

SfM による 3 次元復元結果と 地図データの位置合わせに関する研究

新井大地 鳥居秋彦 奥富正敏
東京工業大学

あらまし 本研究では、SfM (Structure from Motion)を用いた建物の 3 次元復元結果と地図データの位置合わせを行う手法を提案する。はじめに、3 次元復元結果を重力方向へ投影し、点群の位置と、カメラと点群を結ぶ視線情報を用いて建物の境界らしさを表す画像を作成する。次に、地図データから建物の境界を表す画像を生成する。そして、建物の境界を表す 2 つの画像から目的関数を定義し、勾配法を用いて最適化することにより位置合わせを行う。実際に SfM から得られる 3 次元復元結果を用い、建物の初期位置を与えることで、十分な精度で位置推定が可能であることを示す。

1. 序論

コンピュータビジョンの分野において、多視点の画像を用いたオブジェクトやシーンの 3 次元形状復元は盛んに研究されている課題である。多視点の画像群から特徴点の 3 次元位置とカメラの位置・姿勢を推定する SfM (Structure from Motion) は、Bundler[1, 2] や VisualSFM[3] といったシステムが提案され、近年では複数の画像を用いることで、比較的容易に 3 次元復元を行うことが可能になりつつある [9]。

SfM は画像のみを入力として用いるため、別々に作成された 3 次元復元結果はそれぞれ独自の座標系で表される。原理上、画像のみを入力として実スケールを求めることは不可能である。複数の 3 次元復元結果をひとつの座標系に統合し、実スケールを求めることができれば、SfM の有用性が大きく向上する。

そこで本研究では、地図を基準として SfM による 3 次元復元結果の統合することで、復元結果の統合とスケールの推定を同時に行うことを目標とする。

2. 関連研究

Shan らは、航空写真に対して SfM による 3 次元復元結果の位置・方向・スケールを推定することで、3 次元復元結果の実座標系への統合を行った [4]。技術的な特徴としては、SfM による 3 次元復元結果と斜めの視点から撮影された航空写真とのマッチングの手法を提案にある。SfM による 3 次元復元結果から航空写真の視点から見たような画像を作成し、鮮鋭化した航空写真に対して、SIFT[5]を用いた特徴点マッチングを行うことで、通常よりもロバストにマッチングを行っている。

Bruall らは博物館のような大規模な室内のシーンに対して、場所ごとに作成された SfM による 3 次元復元結果を地図に配置する手法 [6] を提案している。注釈のついた室内地図と Google 画像検索を用いて 3 次元モデルの地図上の候補

の位置を推定し、地図のフロアプランの形状に対して 3 次元モデルの方向・スケールが最適になるように目的関数を設定し最適化を行う。また写真の時系列から、撮影者がどの順番で室内を巡回したかを考慮し、3 次元モデルのより良い位置・方向・スケールを推定する。

Kaminsky らは SfM による 3 次元復元結果と衛星画像の位置合わせの手法 [7] を提案している。3 次元復元結果を平面上に投影した 2 次元点群、カメラ座標、また衛星画像を入力として、2 次元点群を衛星画像上に幾何変換し、位置合わせを行う。幾何変換後の 2 次元点群から衛星画像上のピクセルまでの距離を表す Edge-cost、カメラの視線上には物体が存在しないことを考慮した Freespace-cost の 2 つの相違度を表す目的関数を定義し、Coarse-to-Fine の手法を用いて探索を行い、最適な幾何変換行列を推定する。

3. 提案手法

3.1 システムの概要

提案手法の全体の流れについて説明する。SfM により 3 次元復元された建物の 3 次元点群と撮影した際のカメラの位置と方向、また複数の建物が含まれている地図画像を入力とする。

はじめに、入力された SfM のデータを地図画像と比較するために、SfM から得られる 3 次元復元結果から、建物の境界らしい部分を抽出する。入力の 3 次元点群を重力方向に投影することにより 2 次元点群に変換する。次に、2 次元点群を格子状に分割して、格子内部に含まれる点群の数を画素値とする点群画像を作成する。

ここで、SfM に用いる画像には建物の境界でない部分も映っているため、復元された点群も建物の境界でない点群が含まれている。これを取り除き建物の境界のみを検出するために、カメラとそのカメラで撮影した画像に映っている復元点までの間の空間はフリースペースと考えられることを利用する。2 次元点群を格子状に分割し、格子内部を

通る視線の本数を画素値とする視線画像を作成する。点群画像と視線画像を組み合わせることで2次元点群から視線が多く通る部分を除くことにより、建物の境界らしさを表す画像を作成する。

続いて、入力された地図画像から建物の境界の部分から抽出した画像を作成する。その上で、3次元復元結果から算出した境界らしさを表す画像と、地図画像から抽出した境界画像を用いた位置合わせを行うために、2つの画像の類似度を表す目的関数を設定し、ユーザーが入力した初期値に対して勾配法を用いて目的関数の最適化を行う。

各処理については以下で詳細に述べる。

3.2 SfMによる3次元復元結果から建物の境界らしさを表す画像の作成

3.2.1 3次元点群の2次元平面への投影

Ricoh THETAなどのカメラを用いる場合は機器に加速度センサなどの重力方向を記録できる機能が備わっているが、一般的なカメラを用いる場合には重力方向を推定する必要がある。ここでは重力方向を推定する手法として、対象を撮影した際の画像座標系における水平方向ベクトルは重力方向ベクトルに対して垂直であるという仮定のもと、全ての画像に対して画像座標系における水平方向ベクトルを抽出し、その全てのベクトルに対して最も垂直らしいベクトルを重力方向とする Szeliskiの手法[8]を用いる。

推定した重力方向ベクトルを用いて、3次元点群を重力方向に投影し、2次元点群を作成する。

3.2.2 復元点の量を表す点群画像の作成

ある正方形領域に含まれるように2次元点群全体の位置・スケールを正規化し、さらに正方形領域を格子状に分割して正方形小領域を定義する。各正方形小領域を1つの画素と考え、領域内に存在する復元点の数を画素値とした画像を作成する。以下ではこれを点群画像と呼ぶ。実装では400×400ピクセルの画像に変換する。

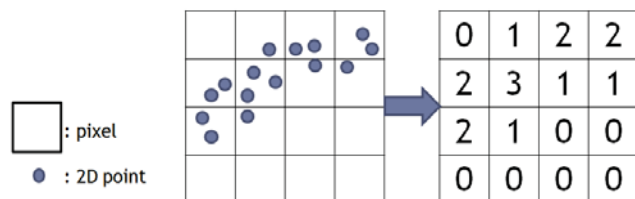


図1 点群画像の作成

3.2.3 フリースペースを表す視線画像の作成

建物を撮影した各画像のカメラの位置から、その画像内に映っている全ての復元点に対してカメラの位置から線分を引き、各正方形小領域を通る線分の本数を画素値とした画像を作成する。以下ではこれを視線画像と呼ぶ。ただし、復元点を含むピクセルについては、フリースペースでないと考えられるので数えないものとする。実装では400×400ピクセルの画像に変換する。

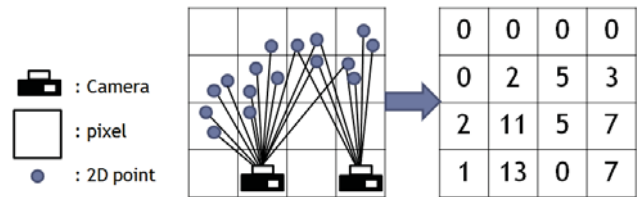


図2 視線画像の作成

3.2.4 境界らしさを表す画像の作成

復元された点の量を表す点群画像とフリースペースを表す視線画像を組み合わせることで、建物の壁らしい部分を抽出した、境界らしさ画像を作成する。

建物の境界らしさを表す部分を抽出するためには、点群が存在しかつ視線が通らない部分、つまり点群画像で高い値かつ視線画像で低い値を持つピクセルを抽出すればよい。よって視線画像の値を反転し、点群画像との積を値とする境界らしさ画像を作成する。実装では点群画像、視線画像ともに画素値を0から1の値に正規化した後、各画素に以下の式を用いて境界らしさ画像を作成する。数式は(1)のように表される。

$$I_{Boundary}(i, j) = I_{point}(i, j) * \{1 - I_{Freespace}(i, j)\} \quad (1)$$

この処理は図3を例にとると、図3(b)の点群画像で黒く、かつ図3(c)の視線画像で白いピクセルのみを抽出することにより、図3(d)のように建物の境界らしい部分を抽出する処理となっている。

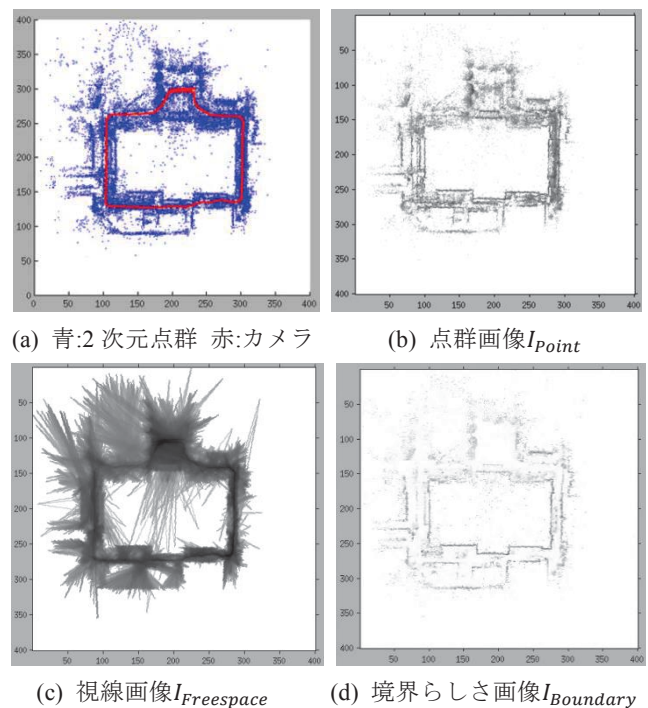
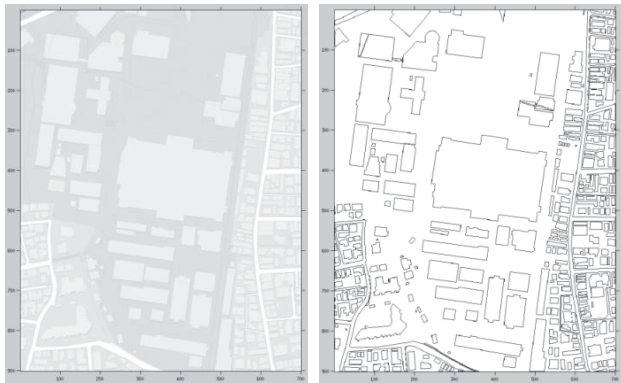


図3 東京工業大学本館の境界らしさ画像の作成

3.3 地図データから建物の境界を表す画像の作成

SfM による 3 次元復元結果から作成された建物の境界らしき画像と比べるために、入力された地図画像から建物の境界を抽出した画像を作成する。

入力された地図画像に対しグレースケール変換を行い、その後適当なしきい値を用いて画像の 2 値化を行う。次に 2 値化された地図画像に対して地図の境界を抽出するために、モルフォロジー演算による輪郭抽出を行い、建物の境界を抽出する。



(a) 入力地図画像 (b) 輪郭抽出画像
図 4 東京工業大学本館の地図画像の輪郭抽出

3.4 境界らしさを表す画像と地図データの位置合わせ

この章では SfM を用いて作成した建物の境界らしき画像と、地図データから作成した建物の境界を表した画像の位置合わせを行うことにより、SfM による 3 次元復元結果の地図上における位置、方向、スケールを推定する。

3.4.1 目的関数の設定

ユーザーが与える位置・方向・スケールの初期値を用いて、SfM を用いて作成した建物の境界らしき画像を入力地図画像上に幾何変換を行う。実装では、図 5 のように初期値を用いて境界らしき画像を地図画像と同じサイズの画像となるように変換する。初期位置 (i, j) 、初期方向 r 、初期スケール s を用いて境界らしき画像上の点 (x, y) を地図画像上の点 (x', y') に幾何変換を行う式は(2)のようになる。

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} s * \cos(r) & -s * \sin(r) & i \\ s * \sin(r) & s * \cos(r) & j \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2)$$

一方、後に用いる勾配法を収束させやすくするために、図 6 のように各画素から最も近い建物までのユークリッド距離を画素値とする距離変換画像を作成する。

地図上に幾何変換された境界らしき画像 I_{SfM} 、距離変換した地図画像 I_{Map} の 2 つの画像は画素値が 0 から 1 に正規化されており、どちらも黒いピクセルほど高い値となる。ここで、2 つの画像 I_{SfM} と I_{Map} の対応する画素値の積をとり、

全てのピクセルに対して和をとった値を目的関数と定義することにより、目的関数が最小となるときに最適な幾何変換となる。数式としては(3)のように表される。

$$Cost = \sum_{i, j} I_{SfM}(i, j) I_{Map}(i, j) \quad (3)$$

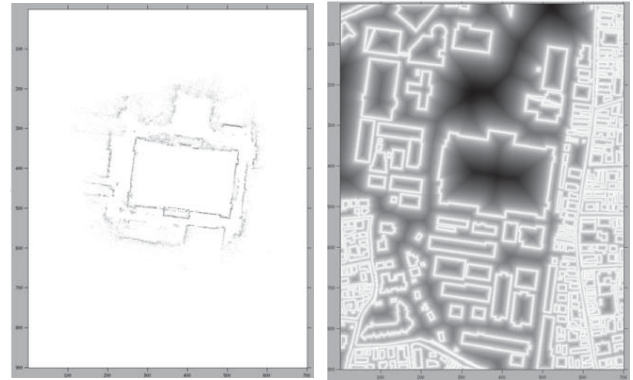


図 5 地図上に幾何変換した境界らしき画像 I_{SfM} 図 6 距離変換された地図画像 I_{Map}

3.4.2 勾配法による目的関数の最適化

位置、方向、スケールの 4 変数に対して勾配法を用いて最適化を行う。勾配法の手法は 2 階微分を計算する必要のないガウスニュートン法を用いる。1 階微分の計算については、中心差分法を用いて計算する。

4. 実験

4.1 実験概要

実験では、横浜赤レンガ倉庫、東京工業大学本館の 2 つのデータセットを用いた。

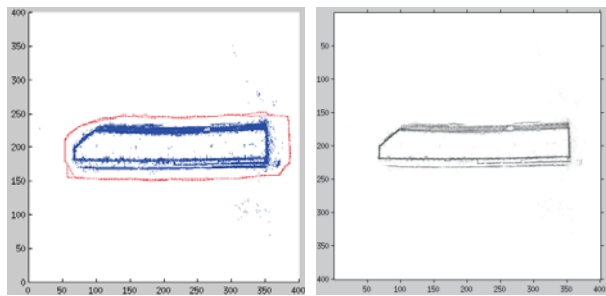
横浜赤レンガ倉庫データセットは透視投影画像で撮影した。三次元復元は VisualSFM[3]を用いて行った。この実験では、主に目的関数の有効性の検証を目的としたため、他の建物が含まれず赤レンガ倉庫のみが含まれる画像を地図として用いた。

東京工業大学本館データセットは、全方位を撮影できるので特徴点の復元が成功しやすく、また重力方向を測定できる Ricoh THETA を用いて全方位天球画像で撮影した。三次元復元は我々の研究室で開発したオンライン SfM システム[9]を用いて行った。地図については GoogleMap から得られる文字が含まれていない地図画像を入力として用いた。

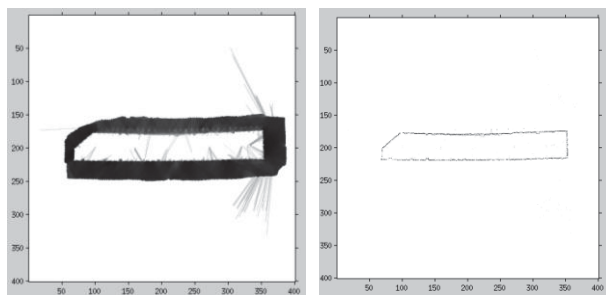
提案システムは Matlab を用いて実装した。

	赤レンガ倉庫	東工大本館
撮影画像枚数	978	2382
復元点数	127180	406705

表1 各データセットにおける三次元復元結果



(a) 青:2次元点群 赤:カメラ (b) 点群画像



(c) 視線画像 (d) 境界らしき画像

図7 横浜赤レンガ倉庫の境界らしき画像の作成

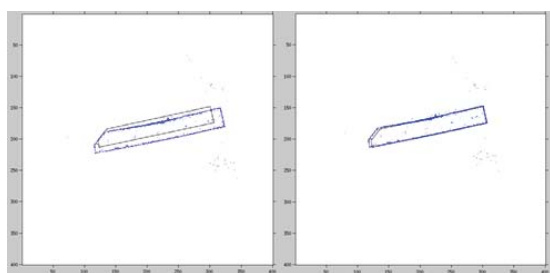


図8 赤レンガ倉庫の実験結果例1

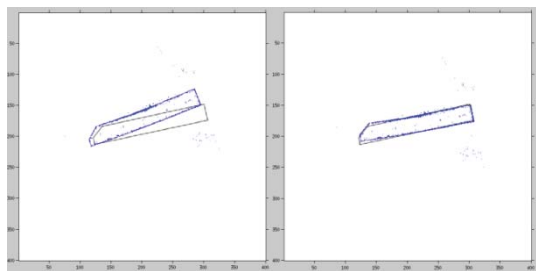


図9 赤レンガ倉庫の実験結果例2

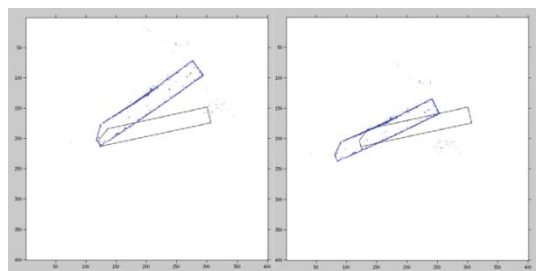


図10 赤レンガ倉庫の実験結果例3

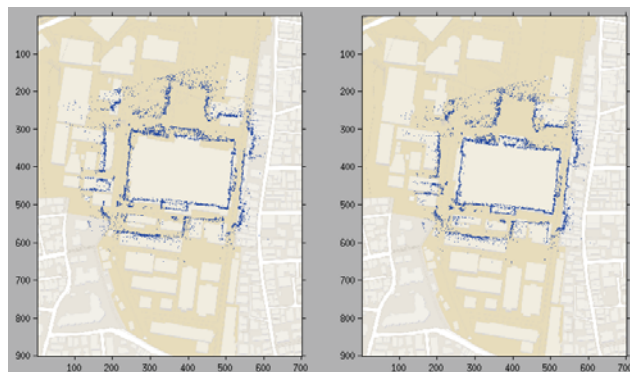


図11 本館の実験結果例1

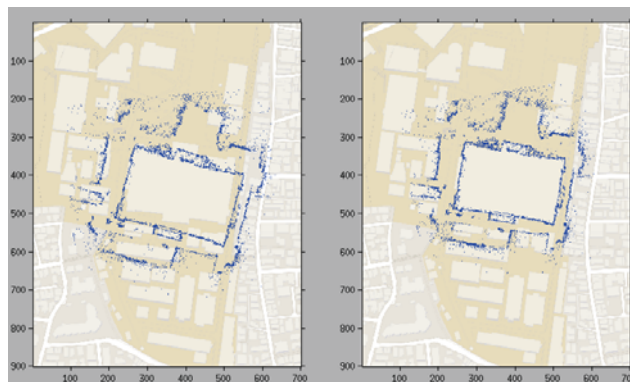


図12 本館の実験結果例2

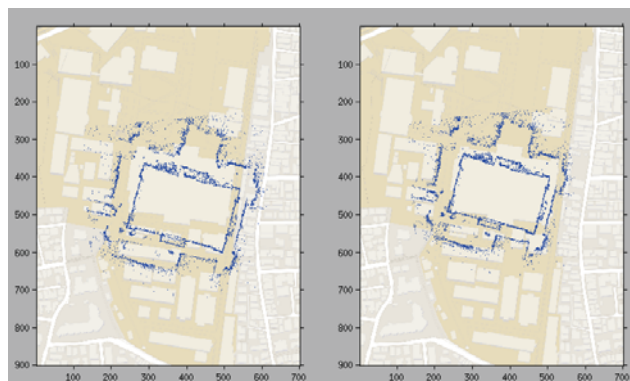


図13 本館の実験結果例3

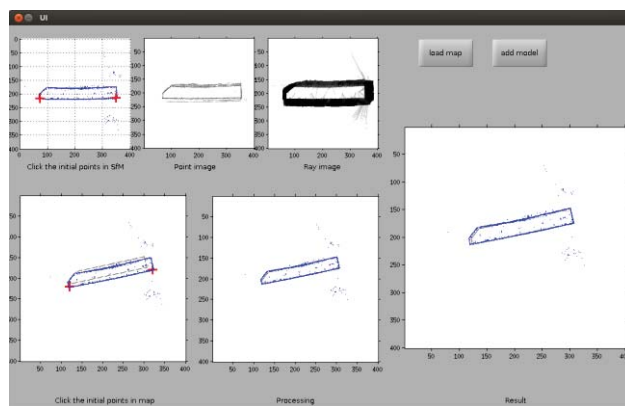


図14 ユーザーインターフェイスを用いた実験例

4.2 ユーザーインターフェイス

実験を行う際の初期値の入力方法としては、図 14 のようなユーザーインターフェイスを用いた。SfM による 3 次元復元結果と地図画像を入力し、境界らしき画像と地図画像の対応する 2 点を選ぶことにより位置・方向・スケールの初期値を入力した後、計算が終了したら勾配法により求めた地図上の建物の正確な位置を表示する機能を実装した。

4.3 実験結果

図 3, 図 7 に東京工業大学本館, 赤レンガ倉庫の点群画像・視線画像・境界らしき画像の作成の実験結果を示す。点群画像に含まれている建物の壁以外の点、視線画像と組み合わせることによって消去され、境界らしき画像では建物の壁が抽出されていることが確認された。

図 8 から図 13 に赤レンガ倉庫, 東京工業大学本館の位置合わせの結果を示す。図 8, 図 9, 図 11, 図 12 のように、初期値に依存するが、おおよその初期値が正しい値であれば、勾配法により真値の解に収束することを確認した。

一方、図 11 や図 14 のように初期値が大幅に真値とずれてしまっている場合においては、勾配法の途中で局所解に収束、あるいは、地図の他の建物に近づいてしまい真値でない別の解に収束する結果となった。

5. 結論

本論文では、SfM による 3 次元復元結果と地図データの位置合わせを行う手法を提案した。カメラの視線が通る空間は freespace となることを考慮して SfM による 3 次元復元結果から建物の境界らしい部分を抽出し、地図と組み合わせ設定した目的関数を最適化することにより地図上での建物の位置、方向、スケールを推定することができた。画像同士の位置合わせの問題を勾配法を用いて解くことにより、既存手法に対して大幅に計算時間が短縮された。

謝辞. 本研究は科研費 25240025 の一部助成を受けたものである。

参考文献

- [1] Noah Snavely, Steven Seitz and Richard Szeliski: “Modeling the world from internet photo collections”, IJCV 80(2), pp. 189–210, 2008.
- [2] Noah Snavely, Steven Seitz and Richard Szeliski: “Photo tourism: exploring photo collections in 3D”, ACM transactions on graphics (TOG), 25(3), pp. 835–846, 2006.
- [3] Changchang Wu: “VisualSFM : A Visual Structure from Motion System”, <http://ccwu.me/vsfm/>.
- [4] Qi Shan, Changchang Wu, Brian Curless, Yasutaka Furukawa, Carlos Hernandez and Steven Seitz: “Accurate geo-registration by ground-to-aerial image matching”, in

3DV, 2014.

- [5] David Lowe: “Distinctive image features from scale-invariant keypoints”, IJCV, 60(2), pages 91–110, 2003.
- [6] Ricardo Brualdi, Yanling He, Bryan Russell and Steven Seitz: “The 3D jigsaw puzzle: mapping large indoor spaces”, in ECCV, 2014.
- [7] Ryan Kaminsky, Noah Snavely, Steven Seitz and Richard Szeliski: “Alignment of 3D point clouds to overhead images”, in CVPR Workshop (Internet Vision), 2009.
- [8] Richard Szeliski: “Image alignment and stitching: A tutorial”, Foundations and Trends in Computer Graphics and Computer Vision, 2(1), 2006.
- [9] 鳥居秋彦, 半澤悠樹, 金紀昌, 阿達大地, 杉本茂樹, 奥富正敏: “オンライン 3 次元復元システム”, 第 18 回画像センシングシンポジウム (SSII2012) 講演論文集, p. DS1-08-1-1, 2012.