

# ローバ搭載型センサフュージョンシステムを用いた高精度、高密度な三次元再構成手法

石川 涼一<sup>1,a)</sup> Roxas Menandro<sup>1,b)</sup> 佐藤 啓宏<sup>1,c)</sup> 大石 岳史<sup>1,d)</sup> 増田 健<sup>2,e)</sup>  
池内 克史<sup>3,f)</sup>

**概要:** 移動体搭載型の三次元センシングシステムは様々なアプリケーション用いられるようになってきている。本論文ではローバに搭載されたプロファイラ型のスキャニングシステムを提案し、高精度高密度な文化財の bas-relief の三次元再構成を行った。全方位カメラと三次元レーザスキャナからハードウェアシステムは構成されている。再構成手法においては安定したモーション推定のため、画像上で応答の良い投影点を選択し、またカメラとレーザスキャナの位置の違いによるミスマッチを全方位カメラから測定点までの距離をもとにした誤差指標で除去している。比較手法に比べて提案手法はより正確な三次元再構成結果を示している。ローカルのモーション推定に加えて、相互情報量を用いた 2D-3D 位置合わせによるグローバルなモーション推定の最適化手法も提案する。

## 1. Introduction

三次元デジタルアーカイブは実世界の三次元モデルを測定によって作成する技術である。この技術はとりわけ遺跡などの保存や修復などで重要な役割を果たしている。図. 1 (a) に示すのはクメール文化の寺に存在するバスレリーフであり、細部にわたって繊細に彫られている。加えてこのレリーフは非常に広い範囲に彫られていて、すべて高密度に測定することは非常に難しい。

従来の手法では対象物体全体を測定するため、レーザスキャナを用いて異なる個所から何度も測定を行う必要があるが非常に労力がかかるという問題がある。そのような従来手法に代わって、移動体にセンサを搭載して測定を行うさまざまな技術が開発されており、測定のプロセスが高速化されている。移動体を用いた手法としては LiDAR のみを用いた方法 ([1], [2]), LiDAR と IMU, GPS を組み合わせた方法 ([3], [4], [5]), LiDAR とカメラを組み合わせた方法 ([6], [7], [8], [9]) がある。我々の研究室の以前の仕事で Zheng らはレールを用いてレリーフを測定する研究を行っているが、レールは重たく運搬に向かない、またセンサの移動が制限されてしまうという問題があった。この問題に対してレールの代わりに、軽量でより自由に動ける

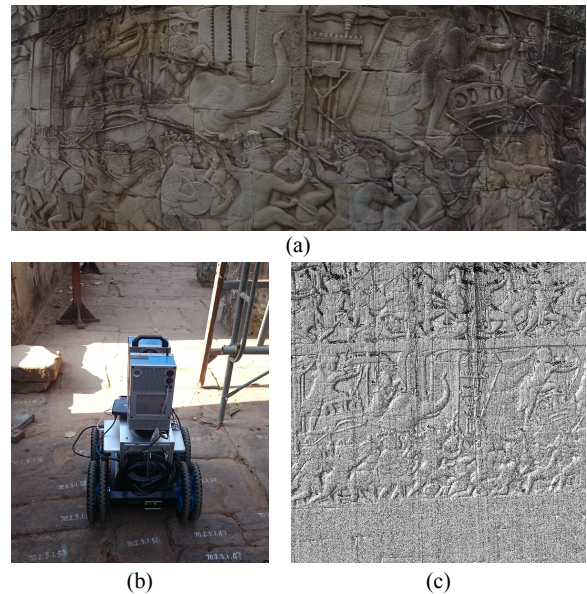


図 1 (a) A bas-relief in a Khmer-style temple. (b) Proposed scanning sensor system mounted on a rover. (c) Obtained reconstruction result of relief by our method.

ローバを用いて、センサをプロファイラモードで動作させて測定を行うシステムを開発した。(図. 1 (b) 参照)

移動体搭載型の測定システムにおいては移動経路を推定する必要がある。本稿では、あらかじめ外部校正された全方位カメラとレーザスキャナから得られる、2D の全方位画像と 3D の点群間の対応を用いた三角法によるセンサの推定手法を提案する。三次元点群を二次元画像上に投影しその点を画像上でトラッキングすることで 2D3D 対応を作成し、さらに画像上で曲率の高い点に投影された 2D3D 対応のみを抽出する。そののち三角法によって、各カメラフ

<sup>1</sup> 東京大学 生産技術研究所

<sup>2</sup> 産業技術総合研究所

<sup>3</sup> Microsoft Research Asia

a) ishikawa@cvl.iis.u-tokyo.ac.jp

b) roxas@cvl.iis.u-tokyo.ac.jp

c) yoshi@cvl.iis.u-tokyo.ac.jp

d) oishi@cvl.iis.u-tokyo.ac.jp

e) t.masuda@aist.go.jp

f) katsuike@microsoft.com

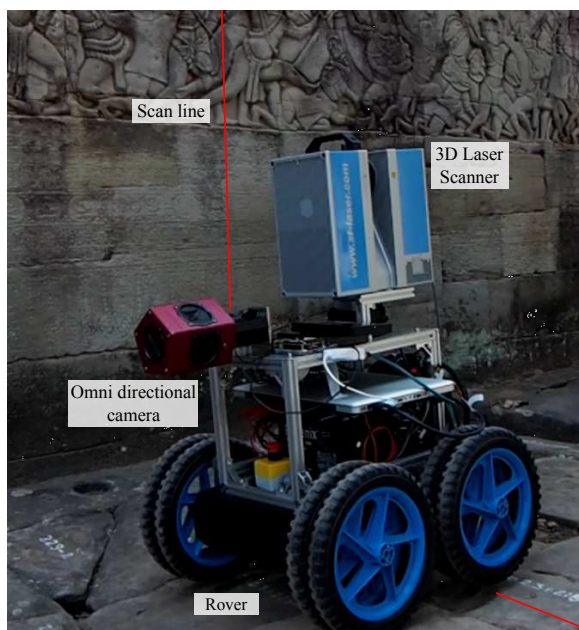


図 2 Hardware system

レーム間の位置姿勢推定を行う。レーザスキャナから測定された点への距離を用いた誤差指標を用いることで、センサ間の位置の誤差を起因とする 2D-3D の mismatches を外れ値として除去している。図 1 (c) に示すように我々の手法によって再現性を持った relief の三次元点群の再構成結果を得ることができる。

三角法による位置姿勢推定に加えて、三次元スキャナのレーザのリフレクタンスとカメラから得られる 8bit 画像間の相互情報量の最大化による multi-modal な 2D-3D 位置合わせ手法を提案する。

本稿のコントリビューションを以下に示す。

- 直接三次元点を二次元の画像に投影し高い曲率の点に投影された 2D3D 対応を用いることによる、安定した三次元再構成手法の提案とデモンストレーション。
- リフレクタンス情報を用いて multi-modal な位置合わせをことによる推定経路の修正手法の提案
- センサ間のポジションのずれによる誤対応を考慮した誤差指標の提案

測定に用いたハードウェアシステムは図 2 に示すように主にローバ、レーザスキャナ、全方位カメラで構成されている。レーザスキャナはローバの動きに合わせて輪切り状に周囲を計測するプロファイラモードで用いられている。データの同期をとるため、スキャン開始時点ではトリガースIGNALをレーザスキャナからカメラに送信している。モーション推定にはこのカメラとレーザスキャナから得られる二次元画像と三次元スキャンを用いる。

## 2. Related Work

様々な三次元測定システムが移動体を用いた三次元再構成システムに用いられている

### Camera system

近年、単眼カメラを用いた Structure from Motion の測定システムが開発されているが [10], [11], [12], [13], 一般的にこれらの手法は他のカメラか、ある種的前提がなければスケールがわからないという問題がある。一方でステレオカメラシステムでは2つのカメラのベースライン [14], [15] を用いてセンサモーションのスケールを得ることができるが、精度に関してはいまだレーザスキャナのレベルには及ばない。

### RGB-D camera system

CMOS RGB-D センサはデプス画像の歪みが少なく、モーション推定に適している ([16], [17], [18]) が、一方で測定対象に集中する光の強さが低いため、Signal-to-Noise Ratio(SNR) も低くなっている [19]。遺跡の測定などでは屋外環境での測定が求められることが多く、概して背景光の強い環境下での長距離の測定の必要があるが SNR の低い RGB-D カメラでは測定が難しい。

### LiDAR-only system

一方で、レーザスキャナシステムは光源出力をより小さい領域に集中させることができる、より大きな SNR を得ることができる [19]。そのため RGB-D カメラと比べて長距離を少ないノイズで測定することができる。しかしながら測定により長い時間を必要とし、その間にセンサが移動すれば得られた三次元点群は移動によって歪んでしまうという問題もある。LiDAR のみ [1], [2] を用いてセンサの移動経路を推定する方法はハードウェアコストがかからず、また、同期やセンサ間の外部校正を必要としないという点で優れているが、測定に長時間を要することにより正確な経路推定は難しい。

### LiDAR + IMU, GPS

LiDAR の経路推定を行う方法の一つに他のセンサを組み合わせる方法がある。IMU や GPS を LiDAR と組み合わせる方法 [3], [4], [5] はシンプルで直接的な経路推定問題の解決法であるが、精度は三次元デジタルアーカイブに用いるには不十分である。

### LiDAR + Camera fusion system

画像や他のデプスセンサを用いた経路推定手法はロボティクスの分野で SLAM を行うためよく設計されている [6]。[5] では LiDAR を RGB カメラと組み合わせて経路推定のロバスト性を向上させている。また [7], [8] ではカメラと組み合わせた LiDAR で周囲の形状をプロファイラしながら測定を行っている。Bokらは大きさが村レベルの遺産の三次元の再構成を、四つの LiDAR と二つのカメラ、さらに GPS を用いて行っている [7]。しかしながら我々の測定ではミリメートル単位の精度が要求され、GPS を用いた経路推定ではそれだけの精度を出すことは難しい。



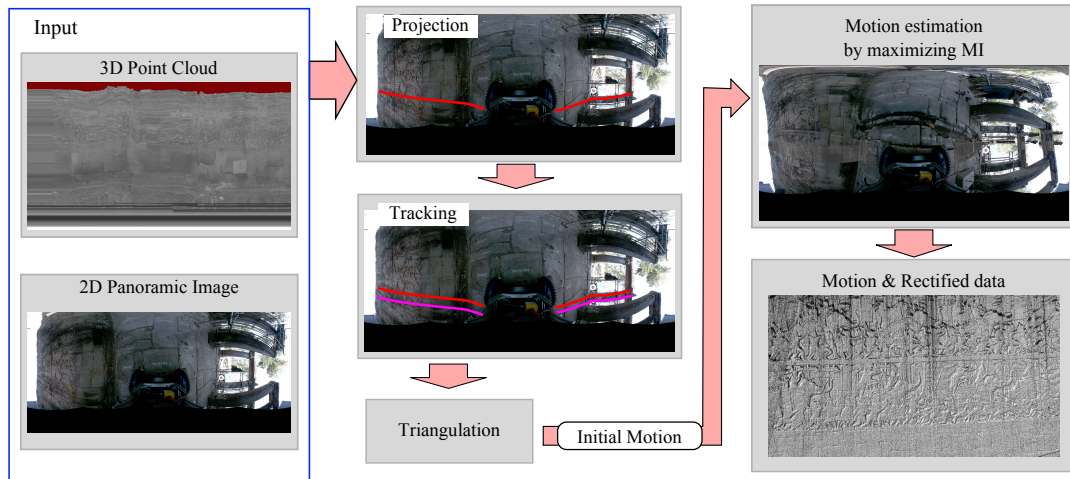


図 3 Method overview. The inputs are the synchronized and calibrated 3D point clouds and 2D image sequence. Initial motion is computed by 3D-2D projection, point tracking and triangulation. Then the motion is optimized using mutual information resulting in a motion and rectified data.

### 3. Local camera pose estimation

レーザスキャナと全方位カメラの融合システムを用いて、スキャン中のセンサの経路の推定を行うことが我々の研究で取り扱う問題である。図. 3 に我々の手法のオーバービューを示す。入力はレーザスキャナから得られた三次元点群と全方位カメラから得られた二次元のパノラマ画像列である。これらの二つのセンサは同期、外部校正されている。測定データはローバを前方に動かしながら測定している。

経路推定では最初に三次元点群を二次元画像に投影することで 2D-3D 対応を構築する。そののち高い曲率の部分に投影された 2D-3D 対応を抽出する。それらの点を画像列状で KLT tracker [20] を用いてトラッキングする。これによって三次元点と対応する二次元点が複数の画像列にわたって得られるので、それらの対応を用いて三角法により、カメラの位置姿勢推定を線形および非線形最適化によって求める。

#### 3.1 Coordinate system

レーザスキャナとカメラはあらかじめキャリブレーションされている。  $i, j$  を 2D のカメラのフレームインデックス、三次元点のインデックス、  $t_i, t_j$  を 2D フレーム  $i$  が撮影された時刻、三次元点  $j$  がスキャンされた時刻、  $\mathbf{X}_j^L$  をレーザスキャナのローカル座標系における点  $j$  の三次元座標、  $\mathbf{x}_i^j$  を三次元点  $j$  に対応するフレーム  $i$  上の二次元座標とする。求めたいものは時刻  $t_i$  におけるセンサの位置姿勢パラメータ  $\mathbf{R}_i, \mathbf{T}_i$  である。  $\mathbf{X}_j^L$  を二次元画像上に投影する操作を以下のように定義する。

$$\mathbf{x}_i^j = Proj(\mathbf{X}_j^L). \quad (1)$$

センサのローカル座標系における点  $\mathbf{X}_j^L$  をワールド座標系  $\mathbf{X}_j^W$  に変換する操作は位置姿勢パラメータ  $\mathbf{R}_j, \mathbf{T}_j$  を用いて以下のように書ける。

$$\mathbf{X}_j^W = \mathbf{R}_j \mathbf{X}_j^L + \mathbf{T}_j. \quad (2)$$

$\mathbf{R}_j, \mathbf{T}_j$  を計算するときは  $\mathbf{R}_i, \mathbf{T}_i$  の線形補完によって求める。

#### 3.2 2D-3D correspondence

2D-3D 対応付けにはまず  $t_i - \Delta t < t_j < t_i + \Delta t$  の間にスキャンされた三次元点をフレーム  $i$  に外部校正によって得られたパラメータを用いて投影する。投影されたすべての点を使うことは計算コストの増大、およびトラッキングに不向きな不安定な点を含んでしまう可能性がある。これを避けるため、曲率の高い点に投影された 2D-3D 対応のみを特徴点として用いる。  $\mathbf{X}_{(i,j)}^L$  を初めにフレーム  $i$  に投影され特徴点として選択された点  $j$  のローカル三次元座標、  $\mathbf{X}_{(i,j)}^W$  を  $\mathbf{X}_{(i,j)}^L$  のワールド座標系における三次元座標とする。  $\mathbf{X}_{(i,j)}^L$  と  $\mathbf{x}_i^j$  がいま投影によって対応付けされている。 KLT tracker を用いて  $\mathbf{x}_i^j$  をトラッキングすることで各フレームにおいて  $\mathbf{x}_i^j$  と対応する点を求めることができる。

#### 3.3 Triangulation

三角法においては [8] にて提案されている Zheng らの手法をもとにしている。フレーム 0 から  $n$  までの位置姿勢があらかじめ計算されて、その各フレーム  $i$  において  $t_i - \Delta t < t_j < t_i + \Delta t$  の間にスキャンされた点  $j$  のワールド座標系における座標  $\mathbf{X}_j^W$  が計算されているとする。フレーム  $n+1$  におけるカメラの位置姿勢パラメータ  $\mathbf{R}_{n+1}, \mathbf{T}_{n+1}$  は 2D-3D 対応  $\mathbf{X}_j^W$  and  $\mathbf{x}_{i+1}^j$  を用いてまず perspective 3-point algorithm [21] と RANSAC [22] に

よって線形に計算する。そののち以下に定義するコスト関数を最小化するように  $\mathbf{R}_{n+1}, \mathbf{T}_{n+1}$  を最適化する:

$$\begin{aligned} & \{\mathbf{R}_{n+1}, \mathbf{T}_{n+1}\} \\ &= \arg \min_{\mathbf{R}, \mathbf{T}} \sum_{i=1}^n \sum_j (\|Proj(\mathbf{R}^T(\mathbf{X}_{(i,j)}^W - \mathbf{T})) \\ & - \mathbf{x}_{n+1}^j\|^2) \\ &+ \sum_{i=1}^n \sum_j (\|Proj(\mathbf{R}_i^T((\mathbf{R}\mathbf{X}_{(n+1,j)}^L + \mathbf{T}) - \mathbf{T}_i)) \\ & - \mathbf{x}_i^j\|^2). \end{aligned} \quad (3)$$

Eq. 3 は世界座標系から現在のフレーム  $n+1$  に投影した時の再投影誤差の和と現在のフレームにおける (時刻  $t_{n+1}$  における) ローカルな三次元点を以前のすべてのフレーム (0 から  $n$ ) に投影した時の再投影誤差の和を示している。

図. 4 に三角法の模式図を示す。

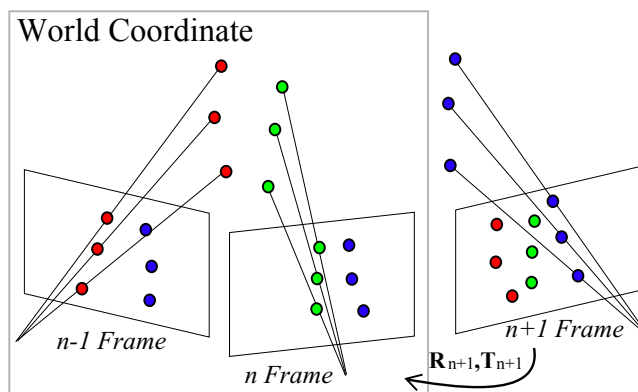


図 4 Triangulation. When 0 to  $n$  frames are worked out (only  $n-1$  and  $n$  frames are shown in figure), pose of  $n+1$  frame is firstly computed by Perspective 3 Point algorithm [21] using red and green points. Then pose is optimized by non-linear optimization with 2D-3D correspondences of red, green and blue points.

### 3.4 Error metric for outlier rejection

図. 5 に示すようにカメラとレーザスキャナの中心位置が違ふとき、その中心位置の違いによる 3D-2D 投影のミスマッチは避けることができない。このミスマッチは図. 5 内の赤い矢印の方向への誤差の蓄積を引き起こす。さらに連続するカメラ間のベースラインが小さいとき、この誤対応を再投影誤差の値を用いて外れ値として除去することは、カメラ中心からスキャンされた点と追跡された点に向かう二つのベクトルがなす角  $\theta$  が小さいため難しい。

カメラとスキャンされた三次元点の距離とそれに対応するデプスの距離差が大きいときであっても再投影誤差を用いてこの誤対応を除去することは難しい。一方でフレーム  $n$  から  $n+1$  へと座標変換された三次元点は追跡された二次元点の方向の延長線上深さ  $d$  のところに存在するはずであり三次元座標系での観測誤差は大きくなる。したがって、再投影誤差の代わりに  $d \sin \theta$  (図. 5 参照) の値を外れ値除去に用いた。この式は図. 5 における赤い直線の距離に相当する。

### 3.5 Bundle adjustment

バンドル調整も経路推定精度を高めるため実装している。エネルギー関数は三次元点をすべての二次元画像上に再投影した時の再投影誤差をもとにしている。エネルギー関数は以下のように定義している:

$$\begin{aligned} & \{\mathbf{R}_0 \dots \mathbf{R}_N, \mathbf{T}_0 \dots \mathbf{T}_N\} \\ &= \arg \min_{\mathbf{R}_0 \dots \mathbf{R}_N, \mathbf{T}_0 \dots \mathbf{T}_N} \sum_i \sum_j \sum_k \\ & \delta_{i,j} \|Proj(\mathbf{R}_k^T((\mathbf{R}_i \mathbf{X}_{(i,j)}^L + \mathbf{T}_i) - \mathbf{T}_k)) - \mathbf{x}_k^j\|^2 \end{aligned} \quad (4)$$

$$\delta_{j,k} = \begin{cases} 1 & (\text{if } \mathbf{X}_j^W \text{ is detected on frame } k) \\ 0 & (\text{else}) \end{cases} \quad (5)$$

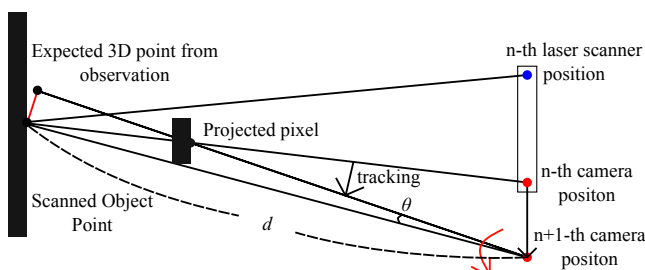


図 5 Mismatch of 3D point and 2D image are caused by sensor position gap. Error accumulates in the red arrow direction.

すべてのカメラの位置姿勢は Levenberg-Marquard 法による再投影誤差の最小化によって最適化される。

## 4. Registration by using MI

2D-3D 対応を用いたローカルな位置姿勢推定によって良い結果を得ることができるが、トラッキングの誤差による蓄積誤差は避けることができない。本章では三角法によって得られた推定結果の精度を高める目的で、リフレクタンスを用いた 2D-3D 位置合わせを行う手法を提案する。

センサーキャリブレーションによって三次元点群の色情報を得ることができるが、3D から 2D への投影誤差は全方位カメラの  $fps$  の低さにより避けることができない。したがって二次元画像から得られた三次元点の色情報の精度は保証されない。一方でレーザのリフレクタンスはレーザの照射点における物体の表面の情報を含み、三次元座標とリフレクタンスの間に誤差はないため、リフレクタンスの情報を経路推定に活用することは価値のあることである。Kurazume らはリフレクタンスと画像の間でエッジを用いる手法を提案している [23]。

我々のリフレクタンスを用いた位置姿勢修正手法は相互情報量の最大化を用いた 2D-3D の位置合わせ手法を用い



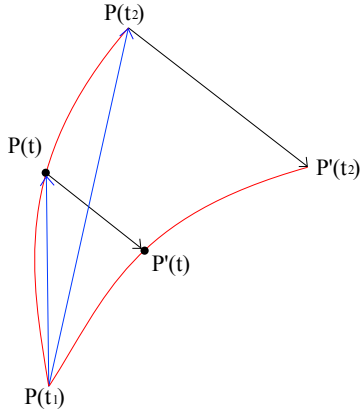


図 6 Deformation model of motion.  $P(t)$  indicates  $(x, y, z)$  or  $(rx, ry, rz)$  in 6-DoF parameters. Motion entire the segment between  $t_1$  and  $t_2$  is rescaled by final position  $P'(t_2)$

ている。この相互情報量を用いた Multi-modal な位置合わせ手法は Viola らによって提案され [24] 主に医学の分野で研究されてきた。

#### 4.1 Motion deformation model

まず MI を用いた最適化段階におけるセンサモーションの変形モデルについて述べる。  $W = (rx, ry, rz, x, y, z)$  は 6-DoF のパラメータを示す。3 にて計算された時刻  $t$  における初期位置を  $W(t)$  とする。初めに二つの異なる時刻  $t_1$  と  $t_2$  におけるカメラフレームを選択する ( $t_1 < t_2$ )。時刻  $t_1$  から  $t_2$  にかけての経路を最適化する。  $W(t)$  を時刻  $t$  における最適化前のもとの位置姿勢パラメータとし、  $W'(t)$  をモーション変形後の時刻  $t$  における位置姿勢パラメータとする。  $W'(t)$  は時刻  $t_1-t_2$  における最後の位置姿勢パラメータ  $W'(t_2)$  を変えることによって決定される。  $P(t)$  をベクトル  $(rx, ry, rz)^T$  か  $(x, y, z)^T$  とする。変形後のパラメータ  $P'(t)(t_1 < t < t_2)$  は以下のように計算される:

$$P'(t) = \frac{|P(t) - P(t_1)|}{|P(t_2) - P(t_1)|} (P(t_2) - P'(t_2)) + P(t). \quad (6)$$

$t_1-t_2$  における位置姿勢パラメータは区間の最終位置姿勢  $W'(t_2)$  によって変形される (図. 6 参照)。MI は変形されたモーション  $W'(t)$  を用いて計算される。

スキャン全体はこのように短いセグメント ( $t_1-t_2$ ) に分けられて最適化される。セグメントを増やせば高い解像度でモーションの変形を行うことができるが、区間が短ければセグメントの MI による位置合わせを行うには情報量として不十分になることが考えられる。このセグメントの長さは前述のトレードオフを考慮して経験的に決定した。

#### 4.2 Mutual information

モーションの最適化のため、RGB 画像をグレイスケール画像に変換したピクセルの値とリフレクタンスの値の間の MI を用いた。MI の計算には Maes らが提案している手法を用いた [25]。画像のピクセルの値  $f$  と三次元点のリフ



図 7 An example of area separation on a panoramic image.

レクタンスの値  $r$  の結合ヒストグラム  $\mathbf{H} = \{h_{r,f}\}$  は三次元点群を二次元画像に投影することで計算される。MI は  $\mathbf{H}$  から以下のように計算される:

$$I = \frac{1}{N} \sum_{f,r} h_{r,f} \log \frac{N h_{r,f}}{h_f h_r} \quad (7)$$

$h_f = \sum_r h_{r,f}$ ,  $h_r = \sum_f h_{r,f}$  であり、また  $N = \sum_{r,f} h_{r,f}$  である。

#### 4.3 Joint histogram construction

ここで残る問題はどのように結合ヒストグラムを作るかである。  $\mathbf{x}_i^j = Proj(\mathbf{R}^T(\mathbf{X}_j^W - \mathbf{T}))$  を三次元点  $j$  をフレーム  $i$  上に投影した時の二次元座標である。  $q_{(i,j),0}$  は  $0 < d_{j,x} < 1$ ,  $0 < d_{j,y} < 1: (d_{j,x}, d_{j,y})^T = \mathbf{x}_i^j - q_{(i,j),0}$  となるピクセルの場所であり、  $q_{(i,j),1} = (1, 0)^T + q_{(i,j),0}$ ,  $q_{(i,j),2} = (0, 1)^T + q_{(i,j),0}$ ,  $q_{(i,j),3} = (1, 1)^T + q_{(i,j),0}$  とする。

結合ヒストグラム  $h_{r,f}$  は以下のように計算される:

$$h_{r,f} = \sum_j \sum_{m=0}^3 w_{j,m} \cdot \delta(r - r_k, f - f_{k,m}), \quad (8)$$

ただし  $w_{j,0} = (1 - d_{j,x})(1 - d_{j,y})$ ,  $w_{j,1} = d_{j,x}(1 - d_{j,y})$ ,  $w_{j,2} = (1 - d_{j,x})d_{j,y}$ ,  $w_{j,3} = d_{j,x}d_{j,y}$  であり、  $\delta$  はパルス関数である。

#### 4.4 Area separation

画像全体を用いて結合ヒストグラムを構築すると、影や過剰な種類の材質が含まれることで、リフレクタンスとピクセルとの間で十分な関連性が得られない可能性がある。この問題を避けるため、図. 7 に示すように、高さがパノラマ画像の高さと等しい四角形の窓をいくつか一列に並べたものを設定する。我々の測定システムにおいては、一つのスキャンラインはパノラマ画像を左から右へ横切るように投影される。従って各小窓内に MI を計算するのに十分な三次元点が投影されることが保証される。MI は各窓内で計算されその総和を最適化に用いる。最終的に、6-DoF の位置姿勢パラメータは Powell 法を用いて MI を最大化することにより最適化される。

### 5. Experimental results

本稿では、三角法による位置姿勢推定と、multi-modal な

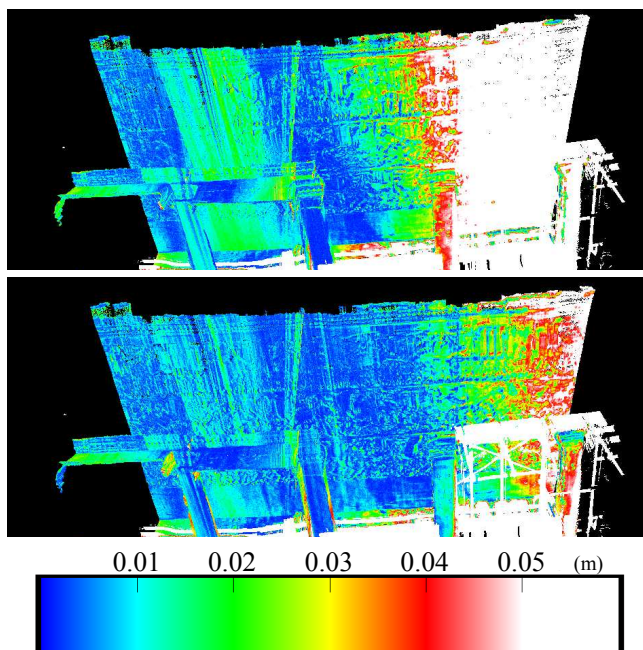


図 9 Error visualized images and a color bar. From the top: Error visualized image reconstructed by Zheng et al.'s method; Error visualized image reconstructed by our method; Color bar indicating error value. Note that scanning scene is the same location but at different time.

位置合わせによる位置姿勢推定の精度向上問題を取り扱っている。Z+F imager 5010 [26] を三次元レーザスキャナとして用い Ladybug 3 [27] を全方位カメラとして用いている。レーザスキャナはプロファイルモードで動作し視野は  $320^\circ$  で一周の解像度は: 10000 *points* である。スキャナのヘッド部分は 50 Hz で回転している。パノラマ画像は  $5400 \times 2700$  の解像度で全方位の映像を撮影している。

測定された三次元点群と二次元画像列はそれぞれレーザスキャナ内の記憶領域とカメラにつながれたラップトップ PC に測定中は保存されている。得られたデータは最新型の PC(Intel Core i7 @ 3.4G, 64GB memory) で C++ で実装されたプログラムで処理される。

### 5.1 Motion Estimation

三角法によるモーション推定の精度を測定するため、同じ個所を静的なスキャンによって測定したデータを用いた。評価のための誤差としては再構成された 3D 点群と、それに対する静的スキャンの中の最近傍点との距離を用いた。また Zheng らの手法 [8] を比較手法とした。

図. 8 に推定された経路, 図. 9 に誤差を可視化した画像, 図. 10 に各手法における誤差の度数分布表を示す。図. 9 と図. 10 から、我々の手法による結果は誤差の小さい点の数が Zheng らの手法より多いことを示している。(図. 9 では誤差の小さい点は青い点で示されている)。図. 8 の上の図は xy 平面上におけるセンサのモーションを示している。この図の青い線は +y に向かう方向にバイアスがかかって

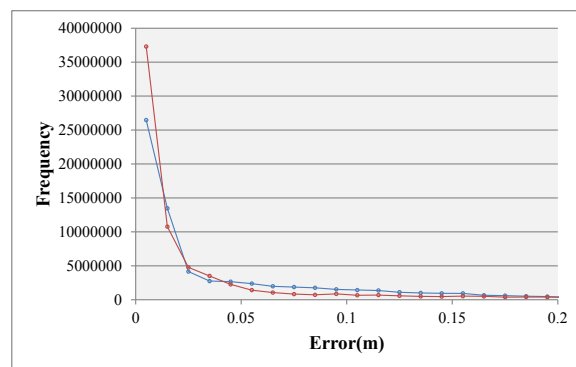


図 10 Frequency distribution of error. (blue) Zheng's method [8]. (red) Our method.



図 12 Occluded 3D points. Several 3D points scanned by laser scanner are not observed by omni-directional camera due to occlusion by the pillar

いることがわかる。これはローバの右側がレリーフしかない左側に比べると複雑な構造をしており、2D3D の誤対応がより多く存在するためと考えられる。我々の手法の方が Zheng らの手法より、これらの誤対応を影響を軽減できているために誤差となる進行方向左側へのバイアスがかかることなく精度の良い結果を示している。

図. 11 に他の二つの三次元再構成結果のリフレクタンス画像を示す。カーブ形状であっても三次元再構成が可能なが示されている。図. 11 の右図では、2D-3D の誤対応が図. 12 に示すように柱によって作られてしまっているが、デプススペースの誤対応除去によって我々の手法では柱による誤対応を受けずに三次元再構成ができている。

### 5.2 MI registration

MI を計算するため、結合ヒストグラムの構築には正規化されたリフレクタンスの値が 0.05 から 0.55 の範囲で、ピクセルのグレイスケールの値が 0.1 から 0.7 の範囲のものを用いた (グレイスケールの値の下限と上限はそれぞれ 0.0 と 1.0 である)。結合ヒストグラムの階級の数とともに 30 である。従って正規化リフレクタンスの階級幅は  $1.66 \times 10^{-2}$ 、グレイスケールの値の階級幅は  $2.0 \times 10^{-2}$  となる。モーションの最適化を行うためのセグメントの長さはおおよそ 14.5 *sec* である。これらの値は経験的に決めている。



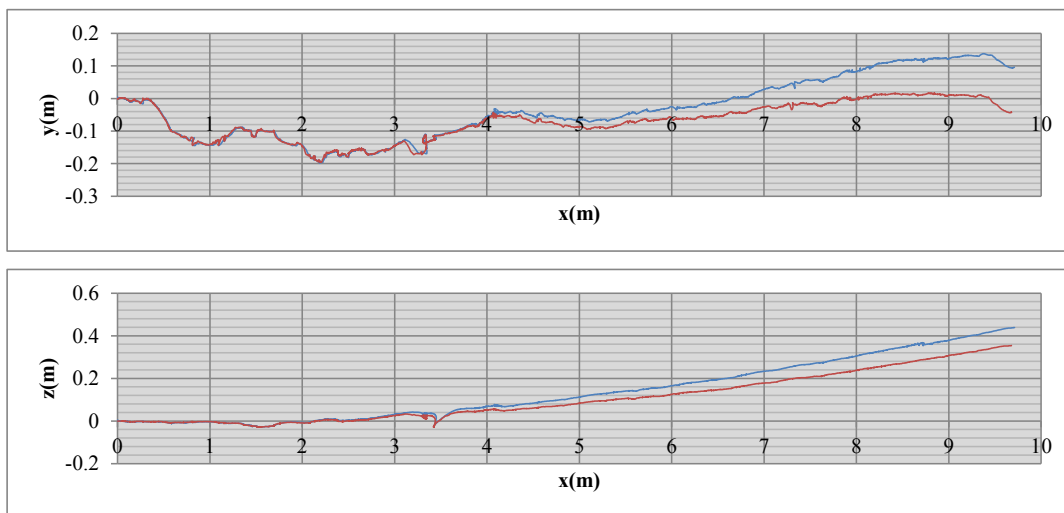


図 8 Estimated motion. (top) trajectory on xy-plane, (bottom) trajectory on xz-plane, blue line is sensor trajectory estimated by Zheng's method [8] and red line is estimated by our method.

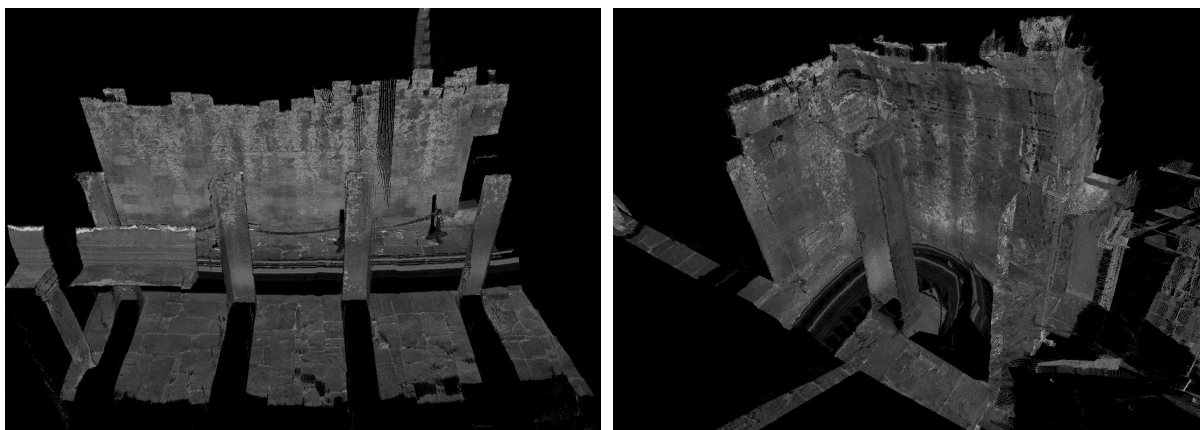


図 11 Other results. Reconstructed results are shown as reflectance image.

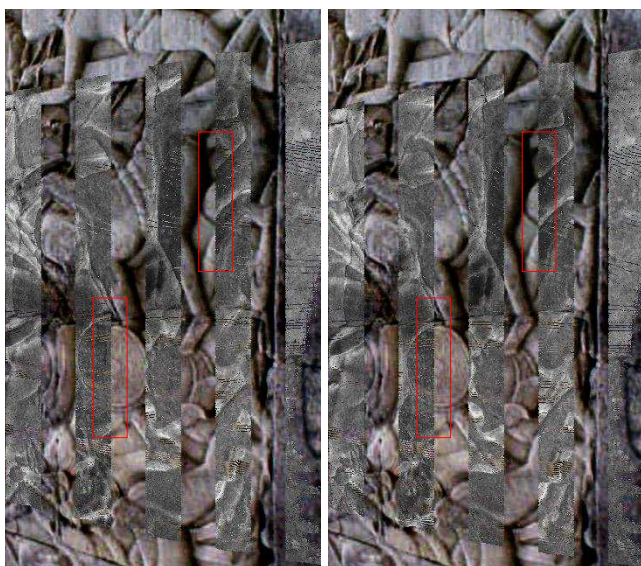


図 13 Images that pixels on where 3D points are projected are colored by reflectance. (left) before registration, (right) after registration. Areas surrounded red windows show that reflectance image are well registered by our method.

図. 13 は三次元点が二次元画像上に投影された画像を示している。三次元点が投影されたピクセルはリフレクタンスの値で色付けされており、三次元点が投影される箇所は二次元画像とリフレクタンス画像のエッジを比較するためストライプ上に制限している。図. 13 では最適化後にリフレクタンス画像と二次元画像のエッジが、とりわけ赤で囲ったエリアにおいて、よりよい連続性をもっていることが示されている。

我々の位置合わせ手法では二次元画像上では良い連続性を持つようモーションが補正されているが、三次元形状においては大きな効果を示せていない。これは2D-3Dの外部校正に誤差が含まれることやモーションの変形モデルが低い解像度であることが原因として考えられる。

## 6. Conclusion

本稿ではセンサ融合システムによる三次元再構成手法を示した。我々の手法はプロファイラタイプのレーザスキャナとカメラの融合システムに向けて設計されており、トラッ



キングに向けた投影点を用いることでモーション推定の安定性を上げ、また、レーザスキャナとカメラの位置の違いによる誤対応を考慮して除去することによって推定の精度を上げている。我々の測定システムによって高密度、高再現性をもったバスレーフの三次元モデルを得ることができ、MIによるモーション推定にはまだ改善の余地があるが、経路の最適化によってリフレクタンス画像と二次元のRGB画像の位置合わせを行った。

我々の手法は他のレーザスキャナ（安価な1軸型のLiDARであっても）とカメラの融合システムにおいても役立つと考えられる。またリフレクタンス情報を用いてのグローバルなモーション修正は価値とやりがいのある課題である。

## Acknowledgment

本研究は、JSPSの科研費（番号16H05864, 16H02851, 25257303, 16J09277）による補助を受けている。

## 参考文献

- [1] C. H. Tong, S. Anderson, H. Dong, T. Barfoot, Pose Interpolation for Laser-based Visual Odometry, *Journal of Field Robotics* 31 (5) (2014) 731–757.
- [2] J. Zhang, S. Singh, LOAM: Lidar Odometry and Mapping in Real-time, in: *Robotics: Science and Systems Conference*, 2014, pp. 109–111.
- [3] S. Scherer, J. Rehder, S. Achar, H. Cover, A. Chambers, S. Nuske, S. Singh, River mapping from a flying robot: state estimation, river detection, and obstacle mapping, *Autonomous Robots* 32 (5) (2012) 189–214.
- [4] M. Bosse, R. Zlot, P. Flick, Zebedee: Design of a Spring-Mounted 3-D Range Sensor with Application to Mobile Mapping, *IEEE Transactions on Robotics* 28 (2012) 1104–1119.
- [5] G. Pandey, J. R. McBride, R. M. Eustice, Ford Campus vision and lidar data set, *International Journal of Robotics Research* 30 (13) (2011) 1543–1552.
- [6] J. Zhang, S. Singh, Visual-lidar Odometry and Mapping: Low-drift, Robust, and Fast, in: *International Conference on Robotics and Automation*, 2014.
- [7] Y. Bok, Y. Jeong, D. Choi, I. Kweon, Capturing Village-level Heritages with a Hand-held Camera-Laser Fusion Sensor, *International Journal of Computer Vision* 94 (2011) 36–53.
- [8] B. Zheng, T. Oishi, K. Ikeuchi, Rail Sensor: A Mobile Lidar System for 3D Archiving the Bas-reliefs in Angkor Wat, *IPSI Tran. on Computer Vision and Applications* 7 (2015) 59–63.
- [9] A. Banno, T. Masuda, T. Oishi, K. Ikeuchi, Flying Laser Range Sensor for Large-Scale Site-Modeling and Its Applications in Bayon Digital Archival Project, *International Journal of Computer Vision* 78 (2-3) (2008) 207–222.
- [10] Y. Furukawa, J. Ponce, Accurate, Dense, and Robust Multiview Stereopsis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (8) (2010) 1362–1376.
- [11] G. Klein, D. Murray, Parallel Tracking and Mapping for Small AR Workspaces, in: *International Symposium on Mixed and Augmented Reality*, 2015.
- [12] R. A. Newcombe, S. J. Lovegrove, A. J. Davison, DTAM: Dense tracking and mapping in real-time, in: *International Conference on Computer Vision*, 2011.
- [13] C. Forster, M. Pizzoli, D. Scaramuzza, SVO: Fast semi-direct monocular visual odometry, in: *International Conference on Robotics and Automation*, 2014.
- [14] D. Nister, O. Naroditsky, J. Bergen, Visual odometry for ground vehicle applications, *Journal of Field Robotics* 23 (1) (2006) 3–20.
- [15] M. Maimone, Y. Cheng, L. Matthies, Two years of Visual Odometry on the Mars Exploration Rovers, *Journal of Field Robotics* 24 (2) (2007) 169–186.
- [16] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, A. D. D. Kim, P. Kohli, J. Shotton, S. Hodges, A. Fitzgibbon, KinectFusion: Real-time dense surface mapping and tracking, in: *International Symposium on Mixed and Augmented Reality*, 2011.
- [17] N. Engelhard, F. Endres, J. Hess, J. Sturm, W. Burgard, Real-time 3D visual SLAM with a hand-held RGB-D camera, in: *RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum*, 2011.
- [18] T. Whelan, H. Johannsson, M. Kaess, J. Leonard, J. Mc-Donald, Robust real-time visual odometry for dense RGB-D mapping, in: *International Conference on Robotics and Automation*, 2013.
- [19] N. Matsuda, O. Cossairt, M. Gupta, MC3D: Motion Contrast 3D Scanning, in: *International Conference on Computational Photography*, 2015.
- [20] B. D. Lucas, T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, in: *International Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.
- [21] R. M. Haralick, C.-n. Lee, K. Ottenburg, M. Nölle, Analysis and solutions of the three point perspective pose estimation problem, in: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, IEEE, 1991, pp. 592–598.
- [22] M. A. Fischler, R. C. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM* 24 (6) (1981) 381–395.
- [23] R. Kurazume, K. Nishino, Z. Zhang, K. Ikeuchi, Simultaneous 2D images and 3D geometric model registration for texture mapping utilizing reflectance attribute, in: *Asian Conference on Computer Vision*, Citeseer, 2002, pp. 99–106.
- [24] P. Viola, W. Wells III, Alignment by Maximization of Mutual Information, *International Journal of Computer Vision* 24 (2) (1997) 137–154.
- [25] F. Maes, D. Vandermeulen, P. Suetens, Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information, *Medical Image Analysis* 3 (4) (1999) 373 – 386.
- [26] Zoller + Fröhlich GmbH, <http://www.zf-laser.com/>.
- [27] Point Gray Research, <http://www.ptgrey.com/>.