

立石賞功績賞の受賞記念講演概要

時空を超える複合現実メディアへの挑戦

～リアルとバーチャルの融合～

奈良先端科学技術大学院大学 名誉教授 横 矢 直 和

学生の受け入れ開始と同時に奈良先端科学技術大学院大学に異動した翌年の1994年度に、立石科学技術振興財団から「画像理解のための並列協調型アルゴリズムの研究」に対して助成をいただき、研究室の立ち上げに支援いただきました。また、この度は、四半世紀余りを経て、当時に構想していた研究内容に対して第6回立石賞功績賞をいただくこととなり、感謝に堪えません。ここに、立石文雄理事長をはじめ立石科学技術振興財団の皆様並びに立石賞選考委員会の皆様に厚く御礼申し上げます。

本講演では、「時空を超える複合現実メディアへの挑戦」と題しまして、受賞対象となった研究の概要についてお話させていただきます。具体的には、最初に、「時空を超える複合現実メディア」の基盤となる複合現実感技術について簡単に述べ、次に、中心的な話題である拡張現実感とテレプレゼンスについて、「タイムマシンと千里眼を実現する」ことを目指して研究室で取り組んできた内容とシステム開発を通じた実証実験をご紹介します。

複合現実の世界

複合現実感(MR: Mixed Reality)は、バーチャルリアリティとコンピュータビジョンの境界領域に、1990年代に現れたメディア技術で、最近では一般的なものになりつつあります。図1の左端が実写映像で表現される現実世界で、右端はすべてCG映像で表現される仮想世界です。複合現実感はこの両者の間に連続的に存在する、リアルとバーチャルの融合空間です。

実写映像にCGを重ね合成したものが「拡張現実」(AR: Augmented Reality)で、現実世界の映像や3次元計測により仮想世界を作り出すのが拡張仮想

想(AV: Augmented Virtuality)あるいは仮想化現実(Virtualized Reality)と呼ばれる技術です。後者の中で、実写映像のみから実際にあたかも遠隔地にいるかのような感覚を与えるメディア技術はテレプレゼンス(Telepresence)と呼ばれます。本日のお話は拡張現実とテレプレゼンスが中心になります。



図1 複合現実の世界

時空を超える複合現実メディアのイメージ

研究室で「時空を超える複合現実メディア」を標榜していた頃に描いていた技術要素と応用展開を図2に示します。このメディアを構成する要素は、広域(屋外)環境の3次元計測とユーザの位置・姿勢の計測からなる実環境のセンシングと、場所に関連した過去・現在・未来



図2 安全・安心・快適に係る生活の質向上を目指す時空を超える複合現実メディア

のユーザに提示したい情報です。これらの情報を現場で提示するのか、遠隔地で提示するかによって、6つの組み合わせができます。

拡張現実感を実現するための技術課題

拡張現実感を実現するための現実世界と仮想世界の合成方法は、光学シースルー方式とビデオシースルー方式の2つに分れます(図3参照)。光学シースルーでは、ハーフミラーを通して現実世界の像とCGが網膜に入り、網膜上で像が合成されます。一方、ビデオシースルーでは、カメラで撮影した現実世界の像とCGがコンピュータ上で合成され、合成像が網膜に入ることになります。我々の研究室では主にビデオシースルー方式を採用してきました。

拡張現実感では、ユーザの視点移動に実時間で追従する、実写とCGのシームレスな合成が必要となり、解決しなければならない課題は、現実世界と仮想世界の位置ずれをなくす幾何的整合性と現実世界と仮想世界の光学的な矛盾を

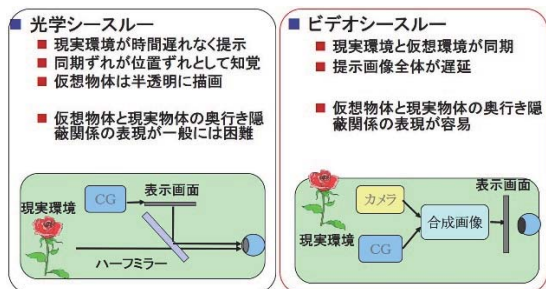


図3 拡張現実感における現実世界と仮想世界の合成方法

表1 拡張現実感の技術課題

	計測の目的	計測する対象
幾何的整合性	現実世界と仮想世界の位置合せ	ユーザ視点の位置と方向
	実物体と仮想物体の奥行き隠蔽表現	ユーザ視点から見た現実世界の奥行き
光学的整合性	実物体と仮想物体の陰影合せ	現実世界の光源環境
	実物体と仮想物体の相互影づけと映り込み表現	現実世界の3次元形状、表面反射特性、光源環境

なくす光学的整合性です(表1参照)。

ビデオシースルー方式において幾何的整合性を確保するためには、ユーザ視点(すなわちカメラ)の位置と姿勢のリアルタイム計測が必要で、光学的整合性の確保には、現実世界の光源環境のリアルタイム計測が必要になります。ビデオシースルー方式での幾何的位置合わせの基本は、カメラで撮影された映像そのものを解析することによってカメラの位置・姿勢を推定する手法です。これには、コンピュータビジョンの分野におけるPnP(Perspective n-Points)問題(図4参照)を解くことによってカメラの位置・姿勢を推定する手法が一般的です。図4において、3次元空間内の特徴点 p の世界座標系での3次元位置 $S_p=(x, y, z)$ が既知で、画像上での検出座標が $x_p=(u, v)$ 、カメラの位置・姿勢を表す変換行列 M による画像平面への投影座標を $x'_p=(u', v')$ とします。PnP問題は、一般に6点以上、平面上の点なら4点以上の既知の点が画像上で検出できれば解けることが知られており、一般には、画像上の検出座標 x_p と投影座標 x'_p の距離の二乗和として定義され

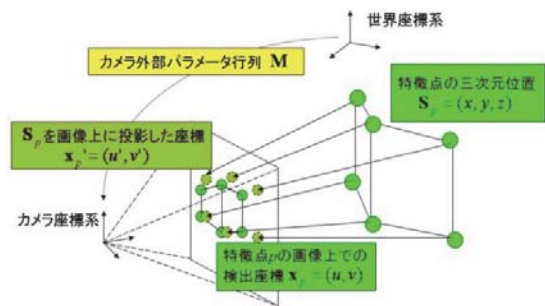


図4 3次元位置が既知の特徴点からのカメラの位置・姿勢推定

る再投影誤差

$$E = \sum_p \|x'_p - x_p\|^2$$

を最小にする行列 M を算出し、カメラの位置・姿勢を決定します。

この原理に基づいて、マーカを用いて幾何的位置合わせを行う最も有名なソフトウェアツールは ARToolKit です。

我々の研究の特徴は、マーカを用いずに、カメラ映像からシーン内の自然特徴点の実時間検出・追跡を行いカメラの位置・姿勢を推定するアルゴリズムを開発し、ARの屋外利用への道を拓いたことです。我々のアプローチでは、事前に、全方位映像から特徴点の3次元位置と周辺の画像パターンなどを格納したランドマークデータベースを構築し、オンライン処理では、特徴点をリアルタイムに検出し、データベース内のランドマークとの対応付けを行うことにより、幾何的位置合わせを実現します（図5参照）。

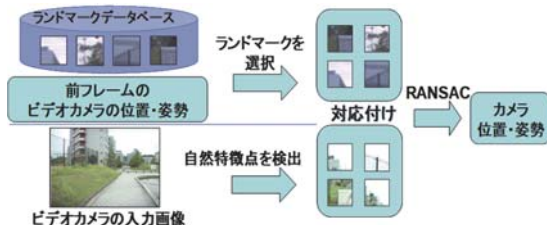


図5 ランドマークデータベースを用いたカメラの位置・姿勢推定

拡張現実感システムの開発：タイムマシン

奈良県内の史跡を対象に、すでに失われた建造物をCGで復元し、先に述べた幾何的位置合わせ法により現場でAR合成を行うことによって、時間を超えてバーチャルに歴史体験を可能とする複数の拡張現実感システムを開発しました。明日香村と東大寺境内で行った一般公開の実証実験のときの様子を図6、図7に示します。



AR映像提示デバイスとしてHMD（左）と小型PC（右）を用いたシステム。この実験では、事前に全方位カメラと全方位レンジファインダを用いて実験地域のランドマークデータベースを構築している。

図6 明日香村・川原寺跡地でのバーチャル飛鳥京の一般公開実験の様子（協力：明日香村、東京大学生産技術研究所・池内研究室）



創建時の大仏殿や喪失して現存しない東塔（七重の塔）などのバーチャル復元を実現した。

図7 東大寺境内での一般公開実験の様子（協力：東大寺、奈良県）

テレプレゼンス

あたかも遠隔地に実際にいるかのような感覚を提示するテレプレゼンスについて述べます。研究の特徴は全方位カメラを用いることです。基本的には、全方位画像でとらえた遠隔地の映像をもとにユーザの視線に追従した平面透視投影画像をリアルタイムに生成・提示するというものです（図8参照）。仮想カメラのパン・チルト・ロール・ズーム操作を実現し、ビデオ映像のインタラクティブ視聴を可能にしました。

この空間を超えるテレプレゼンス研究の初期に利用した全方位カメラ HyperOmni Vision を図9に示します。



図8 全方位画像を用いるテレプレゼンス方式
(撮影協力：NHK エンタープライズ)

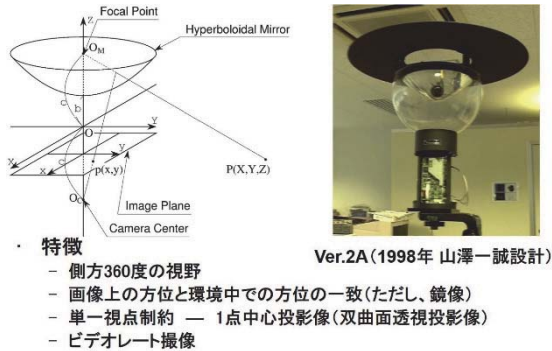


図9 全方位ビデオカメラ HyperOmni Vision



(1) 1本の全方位ビデオを4人が独立に同時視聴できる蓄積再生型システム



(2) 実時間ネットワークテレプレゼンスシステム (NAIST—ハワイ大学間での実証実験, 2007年1月, 協力：ハワイ大学)

図10 HyperOmni Visionを用いたテレプレゼンスシステムの事例

テレプレゼンスシステムの開発：千里眼

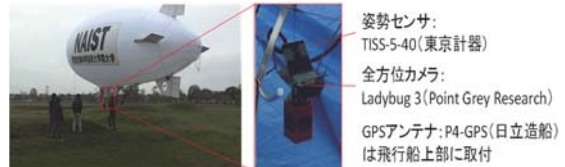
テレプレゼンスシステムの実装方法として、映像提示に関しては蓄積再生型と実時間型、システム構成に関してはスタンドアロン型、ネットワーク型、放送型の3種類を試みました。図10に2つの具体的な事例を示します。

拡張テレプレゼンス：千里眼+タイムマシン

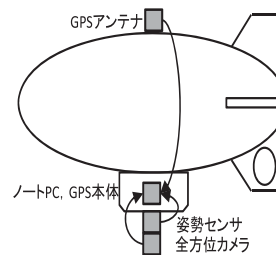
ここでは、飛行船からの空撮全方位映像を用いたバーチャル歴史体験を目的として、前述の拡張現実感とテレプレゼンスを組み合わせた蓄積再生型の拡張テレプレゼンスに関する取り組み「フライスルーバーチャル平城京」を紹介します。

(1) 全方位空撮システム

空撮に使用した無人飛行船と搭載機器を図11に示します。6眼の全方位ビデオカメラユニット Ladybugに加えて、空撮映像に対する



(1) 無人飛行船と搭載機器の外観



(2) 機器の搭載位置

図11 空撮用の無人飛行船と搭載機器

AR合成を行うための現実世界と仮想世界の幾何学的位置合せに必要なカメラの位置・姿勢情報を取得するために、GPSと姿勢センサなどを搭載しています。飛行船の操縦は地上からのリモコン操作により行います。

(2) 前処理

空撮映像に対する CG の AR 合成を行う際には、幾何的位置合わせのために、空撮映像の各フレームでのカメラの位置・姿勢情報が必要であり、ここでは、Structure-from-Motion (SfM) 技法と GPS 測位を併用したハイブリッド手法によりカメラの位置と姿勢を推定します。一般に、SfM では蓄積誤差が発生するとともに、相対的な位置・姿勢情報しか得られずスケールが未知という問題があります。このため、GPS 測地座標系という絶対座標系での位置情報が得られる GPS 測位情報を利用することによって、絶対座標系上でのカメラの位置・姿勢を推定します。この情報を用いて、飛行船の揺れに伴う全方位映像の歪みの補正と死角領域の補完を行った後に、全方位映像の各フレームに CG の AR 合成を行います。

(3) フライスルーバーチャル平城京：タイムマシン＋千里眼

平城遷都 1300 年祭に合わせて、平城宮跡を上空から全方位カメラで空撮し、往時の平城京の建物の CG を AR 合成した映像をもとに、上空から 1300 年前の都を鳥瞰できる拡張テレプレゼンスシステム「フライスルーバーチャル平城京」を開発し、1300 年祭において一般公開を行いました（図 12、図 13、図 14 参照）。



図 12 全方位 AR 画像



図 13 フライスルーバーチャル平城京のユーザ提示画像



(1) 通常のディスプレイ

(2) HMD

図 14 平城遷都 1300 年祭におけるフライスルーバーチャル平城京のデモ風景（協力：社平城遷都 1300 年記念事業協会、凸版印刷（株）、文化庁）

おわりに

本講演では、時空を超える複合現実メディアとして、拡張現実感とテレプレゼンスおよびそれらを統合した拡張テレプレゼンスに関する取組をご紹介しました。現在、これらの技術は本格的な実利用の時代に入ったと感じていますが、ここでは触れなかった厳密な光学的整合性の実時間での確保など、解決が難しい問題が残っています。また、リアルとバーチャルをシームレスに融合する技術の進歩に伴って、フェイク映像などの新たな問題も現れてきますので、今後は、これらの問題に取り組む必要があります。

最期に、本講演内容に関係する研究を中心的に担った研究室の元学生・スタッフおよび協力いただいた奈良先端大の同僚に感謝いたします。