

# 姿勢推定を用いた点群の位置合わせ

石田 将吾<sup>1,a)</sup> 池上 史郎<sup>1,b)</sup> 川口 敦生<sup>1,c)</sup>

**概要:** 物体の3次元モデルを作成する3次元スキャンでは、様々な位置・方向から取得した点群の相対的な位置姿勢を合わせ、一つの点群にまとめる位置合わせ（レジストレーション）の処理が欠かせない。従来の位置合わせ手法は、点群の特徴的な箇所を検出し、複数の点群間で類似している箇所の対応関係を用いて位置合わせを行う。そのため、点群間の重なりが少ない場合や対象物体に特徴的な箇所が少ない場合、対応関係の推定が難しく、正しく位置合わせできないという問題があった。本稿では、物体認識の分野で研究されている姿勢推定を用いることで、この問題を解決する位置合わせ手法を提案する。本手法では点群を取得する際に、同じ方向から対象物体のRGB画像も取得する。そして、位置合わせの際に取得した点群それぞれについてRGB画像から方向を推定し、推定された方向を用いて位置合わせを行う。実験では、実際に3Dスキャナで取得した点群を対象に従来手法と比較を行い、提案手法の有効性を確認した。

## 1. はじめに

3Dスキャナで取得した点群は、一般に単一方向から計測されたデータである。そのため、物体全体の3次元モデルを得るためには、複数の異なる位置・方向から点群を取得し、取得した点群間の相対的な位置姿勢を合わせる必要がある。このような複数の点群間の相対的な位置姿勢を合わせる技術を位置合わせ（レジストレーション）と呼ぶ。ロボットアームやターンテーブルを用いれば、それら機構の位置関係より、取得した点群間の位置姿勢を合わせる事ができる。しかし、こういった方法は対象物の大きさや環境を極端に限定してしまう。

位置合わせは、対象とする複数の点群が重なり合わない状態から互いに重なり合うまで位置姿勢を合わせる粗い位置合わせ（coarse registration）と、ある程度重なりあった状態を初期状態とするより精密な位置合わせ（fine registration）に分類できる[1]。本稿では粗い位置合わせの新たな手法を提案する。

従来の粗い位置合わせ手法は、点群における局所的な特徴量を定義し、点群間で類似する箇所を探索し、対応付けを行う。Johnson[2]らはSpin Imageという頂点の法線ベクトルを軸とする円柱面に他の頂点を投影して得られた画像を、対応付けに用いる手法を提案した。法線相対な座標系を用いることにより、モデル形状について位置姿勢に対

する独立な表現が得られる。Rusuら[3]が提案したFPFHは、 $k$ 近傍におけるそれぞれの点のペアよりヒストグラムを計算し、それを対応付けに用いている。Tombariら[4]は、注目点の近傍点を用いて主成分分析より基準座標であるLocal Reference Frame(RF)を設定し、RFをもとに記述範囲を32次元に分割し、基準点の法線と分割された各スペースの法線ベクトルの内積を計算しヒストグラムを作成するSHOTという手法を提案した。

特徴量として点群ではなくテクスチャを用いる研究もある。Roth[5]はテクスチャの輝度値を用いることで、壺など幾何的に対称的な物体の位置合わせを行った。

従来手法では、点群間の重なりが少ない場合、対応関係の推定が難しいという問題があった。提案手法は、同じ方向からRGB画像と点群を取得し、RGB画像に対して姿勢推定を行い、その結果を用いて位置合わせを行う。例えば対象物を撮影したRGB画像が与えられたとき、それが正面から撮影されたRGB画像なのか、側面から撮影されたRGB画像なのかを推定し、推定結果を位置合わせに用いる。

以下、2章で提案する位置合わせ手法について説明し、3章で既存手法との比較実験により提案手法を評価し、4章でまとめる。

## 2. 姿勢推定を用いた位置合わせ手法

本手法の入力はRGB画像と点群であり、出力は複数の点群を位置合わせした点群である。入力のRGB画像と点群は、同方向から取得したRGB画像と点群を1セットとし、複数の位置・方向から取得した複数セットから成る。

<sup>1</sup> 株式会社リコー ワーク・ソリューション開発本部  
〒224-0035 神奈川県横浜市都筑区新栄町16-1

a) shohgo.ishida@nts.ricoh.co.jp

b) shiroh.ikegami@nts.ricoh.co.jp

c) atsuo.kawaguchi@nts.ricoh.co.jp



図 1 自動車を左前から撮影した写真



図 2 自動車を後ろ斜め上から撮影した写真

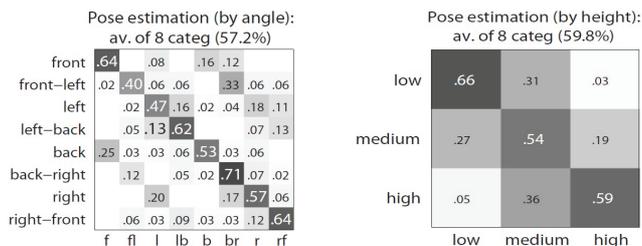


図 4 Savarese らの姿勢推定実験結果

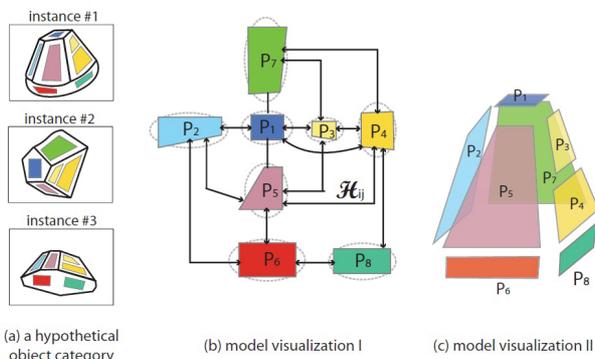


図 3 Savarese らの 3D オブジェクトモデル

本手法は、まず RGB 画像から物体を撮影した方向を推定し（姿勢推定）、次に推定された方向を用いて位置合わせを行う。2.1 節では従来の姿勢推定手法を紹介し、2.2 節では、姿勢推定結果を用いた位置合わせ手法について説明する。

## 2.1 姿勢推定

姿勢推定 (pose estimation) は、ある物体が写った RGB 画像が与えられたとき、物体を撮影した方向を推定する問題である。例えば、図 1 の RGB 画像が与えられたとき、姿勢推定の正しい結果は、自動車に対して左前方向となり、図 2 の RGB 画像が与えられたとき、姿勢推定の正しい結果は、自動車に対して後ろ斜め上となる。

姿勢推定は、物体認識の分野で研究されている問題である。一般的な手法としては、あらかじめ自動車など特定の物体を撮影した RGB 画像を学習データとして多数用意し、どの方向から撮影したもののカラベリングを行い、新たな RGB 画像が与えられたときにどの方向から撮影した RGB 画像と特徴が類似しているかを計算するものである。

代表的な手法としては Savarese らの研究 [6] がある。これは学習データから図 3 のようなパーツとそれらの位置関係から成るモデルを構築する。パーツは KB Detector [7] と SIFT [8] を利用して導出した特徴領域であり、位置関係は射影変換行列で表される。入力の RGB 画像が与えられたとき、同様に特徴領域であるパーツを求め、モデルから最適なパーツとそれら位置関係を選択し、方向を推定する。実験では、自動車、靴、トースターなど 8 カテゴリーの画像に対して姿勢推定が行われた。図 4 は、8 カテゴリーの姿勢推定結果の平均値である。左図は水平方向の推定結果、

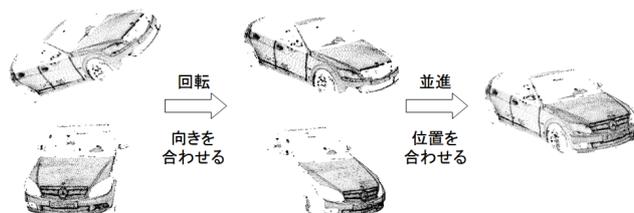


図 6 位置合わせの概略

右図は垂直方向の推定結果を示している。水平方向及び垂直方向において、50%程度の精度で推定できていることが分かる。

また、Payet ら [9] は Savarese らのデータセット [6] を用いて姿勢推定の実験を行った。図 5 は、自動車の画像に対する姿勢推定結果である。概ね 80%以上の精度で推定できていることが分かる。

姿勢推定は、手法や対象物体次第で 80%以上の精度で実現可能である。3 章では、上記を踏まえて位置合わせの実験を行う。

上述のような学習を用いた姿勢推定手法以外にも、重なりのある多数の画像から、バンドル調整 [10] により画像の姿勢パラメータを推定し、それを位置合わせに利用することが考えられる。この場合、学習データは不要となるが、バンドル調整が行えるだけの RGB 画像間の重なりが必要となる。

## 2.2 位置合わせ

提案手法は、2.1 節で説明した姿勢推定で得られた方向を用いて点群の位置合わせを行う。位置合わせの概略を図 6 に示す。まず、姿勢推定結果を用いて点群を回転することで向きを合わせ、その後点群を並進させることで位置を

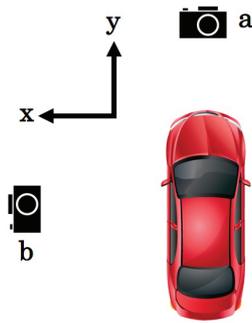


図 7 観測物体とカメラ位置の例

合わせる。

まず、点群の回転について説明する。例えば、 $a$  方向から取得した RGB 画像を  $I_a$ 、点群を  $P_a$  とし、 $b$  方向から取得した RGB 画像を  $I_b$ 、点群を  $P_b$  とする。姿勢推定により、 $a$  方向と  $b$  方向の 2 方向間の角度が  $(\theta_x, \theta_y, \theta_z)$  となったとする。式 (1),(2),(3) の回転行列を用いて点群を回転する。

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (2)$$

$$\begin{pmatrix} x''' \\ y''' \\ z''' \end{pmatrix} = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} \quad (3)$$

次に、点群の並進について説明する。姿勢推定により、 $I_a$  が物体を前方から撮影した RGB 画像であり、 $I_b$  が物体を左側面から撮影した RGB 画像であるという結果が得られたとする (図 7)。このとき、点群  $P_a$  と  $P_b$  は直交の位置関係にある。物体の上下左右が収まるように点群を取得していれば、点群  $P_a$  において  $x$  軸方向に最も大きい点の位置と、点群  $P_b$  において  $x$  軸方向に最も大きい点の位置は重なり、同様に  $y$  軸方向に最も大きい点も重なると考えられる。これより、点が重なるように両点群を  $xy$  平面において並進させることで位置合わせを行う。図 7 には記載していないが、 $x$  軸と  $y$  軸に直交する  $z$  軸に対しても、同様の処理を行う。

得られた位置姿勢を初期位置として、より精密に位置合わせを行うため ICP[11] を用いる。さらに、姿勢推定結果を、ICP 処理中の点群の回転制約に利用することも可能である。例えば、姿勢推定が  $45^\circ$  単位で結果を出力する場合、 $P_a$  に対して  $P_b$  を合わせる場合、 $45^\circ$  以上回転することはない、という制約である。 $45^\circ$  以上回転した場合は、姿勢推定の結果が適切ではないとし、姿勢推定において次にス



図 8 点群の取得位置

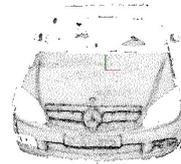


図 9 0 方向から取得した点群



図 10 1 方向から取得した点群



図 11 2 方向から取得した点群



図 12 3 方向から取得した点群

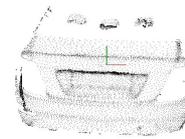


図 13 4 方向から取得した点群

コアが高かった方向を用いて位置合わせを行うなどの利用が考えられる。

### 3. 実験

提案手法の評価のために、実際に 3D スキャナで取得した点群を用いて、従来手法との比較を行った。

#### 3.1 実験データ

1/18 スケールのモデルカーを対象に取得した点群を使用する。点群の取得には、Artec 社の 3D スキャナである Spider<sup>\*1</sup> を利用し、図 8 のように点群は  $45^\circ$  刻みの 8 方向から取得した。

モデルカーより取得した点群の例を図 9, 10, 11, 12, 13 に示す。

#### 3.2 点群の位置合わせ

比較のために従来手法として FPFH[3] を用いた。実装には Point Cloud Library<sup>\*2</sup> を利用した。

##### 3.2.1 0-1 方向の点群の位置合わせ

既存手法による点群の位置合わせの実験結果を図 14 に

<sup>\*1</sup> <http://www.artec3d.com/hardware/artec-spider/>

<sup>\*2</sup> <http://pointclouds.org/>

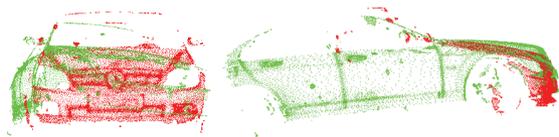


図 14 既存手法による 0-1 方向の点群の位置合わせ結果



図 15 既存手法+ICP による 0-1 方向の点群の位置合わせ結果



図 16 提案手法による 0-1 方向の点群の位置合わせ結果

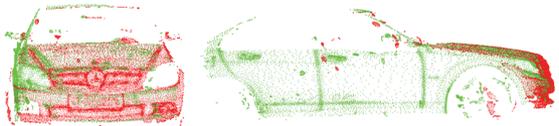


図 17 提案手法+ICP による 0-1 方向の点群の位置合わせ結果

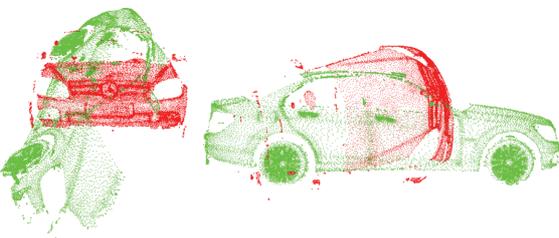


図 18 既存手法による 0-2 方向の点群の位置合わせ結果

示す。0 方向と 1 方向の点群間には重なりがある程度あるため、粗い位置合わせに成功している。より精密に位置を合わせるために、ICP を行った結果は図 15 となる。ICP をかけることにより、1 つの点群として矛盾のないかたちになっている。

提案手法による点群の位置合わせの実験結果を図 16 に示す。既存手法と同様に、粗い位置合わせに成功している。加えて ICP を行った結果を図 17 に示す。こちらも既存手法と同様に、1 つの点群として矛盾のないかたちになっている。

### 3.2.2 0-2 方向の点群の位置合わせ

既存手法による点群の位置合わせの実験結果を図 18 に示す。粗い位置合わせに失敗し、点群が大ききずれる結果となっている。要因として、点群間の重なりが少ないことが考えられる。従来手法では、点群の特徴的な箇所の対応付けより位置合わせを行うため、点群間の重なりが少ない場合



図 19 提案手法による 0-2 方向の点群の位置合わせ結果



図 20 提案手法+ICP による 0-2 方向の点群の位置合わせ結果

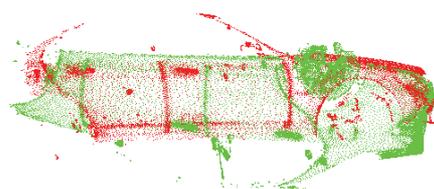


図 21 従来手法+ICP による 1-3 方向の点群の位置合わせ結果

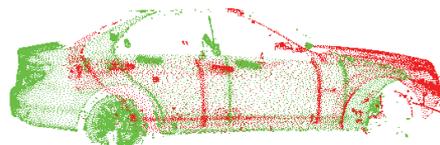


図 22 提案手法+ICP による 1-3 方向の点群の位置合わせ結果

は位置合わせが正しく行われない場合が多い。この状態で ICP による精密な位置合わせを行っても、正しい位置にたどり着かない。

提案手法による点群の位置合わせの実験結果を図 19 に示す。従来手法の結果に比べ、正しい位置関係に近い結果となっている。提案手法により位置合わせを行った点群に、ICP をかけた結果を図 20 に示す。ICP による精密な位置合わせを行っても、正しい位置関係にたどり着いていない。点群間の重なりが少ないため、ICP が正しく処理できていないと考えられる。

### 3.2.3 その他の方向の点群の位置合わせ

従来手法による 1 方向と 3 方向の点群の位置合わせ結果を図 21 に示す。自動車のフロントとリアが重なるような上下が反転する結果となっている。理由として、今回実験に使用したモデルカーが概ね前後対称の形であり、フロント及びリアの凸部が対応付けられたためと考えられる。

提案手法による 1 方向と 3 方向の点群の位置合わせ結果を図 22 に示す。従来手法に比べ、上下は正しい結果となっている。しかし、精密な位置合わせにより正しい位置まで移動できていない。本手法では、点群の端点を合わせるため、1 方向と 3 方向の点群では自動車のフロントあるいは

リアの端点を合わせることができず、粗い位置合わせの精度が伸びなかったためと考えられる。

### 3.3 実験結果の考察

今回使用した実験データに対しては、点群の取得位置の間隔が  $45^\circ$  の場合は従来手法、提案手法ともに正しく位置合わせが行えた。しかし、点群の取得位置の間隔が  $90^\circ$  の場合は、従来手法では正しく位置合わせが行えなかったのに対して、提案手法では粗い位置合わせを行うことができた。従来手法は点群間の重なり大きさによって位置合わせの精度が変化するが、提案手法は点群間の重なりが小さい場合でも位置合わせを行うことができる。

処理時間についても提案手法が優位といえる。本章の実験では、CPU が Intel Core i7 2.93GHz、メモリ 12GB の PC を使用したが、点の数が 30,000 点程度の場合、FPFH の計算に 8 分程度、対応点の探索に 25 分程度を要した。ダウンサンプリングなどの追加処理を行わなければ、点の数が増えれば増えるほど、既存手法は指数関数的に処理時間が長くなる。それに対し提案手法は、位置合わせの処理は点群を回転するだけであるため処理時間は極めて短く、点の増加に対しても処理時間の増加は線形である。

従来手法の多くは、位置合わせに必要なパラメータを調整しなければならない。FPFH であれば考慮する周囲の点の数や、SAC-IA ではイテレーションの回数などが必要となる。対して提案手法は、位置合わせ処理のチューニングの必要はなく、点群の疎密などを考慮する必要はない。

点群間の重なりが少ない場合の ICP については課題がある。提案手法は、点群間の重なりが少ない場合でも粗い位置合わせを行うことができるが、その後段の処理である ICP には相応の点群間の重なりが必要となる。点群間の重なりが少ない場合でも頑健に精密な位置合わせを行うことが課題として残されている。

## 4. おわりに

本稿では、RGB 画像より姿勢推定を行い、その結果を用いた位置合わせ手法を提案した。同方向から取得した RGB 画像と点群を 1 セットとしたとき、入力は複数の位置・方向から取得した複数セットであり、出力は位置合わせした点群となる。姿勢推定により物体を撮影した方向を推定し、それを用いて粗い位置合わせを行い、その後 ICP による精密な位置合わせを行う。

モデルカーより取得した点群を対象に既存手法と提案手法の比較実験を行った。従来手法に比べ、提案手法は点群間の重なりが小さい場合に対しても粗い位置合わせが可能であり、処理時間も短いという結果が得られた。しかし、後段の処理である ICP では相応の点群間の重なりが必要となるため、精密な位置合わせが難しい。今後はこのような課題の検討を進めていく。

## 参考文献

- [1] Salvi, J., Matabosch, C., Fofi, D. and Forest, J.: A review of recent range image registration methods with accuracy evaluation, *Image and Vision Computing*, Vol. 25, No. 5, pp. 578 – 596 (online), DOI: <http://dx.doi.org/10.1016/j.imavis.2006.05.012> (2007).
- [2] Johnson, A. and Hebert, M.: Using spin images for efficient object recognition in cluttered 3D scenes, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 21, No. 5, pp. 433–449 (online), DOI: 10.1109/34.765655 (1999).
- [3] Rusu, R., Blodow, N. and Beetz, M.: Fast Point Feature Histograms (FPFH) for 3D registration, *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pp. 3212–3217 (online), DOI: 10.1109/ROBOT.2009.5152473 (2009).
- [4] Tombari, F., Salti, S. and Di Stefano, L.: Unique Signatures of Histograms for Local Surface Description, *Proceedings of the 11th European Conference on Computer Vision Conference on Computer Vision: Part III, ECCV'10*, Berlin, Heidelberg, Springer-Verlag, pp. 356–369 (online), available from (<http://dl.acm.org/citation.cfm?id=1927006.1927035>) (2010).
- [5] Roth, G.: Registering two overlapping range images, *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on*, pp. 191–200 (online), DOI: 10.1109/IM.1999.805349 (1999).
- [6] Savarese, S. and Fei-Fei, L.: 3D generic object categorization, localization and pose estimation, *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–8 (online), DOI: 10.1109/ICCV.2007.4408987 (2007).
- [7] Kadir, T. and Brady, M.: Saliency, Scale and Image Description, *International Journal of Computer Vision*, Vol. 45, No. 2, pp. 83–105 (online), DOI: 10.1023/A:1012460413855 (2001).
- [8] Lowe, D.: Object recognition from local scale-invariant features, *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Vol. 2, pp. 1150–1157 vol.2 (online), DOI: 10.1109/ICCV.1999.790410 (1999).
- [9] Payet, N. and Todorovic, S.: From contours to 3D object detection and pose estimation, *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 983–990 (online), DOI: 10.1109/ICCV.2011.6126342 (2011).
- [10] Triggs, B., McLauchlan, P. F., Hartley, R. I. and Fitzgibbon, A. W.: Bundle Adjustment - A Modern Synthesis, *Proceedings of the International Workshop on Vision Algorithms: Theory and Practice, ICCV '99*, London, UK, Springer-Verlag, pp. 298–372 (online), available from (<http://dl.acm.org/citation.cfm?id=646271.685629>) (2000).
- [11] Besl, P. J. and McKay, N. D.: A Method for Registration of 3-D Shapes, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 14, No. 2, pp. 239–256 (online), DOI: 10.1109/34.121791 (1992).