

# 解説

## 広域実環境のモデリング

Modeling of Large-Scale Real Environments

横 矢 直 和\* \*奈良先端科学技術大学院大学

Naokazu Yokoya\* \*Nara Institute of Science and Technology

### 1. はじめに

近年、実世界の情報処理が不可欠な自律移動ロボットや実世界とバーチャル世界を融合することにより新しい情報提示手段を提供する複合現実感 (MR: Mixed Reality) において、広域実環境の三次元モデリングへの要求が高まっている。具体的には、移動体による環境認識と自己位置同定や複合現実感におけるユーザの位置・姿勢推定のための自然特徴点を用いたランドマークデータベース構築、デジタルアーカイブに代表される仮想化現実 (AV: Augmented Virtuality) のモデルベースアプローチ [1] [2] において、実環境モデリングが必須となる。これらの応用では、モデリングと位置・姿勢推定が密接に関係している場合が多い。

ロボット分野においては、ロボットが未知環境のモデリング (地図作成) と自己位置推定を同時に行うために、SLAM (Simultaneous Localization and Mapping) と呼ばれる手法が開発されてきた [3] [4]。基本的に、距離センサやカメラなどから得られる情報によってオドメトリを補正するもので、リアルタイムでの地図作成も可能になり、一部は実用的なレベルに達している。しかし、リアルタイム性を要求すると、オフラインの全体最適化処理を取り入れることはできず、長距離移動した際に誤差が蓄積するため、通常の SLAM の広域モデリングへの応用は難しい。

コンピュータビジョンの分野において、環境モデリングは、初期の物体認識研究以来の古典的な課題であり、能動的な距離センサであるレーザレンジファインダや単眼移動カメラで撮影した動画像を用いる Structure-from-Motion の枠組み [5] による三次元環境復元が行われてきた。SLAM や MR におけるユーザ位置・姿勢推定、AV のモデルベースアプローチでもこれらの手法が基本ツールとして用いられている。

以下、本稿では、主として屋外の広域環境を対象として筆者らの研究室で取り組んできた、多視点距離画像計測に

よる三次元環境モデリングおよび Structure-from-Motion 法を用いたランドマークデータベース構築とその MR 応用の事例を紹介する。

### 2. 多視点距離画像計測による広域三次元モデリング

全方位レンジファインダ、全方位カメラ等からなる車載型センサシステムを用いて広域屋外環境のテクスチャ付き三次元モデル生成を行った事例 [6] [7] を紹介する。

#### 2.1 車載型センサシステム

センサシステムの外観を図 1 に示す。機器構成の概要は以下のとおりである。

- (a) 全方位レンジファインダと全方位カメラ：いずれも周囲 360 度を同時に計測・撮影できる全方位センサである。それぞれのデータ取得位置の差が小さくなるように近づけて設置し、事前に、手動で与えた両データの対応点からセンサ間の位置関係を推定し、座標系の統合を行っている。
- (b) 位置・姿勢センサ：レンジファインダとカメラの位置・姿勢情報を取得するための機器であり、ジャイロセンサと RTK-GPS (Real-Time Kinematic GPS) からなる。この二つのセンサは一体型のハイブリッドセンサになっており、両センサの座標系は自動的に統合される。また、センサシステムの移動時には、RTK-GPS の測位値の履歴を用いてジャイロセンサの蓄積誤差を自動的に補正する機能を備えている。

レンジファインダで計測できる位置にジャイロセンサを

全方位カメラ(Point Grey Research, Ladybug) RTK-GPS(ニコントリプル, LogPakII)



図 1 車載型センサシステムの外観

原稿受付 2009 年 12 月 24 日

キーワード: 3D Modeling, Localization, Mapping, Structure from Motion, Landmark Database, Augmented Virtuality

\*〒 630-0192 生駒市高山町 8916-5

\*Ikoma-shi, Nara

設置し、ジャイロセンサ座標系で位置が既知のジャイロセンサ上のマークをレンジファインダで計測することによって (a) と (b) の座標系の関係を決定する。これによって、すべてのセンサ座標系を GPS 座標系へ統合し、位置・姿勢情報付きの全方位距離画像と全方位カラー画像をほぼ同地点で取得することができる。

### 2.2 多視点距離画像の位置合わせ

広域環境のモデリングは、(1) 多地点での距離データとカラー画像の取得、(2) 複数距離データの位置合わせ、(3) 統合された形状モデルへのテクスチャマッピングの三つのステップからなる [6]。 (1) については前節で紹介した車載型センサシステムを用いて環境内を移動しながら、多数の地点で全方位距離画像と全方位カラー画像を同時に取得する (ただし、計測時には車両は停止)。 (2) では三次元点群データの位置合わせ法として考案された ICP (Iterative Closest Point) アルゴリズム [8] を平面同士の対応付けへと拡張した手法を用いる。 (3) では環境内でのオクルージョンを考慮した上で解像度が最も高くなる全方位画像からテクスチャを生成する。

ここで、上記のステップ (2) における距離データの位置合わせは、RTK-GPS とジャイロセンサによって得られている各距離データ取得時の位置・姿勢を初期値として、逐次位置合わせによる位置合わせ誤差の蓄積を防ぐために、以下の反復手順による同時位置合わせを行う。

1. 距離データから平面領域を検出
2. 距離データを GPS 座標系へ座標変換
3. 平面の対応関係を全距離データ間で探索
4. 対応する平面の法線と平面領域の計測点から距離データの回転・平行移動成分を推定し変換行列を更新
5. 変換行列の推定結果が収束していれば終了、収束していなければ 2.へ

### 2.3 多視点計測のためのビュープランニング

広域環境モデリングでは、多数の地点での逐次的な計測が必要になるため、未計測部分を効率よく削減し、最適な距離データ取得位置を決定するビュープランニングが重要である。このため、データ取得の過程で次の計測候補位置に関する情報 (推奨度マップ) を操作者に提示する、以下の機能を有するデータ取得支援システム [7] を開発している。本システムは前述の車載型システムの利用を前提としており、センサシステムの設置位置を道路上に限定している。

1. 計測を行うごとに、距離データとボクセル表現された既計測データとの位置合わせを行い、モデルを更新する。位置合わせには基本的に 2.2 節の手法を用いるが、処理時間短縮のために、新たな距離データと取得位置が近いデータのみを用いて同時位置合わせを行う。
2. センサシステムを設置可能なすべての地点について、対象シーンの計測密度を指標にして、距離計測の推奨

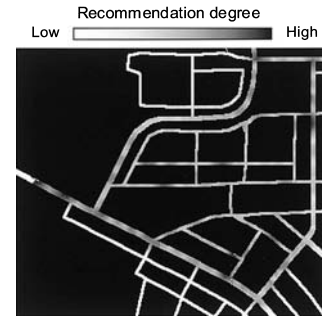


図 2 操作者に提示される推奨度マップの例 (図 3 と同じ地域)

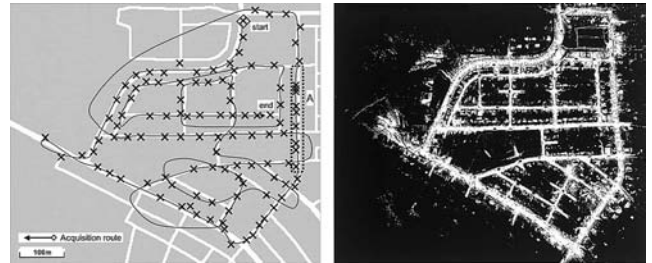


図 3 距離画像計測の位置・経路 (左) と計測結果 (右)

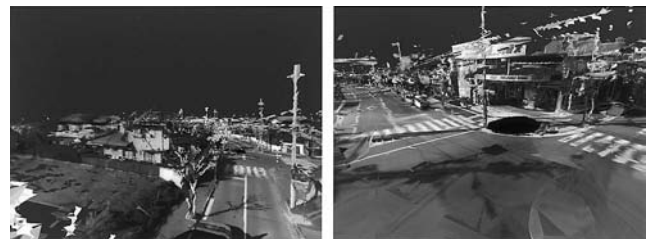


図 4 市街地のテクスチャ付き三次元モデル

度を計算する (詳細については文献 [7] 参照)。

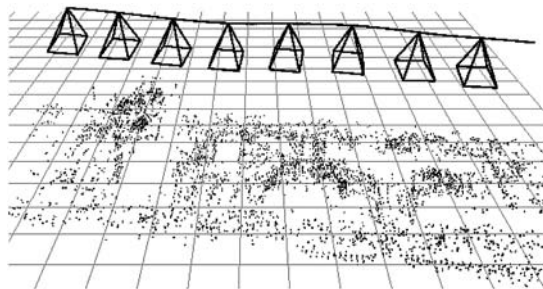
3. 数値地図データを用いて道路の領域に推奨度を描画した推奨度マップを生成し、操作者に提示する (図 2 参照)。センサシステムの操作者は、提示された推奨度マップから、推奨度の高さ、移動経路・時間等を考慮して次の計測地点を決定する。

### 2.4 住宅地域の三次元モデル構築

実際に生駒市内の住宅地域において前述の手法を用いて環境モデリングを行った例を紹介する。図 3 に地図上の計測位置・経路と対象範囲全域での計測結果を示す。この例では、データ取得支援システムを用いて計 112 地点での全方位計測・撮影を行っている。最終的に得られたテクスチャ付き三次元モデルの描画例を図 4 に示す。

## 3. 動画像からの広域三次元モデリング

広域環境を対象とした画像計測や環境モデリングでは、単眼カメラによる移動撮影に基づく Structure-from-Motion 法 (以下、SFM と呼ぶ) [5] を用いることが多い。SFM 法は、動画像上で特徴点追跡を行い、特徴点の三次元座標と



(a) 推定されたカメラの位置・姿勢と特徴点の三次元位置



(b) 生成されたテクスチャ付き三次元モデル

図 5 Structure-from-Motion と動的輪郭モデルによる空撮映像からの三次元モデリング

カメラの位置・姿勢を同時に推定する手法である。動画像の各フレーム撮影時のカメラの位置・姿勢が得られるという特徴から、SFMは様々な画像計測の基本ツールとして用いられている。以下では、SFM法による広域モデリングの事例を紹介する。

### 3.1 SFMによる三次元モデリング

SFM法は、基本的に粗な特徴点上でしか三次元情報が得られないが、後処理としてマルチベースラインステレオ法や動的輪郭モデルを適用することによって、密な三次元形状を得ることができる。図5(a)は、奈良先端大上空からの空撮動画像にSFM法を適用し、カメラの位置・姿勢と自動抽出した特徴点の三次元位置を推定した結果である。この三次元特徴点に対して、地上からレンジファインダを用いて計測した三次元点群データの位置合わせを行い、動的輪郭モデルを用いて面を構成することによって三次元ポリゴンモデルを生成することができる[9]。このようにして生成した三次元モデルに対して空撮画像のテクスチャマッピングを行った結果を図5(b)に示す。ただし、ここでは三次元ポリゴンモデルの各パッチごとに、対応する画像の解像度が最も高いフレーム画像を動画像中から選択して、パッチのテクスチャを生成している。

### 3.2 SFMによる自然特徴点ランドマークデータベースの構築とMR応用

自然環境における移動体の自己位置同定やMRにおけるユーザ視点の位置・姿勢推定では、環境中の自然特徴点をランドマークとして利用するビジョンベースの手法が有効

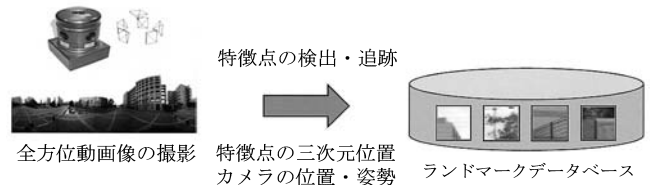


図 6 ランドマークデータベース構築の流れ

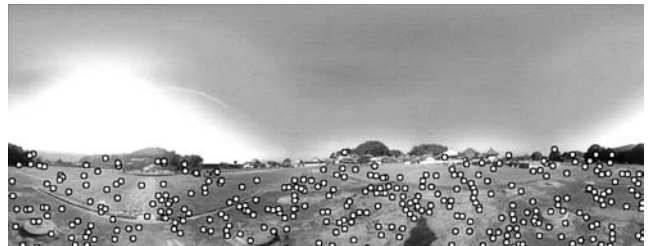


図 7 パノラマ展開された全方位画像上のランドマーク (○印)

である。以下では、全方位動画像を用いたSFMによるランドマークデータベースの構築とMRにおける幾何的位置合わせへの応用について述べる。

#### 3.2.1 ランドマークデータベースの構築

SFM法は原理的にスケール情報を復元できない。すなわち、近くの小さな物と遠くの大きな物を区別できない。環境認識や自己位置推定ではスケール情報が必要になるため、三次元位置が既知の基準点またはGPSから得られる絶対位置情報を用いることによってSFMを補完し、スケール情報付きのランドマークデータベースを構築する。

具体的には、まず、図1のセンサシステムでも用いた全方位カメラの移動撮影により、全方位動画像を取得し、自然特徴点の自動検出・追跡を行い、SFM法によってリアルタイムに、(1)特徴点の画像上の座標と三次元位置および(2)撮影時の全方位カメラの位置・姿勢情報を取得する。次に、三次元位置が既知の特徴点(基準点)またはGPSから得られる絶対位置情報を利用して、累積的な推定誤差を最小化するための動画像全体での最適化処理を行う。この段階でスケール情報が復元される。最後に、特徴点をランドマークとして、各ランドマークについて以下のような情報を保持したデータベースを構築する[10](図6参照)。

- ランドマークの三次元座標
  - 撮影位置に依存した情報(多重スケールの画像テンプレート、SIFT特徴[11]などの画像特徴、撮影位置など)
- 明日香村・川原寺でのランドマークデータベース構築の例を図7に示す。全体では約3,500のランドマークからなるが、同図では、ある時刻の全方位画像上で検出され、ランドマークとして登録された点を○印で示している。

#### 3.2.2 ランドマークデータベースを用いた位置・姿勢推定—MRへの応用—

移動体に搭載されたカメラで撮影した画像からランドマー

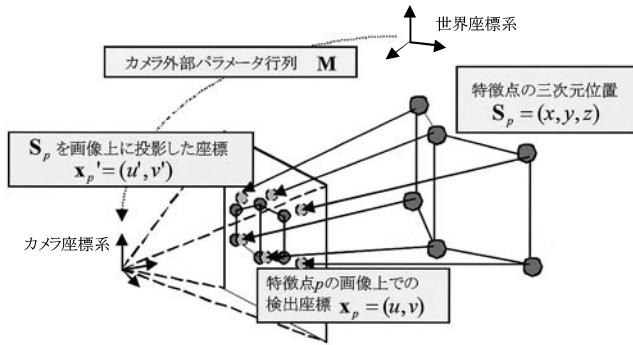


図8 カメラの位置・姿勢を推定する PnP 問題

クデータベースを用いて、移動体の位置・姿勢（厳密にはカメラの位置・姿勢）を推定する方法について述べる。三次元位置が分かっているシーン中の十分な数の点が画像に映っていれば、画像上の座標からカメラの位置と姿勢を推定することができる。このカメラの位置・姿勢推定問題は、シーン中の既知の特徴点を利用して世界座標系からカメラ座標系への変換行列  $M$  を求める PnP (Perspective n-Points) 問題に帰着する。図8において、特徴点  $p$  の世界座標系での三次元位置  $S_p = (x, y, z)$  が既知で、画像上での検出座標が  $x_p = (u, v)$  である。また、 $S_p$  を変換行列  $M$  で画像平面へ投影した点の座標を  $x'_p = (u', v')$  で表している。PnP 問題は、一般には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、画像上での検出座標  $x_p$  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・姿勢を決定する。

具体的な応用として、カメラでとらえた実環境の映像に対して、検出された特徴点とランドマークデータベースの照合によって、実世界とバーチャル世界の幾何的位置合わせを行い、CG 映像を重畳合成してユーザにリアルタイムに提示する MR システムを開発している。図9に、明日香村・川原寺の史跡現場で2種類の MR システムを用いて移動しながら飛鳥時代のバーチャル歴史体験を行っている様子を示す。

#### 4. ま と め

本稿では、広域環境を対象とした多視点距離計測による三次元環境モデリングと SFM 法を用いたランドマークデータベース構築について述べた。具体的な事例として MR 応用を紹介したが、これらの技術はロボットや自動車などの移動体による環境認識や自己位置同定に利用可能である。



図9 MR による史跡現場でのバーチャル歴史体験 (CG の基礎データは東京大学・池内研究室提供)：提示装置として HMD (左) と小型 PC (右) を用いたシステム

#### 参 考 文 献

- [1] K. Ikeuchi and D. Miyazaki (eds.): Digitally Archiving Cultural Objects. Springer, 2008.
- [2] 佐藤智和, 横矢直和：“VR/MR のための画像計測”, 計測と制御, vol.47, no.1, pp.30-35, 2008.
- [3] H. Durrant-Whyte and T. Bailey: “Simultaneous localization and mapping: Part I,” IEEE Robotics & Automation Magazine, vol.13, no.2, pp.99-108, 2006.
- [4] T. Bailey and H. Durrant-Whyte: “Simultaneous localization and mapping (SLAM): Part II,” IEEE Robotics & Automation Magazine, vol.13, no.3, pp.108-117, 2006.
- [5] S. Ullman: The Interpretation of Visual Motion. The MIT Press, 1979.
- [6] 浅井俊弘, 神原誠之, 横矢直和：“全方位距離画像と全方位カラー画像の統合による屋外環境の三次元モデル化”, 画像電子学会誌, vol.34, no.5, pp.529-538, 2005.
- [7] 浅井俊弘, 神原誠之, 横矢直和：“屋外環境の三次元モデル化のための推奨度マップを用いたレンジデータ取得支援システム”, 電子情報通信学会論文誌 (D), vol.J92-D, no.4, pp.531-541, 2009.
- [8] P.J. Besl and N.D. McKay: “A method for registration of 3-D shapes,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.14, no.2, pp.239-256, 1992.
- [9] 北市泰寛, 佐藤智和, 横矢直和：“屋外環境の異種計測データを対象とした動的輪郭法による三次元モデル生成”, 計測自動制御学会第72回パターン計測部会研究会資料, pp.21-28, 2007.
- [10] 武富貴史, 佐藤智和, 横矢直和：“拡張現実感のための優先度情報を付加した自然特徴点ランドマークデータベースを用いた実時間カメラ位置・姿勢推定”, 電子情報通信学会論文誌 (D), vol.J92-D, no.8, pp.1440-1451, 2009.
- [11] D.G. Lowe: “Distinctive image features from scale-invariant keypoints,” Int. Journal of Computer Vision, vol.60, no.2, pp.91-100, 2004.



横矢直和 (Naokazu Yokoya)

1979 年大阪大学大学院博士後期課程修了, 工学博士。電子技術総合研究所 (現産業技術総合研究所) 主任研究官, カナダ・マギル大学客員教授を経て, 1993 年奈良先端科学技術大学院大学教授, 現在に至る。画像処理, コンピュータビジョン, 複合現実感などの研究に従事。情報処理学会論文賞 (1989 年度, 2006 年度) など受賞。電子情報通信学会および情報処理学会フェロー。IEEE, 映像情報メディア学会, 日本バーチャルリアリティ学会などの会員。