

画像列を利用した移動型レンジセンサから得られる 3次元形状修復

阪野 貴彦[†] 長谷川一英[†] 池内 克史[†]

[†] 東京大学 生産技術研究所 〒 153-8505 東京都目黒区駒場 4-6-1

E-mail: †{vanno,k-hase,ki}@cvl.iis.u-tokyo.ac.jp

あらまし 巨大物体の 3次元形状計測では、クレーンやヘリコプタ等を利用した空中からのスキャンが効果的な方法の 1つとして考えられる。ただし、対象物体が貴重な文化遺産である場合、安全性や効率性の面から、このような計測にはさまざまな問題が発生する。そのためわれわれは、気球搭載型のレンジセンサ (FLRS) を利用した。しかし、FLRS では計測中にレンジセンサが運動するため、獲得する形状データに歪みが生じてしまう。そこで、本稿では計測中に移動するレンジセンサから得られるデータを補正する手法を提案する。気球に取り付けたビデオカメラによる画像列から、Full-Perspective Factorization を用いて初期解を求め、次に歪んだ形状データそのものから抽出できる情報を用いて、カメラ運動に関するパラメータの高精度な推定をおこなう。このように推定されたセンサの運動パラメータを用いて、歪んだ形状データを補正した。この手法をカンボジア、パイヨン寺院の計測に適用したところ、FLRS から得られた歪んだ形状データを精度良く復元することができた。

キーワード 3次元計測, 浮遊型レンジセンサ, 大規模文化財, カメラ運動推定

Shape Recovery of 3D Data Obtained from a Moving Range Sensor by using Image Sequences

Atsuhiko BANNO[†], Kazuhide HASEGAWA[†], and Katsushi IKEUCHI[†]

[†] Institute of Industrial Science, University of Tokyo Komaba 4-6-1, Meguro-ku, Tokyo, 153-8505 Japan

E-mail: †{vanno,k-hase,ki}@cvl.iis.u-tokyo.ac.jp

Abstract For a large object, scanning from the air is one of the most efficient methods of obtaining 3D data. In the case of large objects valuable for their cultural heritage, there are some difficulties in scanning them with respect to safety and efficiency. To remedy these problems, we use the Floating Laser Range Sensor (FLRS), in which a range sensor is suspended beneath a balloon. The obtained data, however, have some distortion due to movement during the scanning process. In this paper, we propose a method to recover 3D range data obtained by a moving laser range sensor. This method is applicable not only to our FLRS, but also to a general moving range sensor. Using image sequences from a video camera mounted on the FLRS enables us to estimate the motion of the FLRS without any physical sensors such as gyros and GPS. At first, the initial values of camera motion parameters are estimated by full perspective factorization. The next stage refines camera motion parameters using the relationships between camera images and the range data distortion. Finally, by using the refined parameters, the distorted range data are recovered. In addition, our method is applicable with an un-calibrated video camera and range sensor system. We applied this method to an actual scanning project and the results showed the effectiveness of our method.

Key words 3D Measurement, Shape from Motion, Moving Range Sensor, Large Cultural Heritage, Motion Estimation

1. はじめに

最近の 3次元形状計測技術の発達により、実物体のモデリングに関する研究がさかんにおこなわれている。実物体モデリングの技術は、学術、産業、エンターテインメントなど多くの分

野で必要とされ、また波及効果が期待できる。

その中でも、巨大文化遺産のモデルリングはもっとも重要で、包括的な内容を含んだアプリケーションのひとつである。これら文化遺産のモデル化は、多くの分野で重大な意義をもたらす。まず、モデル化をおこなうことで、文化遺産の形状をデジタル

化されたデータとして、アーカイブ化することができる。次に、文化遺産の形状をデジタル化という劣化しないデータとして保存できるため、自然災害、火災、戦争等などによって、たとえば破壊されたとしても、修復・復元が可能となる。また、インターネットやDVDなどの媒体を通して、自宅に居ながらにして、その文化遺産を訪れたような疑似体験を提供できるシステムを構築することも可能となる。

このように、実物体のモデリングの技術は多くのアプリケーションに利用することができる。コンピュータビジョンの分野ではこれまでに、彫刻や大仏、歴史的建造物、街並などの巨大文化遺産を対象としたモデリングがおこなわれてきた[10][12][7]。このようなモデル化をおこなうとき、まず対象物体の3次元形状を計測することから始まる。基本的に、計測する際には、計測機器を三脚に取り付け、安定したところに置く。しかし、対象物体が巨大になると、地上においた計測機器から計測できない部分が生じることになる。そのため、その近くに一時的に櫓を組んで巨大物体を計測してきた。ところが、この方法はコストがかかるうえ、櫓を組むにも時間がかかってしまう。さらに、対象物体が巨大になれば、周りの櫓からは計測できない部分が大きくなり、この計測方法での限界が訪れる。

著者らは現在、カンボジア王国・アンコール遺跡群のひとつバイヨン寺院[19]をデジタルコンテンツ化するプロジェクトをおこなっている[8]。この寺院の大きさは、およそ100メートル四方もあり、これまでようにいくつか櫓を組んだだけでは、全体を計測することは不可能である。このような超巨大物体を計測する場合、これまでいくつかの方法が考えられてきた。たとえば、ヘリコプターや航空機にレンジセンサを搭載しての3次元計測である[17]。しかし、この場合、ヘリコプターなどのエンジンによって、高周波数の振動が発生してしまい得られる形状データに影響を与えてしまう、近くに接近できない、等の問題が起こる。また、対象物体が文化的に非常に価値が高いことを考慮すれば、ヘリコプターはおろかクレーンなどの重機の使用も、安全面を考慮して避けるべきであろう。

以上のことを踏まえて、[6]では巨大文化遺産を計測するための新しい方法を考案し、これを Floating Laser Range Sensor(以下、FLRS、図1参照)と名付けた。このシステムは、形状計測機器を気球に吊るして、空中から計測をおこなうものである。この方法により、高周波数の振動を発生させずに、広範囲にわたって形状計測が可能となった。ただし、問題点も発生する。計測時間中に計測機器そのものが運動するため、結果として得られた形状データが歪んでしまうことである[5]。

そこで本稿では、このようなFLRSから得られた歪んだデータを復元する手法を提案する。レンジセンサが動くことによって歪んだレンジデータが得られるが、同時にFLRSに搭載したビデオカメラによって画像列も獲得する。ここでは、GPSやジャイロ等のセンサを用いずに、この画像列からFLRSの動きを推定する。FLRSの動きを推定するにあたっては、歪んだレンジデータそのものからも条件を抽出して、いくつかの制約条件下でのコスト関数の最小化をおこなった

このような非線型関数の最適化問題を解く際、ローカルミニ



図1 FLRSとバイヨン寺院

マムに陥るのを避けるために、われわれは初期解として Factorization [18][2][13][14][4]による解を用いた。特に今回、実際のカメラモデルである透視投影下での Perspective-Factorizationの解を採用した。このようにして推定されたFLRSの運動から、最終的に歪んだレンジデータの修正をおこなう。さらに、ここで提案する手法は、気球搭載型のFLRSに特化した手法ではなく、一般的に滑らかに運動するレンジセンサに対して適用可能である。

本稿の構成は以下の通りである。2章では、われわれが提案するセンサ運動推定のための手法について述べる。まず、Full-Perspective Factorizationについての簡単な説明をおこなう。次に、カメラ運動のパラメータ推定に関するリファインメント手法について述べる。3章では、本手法をバイヨン寺院に適用した実験とその結果について述べる。また、本手法を評価するため、修正したレンジデータと、地上に固定したレンジセンサから得られたレンジデータとの比較をおこなった。ここまでの手法は、ビデオカメラとレンジセンサとの間でキャリブレーションがされていることが前提であったが、4章では、ビデオセンサ間でのキャリブレーションができていない場合でも、カメラ運動のリファインメントがおこなえることを示す。そして最後にまとめとする。

2. 運動パラメータ推定手法

本研究では、FLRSから得られる画像と歪んだ形状データを用いて、センサの動きを推定することによって、歪んだ形状データを修復することである。そこで、われわれは3つの制約条件を課し、最適化問題のフレームワークによって、カメラ位置・姿勢の推定をおこなう。この条件のもと、コスト関数の最小化の枠組みで最適化問題を解くことで、カメラパラメータの高精度な推定をおこなった。われわれの用いたコスト関数は非線型であり、ローカルミニマムに陥るのを避けるために、あらかじめ精度の良い初期解を必要とする。

そこで、ピンホールカメラモデルを忠実に再現した Full-Perspective Factorizationによる解を初期解とすることで、最適化計算の安定化をはかった。

2.1 Full-Perspective Factorization

ビデオカメラでは、 F 枚のフレームにわたって、 P 個の特徴点をトラッキングする。フレーム番号 f において、トラックされた特徴点 p は画像座標系で、 $\{(u_{fp}, v_{fp}) | f = 1, \dots, F, p = 1, \dots, P, 2F \geq P\}$ に表示されているとする。

画像フレーム f を取得するときのカメラ位置を、世界座標系で \vec{t}_f とする。また、そのときのカメラの姿勢を 3 つの単位ベクトル $\vec{i}_f, \vec{j}_f, \vec{k}_f$ で表す。この 3 つのベクトルは、カメラ座標系でのそれぞれの座標軸を表していて、 \vec{i}_f, \vec{j}_f は画像面に平行な 2 つの軸であり、 \vec{k}_f は光軸に平行な軸とする。

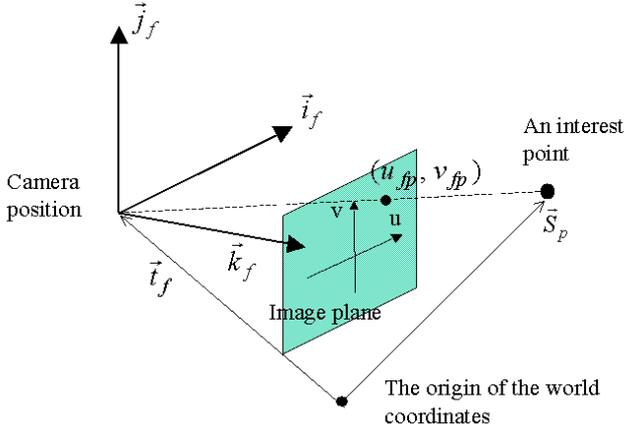


図 2 カメラ運動のパラメータ

したがって、推定すべきパラメータは、 P 個の 3 次元座標と F 個のカメラパラメータ (位置, 回転) であり、すべてで $(3 \times P + 6 \times F)$ 個ある。

ピンホールカメラモデルでは、3 次元世界座標系において \vec{s}_p にある点は、次の式によってフレーム f の画像面で (u_{fp}, v_{fp}) に投影される。

$$u_{fp} = f \frac{\vec{i}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \quad (1)$$

$$v_{fp} = f \frac{\vec{j}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \quad (2)$$

簡単な式変形により、右辺を弱透視投影での式と同じ形で、次のようになる。

$$\frac{\vec{k}_f(\vec{s}_p - \vec{t}_f)}{z_f} u_{fp} = \frac{f}{z_f} \vec{i}_f(\vec{s}_p - \vec{t}_f) \quad (3)$$

$$\frac{\vec{k}_f(\vec{s}_p - \vec{t}_f)}{z_f} v_{fp} = \frac{f}{z_f} \vec{j}_f(\vec{s}_p - \vec{t}_f) \quad (4)$$

$z_f = -\vec{k}_f \vec{t}_f$ と置き換えることにより、

$$(\lambda_{fp} + 1)u_{fp} = \frac{f}{z_f} \vec{i}_f(\vec{s}_p - \vec{t}_f) \quad (5)$$

$$(\lambda_{fp} + 1)v_{fp} = \frac{f}{z_f} \vec{j}_f(\vec{s}_p - \vec{t}_f) \quad (6)$$

$$\lambda_{fp} = \frac{\vec{k}_f \cdot \vec{s}_p}{z_f} \quad (7)$$

となる。つまり、2 次元画像上における各特徴点の座標値 (u_{fp}, v_{fp}) にある数 λ_{fp} を掛けることにより、透視投影空間からアフィン空間である弱透視投影空間に変換することができる。したがって、この変換に必要な λ_{fp} を繰り返し計算によって求めることで、弱透視投影下での Factorization を用いて、透視投影空間での物体形状、カメラの運動パラメータを求めることができる。以下にそのアルゴリズムを示す。

Input 各フレームにおいて P 個の特徴点をトラッキングした F 枚の画像

Output P 個の特徴点の 3 次元座標 \vec{s}_p およびフレーム番号 f を撮影した時のカメラ位置 \vec{t}_f と姿勢ベクトル $\vec{i}_f, \vec{j}_f, \vec{k}_f$ の各成分

(1) $\lambda_{fp} = 0$ と仮定する。

(2) $(\lambda_{fp} + 1)u_{fp} = f \vec{i}_f(\vec{s}_p - \vec{t}_f)/z_f, (\lambda_{fp} + 1)v_{fp} = f \vec{j}_f(\vec{s}_p - \vec{t}_f)/z_f$ と置き、弱透視投影での Factorization をおこない、 $\vec{s}_p, \vec{t}_f, \vec{i}_f, \vec{j}_f, \vec{k}_f, z_f$ を求める。

(3) $\lambda_{fp} = \vec{k}_f \cdot \vec{s}_p / z_f$ を計算する。

(4) step3 で求めた λ_{fp} を step2 に代入し、再び Factorization を解いて新たに λ_{fp} を求める。これを繰り返す。

Until λ_{fp} の更新する値が充分小さくなるまで。

2.2 高精度パラメータ推定のためのリファインメント

2.2.1 トラッキングにおける制約

最も基本的な制約条件として、3 次元空間中の特徴点 \vec{s}_p は、各フレームで画像中の (u_{fp}, v_{fp}) に投影されていなければならない。つまり、実際にえられた画像中の特徴点と、再構成された画像面に投影された特徴点との誤差が最小となるように制約を課すこの制約条件によって、次のコスト関数を最小化しなければならないことがわかる。

$$F_A = \sum_{f=1}^F \sum_{p=1}^P \left(\left(u_{fp} - f \frac{\vec{i}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \right)^2 + \left(v_{fp} - f \frac{\vec{j}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \right)^2 \right) \quad (8)$$

F_A を最小化することによって、3 次元空間で固定している特徴点を、移動するカメラによって正しくトラックする様子が再現できる。ただし、実際には微分計算を簡単にするため、次のように変形したものを F_A として実装する。

$$F_A = \sum_{f=1}^F \sum_{p=1}^P \left(\left(\vec{k}_f(\vec{s}_p - \vec{t}_f) u_{fp} - f \vec{i}_f(\vec{s}_p - \vec{t}_f) \right)^2 + \left(\vec{k}_f(\vec{s}_p - \vec{t}_f) v_{fp} - f \vec{j}_f(\vec{s}_p - \vec{t}_f) \right)^2 \right) \quad (9)$$

2.2.2 カメラの運動における制約

[6] では気球を選択した理由のひとつが、気球によって高周波数の振動の影響を受けない、というものであった。裏を返せば、気球は低周波数のみの滑らかな振動の影響を受ける、といえる。気球を用いて計測をおこなう場合、地上から人手でロープによって支えておこなっているため、気球の運動に影響を与

えるものは、風や、ロープを伝う人の力のみである。したがって、気球の運動は非常に滑らかであり、急加速、急減速、折れ曲がった軌道を描くような運動は実際に発生することがない。つまり、気球の運動に特異点や不連続性はないものと仮定できる。このことから、つぎのコスト関数を設定する。

$$F_B = \int \left(w_1 \left(\frac{\partial^2 \vec{t}_f}{\partial t^2} \right)^2 + w_2 \left(\frac{\partial^2 \mathbf{q}_f}{\partial t^2} \right)^2 \right) dt \quad (10)$$

ここで、 \vec{t}_f はカメラの位置を表し、 t は時間、 w_1, w_2 は重み係数を表す。カメラの回転については、単位クォータニオン \mathbf{q}_f で評価している。このクォータニオンは $\vec{i}_f, \vec{j}_f, \vec{k}_f$ から直ちに求めることができる。被積分関数の第1項は、カメラの位置に関する移動が滑らかであることの制約となっており、第2項はカメラの回転に関して滑らかであることの制約になっている。したがって、カメラの運動が滑らかであるとき、関数 F_B は小さな値をとる。ただし、実装において積分は離散的におこなわれるため、次の式を用いる。

$$F_B = \sum_{f=1}^F \left(w_1 \left(\frac{\partial^2 \vec{t}_f}{\partial t^2} \right)^2 + w_2 \left(\frac{\partial^2 \mathbf{q}_f}{\partial t^2} \right)^2 \right) \quad (11)$$

2.2.3 歪んだレンジデータに関する制約

FLRS から得られる距離画像は、センサそのものが運動しているため、結果として正しくない形状が得られる。ただし、“正しくない” というのは距離画像全体を見た場合であって、局所的には正確な情報を含んでいる。以下の制約では、この局所的に正確な距離情報を用いた制約条件を示す。

FLRS では、距離画像面で見るとラスタスキャン順にレーザーを発している。1回の計測に要する時間は既知であるので、距離画像の各ピクセルでの距離値を計測したときの、計測開始時からの時刻が求められる。つぎに、気球に搭載したレンジセンサとビデオカメラの同期がとれていれば、その距離画像の各ピクセルを計測したときのビデオ画像のフレーム番号 f を見つけることができる。さらに、レンジセンサとビデオカメラとのキャリブレーションができていれば、その計測された距離画像中のピクセルが、ビデオ画像 f のどの部分を計測しているのかがわかる。このとき、ビデオ画像に対応する部分が、トラックしている特徴点に一致していると、ビデオ画像 f を撮影している時刻でのセンサと3次元空間中の特徴点との相対位置関係が特定できる。したがって、第3の制約条件として、以下の関数を最小化をおこなう。

$$F_C = \sum_{p=1}^P \left\| \mathbf{x}_{fp} - \|\vec{s}_p - \vec{t}_{fp}\| \right\|^2 \quad (12)$$

ここで、添え字 fp は、特徴点 p がレンジセンサによってスキャンされたときのフレーム番号を示す。また、 \mathbf{x}_{fp} はこのときのレンジセンサの出力である。

あるいは、 $\mathbf{x}_{fp} = (x_{fp}, y_{fp}, z_{fp})$ とすれば、上の制約をさらに厳しくしたものとして、

$$F_C = \sum_{p=1}^P \left((x_{fp} - \vec{i}_{fp}(\vec{s}_p - \vec{t}_{fp}))^2 \right.$$

$$\left. + (y_{fp} - \vec{j}_{fp}(\vec{s}_p - \vec{t}_{fp}))^2 + (z_{fp} - \vec{k}_{fp}(\vec{s}_p - \vec{t}_{fp}))^2 \right) \quad (13)$$

が得られる。

2.2.4 全コスト関数

以上の3つの関数の重みつき加算をおこなった全コスト関数

$$F = w_A F_A + w_B F_B + w_C F_C \quad (14)$$

を最小化することによって、特徴点位置とカメラの運動パラメータに関しての推定をおこなう。

この全コスト関数の最小化計算には、 F_B の計算が、いわゆる関数の当てはめではないため、共役勾配法 [15] [16] を用いた。各共役方向に関する1次元の最小化探索には、黄金分割法を採用することで、計算の効率化を図った [9]。

この非線型最小化のためには、ローカルミニマムに陥るを防ぐのに、真の解に近い初期解が必要である。そのため、前述したように最小化計算の初期解として、Full-Perspective Factorization による解を用いた。

3. 実験

このバルーン型プラットフォームを用いて、カンボジア王国のバイヨン寺院の計測をおこなった。この寺院の大部分は地上に置かれたセンサから計測可能であるものの、屋根や塔の最上部に関しては地上からのセンサでは計測不能であるため、主にこれらの部分を空中から計測した。

われわれが用いたシステムでは、1回の計測に72枚の画像が得られる。各画像からの特徴点抽出には、スケール変化に関してロバストな SIFT key [11] を用いた。SIFT key から得られる128次元特徴量のマッチングをとることで、全フレームに渡っての特徴点のトラッキングをおこなった。これにより、72フレームを通して、各計測で平均100点ほどの特徴点をトラックすることができた。

形状復元の実験結果を図3に示す。同図の左(a)はFLRSから得られたそのままの歪んだデータ(以下、オリジナルデータ)である。このデータセットを獲得したときの気球の動きが大きかったため、得られた形状データの歪みは大きい。図3の中央(b)は、Full-Perspective Factorization のみによって復元した形状である。Factorization のみによる運動パラメータ推定では、スケールの曖昧性が残ってしまうため、手で最適なスケールを与えた。このモデルを一見すると、Factorization のみによるものでも、適切に復元できているように見える。しかし、全体的に見ると、S字型の歪みがあり、長方形となるべき入り口の形状も歪んでいるのが分かる。一方、図3の右側(c)は、われわれの手法によって復元した形状である。同図(b)で見られたようなS字型の歪みが除去されており、入り口の形状も正しく長方形に復元されている。

次に、この手法によって復元した形状の精度を評価するため、地上に置かれたCyrax2500によって得られた形状データとの比較をおこなった。この地上固定センサによって得られた形状を正確なデータとして、以降の考察を進める。提案手法で復元

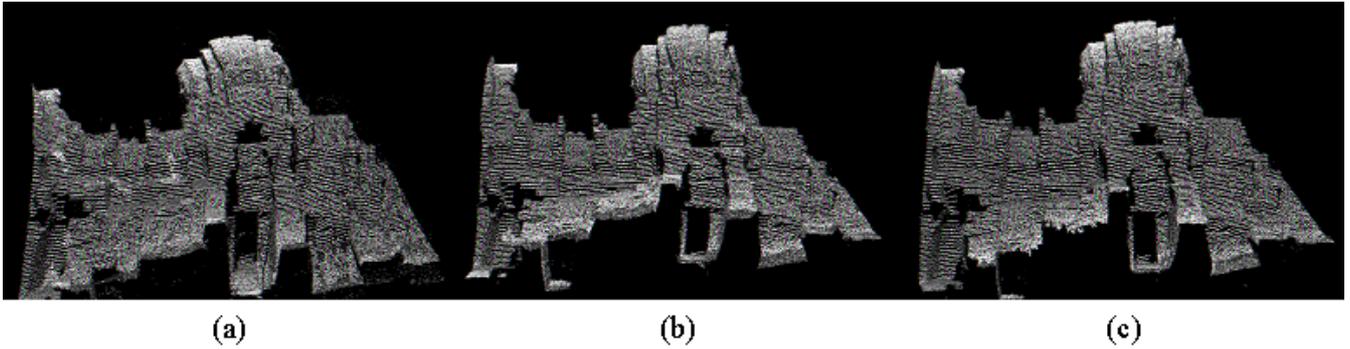


図3 (a):FLRS から得られる歪んだオリジナルデータ. (b):Full-Perspective Factorization による復元モデル (スケールの曖昧性は除去済) (c):提案手法による復元モデル.

した形状と、地上固定センサによって得られた形状との間で、ICP アルゴリズム [1] [20] を適用し、位置合わせをおこない、重なり合った部分に注目する。位置合わせの結果を図 4 に示す。図 4(a) に計測した領域を写した画像、(b) には位置合わせの結果を示す。図 4(b) で、白色の密な形状データは、地上に固定されたセンサによるモデルであり、一方の紫色の粗い形状データは、提案手法によって復元されたモデルである。このように、位置合わせされた両者のデータより、本手法によって適切に形状が補正されていることが分かる。

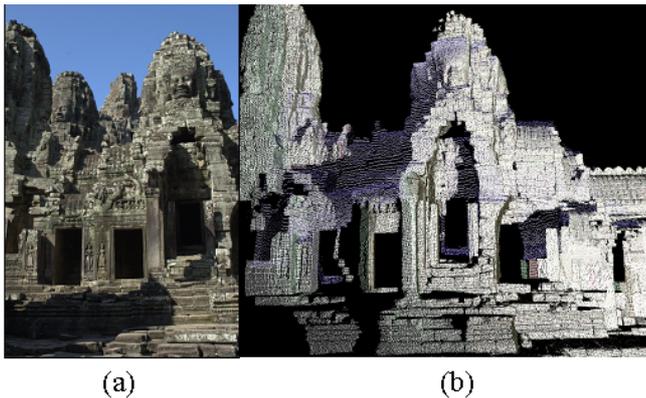


図4 (a):計測した部分を撮影した実画像. (b):地上固定センサから得られたデータに、提案手法によって復元した形状モデルを ICP アルゴリズムによって位置合わせしたもの

図 5 は、図 3 で示した形状と地上固定センサによって得られた形状データとをそれぞれ ICP アルゴリズムによって位置合わせしたときの、対応点間距離をそれぞれ示したものである。図 5 の左側から、(a)FLRS から得られたオリジナルデータ、(b)Full-Perspective Factorization のみによって復元した形状 (ただし、スケールの曖昧性については除去済み)、(c) 提案手法によって復元した形状、と地上固定センサによる形状とを比較したものである。緑色に色付けされた領域は、ICP での対応点間距離が閾値 (6.0cm に設定) 以下の領域を示しており、比較している両者の形状が一致していると見なすことができる。青色の領域は、対応点間距離が閾値以上の領域であり、形状が一致していないと見なすことができる。右側に行くにしたがって、

緑色の領域が広がっていることがわかり、本手法によって復元された形状が、地上固定センサによって得られた形状と一致しているようすがわかる。なお、復元した形状モデルの中央付近にあたる部分が青色になっている (形状が一致していない) のは、地上固定センサによる形状モデルにおいて対応する部分にデータがないからである。

最後に、推定されたカメラの運動パラメータについて考察する。図 6 は、推定されたカメラ運動の並進成分について、カメラの画像面に平行な 2 成分を時間軸に沿って示したものである。図 6(a) は Full-Perspective Factorization のみによって推定されたカメラの運動パラメータからスケールの曖昧性を取り除いたものを示しており、(b) は提案手法によって推定された運動パラメータを示している。このデータセットでは、Factorization での収束が良くなかったためか、図 6(a) のグラフは滑らかでない不連続なものとなっており、現実の気球の運動としては考えられない。一方、図 6(b) でのグラフは (a) の結果を初期解として得られたものであるが、滑らかに変化しており、気球の運動らしいものになっている。図 6(b) による運動パラメータをもとに復元した形状が正しく復元されていることから、結果として、運動パラメータが正しく推定できていると考えられる。

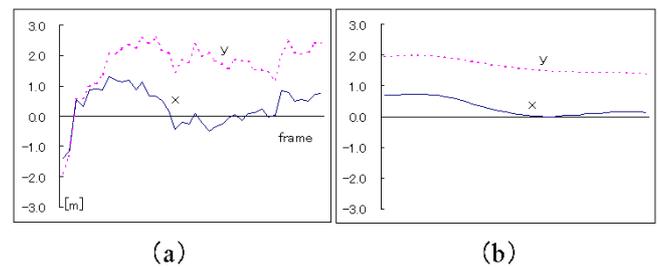


図6 カメラ運動の並進成分の軌跡: (a) Full-Perspective Factorization によって得られたパラメータから手動によってスケールの曖昧性を除去したもの. (b) 提案手法によって推定された軌跡

4. キャリブレーションされていないレンジセンサ - ビデオカメラ

これまでに述べてきた手法は、レンジセンサでスキャンした領域と、ビデオカメラによって撮影された領域との位置関係が

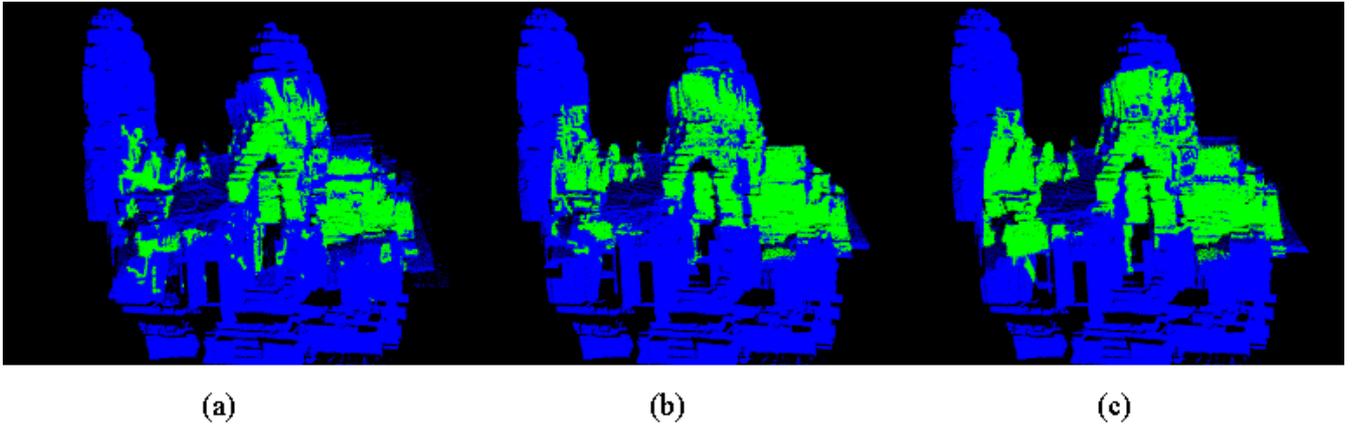


図5 地上に固定されたレンジセンサ (Cyrax2500) から得られた形状モデルとの比較。(a):FLRSから得られる歪んだオリジナルデータ,(b):Full-Perspective Factorizationによる復元モデル(スケールの曖昧性は除去済),(c):提案手法による復元モデル,のそれぞれをCyrax2500によって得られた正確なデータとICPアルゴリズムによって位置合わせし,そのときの対応点間距離を示したもの.緑色の領域は対応点間距離が6.0cm以下の領域であり形状が一致していることを示している.青色は対応点間距離がそれ以上,もしくはどちらかのデータが欠損している領域.

あらかじめ既知の場合であった。つまり,レンジセンサ座標系とビデオカメラ座標系との変換が既知であるキャリブレーションされたシステムであった。

ここでは,レンジセンサに対して,ビデオカメラをどのように取り付けたかが未知であるような,キャリブレーションされていないシステムであっても,われわれの手法が適用できることを示す。レンジセンサとビデオカメラがキャリブレーションされていないとき,2.2.3で述べたようなレンジセンサからカメラ運動を推定するための情報を抽出することができない。したがって,両者をキャリブレーションする必要がある。まず,ビデオ画像列のみを利用して,対象物体のユークリッド復元と,カメラ運動パラメータの推定をおこなう。次に,推定された形状を,カメラ運動に従って変形する。このように変形された推定形状は,運動するレンジセンサから得られたレンジデータと相似形をなしているはずである。そのため,両者をスケールパラメータを含んだ位置合わせをおこなうことによって,レンジセンサ座標系とビデオカメラ座標系とのキャリブレーションができる。このように,両者のキャリブレーションができれば,前章で述べた手法が適用できる。以下には,この手法についての詳細を述べる。

4.1 追跡点の増加

まず,ビデオ画像列のみを利用して,対象物体のユークリッド復元と,カメラ運動パラメータの推定をおこなう。前章と同様に,ここでも,Full-Perspective Factorizationを用いて,画像列のみから形状とカメラ運動を取り出す。このときの解を初期解として,リファインメントをかけるのだが,ここではレンジセンサとビデオカメラとの間でキャリブレーションができていないので,2.2.3で述べた“歪んだレンジデータに関する制約”を課することはできない。したがって,“カメラの運動における制約”(bundle adjustment)と“カメラの運動における制

約”(滑らかな運動を保証)の2つの制約条件の下でカメラ運動推定の高精度化をおこなうが,ここではまだ,スケールの曖昧性が残ることになる。

この段階で,画像列から復元できている形状は,画像列の全フレームに渡って追跡できている特徴点のみである。仮に,このような疎な形状データと,レンジセンサから得られた密な形状データとの間でスケールを含めた位置合わせをおこなおうとする場合,間違っただけに収束してしまうことが容易に想像される。このため,形状復元するための特徴点数を増やす。本研究では,ある特徴点が一定の連続したフレーム数以上で追跡されていれば,3次元復元をおこなうことにした。このようにして付け加えられた特徴点の3次元復元には,最尤推定法(ML法)を利用した。

$$\vec{s}_{ML} = \max_{\vec{s}} p(U|\theta, \vec{s}) \quad (15)$$

ML法では,カメラの運動パラメータ θ と3次元特徴点 \vec{s} が与えられたとき,画像列上で特徴点が位置 U で観測される確率が最も高くなるように,3次元座標 \vec{s} を推定する。われわれの実装では,ガウシアン関数をノイズモデルとして採用した。これにより,画像列の全フレームに渡って追跡されていない大量の点を3次元復元することができる。

4.2 スケール付きICPによるキャリブレーション

次に,レンジセンサによって得られた歪んだ形状と,3次元復元された特徴点を用いて,レンジセンサ座標系とビデオカメラ座標系のキャリブレーションをおこなう。フレーム番号 f を時刻を示すパラメータとして用いると,センサ全体の回転行列は各フレームで次のように表すことができる。

$$R_f = (\vec{i}_f \vec{j}_f \vec{k}_f) \quad (16)$$

次に,両座標系の変換関係を記述するための回転行列を R_{intra} ,並進ベクトルを \vec{t}_{intra} とする。スケールをあわせるための変数

を s とすれば、ロバスト推定 [3] を考慮した次式を最小化することによって、両座標系のキャリブレーションをおこなうことができる。

$$\min \sum_{p=1}^P \log\left(1 + \frac{z_{fp}^2}{2\sigma^2}\right) \quad (17)$$

$$z_{fp} = \left\| sR_f^T(\vec{s}_p - \vec{t}_f) - R_{intra}\vec{x} - \vec{t}_{intra} \right\|^2 \quad (18)$$

式 (18) 中の第 1 項は、3 次元復元された特徴点の point cloud [21] をカメラ運動に従って変形させたものに対応しており、これをレンジセンサから得られた歪んだ形状と位置合わせをおこなうことで、両座標系の関係 ($R_{intra}, \vec{t}_{intra}$) を求めることができる。 $R_{intra}, \vec{t}_{intra}, s$ を推定することができれば、2.2.3 で述べた“歪んだレンジデータに関する制約”を課すことができるので、第 2 章で述べたリファインメントを用いて、カメラ運動の高精度な推定をおこなうことができる。

4.3 実験結果

図 7(a) には画像列のサンプル画像を、図 7(b) にはキャリブレーションされていないシステムで得られた歪んだオリジナルデータを示す。図 7(a) と (b) とを見比べてみて、レンジデータの歪みの度合いが大きいことが分かる。このデータセットでは、気球の運動が激しかったため、画像での特徴点追跡も難しく、全フレームに渡って追跡できた特徴点はわずか 18 点であった。そのため、前述したように、連続 30 フレーム以上追跡できた点を復元するようにしたところ、2431 点を復元することができた。

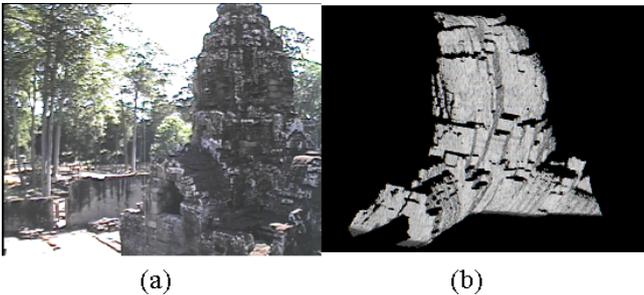


図 7 (a): 画像列の 1 例 (b): FLRS から得られた歪んだオリジナルデータ

図 8(a) には、この章で提案した手法によって復元した形状を、(b) には地上固定センサから得られた形状と (a) の形状を位置合わせしたものを示す。図 8(a) では、わずかに歪みが残っているものの、図 7(b) のオリジナルデータと比較すると、適切に復元できていることがわかる。また、図 8(b) を見てみると、地上固定センサによる形状と位置合わせできており、キャリブレーションされていないシステムにおいても、われわれの手法が適用できることを示している。

5. まとめ

われわれは、運動するレンジセンサから得られる歪んだレンジデータと、画像列を利用することで、運動パラメータを推定する手法について述べてきた。まず、ビデオカメラから得ら

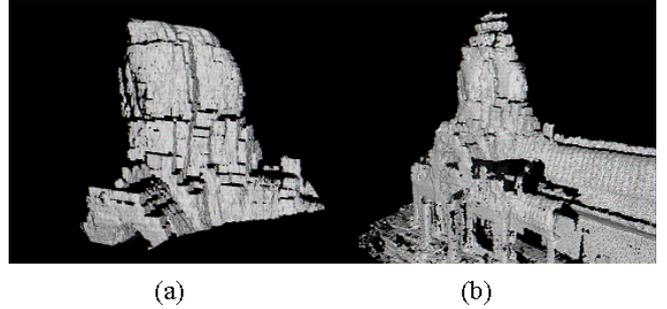


図 8 (a): キャリブレーションされていないシステムでの復元モデル (b): 地上固定センサから得られたモデルと復元モデルとを位置合わせしたもの

れた画像列を用いて、カメラの動きの推定を Full-Perspective Factorization によっておこなった。次に、3 つの制約条件の下でコスト関数を設定し、Factorization による解を初期解として、非線形最適化のフレームワークでカメラ運動推定のためのリファインメントをおこなった。このとき、歪んだレンジデータそのものからも、カメラ運動の推定を高精度化するための情報を抽出した。最後に、このように推定された運動パラメータから、歪んだ形状データを正しく復元することができた。図 9 にバイヨン寺院全体のモデルを示す。

また、本手法は、あらかじめキャリブレーションされていないシステムであっても適用可能であることを示した。画像列から 3 次元復元したものをカメラ運動に従って変形させたモデルと、運動するレンジセンサから得られた歪んだオリジナルデータを位置合わせすることによってキャリブレーションをおこなう手法を示した。

この研究で示したフレームワークでは、レンジセンサとビデオカメラとの同期が取れていて、レンジセンサが滑らかに運動することが保証できていれば、運動するレンジセンサ全般に適用が可能であり、気球を用いたプラットフォームに特化された手法ではない。

本研究では、気球を用いて、計測中に移動するレンジセンサから形状を復元し、6.0cm を閾値とした形状マッチングにおいて、良好な形状復元が実現できた。また、移動しながらレンジセンサで計測する手法が実現可能であり、大規模物体の計測に効果的であることを示すことができた。今後の課題としては、さらに高い精度での復元形状のための改良が挙げられる。

謝 辞

本研究の一部は、文部科学省科学技術振興調整費リーディングプロジェクト「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」および日本学術振興会特別研究員制度による支援を受けています。



図 9 バイヨン寺院の全体モデル

文 献

- [1] P.J.Besl and N.D.McKay, "A method for registration of 3-D shapes," *IEEE Trans. on PAMI*, vol. 14, pp. 239-256, 1992.
- [2] J.Costeira and T.Kanade, "A multi-body factorization method for motion analysis," *Proc. of ICCV*, pp. 1071-1076, 1995.
- [3] F.R.Hampel, E.M.Ronchetti, P.J.Rousseeuw and W.A.Stahel, "Robust Statistics: The Approach Based on Influence Functions," John Wiley, 1986.
- [4] M.Han and T.Kanade, "Perspective factorization methods for Euclidean reconstruction," *CMU-RI-TR-99-22*, 1999.
- [5] 長谷川, 廣田, 小川原, 倉爪, 池内, "気球搭載型レーザーレンジセンサ," *MIRU2004, 論文集 I*, pp. 739-744, 2004.
- [6] Y.Hirota, T.Masuda, R.Kurazume, K.Ogawara, K.Hasegawa and K.Ikeuchi, "Flying Laser Range Finder and its data registration algorithm," *Proc. of ICRA*, pp. 3155-3160, 2004.
- [7] K.Ikeuchi, A.Nakazawa, K.Hasegawa and T.Ohishi, "The Great Buddha Project: Modeling Cultural Heritage for VR Systems through Observation," *The second IEEE and ACM International Symposium on Mixed and Augmented Reality(ISMAR2003)*, 2003
- [8] K.Ikeuchi, K.Hasegawa, A.Nakazawa, J.Takamatsu, T.Oishi, T.Masuda, "Bayon Digital Archival Project," *10th International Conference on Virtual Systems and Multimedia(VSMM 2004)*, 2004
- [9] D. A. Jacobs, "The State of the Art in Numerical Analysis," Academic Press, London, 1977.
- [10] The Digital Michelangelo Project, <http://graphics.stanford.edu/projects/mich/>
- [11] D.G.Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, 2004.
- [12] D.Miyazaki, T.Oishi, T.Nishikawa, R.Sagawa, K.Nishino, T.Tomomatsu, Y.Takase and K.Ikeuchi "The great buddha project: Modelling cultural heritage through observation" *Proc. of VSMM*, pp. 138-145, 2000.
- [13] T.Morita and T.Kanade, "A sequential factorization method for recovering shape and motion from image streams," *IEEE Trans. on PAMI*, vol. 19, No. 8, pp. 858-867, 1997.
- [14] C.Poelmann and T.Kanade "A paraperspective factorization method for shape and motion recovery," *IEEE Trans. on PAMI*, vol. 19, No. 3, pp. 206-218, 1997.
- [15] E. Polak, "Computational Methods in Optimization," Academic Press, New York, 1971.
- [16] J. Stoer and R.Bulirsch, "Introduction to Numerical Analysis," Springer-Verlag, New York, 1980.
- [17] S.Thrun, M.Diel and D.Haehnel, "Scan alignment and 3-D surface modeling with a helicopter platform," *The 4th International Conference on Field and Service Robotics*, 2003.
- [18] C.Tomasi and T.Kanade, "Shape and motion from image streams under orthography: a factorization method," *International Journal of Computer Vision*, Vol. 9, No. 2, pp. 137-154, 1992.
- [19] J.Visnovecova, L.Zhang and A.Gruen, "Generating a 3D model of a bayon tower using non-metric imagery," *Proc. of the International Workshop Recreating the Past - Visualization and Animation of Cultural Heritage*, 2001.
- [20] Z.Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, Vol. 13, pp. 119-152, 1994.
- [21] W.Zhao, D.Nister and S.Hsu, "Alignment of Continuous Video onto 3D Point Clouds," *Proc. of CVPR*, Vol. 2, pp. 964-971, 2004.