

# 没入型通信環境における合成人物像の利用

## Video Avatar in the Networked Immersive Environment

小木哲朗, 酒井満隆/筑波大学大学院 システム情報工学研究科  
Tetsuro Ogi, Mitsutaka Sakai/University of Tsukuba

In this study, several immersive projection displays such as the CG Gallery, the CABIN and the CAVE were connected through the JGNII (Japan Gigabit Network 2) network, and the video avatar technology was experimentally applied for various kinds of communications in the networked immersive environment. The video avatar was used for the mutual communication in the networked CAVE among three sites, and the positional relationship among the users could be represented in the shared virtual world. This technology can also be used in the real world environment by displaying it in the augmented reality system. In this study, the immersive augmented reality display that uses a large half mirror film was developed, and it was used to superimpose the video avatar image on the real scene. In addition, the video avatar technology was used to record the body movement of the Shorinji Kempo, and the recorded image was replayed in the immersive projection display with the force feedback, so that the user can practice the movement of the kata effectively.

### 1. 緒言

近年、CAVE や CABIN に代表される没入型映像提示装置 (IPT: Immersive Projection Technology) が各地に設置され、種々の研究用途に使われるようになってきた[1]。没入型映像提示装置では、広視野の立体視映像を提示することで臨場感の高い仮想体験を行うことが可能である。没入型映像提示装置が各地で利用可能になってくると、次の段階としてこれらを通信網で接続し、臨場感の高い仮想空間を共有するための手法が要求されてくる[2]。

没入型通信環境において利用者が臨場感の高い対話を行うためには、利用者自身の姿を現実感の高い手法で表現することが必要である。このような要求に対する人物の表現手法として、筆者等はこれまでに人物像合成技術 Video Avatar について検討を行ってきた[3]。Video Avatar とは利用者の姿を実寸大の実写映像として仮想空間に合成することで、一人称体験としての相互対話を行う方法である。人物像の表現形態としては、平面映像を用いる 2 次元 Video Avatar、表面形状映像を用いる 2.5 次元 Video Avatar、立体形状映像を用いる 3 次元 Video Avatar 等の種々の方法が考えられている。正確な身体動作の表現が要求される応用分野では 2.5 次

元形状や 3 次元形状を用いた映像表現が必要であるが、対面型の対話等に限定される場合には処理速度を優先させた 2 次元形状の合成人物像の方が有効な場合も多い[4][5]。

本研究では多地点に存在する CAVE 型没入映像提示装置を JGNII (Japan Gigabit Network 2) 通信網で接続した没入型通信環境を構築し、広帯域通信網上での Video Avatar の利用方法について検討を行った。合成人物像を遠隔地間で実時間で送受信することで臨場感の高い対話手段として利用することができる。また人物の動作を Video Avatar 情報として記録・再生することで、身体動作に関する教示手段等としての応用が期待される。

### 2. 没入型通信環境

本研究では、JGNII、つくば WAN の広帯域通信網を介して、多地点間の没入型映像提示装置を相互に接続した通信環境を構築した。JGNII は情報通信研究機構が運用する超高速・高機能の研究開発用通信網で、幹線部分は最大 20Gbps の帯域を有している。通信網への接続地点は全国に 64 箇所設置され、筑波地域はつくば JGNII Research Center、東京、京都はそれぞれ東京大学、京都大学が接続地点となっている。ま

た、つくば WAN は筑波研究学園都市の各研究機関を 10Gbps の超高速 Access Ring で結ぶ研究開発用通信網であり、つくば JGNII Research Center において、つくば WAN と JGNII の両通信網が接続されている。本研究では筑波大学から、つくば WAN を経由して JGNII 通信網に接続し、つくば JGNII Research Center、東京大学、京都大学、岩手県立大学、北陸先端科学技術大学院大学の間で、それぞれが所有する没入型映像提示装置間を通信網で接続した通信実験環境を構築した。図 1 は、本研究で構築した没入型通信環境の構成を示したものである。

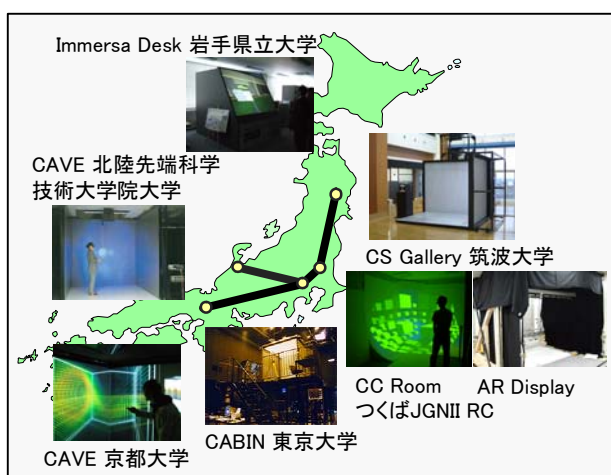


図 1 実験用の没入型通信環境

筑波大学では、正面、右面、床面の投影面で構成された 3 面構成の CAVE 型映像投影装置 CS Gallery (Cyber Space Gallery) が設置されている。CS Gallery は正面と床面は 2.10m x 2.63m、右面は 2.10m x 2.10m の投影面で構成され、各投影面に対して 2 台ずつの DLP 式映像投影装置 NEC ViewLight LT245J を用い、円偏光方式による立体視映像の提示を行っている。また映像を生成するための計算機としては、画像描画装置 NVIDIA Quadro FX3400、画像合成装置 ORAD DVG を備えた 6 台の HP XW6200 による PC Cluster を構成し、描画負荷の変動に応じた動的負荷分散描画機能を実現している[6]。

また東京大学には正面、右面、左面、天井、床面の投影面で構成された 5 面構成の CABIN、京都大学には正面、右面、床面の 3 面構成の CAVE 装置がそれぞれ設置されている。岩手県立大学では 1 面投影面の ImmersaDesk、北陸先端科学技術大学院大学では 4 面構成の CAVE をそれぞれ所有している。

また、つくば JGNII Research Center では、部屋の角部分を利用した曲面投影面と魚眼式映像投影装置で構成された CC Room、大画面の半透明鏡を利用した没入型拡張現実感提示装置が設置されている。本研究では、これらの没入型映像提示装置を相互に広帯域通信網で接続することで、臨場感の高い対話環境としての利用を可能にした。

### 3. Video Avatar の生成方法

没入型通信環境における人物像の表現手法としては、ここでは 2 次元形状の映像による Video Avatar を使用した。Video Avatar は、実写撮影装置で撮影された人物の撮影映像から背景を切り抜いた人物像を作り、これを仮想世界に合成することで生成される。

人物映像の撮影には、IEEE1394 非圧縮出力の映像撮影装置 SONY DFW-X710 に広角固定焦点の光学系 PENTAX H-416 を装着した実写撮影装置を使用した。撮影時の画像取込み速度は XGA (1024x768) の解像度で 15fps まで可能であるが、ここでは後の画像処理の効率を上げるため VGA (640x480) の解像度で計算機への画像取込みを行った。

また人物像の切り抜きに関しては、それぞれの映像提示環境で利用可能な背景差分法を使用した。背景差分法とは、あらかじめ背景となる画像を静止画として撮影しておき、実写撮影装置で撮影された人物映像との差分を 1 画面の取込みごとに計算することで、人物像だけが切り出された動画を得る方法である。ここでは、取込み画像は 24bit の RGBA 画像で表現し、差分計算は撮影画像と背景画像の R (赤) G (緑) B (青) 値の差の絶対値の合計を計算し、これが閾値より大きければ  $\alpha$  値を不透明、小さければ透明とした。

切り抜かれた Video Avatar の映像は画面の更新ごとに Texture 情報として扱い、Texture Mapping の手法によって仮想世界に合成する。人物像の配置位置は撮影に用いた映像撮影装置の位置を基準にした相対座標で決まるため、仮想世界内に仮想の撮影位置を設定し、そこから一定距離だけ離れた場所に四角形の画像領域を設定する。この画像領域に Video Avatar を Texture Mapping により貼り付けると、切り抜かれた背景部分は  $\alpha$  値により透明となるため、人物像だけが仮想世界に合成された形で表示される。図 2 は、以上の背景差分法による Video Avatar の生成方法を図示したものである。

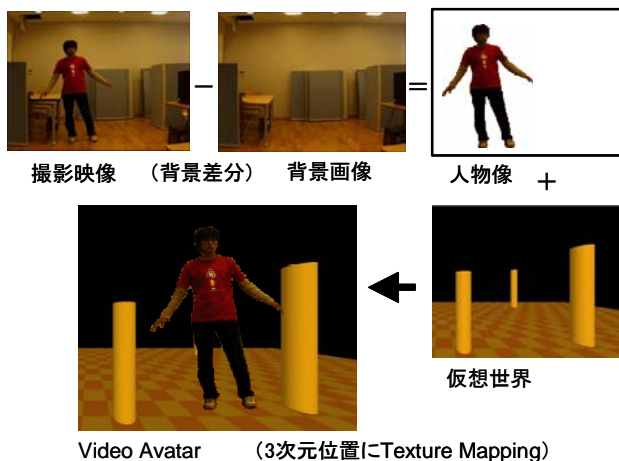


図2 背景差分法による Video Avatar の生成方法

#### 4. 多地点 CAVE 間通信

Video Avatar 技術の利用方法として、まず筑波大学、東京大学、京都大学間の CAVE 間通信環境を使用した多地点間での相互対話への適用実験を行った。Video Avatar 映像を多地点間で送受信する際に全体の動作効率をどこまで保てるか、また共有仮想世界の中で利用者がお互いの位置関係を正しく表現できるか等を検証することが、ここでの目的である。

各地点では、それぞれの CAVE 装置を JGNII 通信網に接続するため、通信用 PC を設置した。人物映像を撮影するための映像撮影装置は通信用 PC に接続され、JGNII を用いた実写人物像の送受信はこの通信用 PC を介して行われる。通信用 PC は、相手から受け取った人物像に対して、背景差分処理を行い、切り抜かれた人物像を CAVE を構成する各 PC へ配信する役割も持つ。

Video Avatar を用いた相互対話環境では情報の損失よりも実時間性が重視されるため、通信方法としては UDP を用いている。また本手法では1枚の人物画像は RGB 成分により  $640 \times 480 \times 3 = 921,600$  byte で構成されているが、UDP で一度に送れる Packet の大きさには制限があるため、1枚の画像情報を分割し Packet 番号とともに分割画像を送信し、受信側で1枚の更新映像として再合成を行っている。

また Video Avatar を通信網を用いて送信する場合、合成人物像の生成過程のどの段階で画像情報を送信するかは、全体の動作効率を考える上で重要な問題である。ここでは予備実験として、背景差分処理を行う前に画像情報を送信する場合と、行った後に送信する場合とで動作効率の比較を行った。その結果、撮影映像の送信後に受信側で背景差分処理を行っ

た場合は 14.9fps となり、撮影装置からの取り込みとほぼ同じ更新速度で全体の処理を行うことができたが、送信前に撮影側で背景差分処理を行った場合は 12.4fps となり、全体の動作効率が低下した。これは Video Avatar 映像の通信にかかる負荷よりも撮影装置からの画像取込みと背景差分処理による切り抜きの負荷の方が全体に占める割合が大きいことを意味している。そのため、ここでは映像の取込みと背景差分処理を分け、背景差分前の撮影映像の画像を1枚ずつ JGNII 通信網に送信する方法を用い、3地点間での相互対話実験を行った。

PC Cluster を用いた CAVE の通信環境では、CAVE 間の通信と CAVE を構成する PC 間の通信が存在する。ここでは JGNII を用いた CAVE 間での通信負荷が CAVE を構成する PC 間の通信に影響を及ぼすことを防ぐため、両者の通信は処理工程を分け、Shared memory を介して情報を受け渡す方法を用いた。表1は、筑波大学と東京大学、京都大学のそれぞれの間で2地点間の Video Avatar 通信を行った場合、表2は筑波大学、東京大学、京都大学の3地点間で相互に Video Avatar 通信を行った場合の通信速度を示したものである。また、表3は3地点間通信時の各 CAVE 装置における描画更新速度を示したものである。

この結果から、筑波大学と東京大学の間では十分な通信速度を保てた (14.9fps) が、京都大学は学内通信網を経由して CAVE を接続しているため、筑波大学と京都大学間の通信速度は低かった (4.22fps)。そのため、3地点間通信では最も遅い地点に引きずられ、全体の動作効率は低下した (3.21~3.42fps)。この問題を解決するためには、全体の通信を管理する映像通信制御装置 (Video Avatar Server) を設け、お互いが相手の動作効率に影響を及ぼさないようにすることが必要である。一方、各 CAVE 内での描画更新速度は筑波大学で 15.3Hz、東京大学で 7.15Hz、京都大学で 9.92Hz と、映像提示装置間の通信速度に引きずられることなく各 CAVE の性能に従った描画速度が実現された。

図3は上記の方法に従って、筑波大学、東京大学、京都大学間で Video Avatar 通信を行っている様子を示したものである。筑波大学の CS Gallery から見た場合、正面の投影面には京都大学の利用者、右面の投影面には東京大学の利用者の姿が表示されている。これは筑波大学の利用者の右前方に東京大学の利用者が立ち、左前方に京都大学の利用者が立つ

ている状態を示している。一方、東京大学の利用者から見た場合は、左前方に筑波大学の利用者が見え、右前方に京都大学の利用者が見えている。このように、各地点から送られてくる Video Avatar 映像を共有仮想世界の中で相手の存在する位置に合成することで、共有仮想空間の中でお互いの位置関係を表現することができ、空間を共有した多地点間での相互対話を実現することができた。

表 1 2 地点間の映像通信速度

	東京大学	京都大学
筑波大学	14.9	4.22

(fps)

表 2 3 地点間の映像通信速度

	筑波大学	東京大学	京都大学
筑波大学	—	3.28	3.29
東京大学	3.42	—	3.42
京都大学	3.21	3.21	—

(fps)

表 3 各映像提示装置の描画更新速度

	筑波大学	東京大学	京都大学
描画速度	15.3	7.15	9.92

(Hz)

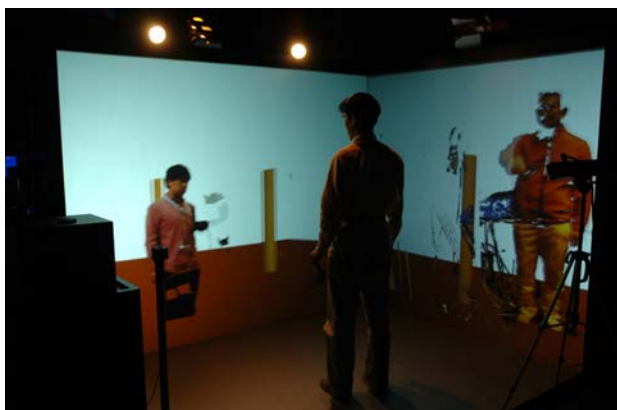


図 3 筑波大学—東京大学—京都大学の 3 地点間通信実験

## 5. 実空間での Video Avatar 合成

遠隔地の人間と Video Avatar 技術を用いた対話を行う場合、CAVE 等の仮想環境の中だけではなく実空間での利用方法も考えられる。仮想空間同士を接続するのではなく、実際

の作業現場等を通信網に接続し、遠隔地の人間と対話を行う方法である。この場合、実空間に Video Avatar を合成表示するため、没入感の高い拡張現実環境を構築することが必要である。

本研究では、大画面の投影面を用いた拡張現実空間を実現するため、2.0m×2.0m の透過率の高い薄膜状の半透明鏡を用いた没入型拡張現実感映像提示装置を構築した。図 4 は構築した映像提示装置の構造を図示したものである。半透明鏡は 45 度の角度で設置され、映像投影装置による映像は天井から床面に配置された投影面上に投影される。利用者は半透明鏡で反射された虚像と半透明鏡の背後の実空間を光学的に融合し、広視野の拡張現実空間を見ることができる。

映像投影装置としては Infocus DepthQ を使い、時分割方式による立体視映像を提示し、利用者は液晶切換式眼鏡 StereoGraphics Crystal Eyes を使用して立体視映像を見る。視点位置の計測には磁気式位置計測装置 Polhemus Liberty を使い、視点位置の変化に追従した映像が提示される。利用者は半透明鏡で反射された仮想世界の映像と半透明鏡を透過した現実世界を重畳して見るが、この際、仮想世界と現実世界の像の明るさが同等になるように、半透明鏡の背面には照明を取り付けている。

Video Avatar 映像を実空間に合成表示する場合は、仮想世界の映像との合成処理を行う必要はなく、Video Avatar が映像提示空間を越えて移動することもない。そのため、背景差分処理を行った人物像をそのまま映像提示装置に提示することで実空間と合成されるが、この際、現実空間と正しい位置関係で合成するためには、相互の位置関係の正確な調節を行う必要がある。具体的には、実空間に対する投影面の位置合わせ、投影面に対する映像投影装置の位置合わせ、および視点位置計測に応じた正しい視差映像を提示するための位置計測装置の基準位置調整等を行う。

図 5 は半透明鏡の背後に立った実空間上の人物と Video Avatar として表示された遠隔地の人物像が並んで立っている状態を示したものである。

拡張現実感型の映像提示装置を用いる場合、現実空間そのものを遠隔地間で共有することはできないため、両地点で同等の環境を用意する、あるいは仮想空間と実空間の間で情報共有を行う等の対話形態となる。また、利用者は拡張現実空間と対話を行うだけではなく、現実世界と融合された拡張現

実世界の映像を利用者に見せることで、遠隔授業や遠隔講演、博物館における展示物の説明等の分野における利用が考えられる。

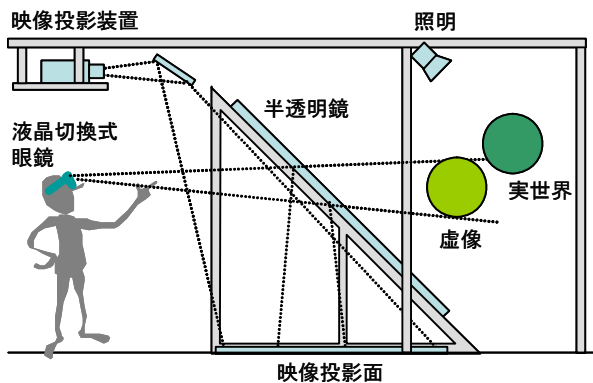


図4 没入型拡張現実感映像提示装置の構成

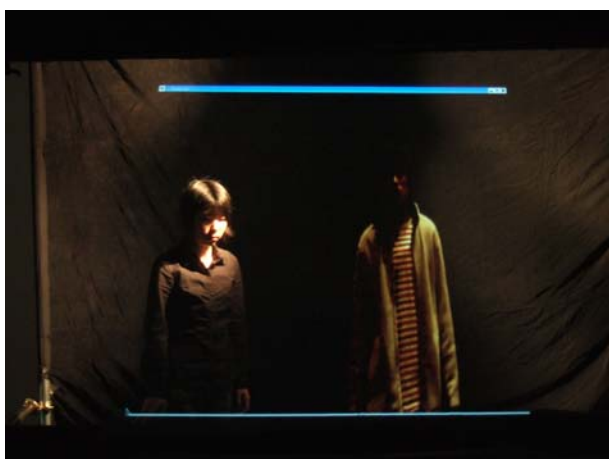


図5 実空間と融合表示されたVideo Avatar

## 6. Video Avatar による動作教示

Video Avatar は実時間で送受信して相互対話に使用するだけでなく、Video Avatar 情報を記録・再生により利用する方法も考えられる。人間の動作に関する情報を計算機内で3次元映像情報として記録しておくことで、仮想空間の中で繰り返し再生し、体験を行うことができる。この方法は、模範的な身体動作を伝承していくことが必要な武道や舞踏等における動作の記録、伝達手段として利用することができる。また記録されたVideo Avatar 映像はDatabase化して通信網上に蓄積していくことで、複数の利用者の協調的な練習や学習に利用することが可能になる。

本研究では、少林寺拳法における基本的な身体動作である型の動きをVideo Avatar 情報として記録し、没入型提示装置の中で対話的に再生することで、初心者向けの型の練習に利用する方法を試みた。空手や少林寺拳法等の武道では、練習生は師範と向き合い、師範の動きを見ながら自分の動きを練習する。この模範となる師範の動作をVideo Avatar 情報として記録して仮想空間の中で再生する、あるいは練習生の動作を記録して仮想空間の中で師範の動きと比べて提示することで、効率的な動作の習得ができることが期待される。本実験では、Video Avatar の映像情報とともに、動作計測装置(motion capture)によって身体の動きを記録し、模範動作を視覚情報とともに触覚情報によっても提示する方法を試みた。

実験ではまず模範となる動作を映像撮影装置で撮影すると同時にCyVerse Stereo Labeling Camera SLC-G02を用いて動作計測情報を記録した。Stereo Labeling Cameraは、再帰性反射材を使用した目印(Marker)に赤外光を照射し、反射光の撮影映像から、目印の3次元位置を計測する装置である。ここでは腕の動きにだけ注目したため、模範者は両腕の各関節に目印を付け、2台のStereo Labeling Cameraを用いて2方向から3次元動作を記録した。記録されたVideo Avatarは、CS Galleryの仮想空間の中で再生表示され、練習者は目の前に表示された実寸大の大きさのVideo Avatar映像を見ながら、繰り返し型の練習を行うことができる。

またこの際、練習者は模範者と同じ両腕の関節位置に目印を装着し、身体の動きを計測することで、模範動作との動きのずれを計測した。この計測された身体動作のずれは、視覚映像として練習者に提示する他、振動子を用いた触覚提示装置およびGyro-momentによる力覚提示装置を用いて、練習者に触力覚情報として提示する方法を行った[7]。

図6は本装置を用いて模範者が少林寺拳法の型の動作を記録している様子、図7は練習者がVideo Avatarによる視覚情報と触覚提示装置を用いて型の練習を行っている様子を示したものである。本実験では、2次元のVideo Avatar映像を用い、動作計測によって3次元的な動作の記録を補ったが、より正確な身体動作を表現するためには、自由視点での表示が可能で3次元形状の映像を用いたVideo Avatarの記録が必要であると考えられる。また実験では、Video Avatar映像と同時に、触力覚情報の提示を用いて身体動作の教示を行

ったが、体験型学習における効果的な五感情報の利用法としての可能性を示すことができた。



図6 実写映像と動作計測装置による少林寺拳法の型の記録



図7 Video Avatar と触覚情報を用いた身体動作の教示

## 7. 結論

本研究では、各種の没入型映像提示装置を広帯域通信網で接続した没入型通信環境を構築し、Video Avatar の利用方法について検討を行った。合成人物像を多地点間通信に展開したときの利用者間の対話通信効率、拡張現実感映像提示装置を用いた実空間での人物像の表現方法、記録・再生方式による教示装置としての利用方法等について実験装置を用いた検討を行った。Video Avatar の応用分野としては、遠隔地間での相互対話の他、展示や講演、動作教示等、種々の分野での利用が想定される。今後は、種々の応用分野で効果的かつ容易に利用可能にするため、多地点間での Video Avatar 配信用の通信管理技術や、Video Avatar 情報を蓄積・管理するための Database 技術等について研究をすすめていく予定である。

## 謝辞

本研究は、情報通信研究機構の民間基盤技術研究促進制度「Tele-Immersive Conference System に関する研究開発」、および総務省の戦略的情報通信研究開発推進制度「実空間による人物像の記録・伝送・再生に関する研究開発」の一部として行った。JGNII の利用にあたっては、つくば JGNII Research Center の協力をいただいた。また CAVE 間通信の実験では東京大学立山義祐氏、京都大学江原康生助手、拡張現実感映像提示装置の構築では株式会社 SPIN 斉藤康太氏、小山尚英氏、Video Avatar の記録再生実験では、筑波大学岩田洋夫教授、矢野博明助教授、松本睦樹氏に協力をいただいたことを感謝します。

## 参考文献

- [1] C. Cruz-Neira, D. Sandin, T. DeFanti: Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE, Proc. of SIGGRAPH' 93, pp. 135-142 (1993)
- [2] J. Leigh, T. DeFanti, A. Johnson, M. Brown, D. Sandin, "Global Tele-Immersion: Better than Being There", ICAT' 97, pp. 10-17 (1997)
- [3] T. Ogi, T. Yamada, Y. Kurita, Y. Hattori, M. Hirose: Usage of Video Avatar Technology for Immersive Communication, Proc. of LUAR 2003, pp. 24-31 (2003)
- [4] 栗田、國枝：多視点カメラによる3次元モデル生成とネットワーク伝送システム、日本バーチャルリアリティ学会第7回大会論文集、pp. 489-492 (2002).
- [5] J. Insley, J. Sandin, T. DeFanti, Using Video to Create Avatars in Virtual Reality, Visual Proceedings of 1997SIGGRAPH, pp. 128 (1997)
- [6] 小木、内野：動的負荷分散型 CAVE システムの開発、日本バーチャルリアリティ学会 10 回記念大会論文集、pp. 117-120 (2005)
- [7] 吉江、矢野、岩田：ジャイロモーメントを用いた力覚呈示装置、ジャイロモーメントを用いた力覚呈示装置、日本バーチャルリアリティ学会論文誌、Vol. 7, No. 3, pp. 329-337 (2002)