

# 三次元 画像センシングの 新展開

—— リアルタイム・高精度に向けた要素技術から産業応用まで

第3編 実用事例

第1章 映像・サービス分野

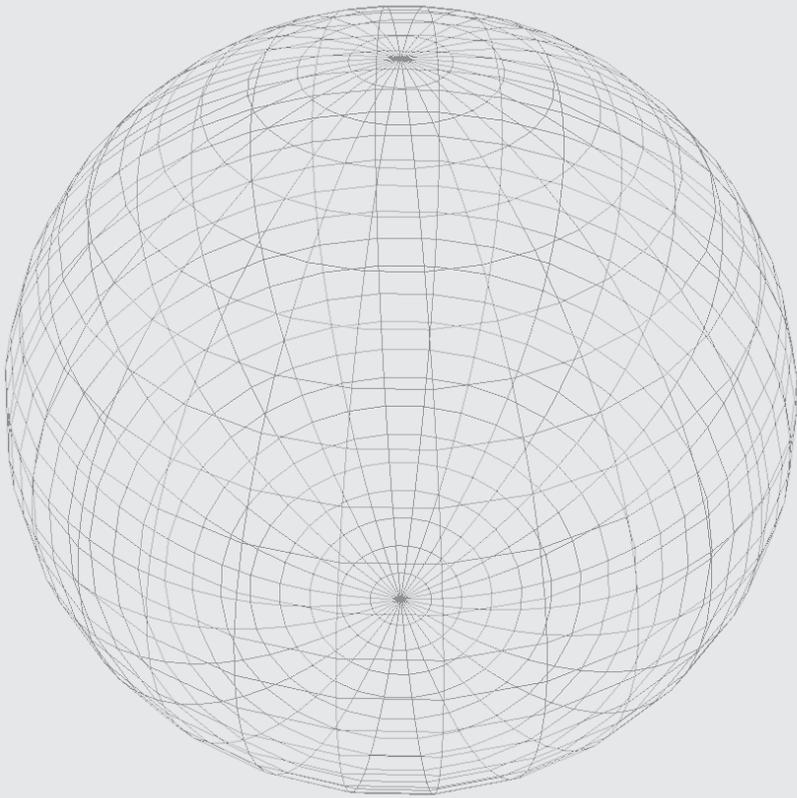
第1節 三次元計測と複合現実への応用

奈良先端科学技術大学院大学 佐藤 智和

奈良先端科学技術大学院大学 横矢 直和

## 第 3 編

# 実用事例





## 第1節 三次元計測と複合現実への応用

奈良先端科学技術大学院大学 佐藤 智和 奈良先端科学技術大学院大学 横矢 直和

### 1 はじめに

複合現実（Mixed Reality；MR）とは、カメラで撮影される実世界の情報とCGで描かれる仮想世界の情報を融合し、これをディスプレイを通じてユーザーに提示する新たな情報提示手段である。MRによって構築される世界は、図1に示すように、現実世界と仮想世界の間に存在する連続的な概念である。MRは、技術的には実写画像にCGを重畳合成する拡張現実（Augmented Reality；AR）と、実世界の計測情報を取り込むことで仮想世界を構築する仮想化現実（Augmented Virtuality；AV）に大別できる。

このようなMRの実用事例として、スマートフォンをかざすことで利用者周辺の各種案内情報などをカメラ映像に合成提示するARアプリケーションや、Google Earth, Google Street Viewに代表される実写画像群により仮想世界を構築・再現するAVアプリケーションが開発され、一般に利用されはじめている。以下では、高品位なMRを実現するために筆者らのグループが実施した研究事例について、AR, AVに分けて紹介する。また、新たなMRの形態として近年注目されている隠消現実（Diminished Reality；DR）についても紹介する。

### 2 三次元計測と拡張現実（AR）への応用

ARを実現するためには実世界と仮想世界（CG）の位置合わせが必要不可欠であり、一般的にこの位置合わせには実写画像を取得するカメラの撮影位置・姿勢情報が用いられる<sup>1)</sup>。AR実現のためのカメラ位置・姿勢推定手法としては、その簡便さから、カメラで撮像されたマーカを



図1 複合現実感が創り出す世界

手がかりとして用いる方法<sup>2)</sup>やGPS・ジャイロ・コンパスなどのセンサ情報を用いる手法が広く利用されている。前者は、最近の携帯ゲーム機におけるARアプリケーションなどに採用されているが、大量のマークを配置することが困難な広域環境での利用は難しい。これに対して、広域環境での利用が前提となるスマートフォン用のARアプリでは、一般に各種内蔵センサによってカメラの位置・姿勢を推定するが、計測誤差が大きく合成対象の位置ずれが目立つという欠点がある。ここでは、これら従来型システムの問題点の解決を目指して開発された、自然特徴点ランドマークを用いたARシステムと、Indirect-ARシステムを紹介する。

## 2.1 自然特徴点ランドマークデータベースを用いたARシステム

筆者らのグループは、システムの利用対象となる環境をあらかじめ全方位動画画像として撮影し、これを解析することにより得られる自然特徴点の三次元情報をランドマークデータベースとしてマークの代わりに用いることで、マークを用いずにカメラの絶対位置・姿勢を推定するシステムを開発した(図2)<sup>3)</sup>。本手法では、まずオフライン処理として、シーンを撮影した全方位動画画像に対してStructure from Motion法を適用することで画像上に存在する自然特徴点の三次元位置と全方位カメラの動きを復元する。ここでは、GPSや三次元位置が既知の外部指標を用いることで、三次元復元における誤差の蓄積を抑止する。復元された自然特徴点の三次元位置および特徴点の見え方を表す画像パターンは、ランドマーク情報としてデータベースに格納される。オンライン処理では、モバイル端末上で撮影された実画像上の自然特徴点を抽出し、データベースに格納されたランドマークとの照合を行うことで対応点を決定し、対応するランドマークの三次元座標を用いてデータベース構築時に定義された座標系におけるモバイル端末のカメラの位置・姿勢を推定する。図2(b)、(c)は、奈良県明日香村において実施した実証実験において、本手法を実装した試作システム利用時に入力画像上で検出されたランドマークと、これにより推定さ

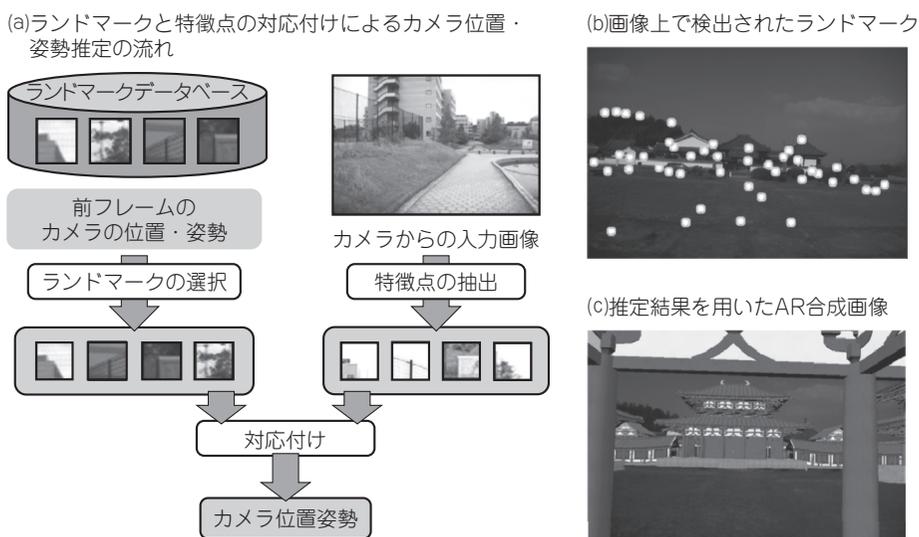


図2 自然特徴点ランドマークデータベースを用いたARシステム



図3 Indirect-ARにおける幾何位置合わせ

れたカメラ位置・姿勢を用いて生成されたAR合成画像の一例である。

## 2.2 Indirect-AR システム

近年、Indirect-AR と呼ばれる新たな AR 画像提示手法が提案されている<sup>4)</sup>。Indirect-AR では、対象となるシーンの情景を事前に全方位カメラなどにより取得し、その画像上にあらかじめ仮想物体を合成しておく。AR システム利用時には、コンパス・ジャイロなどのセンサ情報を用いてオンラインで端末姿勢を取得し、事前に生成した AR 合成画像の適切な箇所をディスプレイ上に提示する。この方式では、図3に示すように、ユーザーに提示される画像上においては原理的に事前撮影画像と仮想物体の間の位置ずれが生じず、またオンラインでの処理コストが小さく頑健性が高いためモバイル型の AR システムに適している。なお、提示される事前撮影画像と実環境の位置合わせにはセンサを用いるため比較的大きな位置ずれが生じるが、図4に示すように、多くのモバイル端末の撮影インタフェースでは、すでにユーザー視点とカメラ視点の間に大きなずれが生じており、このような撮影方式に慣れたモバイル端末の利用者は、事前撮影画像と実環境の位置ずれにほとんどの場合気づかない<sup>4)5)</sup>。ただし、本手法は視点移動に対応しておらず、ユーザーが事前撮影画像取得地点から大きく移動した場合には、実環境と提示画像間に大きな不整合が生じるという課題がある。

図5は、試作した Indirect-AR システムを用いて東大寺で実施した一般公開実験の様子である<sup>5)</sup>。本システムでは、AR 画像上に存在する動物体をあらかじめ消去しておくことで、事前の画像撮影時およびシステム利用時における他者の存在を気にすることなくコンテンツを鑑賞できる。なお、Indirect-AR では、事前の画像取得時とシステム利用時において、天候が異なる場合など画像上に大局的な変化が生じる場合において臨場感が低下する。筆者らが提案したシステム<sup>6)</sup>では、図6に示すように、あらかじめ様々な天候に対応した AR 画像を作成しておき、システム利用時にはカメラから得られる画像に照明条件に近い AR 画像を選択提示することで、臨場感の低下を防いでいる。

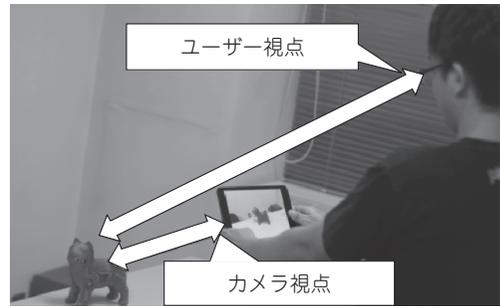


図4 一般的な撮影インタフェースにおける視点位置のずれ



図5 東大寺での一般公開実験の様子

### 3 三次元計測と仮想化現実（AV）への応用

Google Earth, Microsoft Virtual Earth などの実環境の計測データに基づいて構成された三次元地図アプリケーションは AV の実用事例であるが, 一般に広域環境を対象として高精度な三次元計測を行うことは難しく, 自動での三次元復元を行う場合には形状の欠損・変形, テクスチャの歪み等の問題が生じ, 任意の視点位置における写実的な映像再現を行うことが難しい<sup>7)</sup>。一方, Street-view などの全方位画像を用いたテレプレゼンスシステムは, 地図上で選択された撮影地点における実写画像をそのままユーザーに提示することで実写に基づく仮想空間を構築しているといえるが, 三次元地図アプリケーションのように視点位置を自由に変更することはできない。これら双方の問題を解決すべく, 筆者らのグループは全方位画像に加えて三次元形状情報を用いることで写実的な任意視点の画像生成を実現し, 仮想世界を自由な視点から観察できるシステムを構築している。以下, シーンの三次元形状を利用した自由視点画像生成による AV の実現例として, 視覚的タイムマシンを実現する自由視点型拡張テレプレゼンスシステムと, 現在の遠隔地の様子を自由な視点から観察する千里眼型ロボット操縦支援システムについて紹介する。

#### 3.1 自由視点型拡張テレプレゼンスシステム

ここでは, 飛行船からの全方位空撮映像を用いたバーチャル歴史体験システムとして構築した「フライスルーMR 平城京」<sup>8)9)</sup> について紹介する。本システムでは, 現在の平城宮跡の全方位映像に奈良時代の建物を合成し, 空から当時の情景を仮想体験できる。これを実現するために, 現在の平城宮跡の情景を飛行船に搭載した全方位カメラを用いてビデオ映像として取得し, Structure from Motion 法と GPS を併用することでカメラの位置・姿勢情報を復元した。また, Multi-view Stereo 法の最新の実装である CMPMVS<sup>10)</sup> を用いてシーンの形状情報を復元している (図7)。

本システムでは, 背景となる実写画像について, 仮想視点のカメラ位置に依存して最適なテクスチャを選択的に貼り付ける視点依存テクスチャマッピング (View-dependent Texture Mapping: VDTM) 法<sup>11)</sup> を用い, 図8上に示すような写実的な自由視点画像を生成し, これにCGで再現された平城宮を重畳合成することで, 同図下に示すような平城宮跡の様子を仮想空間内に視覚的に再現している。

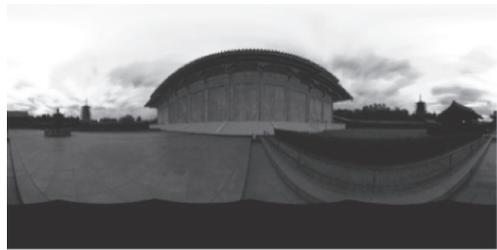


図6 異なる天候に対応する事前生成 AR 画像の例

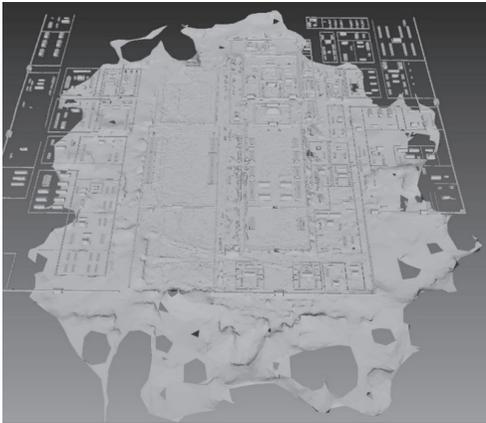


図7 空撮映像から推定されたシーンの形状とそれに位置合わせされた仮想物体



図8 仮想空間内に再現された平城宮  
上：建物合成前，下：合成後

### 3.2 自由視点画像生成による遠隔ロボットの操縦支援システム

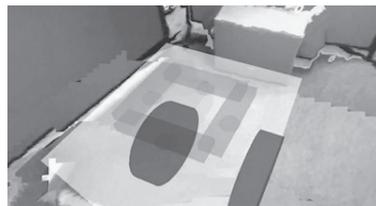
遠隔地から送られてくるリアルタイムの実映像を用いたテレプレゼンスシステムは、遠隔会議や遠隔ロボットの操縦などに用いられる。筆者らはこのようなテレプレゼンスシステムを拡張し、三次元情報を用いることで、遠隔地のロボット周辺の状況を自由な視点から観察できる新たな遠隔ロボット操縦支援システムを開発した(図9)<sup>12)</sup>。本システムで用いるロボットには、4台のRGB-D距離画像センサと全方位カメラが搭載されており(図9(a))、これらを用いて全方位画像と三次元の奥行き情報を取得する。取得した情報は、インターネット経由で遠隔地にいる操縦者に送られ自由視点画像生成に用いられる。操縦者はヘッドマウントディスプレイを装着しており、これに取り付けられた磁気センサによって追跡される操縦者の頭の動きに応じて仮想視点位置が決定され、映像がレンダリングされる。ここでは、図9(b)に示すように、自由視点画像上にロボットを表すCGをAR合成しており、本システムを用いることで、カメラ位置から死角となる領域の確認や、ロボットと周辺物体の位置関係を把握することが容易となる。

本システムでは、ユーザーの頭部の位置・姿勢に応じて決定される仮想視点に対応する自由視点画像を実時間で生成しているが、ここでは欠損のない自由視点画像を生成するために、計測情報から視点位置に依存した奥行き画像を逐次生成する視点依存ジオメトリ(View-dependent

(a)遠隔地のロボット



(b)拡張自由視点画像



(c)操縦風景



図9 自由視点画像生成による遠隔ロボットの操縦支援システム

Geometry ; VDG) 法と, [3.1] で述べた VDTM 法を組み合わせて利用している。ただし, 実時間での描画を実現するために VDG で用いる奥行き画像を低解像とする必要があり, 試作システムによる提示画像にはブロック状のアーティファクトが生じている。このような問題を解決すべく, 筆者らのグループでは実時間処理を実現しながら, より高品位な自由視点画像を生成する手法について開発を継続している。

#### 4 三次元計測と隠消現実 (DR) への応用

近年, 不要な現実物体を視覚的に消去する隠消現実 (DR) に関する研究が行われている。DR を用いることで, 例えば図 10 に示すような, AR システムにおけるマーカの視覚的な隠蔽・消去や, 図 11 に示すような, 画像中の不要物体の視覚的消去が実現できる。筆者らのグループでは, 特別な事前準備なしに利用可能な DR システムを実現するために, インペインティング手法を用いる隠消現実システムを開発している<sup>13)14)</sup>。インペインティングは, 画像上で与えられた対象領域をそれらしいテクスチャで埋めることで視覚的に消去・修復する技術であり, これまで二次元の写真を対象としたさまざまな手法が提案されてきた。

筆者らが開発した手法では, マーカを消去する場合にはマーカが存在する平面を用い, マーカが存在しない場合には Visual SLAM 法<sup>15)</sup> によって推定される三次元構造を用いてシーンを複数の平面に分割し, 平面上のテクスチャに対するインペインティング処理を適用することで, 高品位な DR 画像の生成を実現している。具体的には, システム実行時の特定のキーフレームにおいて, シーン中の平面を鉛直方向から見た画像を射影変換により生成し, 変換された画像上に射影される対象領域を従来のインペインティング手法<sup>16)</sup> を用いて修復する。このようにして修復された画像を, 復元されたシーンの構造を考慮しながら不要物体上に逐次テクスチャとして貼り付けることで, 視点変更時においても幾何的に整合した DR 画像の生成を実現している。

#### 5 おわりに

本稿では, カメラや各種センサにより得られる三次元情報を用いた MR システムの構築例について紹介した。スマートフォン上で動作する MR アプリケーションの普及により, MR は身近なものとなりつつあるが, 本稿で紹介した手法を応用することで, より高品位・高機能な MR アプリケーションの実現が可能であると考えら

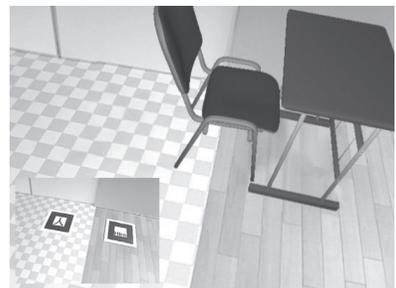


図 10 AR 家具配置シミュレーションシステムにおけるマーカの視覚的消去の例

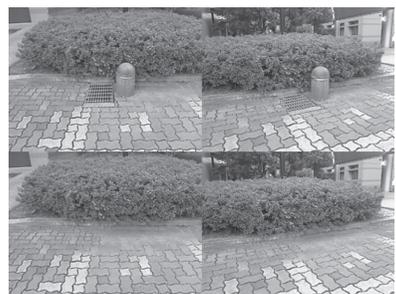


図 11 一般的なシーンを対象とした DR の実現例

上: 入力画像, 下: DR 画像

れる。また、近年の距離画像センサの小型化・低価格化などに伴って、将来これらがモバイル機器に内蔵され、三次元情報の取得がより安定かつ容易となることで、三次元情報を用いたさまざまなMRアプリケーションが登場することが期待される。筆者らは今後、より臨場感の高いMRシステムを開発するために、実時間での高品位な自由視点画像生成手法の開発に取り組むとともに、Indirect-ARに自由視点画像生成手法を組み込むことで、視点移動に対応したより実用的なARシステムを開発する予定である。

## 文 献

- 1) 佐藤智和, 横矢直和: VR/MRのための画像計測, 計測と制御, **47** (1), 30-35 (2008).
- 2) H. Kato and M. Billinghurst: Marker tracking and HMD calibration for a video-based augmented reality conferencing system, Proc. IEEE/ACM Int. Workshop on Augmented Reality (IWAR99), 85-94 (1999).
- 3) T. Taketomi, T. Sato and N. Yokoya: Real-time and accurate extrinsic camera parameter estimation using feature landmark database for augmented reality, *Int. Journal of Computers and Graphics*, **35** (4), 768-777 (2011).
- 4) J. Wither, Y. Tsai and R. Azuma: Indirect augmented reality, *Int. Journal of Computers and Graphics*, **35** (4), 810-822 (2011).
- 5) T. Akaguma, F. Okura, T. Sato and N. Yokoya: Mobile AR using pre-captured omnidirectional images, Proc. Symposium on Mobile Graphics and Interactive Applications, 26:1-26:4 (2013).
- 6) F. Okura, T. Akaguma, T. Sato and N. Yokoya: Indirect augmented reality considering real-world illumination change, Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR2014), 287-288 (2014).
- 7) 佐藤智和, 横矢直和: 仮想化現実空間構築のための画像計測, 光技術コンタクト, **9** (47), 36-41 (2009).
- 8) 横矢直和: 時空を越える拡張テレプレゼンス〜フリスルーMR 平城京〜, JACIC 情報, **26** (3), 62-67 (2011).
- 9) F. Okura, M. Kanbara and N. Yokoya: Mixed-reality world exploration using image-based rendering, *ACM Journal on Computing and Cultural Heritage*, **8** (2), 9:1-9:26 (2015).
- 10) M. Jancosek and T. Pajdla: Multi-view reconstruction preserving weakly-supported surfaces, Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR2011), 3121-3128 (2011).
- 11) P. Devevec, C. Taylor and J. Malik: Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach, Proc. ACM SIGGRAPH, 11-20 (1996).
- 12) F. Okura, Y. Ueda, T. Sato and N. Yokoya: Free-viewpoint mobile robot teleoperation interface using view-dependent geometry and texture, *ITE Trans. on Media Technology and Applications*, **2** (1) 82-93 (2014).
- 13) N. Kawai, M. Yamasaki, T. Sato and N. Yokoya: AR marker hiding based on image inpainting and reflection of illumination changes, Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR2012), 293-294 (2012).
- 14) N. Kawai, T. Sato and N. Yokoya: Diminished reality considering background structures, Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR2013), 259-260 (2013).
- 15) G. Klein and D. Murray: Parallel tracking and mapping for small AR workspaces, Proc. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR2007), 13-16 (2007).
- 16) N. Kawai, T. Sato and N. Yokoya: Image inpainting considering brightness change and spatial locality of textures and its evaluation, Proc. Pacific-Rim Symposium on Image and Video Technology (PSIVT2009), 271-282 (2009).