、後方にあるときは後方から聞こえてくるように音を出力した、また、距離に対する量的表現を考えると、提示すべき情報を演者と仮想物体間の距離の変化とした。このとき、その距離の変化に応じて音量を変化させ、距離が小さいほど音量は大きく、距離が大きいほど音量は小さく出力した。

手法2について、奥行き方向に対する記号的表現に関して、提示する情報を奥行き方向の演者と仮想物体の位置関係とした。この位置関係とは、カメラに向かって仮想物体が a. 演者より前方にあるか、b. 演者と同じ奥行きにあるか、c. 演者の後方にあるかの三通りである。今回はそれぞれの位置関係に対して a. ピアノの音色、b. ピアノの和音の音色、c. ギターの音色を出力した。また、距離に対する量的表現としては、手法1と同様に演者と仮想物体間の距離の変化に応じて音量を変化させた。

### 3.5.3 聴覚情報による三次元位置情報のデザイン

提示する聴覚情報として、3.5.2と同様に信号音を利用する。

演者に対する仮想物体の三次元位置は連続的に変化する値として考えられる。そのため、その変化する値に対して音量や音の高さを変化させることで、量的に三次元位置を表すことが可能である。一方、その位置に対して記号的にルールを定めていてはルールの数が膨大になると考えられるため、本研究では三次元位置の提示に関しては記号的表現は用いない。

このことから、聴覚情報によって仮想物体の三次元位置を提示するデザインとして、[手法3] 量的表現が考えられる。本研究では仮想物体が存在する位置で音が鳴っているかのように演者が知覚できるように、仮想物体の位置に仮想的な音源を設置し、演者の耳に到達する音を再現してイヤホンから出力した。

[手法1]~[手法3]の音を出力するシステムとして聴覚ディスプレイシステムを利用する。聴覚ディスプレイシステムとは、その音源の距離や方向に発生する両耳に届く音の差異といった音の変化を再現することで、様々な方向にある音像を仮想的に出現するシステムのことである。本稿ではこの聴覚ディスプレイシステムを利用した機能のことを信号音提示機能と呼ぶこととする。

このことから、この信号音提示機能と3.5.1で述べた奥行き可視化機能を提示

機能として利用することで、演者が仮想物体の把持を行う上で必要な情報が獲得できる。取得機能としては尾原らの提案した把持検出機能を利用する。これらことから、本実験でシステム構成として図9のようなシステムを提案する。

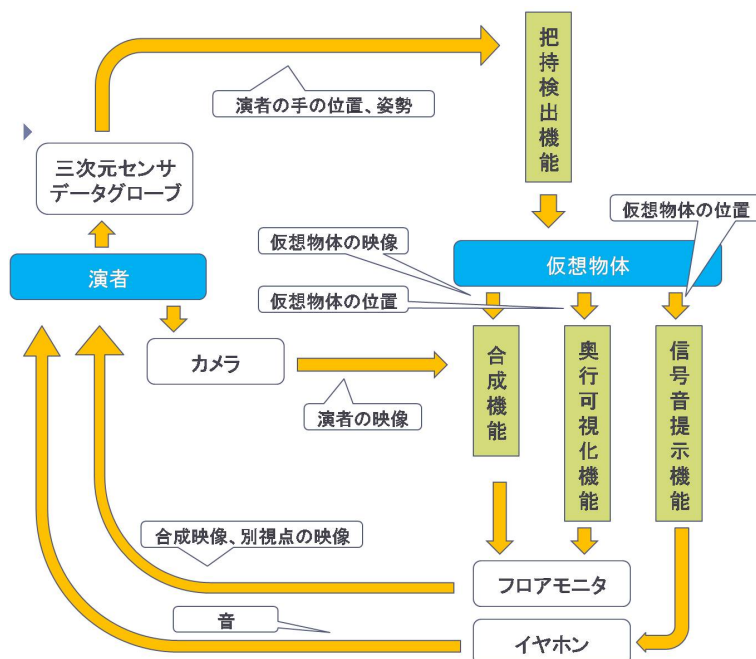


図9: 提案するインタラクティブバーチャルスタジオ

## 第4章 実験

### 4.1 把持を行う手の奥行きのばらつきの検証

#### 4.1.1 実験の目的

3.5.2では仮想物体の奥行き提示方法として、演者を仮想物体を把持しやすい奥行きまで誘導することを提案した。しかし、演者にとって仮想物体を把持しやすい奥行きというのは自明ではなく、そのため演者をどの奥行きに誘導すべきかがわからない。そこで、本実験では演者の立ち位置を決めたときに把持を行おうと差し出す手の位置を調査し、演者を誘導すべき位置を明らかにすることを目的とする。

#### 4.1.2 実験の方法

被験者として20歳代の男女3名(男性2名, 女性1名)に対して実験を行った。被験者の身長と利き腕は150cm:右利き, 160cm:左利き, 170cm:右利きである。



仮想物体把持はバーチャルスタジオにおいて行うので、本実験もバーチャルスタジオシステムを使用する。実験で使用するバーチャルスタジオは壁と床が青色で統一されている。被験者が仮想物体の把持を行う位置を固定し、体の向きを指示するために床上には図 10 のようなボードを設置した。カメラ横にはフロアモニタが設置されており、演者と仮想物体が合成された映像を提示する。被験者の手の三次元位置の計測には U3DTracker を使用した。

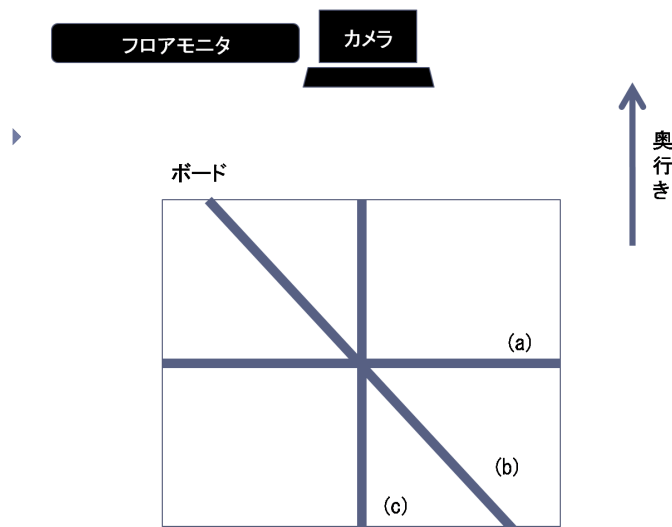


図 10: 使用するボード

#### 4.1.3 実験手法

以下に被験者に行ってもらおうタスクを説明する。

1. 被験者の頭を中心がボードの中心と重なるようにボード上の (a) の線の上に両足を乗せて、フロアモニタを参照しながら仮想物体の把持を行うような姿勢を取る。
2. 同様に (b), (c) の線の上に両足を乗せて、把持を行うような姿勢を取る。
3. 次に、一度適当な奥行き場所へ移動し、そこからボードへと歩いていき、把持を行うような姿勢を取る。このときも、ボードと同じ奥行き場所に仮想物体があるということは事前知識として与えておく。また、把持を行う時の両足の位置は特に指定しない。
4. 1~3 の試行を 1 サイクルとして、被験者ごとに 3 サイクルずつ実行する。

仮想物体は教材としても幅広く利用できる球を採用した。仮想物体の半径は10cmに設定した。奥行きは10のように、カメラに対する方向を正とする。合成映像において、仮想物体は被験者の右手の方に提示するものとする。また、このとき、システムによって把持の判定が行われると、被験者はその把持が成功するまで手を探るように動かす可能性があるのでシステムによる把持の判定は行わなかった。

上の各タスクにおいて、被験者が仮想物体に対して把持を行っていることを認識している状態で被験者に合図を送ってもらい、その時の奥行きを計測した。その結果から体の向きによって差し出した手の奥行きの平均と標準偏差を求めた。

表 3: 体勢における奥行きのばらつき (cm)

	全試行	(a) の体勢	(b) の体勢	(c) の体勢	歩行
平均	10.76	15.85	9.70	4.29	11.54
標準偏差	5.83	3.74	4.53	6.94	1.74

表 4: (c) の体勢を除く奥行きのばらつき (cm)

	(c) を除く全試行
平均	12.55
標準偏差	4.25

#### 4.1.4 結果と考察

全試行における奥行きのばらつき、各体勢における奥行きのばらつきの平均と標準偏差を3に示す。各体勢における試行に着目したとき、(a)と(c)の体勢では、差し出す手の奥行きの平均が10cm以上異なる。これは(a)のようにカメラに対して正面を向いている状態から、(c)のように真横になるにつれて被験者の肩の位置が変化することが原因であると考えられる。

しかし、本実験は被験者を仮想物体と同じ奥行きの位置へと誘導した場合を想定しており、その時(c)のようにカメラに対して真横に向いた状態で把持することは考えづらい。実際に実験において歩行の場合、把持の際に体の向きを(c)のようにカメラに対して真横に向ける例は見受けられなかった。このことから、今回は(c)の体の向きにおける奥行きのばらつきは考えないこととする。

表4に(c)を除いた場合の奥行きのはらつきの平均と標準偏差を示す。このときの平均は12.55cmであることから、被験者は仮想物体が被験者と同じ奥行きに存在しているという事前知識があった場合に、演者は自身の奥行きよりも正の方向におよそ12cm異なる奥行きに手を差し出すということが分かる。このことから、演者を仮想物体を把持しやすい奥行きまで誘導する場合、仮想物体の存在する奥行きから負の方向に12cm異なる奥行きに演者を誘導すべきであるということがわかる。

また、この奥行きのはらつきが正規分布に従うと仮定すると、5パーセンタイル値は5.55cmであり95パーセンタイル値は19.55cmであり、被験者の差し出す手の奥行きの値は、被験者の頭の中心から5.55~19.55cmという14cmの区間に90%の値が存在していることになる。このことから、被験者は仮想物体が被験者と同じ奥行きに存在しているという事前知識があった場合、その仮想物体を把持しようと差し出した手の奥行きは90%の確率でこの14cmの範囲に入る。このことから、差し出す手のはらつきもおよそ掌一つ分であり、仮想物体の把持を行う上で十分有効であると言える

## 4.2 視聴覚情報の提示による仮想物体把持への影響の検証

### 4.2.1 実験の目的

第3章では、従来のバーチャルシステムに加えて新たな視覚情報、聴覚情報を利用することを検討し、それぞれ映像と音のデザインを行った。本実験では、フロアモニタの提示する合成映像に加えて仮想物体把持の際にこれらの情報を提示することで、試行におけるそれらの情報提示の影響を検証する。

### 4.2.2 実験の方法

被験者として20歳代の男女7名(男性6名、女性1名)に対して実験を行った。本実験もバーチャルスタジオシステムを利用する。バーチャルスタジオの概要は4.1の実験と同様であるが、床上には目印となるものは設置しない。被験者の手の三次元位置の計測にはU3DTrackerを用いた。演者の頭の向きを計測するために、U3DTrackerを設置した帽子を用いた。また、演者に聴覚情報を提示するために利用するイヤホンにはSENNHEISER社のstyle wireless MX W1を用いた。また、本研究は聴覚ディスプレイシステムとしてSLABを用いた。SLABはNASAの提供するオープンソースアプリケーションであり、対象OSはWindowsである。バッファによって生じる遅延は5ms程度であり、処理遅延も6.1msと

短い。

演者役の被験者に仮想物体に向かって自らの手を伸ばし、仮想物体の把持をするタスクを行ってもらった。仮想物体は 4.1 の実験に用いたものと同じものを使用する。このとき、仮想物体の位置は事前には知らされず、試行によって提示される仮想物体の奥行き位置はランダムに変化する。提示する情報は以下のものを使用した。

表 5: 提示する情報

	提示する情報	
A	視覚のみ (合成機能のみ)	
B	視覚 + 視覚 (合成機能 + 奥行き可視化機能)	
C	視覚 + 聴覚 (奥行き) (合成機能 + 信号音提示機能)	量的表現 × 量的表現
D	視覚 + 聴覚 (奥行き) (合成機能 + 信号音提示機能)	記号的表現 × 量的表現
E	視覚 + 聴覚 (三次元) (合成機能 + 信号音提示機能)	

ここで、C、D では 3.5.2 で述べたように被験者と仮想物体との奥行きの位置関係によって出力する音を変化させている。把持しやすい奥行きと言うのは 4.1 の実験の結果から、実際に仮想物体が存在する奥行きにおよそ 12cm 奥であるということがわかっている。本実験では、差し出す手の奥行きの平均の位置が仮想物体の直径に含まれる 20cm の範囲へと被験者を誘導することとした。このことから、演者を誘導すべき範囲は図 11 において、実際に仮想物体の存在する直径 20cm の範囲よりも 12cm 奥の範囲であるアとイの間である。

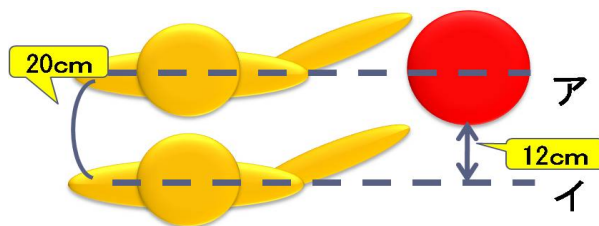


図 11: 誘導する範囲

実験はこれら A~E の 5 種類の場合での試行をランダムに一通り行うことを 1 サイクルとして、各被験者に 3 サイクルずつ行ってもらった。1 サイクル目は被験者が本システムに慣れることを目的とし、把持に成功した場合は成功した

ことを知らせた。ここでいう把持の成功とは演者の手が仮想物体の表面から誤差として 5cm 範囲に演者の手の三次元位置が存在した場合と定義した。1 サイクル目が終わったあと、引き続き 2 サイクル目と 3 サイクル目を行った。1 サイクル目では把持が成功したときは演者に対して把持を成功したことを知らせたが、2 サイクル目と 3 サイクル目では知らせないこととした。これはシステムで把持の成功を知らせてしまうと、その成功のサインが得られるまで被験者が手を動かしてしまう可能性があるからである。2 サイクル目と 3 サイクル目は以下の手順で行った。

1. A~E のいずれかの情報を基に仮想物体に対して把持を行う。
2. 被験者が仮想物体を把持していると認識した時点で被験者はその旨を伝える合図を送る。
3. 1, 2 を A~E の場合ですべて行う。

聴覚情報を提示することで演技に対して阻害が発生しているかどうかを調べるため、上の手順を 2 サイクル目では周囲の環境は無音の状態で行い、3 サイクル目では常に会話をしながら行った。これらの試行に共通して、把持を行う際に手を差し出すことは一度のみと限定し、さらに差し出した手を前後に動かすような仕草を見せた場合は目視で無効と判定した。明らかに手が把持の姿勢でない場合も目視で無効と判定した。また、三次元位置センサとして使用する U3DTracker には 100~300ms 程度の遅延が発生していたため、被験者にはその旨も伝えた。

各試行において、演者が把持をしていると認識し、合図を送った時点での演者の手の奥行きと、仮想物体の重心からの距離を測定した。さらに全試行終了後に A~E の情報提示に関するアンケートをとった。

#### 4.2.3 結果

以下の表 4.2.3 には A~E ごとの、演者が把持と認識した時の演者の手と仮想物体の重心との奥行きの差の平均と標準偏差について、会話をしない場合、会話をした場合の値を示す。表 4.2.3 には演者の手と仮想物体の重心との距離に関しても同様に会話をしない場合、会話をした場合の平均と標準偏差の値を示す。また、距離は仮想物体の重心と演者の手の距離であり、重心と表面を距離は 10cm であるため、表の結果から 10cm を引いた値が仮想物体の表面からの演者の手の距離となる。

表 6: 奥行き之差 (会話なし)(cm)

	平均	標準偏差
A	40.8	23.96
B	9.71	3.30
C	18.88	11.58
D	6.18	5.84
E	21.16	18.03

表 7: 奥行き之差 (会話あり)(cm)

	平均	標準偏差
A	35.56	14.51
B	8.69	6.85
C	13.51	8.48
D	8.81	6.04
E	20.90	16.30

表 8: 距離の大きさ (会話なし)(cm)

	平均	標準偏差
A	37.95	18.42
B	11.91	6.45
C	22.73	10.32
D	13.25	6.19
E	21.92	13.53

表 9: 距離の大きさ (会話あり)(cm)

	平均	標準偏差
A	31.54	12.8
B	16.01	3.33
C	13.59	12.84
D	15.16	10.59
E	22.85	12.62

#### 4.2.4 考察

仮想物体の重心と演者の手の奥行き之差に関して考察する。会話なしの各試行において、A の試行、つまり従来の合成映像のみの場合に比べてすべての場合において平均や標準偏差が小さくなっていることから、被験者に対して情報を提示することで仮想物体把持において支援が行えていると考えることができる。中でも、平均に関しては奥行き方向と距離をそれぞれ記号的表現と量的表現によって提示した際の試行である D の試行が最も誤差が少なく把持が行えた。次に小さかったのは別視点として被験者と仮想物体を上から見た位置関係の映像を提示した B の試行であった。これらの試行は標準偏差も小さく、安定して被験者の支援が行えたと考えられる。

さらに、会話をしながら行った場合の結果を見たとき、B の試行と D の試行は標準偏差が大きくなっており、支援効果が小さくなっている。これは、会話をしながら仮想物体の把持を行うというタスクを行うことで、タスクに対して集中できなかったことから B の試行と D の試行に対して影響が現れたと見ることができる。

次に、仮想物体の重心と被験者の手の距離に関して考察する。会話なしの各試行においても、奥行き之差の場合と同様に、B と D の試行が誤差が少なく把持が

行えた。その他の試行に対しても奥行きの場合と同様の傾向がうかがえる。

ただ、会話をしながら行った場合の結果を見たとき、Dの試行の標準偏差が大きくなっていることから、会話をすることでBよりも大きな影響が現れたと見ることができる。実験後にとったアンケートでは、BとDは演技を阻害せず、かつ位置を特定しやすいという意見が多数あった。このことから、会話をしながらDのような聴覚情報を提示した場合に、被験者自身は阻害されたという印象は持っていなくても、その影響は少なからず発生するということと言える。ただ、奥行きの提示に関しては視覚で提示するBの場合と同等の精度を保っていることから、誘導を行う際に発生する音のデザイン次第では、距離に関してもより精度を向上できるのではないかと考えられる。

以上のことから、すべての情報を視覚によって提示しなくても、視覚と連携し、適切なデザインをおこなうことで位置情報の提示に関して聴覚も十分なモダリティとなることが考えられる。

## 第5章 結論

本研究では、視聴覚情報の提示による演者のバーチャルスタジオにおける仮想物体把持支援をおこなった。本手法では、演者がバーチャルスタジオにおいて仮想物体の把持を行う際に、従来のバーチャルスタジオでも用いられているフロアモニタの提示する合成映像の視覚情報に加えて、別の視点から見た視覚情報とイヤホンを利用して提示する聴覚情報によって仮想物体の三次元位置を提示することで実現した。

本手法を実現するうえで、まずフロアモニタとイヤホンを利用した仮想物体の三次元位置の提示方法を検討した。その提示方法の中で、合成映像とは別の視覚情報によって奥行きを提示する方法として、カメラの視点とは別の視点から演者と仮想物体の位置関係を反映した映像を提示した。また、聴覚情報によって奥行き情報を提示する方法として、演者を仮想物体把持しやすい奥行きまで誘導することでその位置を提示した。このときの把持しやすい奥行きの範囲は実験によって確認した。

次に、それらの提示方法を実現する聴覚情報のデザインを行い、最後に、フロアモニタの提示する視覚情報に加えて、新たな視覚情報や聴覚情報を用いて仮想物体把持の試行を行うことで、バーチャルスタジオでの仮想物体把持におけ

る視聴覚情報の有用性を確認した。

本研究では、仮想物体の把持を行う際に、演者の横に単一の仮想物体が存在することを想定して仮想物体の三次元位置の提示を行った。しかし、演者が仮想物体の把持を行う際に、必ずしも演者は仮想物体の横から把持を行うとは限らず、かつそれも単一とは限らない。このため、今後の課題としては、例えば演者の正面にあるような物体に対して把持を行うなど、仮想物体のより自由度の高い三次元位置の知覚の支援が挙げられる。さらに、例えば複数の物体の情報を提示しようとする際には、提示すべき情報が大幅に増えることが考えられる。聴覚は視覚に比べて一度に示せる情報の数が限られるため、目的に応じた聴覚情報の最適なデザインも課題と言えよう。

また、本研究では三次元位置の提示に新たな感覚である聴覚を利用したが、例えば演者が仮想キャラクターと会話を行うような合成映像において、仮想キャラクターの三次元位置に仮想音源を設置することで、実際に演者自身が仮想キャラクターと会話しているような感覚で、演技を行うことも可能であるように、聴覚情報の別の利用方法も考えられる。

今後は、視覚情報や聴覚情報を付与することで提示される三次元位置の精度や自由度の向上だけでなく、バーチャルスタジオにおいて聴覚情報を付与することで、演者の演技を支援しうる領域を調べていくことも考えている。

## 謝辞

本研究を進めるにあたり、多くのご教授を賜りました美濃導彦教授、椋木雅之准教授、飯山将晃准教授に深謝いたします。また、日頃より研究に関してあらゆる面で熱心に直接ご指導くださいました船富卓哉助教に厚く御礼申し上げます。グループミーティングなどでアドバイスをしてくださいましたモデルグループの皆様、コンテンツ作成室の皆様、ならびに美濃研究室の皆様に深く感謝いたします。

## 参考文献

- [1] 教育機器編集委員会（編）. 産学教育機器システム便覧. 日科技連出版社, 1972.
- [2] 大島康介, 船富卓哉, 飯山将晃, 角所考, 美濃導彦. バーチャルスタジオにお



- ける現実物体を用いた仮想物体の直接操作. 電子情報通信学会総合大会講演論文集, p. 286, 2006.
- [3] 尾原秀登, 角所考, 美濃導彦. バーチャルスタジオにおける演者の手を用いた仮想物体の直接操作の実現. 信学技法:MVE, Vol. 106, No. 91, pp. 103–109, MVE2006-21.
- [4] 和氣早苗. Sui (sound user interface) : サイン音を用いた情報表示とそのデザイン. ヒューマンインタフェースシンポジウム 2005, pp. 105–110, 2005.
- [5] 和氣早苗, 上窪真一, 福住伸一, 旭敏之, 広明敏彦. 音響インタフェース設計手法' 報知音多次元設計手法' 確立への一考察. 電子情報通信学会論文誌 D- , J82-D- , No. 10, pp. 1721–1728, 1999.