

# 投射型立体映像との3Dインタラクションを 客観視点映像化する複合現実感システム

片岡 直哉<sup>†</sup> 鶴田 剛史<sup>†</sup> 木村 朝子<sup>†</sup> 柴田 史久<sup>†</sup> 田村 秀行<sup>†</sup>

<sup>†</sup>立命館大学 情報理工学部 〒525-8577 滋賀県 草津市 野路東 1-1-1

E-mail: <sup>†</sup> {kataoka, tsuruta}@rm.is.ritsumei.ac.jp, {asa, fshibata}@is.ritsumei.ac.jp, HideyTamura@acm.org

**あらまし** 我々はこれまで、立体映像表示のできる投射型の大型画面を利用した「広視野電子作業空間」の研究開発を推進してきた。本稿では、このよう立体映像表示システムにおいて利用者の眺めている立体映像と、その映像に対する利用者によるインタラクションの様子を同時に記録・提示するシステムを提案する。立体映像表示装置と同じ座標系内に単眼のカメラを設置し、カメラの位置姿勢情報を与え、スクリーンに投影されている映像を除去するとともに、立体映像表示生成する映像とカメラから得た現実世界の映像を合成する。本システムは、単に成果発表等に用いるだけでなく、3Dインタラクションを含むヒューマンインタフェース（HI）分野の研究ツールとしても極めて有用である。

**キーワード** 客観視点映像, 投射型立体映像, 3Dインタラクション, 複合現実感

## Mixed Reality System Visualizing 3D Interaction with Projected Stereoscopic Image from Objective Point of View

Naoya KATAOKA, Takeshi TSURUTA, Asako KIMURA  
Fumihisa SHIBATA, and Hideyuki TAMURA

College of Information Science and Engineering, Ritsumeikan University

1-1-1 Nojihigashi Kusatsu-shi, Shiga, 525-8577 Japan

E-mail: {kataoka, tsuruta}@rm.is.ritsumei.ac.jp, {asa, fshibata}@is.ritsumei.ac.jp, HideyTamura@acm.org

**Abstract** We have developed a system named "Wide-view electronic working space" using a projection type wide-view stereoscopic image display. In this paper, we propose a system which displays and records both the 3D CG images that the user is looking at and the interaction between the CG and her/him at the same time from Objective Point of View. A camera can be set up in any place of the same coordinate system with the stereoscopic image display. Using the position and posture of the camera, the 3D CG images are drawn on the real world image from the camera. This system is extremely useful for a research tool in human interface (HI) field including 3D interaction, as well as for preparing images and videos for presentation of such a stereoscopic system.

**Keyword** Objective Point of View, Projection type Stereoscopic Image Display, 3D Interaction, Mixed Reality

### 1. はじめに

コンピュータによる作業効率の向上、対話型電子作業の多様化とともに、デスクトップ型モニタの限られた画面サイズでは満足できず、広い作業を確保できる大型スクリーンや壁面ディスプレイの利用が進行している[1]。電子作業空間をより魅力的なものにするには、その画面の表示は単眼画像に留まらず、両眼立体視を利用した立体映像表示へ発展するのも自然な成り行きである。大型の立体映像表示システムとしては、バーチャルリアリティ分野で、立方体形状をした壁面に囲まれる CAVE[2]やその変形システム、アーチ状や球面状のスクリーンで広視野をカバーするイマーシブ・デ

イスプレイの開発・導入が進んでいた。筆者らは、そうした立体映像表示できる投射型の大型画面を利用して「広視野電子作業空間」の研究開発を推進してきた[3]。ここでは、大画面に向かって、利用者がジェスチャ操作で作業することを想定している。

立体映像の導入は、利用者・体験者にとっては魅力的なものであっても、機材は一般的でなく、その内容を成果報告等で広く他人にアピールすることは容易ではない。たとえ、利用者の眺めている立体映像（スクリーン上に偏光方式や時分割方式で表示される一対の映像）を専用装置を準備して再現することはできても、その映像に対してジェスチャ操作を行なうインタラク

ションの様子も同時に記録し、提示することはできなかった。

この問題を解決するため、我々は単眼画像でありながら、鳥瞰視点からの透視図法で、スクリーン面から3D物体が飛び出したかのような印象を与え、かつ体験者の挙動も同時に静止画または動画で記録・提示することが可能なシステムを開発した。同様なアイデアは、複合現実感分野で「客観視点映像」と呼ばれ、現実世界の映像と仮想世界の映像を合成して表示することが実現されている[4]。複数人がシースルーHMDを装着し、複合現実空間を共有しながら同時体験できるシステムでは、客観視点映像を得るカメラを1ユーザと見なして増やすだけで容易に実現できる。しかしながら、スクリーン面に両眼視差のある立体映像を表示する投射型システムでは、別の問題が生じ、実現は単純ではない。

本稿では、この問題に対して、我々がどのような方法でこれを解決したかについて述べる。これもある種の複合現実感システムとなっている。このシステムは、単に成果発表等に用いるだけでなく、3Dインタラクションを含むヒューマンインタフェース(HI)分野の研究ツールとしても極めても有用であることにも言及する。

## 2. HI 研究ツールとしての客観視点映像に求められる要件

本研究では、大型スクリーンへの投射型の立体映像装置を利用する場合の体験者と立体映像の3Dインタラクションを客観視点映像化し、それをHIツールとして利用することを想定している。具体的には、以下のような用途を想定している。

- (a) 未体験者へのシステムおよびその操作方法の教示
- (b) 体験者の操作方法、内容の確認
- (c) 各体験者間での操作方法、内容の違いの分析
- (d) 体験者自身が自分の動作を客観的に確認

図1のような視点から、単純にビデオ撮りした場合、操作者の挙動は捉えることはできるが、スクリーン内の映像は二重映しになり、第3者には何が表示され、何をどのように操作しているのかは正確には把握できない。複数人が複合現実空間を共有できる複合現実感システムの場合、図2のような鳥瞰位置からの客観視点映像を利用することができる。これは、同じ世界座標内に単眼のカメラを設置し、カメラの位置姿勢を与えて、現実世界の画像と仮想世界の画像を実時間で合

成しているからである。

この方式にヒントを得て、図3のように操作対象の3Dオブジェクトがスクリーンから飛び出したかのような印象を与える映像を得る方法を考える。それには、スクリーン位置を世界座標内で表現し、かつ両眼立体視映像を生成するコンピュータから表示すべき映像を取り出して、この世界座標内で単眼映像として配置できるシステムを別途構築する必要がある。



図1 立体映像提示装置のスクリーンをそのままビデオ撮影した場合

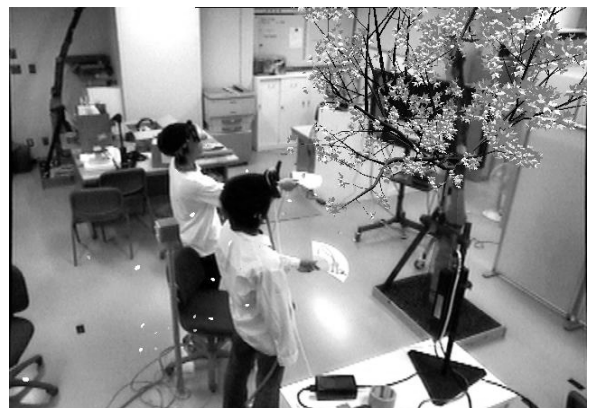


図2 HMDを利用した複合現実感システムにおける客観視点映像(桜の花がCG)



図3 3D物体の飛び出し感を与える表示(パースを合わせた嵌め込み合成図)

さらに、上記の(a)-(d)のような用途で利用するためには、以下の要件を満たすことが望ましい。

(1) 客観視点位置（カメラ位置）の可変性

体験者の操作を観察する際、目的によって体験者に近い位置から観察したい場合もあれば、体験者の背後から広角で観察したい場合もある。そこで、客観視点位置（カメラの位置）を容易に変更可能であることが望ましい。

(2) 立体映像表示装置との映像同期

体験者が立体映像を操作している様子を、リアルタイムで観察するためには、プロジェクタで投影されている映像と、客観視点映像で生成される映像がずれることなく同期して生成される必要がある。

(3) 客観視点映像の録画機能

後で改めて各体験者の操作風景を比較したり、体験者自身が自分の動作を客観的に確認するためには、客観視点映像を静止画や動画で保存する必要がある。

(4) スクリーンに投影された映像の消去

図2のような場合と異なり、プロジェクタ投影方式の立体映像装置では、スクリーンにはすでに体験者が両眼立体視を行うための映像が投影されている。よって、客観視点カメラからの映像にそのままCGを合成すると、スクリーンに投影された映像と、新たに合成したCGの両方が表示されることになる。体験者が見ている映像と同じものを客観視点映像化するためには、まず客観視点カメラの画像からスクリーン上に投影された画像を“消す”作業を行った後に、CG合成を行う必要がある。

### 3. 対象となるシステムと客観視点合成

#### 3.1. システム構成

図4に本研究が対象としたシステムの構成を示す。広視野没入型ディスプレイとしては、高さ1.8 m、弧長6.1 mの大型アーチ状スクリーンを用いている。計3台のプロジェクタ（CDS社製のMirage6000 DLP）を用いることで、アーチ状スクリーンの正面、左、右の部分に投影している。正面、左、右の映像が重なる部分は湾曲補正装置を用い、境目をブレンディング処理している。Mirage6000 DLPは時分割立体視に対応している。このステレオ映像を、液晶シャッター眼鏡（CrystalEYES社CrystalEYES3）を用いて見ると、赤外線エミッタによる信号に応じて、液晶メガネのシャッターが開閉し、右目と左目用の映像が交互に映し出され、両眼立体視が可能である。3台のプロジェクタで出力される映像はそれぞれ右面、正面、左面用のグラフィックス・ワークステーション（HP社xw8000, OS: Windows2000 Professional, CPU: Intel Xeon 3.06GHz×2, メモリ: 4.0Gbyte, グラフィックボード: 3Dlab

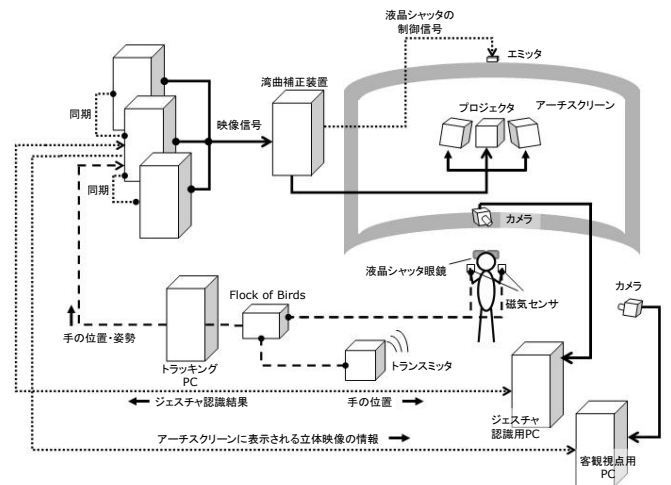


図4 システム構成

Wildcat4 7210) で同期を取りながら生成される。

本システムでは立体映像と3Dインタラクションを行う方法として、ジェスチャの認識入力を採用している。このジェスチャ入力の詳細に関しては、文献[3]を参照されたい。

客観視点映像の合成用には、DELLのDimension8300（OS: Windows XP Professional, CPU: Intel Pentium4 2.60GHz, メモリ: 1.0Gbyte, グラフィックボード: NVIDIA GeForce FX 5200）を用いてシステムを構成した。また、客観視点から見た実世界映像の入力には、SONYのHandycam DCR-HC90を用いた。ビデオカメラとPCはIEEE1394ケーブルで接続しており、動画はリアルタイムでPCに入力される。

客観視点映像用PCと立体映像表示用グラフィックス・ワークステーションは、LANケーブルで接続し表示データの同期をとる。客観視点映像用PCの開発環境はVisual C++で、映像の合成はOpenGLライブラリを用いて実行する。

#### 3.2. キャリブレーション方法

前述したように客観視点映像を生成するためには、客観視点カメラから捉えた映像に、立体映像装置に表示するデータを、幾何学的に変換して重畳する。そのためには、事前に立体映像装置の座標系とカメラ座標系とのキャリブレーションを行うことが必須である。図5に本研究で取り扱う座標系を示す。

本研究では客観視点の位置を容易に変更可能とするために、実世界にマーカを配置し、まず世界座標系とマーカ座標系を合わせ、次にカメラ座標系とマーカ座標系を合わせることで、カメラ座標系と世界座標系の位置合わせを実現する。

世界座標系は、立体映像表示装置が描画するCG座標系と原点を一致させ、鉛直上向きをYw軸、床面をXwZw平面とする。マーカ座標系は右手系で表現され、回転・平行移動が可能でマーカの大きさ・位置をユー

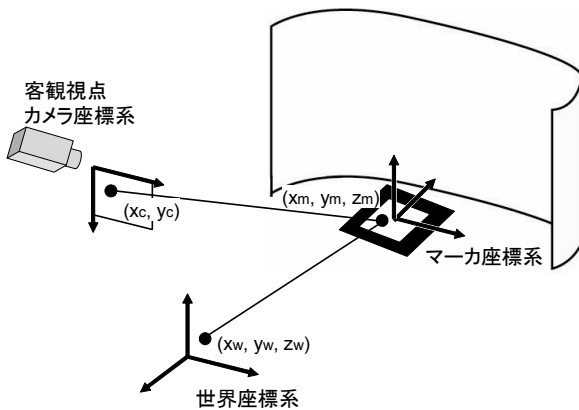


図 5 キャリブレーションで扱う座標系

ザ入力によって与えることができる。よって、世界座標系での配置場所を与え、座標系を回転させて世界座標系との位置合わせを行う。

カメラ座標系からマーカ座標系への変換行列は、加藤等の提案する ARToolKit ライブラリ[5]を用いて求める。カメラの入力画像からフレームを静止画として切り出し、ARToolKit のマーカ検出を行う。マーカはカメラ位置が遠い場合でも認識できるように 70cm×70cm の大きめのものを採用し、スクリーン中央下に配置した。ARToolKit ライブラリでは、まずテンプレート・マッチングで検出したマーカの種類・回転を求め、その後、カメラから見たマーカの位置姿勢を取得し、その逆行列を計算することでマーカから見たカメラの位置姿勢が計算される。そこに世界座標系のマーカの位置姿勢行列を乗算することで、世界座標系でのカメラの位置姿勢を表わす行列を求めることができる。

一旦このキャリブレーションを行えば、客観視点カメラを固定している間は安定した幾何学的整合が保たれるので、操作体験時にはマーカは必要でない。しかし、カメラ位置をしばしば移動して客観視点映像を得る場合には、瞬時に位置合わせが実行できるよう、このマーカは常置しておく方が望ましい。

## 4. 遭遇した問題への対処

### 4.1. スクリーン投射映像の除去処理

こうして得られた変換行列を利用して、体験者が見るスクリーン上の映像と同じ情報を、カメラ座標系の 2 次元平面に重畳描画することで客観視点映像を生成する。しかし、ビデオカメラから取得した映像にそのまま CG 映像を重ね合わせると、図 6 のようにスクリーンに投影されたステレオ映像がそのまま映ってしまうという不都合が生じる。これは、図 2 のような HMD を用いるシステムにはなく、スクリーン投射型のシステムゆえに生じる問題である。



図 6 スクリーンに投射映像が残っている場合

この問題を解決するためには、スクリーン面には何も映っていないかのように、投射映像を除去する処理を施す必要がある。我々は、プロジェクタからスクリーンに何も投影されていない状態のカメラ画像からスクリーン領域を検出・保存し、スクリーン全体を意図的にこの状態に保つ方法を採用することにした。

この処理は、以下の手順で達成される。

- (1) 客観視点映像でスクリーン面に相当する領域内をクリックすると、クリックされた画素とその 4 近傍で色の平均が計算される。
- (2) その平均された色とウィンドウ上の画素ごとに色の差分を取り、それが規定の閾値未満ならその画素を記憶する。
- (3) その記憶した各画素で構成される領域に対し収縮・膨張演算を施し、連結成分のラベリング処理を実行する。ここで収縮・膨張演算を行うのは、記憶された画素での雑音処理の役目とラベリングの処理軽減のためである。
- (4) ラベリング処理で得られた領域の中で、クリックした点を含む領域をスクリーン領域と見なしてフラグを立てる。
- (5) カメラ映像をキャプチャして客観視点映像を更新する際に、このフラグのない画素のみを映像更新することで所望の除去処理が達成される(図 7)。これは、スクリーン面のみを「白抜き」にした効果を与えることに相当する。

### 4.2. 処理の手順

次に客観視点映像を合成する手順を説明する。

- (1) マーカをスクリーン手前の世界座標系で規定した位置に置く。現システムでは、スクリーンをどの方向から写したときもマーカが捉えられるよう、マーカの中心からスクリーンまでの距離が 85.0cm の位置、 $(x_w, y_w, z_w) = (0.0, 0.0, -211.0)$  [cm]を採用している。
- (2) 客観視点カメラの映像を静止画として切り出す。この画像を用いて 3.2 で述べたキャリブレーションを行ない、4.1 で述べたスクリーン投射映像の除去処理



図 7 スクリーン投射映像の除去処理後



図 9 他の視点からの客観視点映像

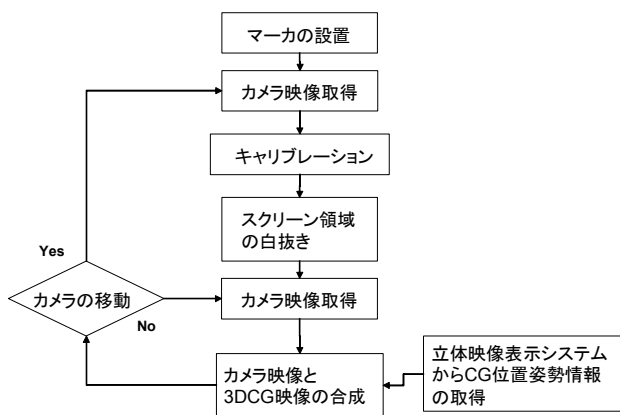


図 8 客観視点映像の合成手順

を実行する。

以上の (1) (2) の手順は、カメラの設置またはカメラ位置の変更をした時に、最初に 1 度だけ行えば十分である。これは、操作体験中は原則として客観視点カメラを移動させないことを前提としていて、キャリブレーションで求めた世界座標系でのカメラの位置姿勢を表す行列の更新をする必要がないからである。

(3) この前処理の後には、客観視点位置での映像合成処理が自動的に図 8 のように繰り返し実行される。

カメラの視点位置を変えたい場合は、(2) の手順を繰り返せば良い。体験者の体形には個人差があるので、図 9 に示すように、ジェスチャ操作やその反応が見やすい位置にカメラを移動させることができる。

また、これらの手順により合成された客観視点映像は OpenGL のウィンドウ上に表示され、キーボード操作によりそのまま BMP の静止画か AVI 形式の動画ファイルとして保存することが出来る。

### 4.3. 3D オブジェクトの強調表示

図 7, 図 9 の画像は、いずれも世界座標系内に 3D オブジェクト (これらの図では、複数の 2 次元画像が 3 層構造の表示面に貼り付けられている) を正しい透視

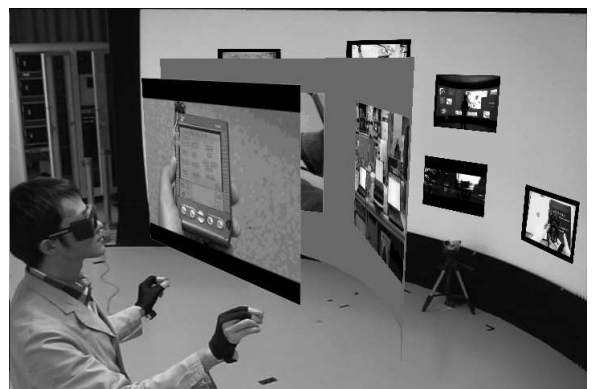


図 10 3D オブジェクトを強調表示した結果

図法で描いた結果である。しかしながら、一般的感覚としては、飛び出した 3D オブジェクトが随分小さく感じてしまう。体験者自身は、目の前に 3D オブジェクトが迫って来ると、視野の大半を占めるので遠近感を強く感じるが、客観視点でその様子を見てもその実感はない。

図 3 のイメージ図と比べれば、その感覚の違いが分かるだろう。これはパースのみを合わせて、大きさは考えずに嵌め込み合成した画像であるが、一般的にはこうした描きの方が、飛び出し感を与えることができる。立体ディスプレイの宣伝資料でも同様な手法が採られていることが少なくない。

これは、スクリーン面に投射される二重像がスクリーン上で占める大きさが、そのまま平行移動して描かれる方が、印象としては心地良いということを意味している。ある種の錯視現象である。ならば、その錯覚を利用し、正しい当時図法での大きさよりも強調して 3D オブジェクトを表示することが考えられる。

本システムでは、3D オブジェクトの重心位置とスクリーン中心を結ぶ線を中心軸として、3D オブジェクトの拡大強調処理を行なった結果を表示する機能をもたせた。拡大の倍率に関しては、パラメータで調整できるが、スクリーン面から体験者方向に飛び出す位置

に応じて大きさに矛盾がないよう、整合性は保たれている。図 10 にこの強調表示結果を示す。

## 5. 本システム運用と考察

従来の複合現実感システムでは、客観視点カメラは定位置で運用という固定観念があった。本システムでは即時キャリブレーションにより簡単に移動できる機能をもたせたことから、運用面での利便性が格段に向上した。

ただし、全く任意の位置で客観視点映像を取得して意味がある訳ではない。図 11 に示すような位置では、白抜きしたスクリーン面に操作者の手の挙動が隠されてしまうし、表示オブジェクトとも交錯する。これは、映像合成時に CG 表示が必ず前面描画され、オクルージョンの前後判定を行っていないためである。実時間レンジファインダを導入すれば、操作者のジェスチャの幾何形状データが得られるので、この問題が解決できるがコスト的に難がある。カメラ設置位置を上手く選択する方が得策である。

同様に、この客観視点映像を撮影している場合は、照明条件に関しても配慮が必要である。画像の視認性だけを考えれば、体験者はもっと暗い環境を好むが、プレゼンテーションや 3D インタラクションの様子を分析することが目的であれば、ある程度の犠牲は止むを得ない。ただし、3D-CG で描くオブジェクトは照明条件に関わらず明るく表示できるので、操作者をスポットライトで局所的に照らすことで、少し暗い環境にすることも可能である。

## 6. むすび

投射型立体映像ディスプレイを用いた 3D インタラクションを客観視点からの 2D 映像として可視化・記録するシステムに関して報告した。

本システムは、当初研究成果の外部でのプレゼンテ

ーション用に開発したものであるが、実際にはむしろ HI 研究のツールとしての価値が高いことが判明した。ジェスチャ・インタフェースの評価実験時に被験者に手本を教示したり、各体験者の挙動の違いや操作の円滑さを分析するのにも極めて有用である。本システムを活用して、今後、本来の「広視野電子作業空間のヒューマンインタフェースに関する研究」を加速させ、新たな知見が得られ、より良いインタフェースを産み出すことを促進したい。

## 謝辞

本システムの開発・運用・実験等に関しては、研究室在籍の学生、鬼柳牧子・上野由佳・前野恭平・平沼真吾らの諸君の協力を得た。ここに感謝の意を表します。本研究の一部は、(財)国際コミュニケーション基金と(財)大川情報通信基金の研究助成による。

## 文 献

- [1] G. Kurtenbach and G. Fitzmaurice (eds): Special Issue on Application of Large Display, IEEE Computer Graphics and Applications, Vol.25, No.4, pp.22-65, 2005
- [2] 中嶋正之: “CAVE -新しい立体映像空間”, 電子情報通信学会誌, Vol.80, No.8, pp.888-890, 1997
- [3] 木村朝子, 鶴田剛史, 酒井理生, 鬼柳牧子, 柴田史久, 田村秀行: “広視野電子作業空間に関する考察とシステム試作～マイノリティ・レポート型 I/F とその発展形”, インタラクション 2005 論文集, pp. 143-150, 2005
- [4] 田村秀行, 内山晋二: “MR プラットホーム・システム”, 日本 VR 学会誌, Vol.8, No.1, 2003
- [5] 加藤博一, Mark Billinghurst, 浅野浩一, 橋啓八郎: “マーカー追跡に基づく拡張現実感システムとそのキャリブレーション”, 日本 VR 学会論文誌, Vol.4, No.4, pp.607-616, 1999



図 11 あるカメラ位置では操作者の手が隠蔽される