

画像列を利用した気球センサーからの3次元形状補正

阪野 貴彦 長谷川 一英 池内克史
東京大学生産技術研究所
{ vanno, k-hase, ki } @cvl.iis.u-tokyo.ac.jp

Abstract

For a large scale object, scanning from the air is one of the most efficient methods of obtaining 3D data. In the case of large cultural heritage objects, there are some difficulties in scanning them with respect to safety and efficiency. To remedy these problems, we have been developing a novel 3D measurement system, the Flying Laser Range Sensor (FLRS), in which a range sensor is suspended beneath a balloon. The obtained data, however, have some distortion due to the intra-scanning movement. 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 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 parameter, the distorted range data are recovered. We applied this method to an actual scanning project and the results showed the effectiveness of our method.

1. はじめに

最近の3次元形状計測技術の発達のおかげで、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。われわれはこれまでに、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をおこなってきた [6]。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。

われわれは現在、カンボジア王国においてバイヨン寺院 [12] 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。この寺院の大きさは、およそ100メートル四方もあり、これまでどの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 [10][15]。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、との問題が起こる。また、対象物体が文化的に非常に価値が高いため、クレーンなどの重機を使用することは、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、われわれは巨大文化遺産を計測するための新しい方法を考案した。これを Flying Laser range Sensor (以下、FLRS, Figure 1 参照) [4] 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。



Figure 1: Flying Laser Range Sensor and Bayon temple

そこで本論文では、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。ここで提案する手法は、われわれの 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。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。次に、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、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い、

また歪んだレンジデータそのものからも条件を抽出した。このような非線型関数の最適化問題を解く際、ローカルミニマムを避けるために、われわれは初期解として Factorization[11][2][7][8] による解をもちいた。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[3] の解を採用した。このようにして推定された FLRS の運動をもちいて、最終的に歪んだレンジデータの修正をおこなう。

本論文の構成は以下の通りである。第 2 章では、われわれが初期解を求めるために利用した Perspective-Factorization について簡単に説明をおこなう。第 3 章では、われわれが提案するセンサ運動推定のためのリファインメント方法について述べる。第 4 章では、本手法をバイオン寺院に適用した実験とその結果について述べる。ここでは、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ている。そして最後に、まとめとする。

2. Perspective-Factorization

まず簡単に、Weak-Perspective Factorization について説明し、次にこれをもちいて、Perspective Factorization への拡張法について説明する。Perspective Factorization よって得られた解は、第 4 章で述べる非線型最適化のための初期解として利用する。

2.1 Weak-Perspective Factorization

得られた画像列において、全フレーム (F 枚) にわたって、 P 個の特徴点がトラッキングされているとする。各特徴点 p は、3次元物体上の点 s_p に対応している。画像上での座標系では、フレーム番号 f において、特徴点 p は $\{(u_{fp}, v_{fp}) | f = 1, \dots, F, p = 1, \dots, P, 2F \geq P\}$ に表示されているとする。

画像特徴点の水平方向成分 u_{fp} をもちいて、 $F \times P$ の行列 U を定義する。同様に、画像特徴点の垂直方向成分 v_{fp} をもちいて、 $F \times P$ の行列 V を定義する。両者を合わせて、 $2F \times P$ の観測行列 W を次のように定義する。

$$W = \begin{pmatrix} U \\ V \end{pmatrix}$$

画像番号 f を取得するときのカメラ位置を、世界座標系で 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 i_f, j_f, k_f で表す。この3つのベクトルは、カメラ座標系での座標軸を表していて、 i_f, j_f は画像面に平行な2つの軸であり、 k_f は光軸に平行な軸である (Figure 2 参照)。

弱透視投影カメラモデルでは、3次元世界座標系において s_p と示される点は、フレーム f の画像面で (u_{fp}, v_{fp}) に投影される。

$$u_{fp} = \frac{f}{z_f} i_f^t (s_p - t_f)$$

$$v_{fp} = \frac{f}{z_f} j_f^t (s_p - t_f)$$

$$\text{where } z_f = k_f^t (\bar{c} - t_f)$$

ここで、 \bar{c} はトラックした特徴点の重心とする。世界座標系の原点はどこに設定しても一般性を失わないため、この重心を原点、つまり $\bar{c} = 0$ とする。すると、 $z_f = -k_f^t t_f$ となるため、上式は次のようになる。

$$\begin{cases} u_{fp} = m_f^t s_p + x_f \\ v_{fp} = n_f^t s_p + y_f \end{cases}$$

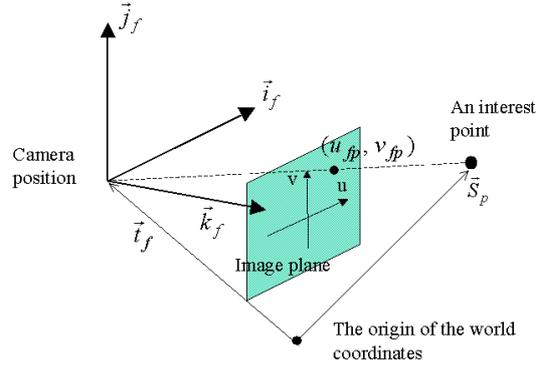


Figure 2: the coordinate system: t_f denotes the position of the camera at time of frame f . The camera pose is determined by three unit basis vectors.

$$\text{where } m_f = \frac{f}{z_f} i_f, \quad x_f = -\frac{f}{z_f} i_f^t t_f$$

$$n_f = \frac{f}{z_f} j_f, \quad y_f = -\frac{f}{z_f} j_f^t t_f$$

行列形式で表現すれば、

$$\begin{pmatrix} u_{11} & \dots & u_{1P} \\ u_{21} & \dots & u_{2P} \\ \vdots & \vdots & \vdots \\ u_{F1} & \dots & u_{FP} \\ v_{11} & \dots & v_{1P} \\ \vdots & \vdots & \vdots \\ v_{F1} & \dots & v_{FP} \end{pmatrix} = \begin{pmatrix} m_1^t \\ m_2^t \\ \vdots \\ m_F^t \\ n_1^t \\ \vdots \\ n_F^t \end{pmatrix} (\bar{s}_1 \dots \bar{s}_P) + \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_F \\ y_1 \\ \vdots \\ y_F \end{pmatrix}$$

また、原点と重心とを一致させたことにより、観測行列の要素を各行で加算すると、

$$\sum_{p=1}^P u_{fp} = \sum_{p=1}^P m_f^t \bar{s}_p + \sum_{p=1}^P x_f = P x_f$$

同様に、

$$\sum_{p=1}^P v_{fp} = P y_f$$

となる

そこで、 $\tilde{W} = W - (x_1 \ x_2 \ \dots \ x_F \ y_1 \ \dots \ y_F)^t$ と変換することで、計量行列 \tilde{W} は2つの行列 M と S の積となる。

$$\tilde{W} = M \cdot S$$

where $M : 2F \times 3$ Matrix $S : 3 \times P$ Matrix

次に、まだ M と S は未知であるため、 \tilde{W} を SVD (Singular Value Decomposition 特異値分解) を用いて以下のように分離する。

$$\tilde{W} = M \cdot S = O_1 \Sigma O_2$$

O_1 は $2F \times P$ の行列で、 Σ は $P \times P$ の対角成分のみからなる行列で、 O_2 は $P \times P$ の行列である。また、

$$O_1^t O_1 = O_2^t O_2 = O_2 O_2^t = I \quad (I : P \times P \text{ Unitmatrix})$$

が成り立つ。対角行列 Σ の対角成分には特異値 ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_P$) が入る。ここで、行列 M と S の rank は 3 であるため、両者の積である行列 \hat{W} の rank は高々 3 である (Rank theorem)。そこで、 Σ で特異値が大きいものから並んでいることに注目して、各行列を次のように分解する。

$$\begin{aligned} O_1 &= (O_1' | O''_1) \\ \Sigma &= \begin{pmatrix} \Sigma' & 0 \\ 0 & \Sigma'' \end{pmatrix} \\ O_2 &= (O_2' | O''_2) \end{aligned}$$

O_1' は、行列 O_1 の左から第 1 列から第 3 列までによって構成される $2F \times 3$ の行列であり、残りの第 4 列以降から構成される $2F \times (P-3)$ の行列が O''_1 である。同様に、 Σ' は対角行列 Σ の左上の 3×3 の部分であり、 O_2' は O_2 の第 1 行から第 3 行までによって構成される $3 \times P$ の行列である。仮に行列 \hat{W} にノイズが含まれていないとすると、特異値 σ_4 以降は 0 になるため、近似式として次式が得られる。

$$\hat{W} \simeq \hat{W}' = O_1' \Sigma' O_2' + O''_1 \Sigma'' O''_2$$

ここで、

$$\begin{aligned} \hat{M} &= O_1' [\Sigma']^{1/2} \\ \hat{S} &= [\Sigma']^{1/2} O_2' \end{aligned}$$

と置き換えることによって、

$$\hat{W}' = \hat{M} \hat{S}$$

計量行列は、 $2F \times 3$ と $3 \times P$ の 2 つの行列の積の形式に分解することができた。ただし、このままでは、ただの分解ができただけであり、任意の 3×3 の正則行列 Q を用いると

$$(\hat{M}Q)(Q^{-1}\hat{S}) = \hat{M}(QQ^{-1})\hat{S} = \hat{M}\hat{S} = \hat{W}'$$

が成立してしまい、 \hat{M} が解であるなら、 $\hat{M}Q$ も解になり、解が一意に定まらない。そのため、カメラの姿勢をあらわすパラメータからなる行列 $\hat{M}Q$ が満たすべき拘束条件

$$|\vec{m}_f| = |\vec{n}_f|, \quad \vec{m}_f \cdot \vec{n}_f = 0$$

を課すことにより、スケールの曖昧性を残して、カメラの姿勢に関する行列 $M \simeq \hat{M}Q$ が求められる。また、形状を示す行列 S は $Q^{-1}\hat{S}$ として求められ、カメラ位置 \vec{t}_f も先の式により求められる。

このようにして、動画画像から得られる計量行列のみによって、特徴点の 3 次元位置と、カメラの位置・姿勢が計算のみによって求められる。

2.2 Perspective Factorization への拡張

以上は、本来非線型である透視投影カメラモデルを線型近似した弱透視投影下での定式化であった。次に、この線型モデルを用いて繰り返し手法によって、非線型モデルでの近似解を求める。透視投影において、3 次元中の点 $\vec{s}_p = (X, Y, Z)$ は次の変換によって、2 次元画像上 (u_{fp}, v_{fp}) に射影される。

$$\begin{aligned} u_{fp} &= f \frac{\vec{i}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \\ v_{fp} &= f \frac{\vec{j}_f(\vec{s}_p - \vec{t}_f)}{\vec{k}_f(\vec{s}_p - \vec{t}_f)} \end{aligned}$$

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。

$$\begin{aligned} \frac{\vec{k}_f(\vec{s}_p - \vec{t}_f)}{z_f} u_{fp} &= \frac{f}{z_f} \vec{i}_f(\vec{s}_p - \vec{t}_f) \\ \frac{\vec{k}_f(\vec{s}_p - \vec{t}_f)}{z_f} v_{fp} &= \frac{f}{z_f} \vec{j}_f(\vec{s}_p - \vec{t}_f) \end{aligned}$$

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

$$\begin{aligned} (\lambda_{fp} + 1) u_{fp} &= \frac{f}{z_f} \vec{i}_f(\vec{s}_p - \vec{t}_f) \\ (\lambda_{fp} + 1) v_{fp} &= \frac{f}{z_f} \vec{j}_f(\vec{s}_p - \vec{t}_f) \end{aligned}$$

$$\text{where } \lambda_{fp} = \frac{\vec{k}_f \cdot \vec{s}_p}{z_f}$$

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}) がある数 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 $\vec{i}_f, \vec{j}_f, \vec{k}_f$

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 をおこなう。 $\vec{s}_p, \vec{t}_f, \vec{i}_f, \vec{j}_f, \vec{k}_f, z_f$ 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。

3. カメラパラメータの Refinement

入力画像にノイズがないとすれば、以上に述べた Factorization によって良好なカメラパラメータが得られる。結果として、歪んだ 3 次元形状は正しく復元される。しかし、実際に得られる実画像にはノイズが含まれるため、Factorization のみによって得られたカメラパラメータだけでは、FLRS によって得られた形状データを正しく復元するのに不十分である。そこで、カメラパラメータの推定精度を向上させるため、われわれはさらに 3 つの制約条件を課すことにした。これによって、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精度向上を図った。コスト関数を最小化するにあたっては、Perspective Factorization によって得られた解を初期解として与えた。

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。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。

$$F_A = \sum_{f=1}^F \sum_{p=1}^P \left(\left(u_{fp} - f \frac{i_f(\vec{s}_p - \vec{t}_f)}{k_f(\vec{s}_p - \vec{t}_f)} \right)^2 + \left(v_{fp} - f \frac{j_f(\vec{s}_p - \vec{t}_f)}{k_f(\vec{s}_p - \vec{t}_f)} \right)^2 \right)$$

カメラの運動における制約

われわれが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。裏返せば、気球は低周波数の振動の影響を受ける、といえる。われわれが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。このことから、つぎの制約条件を与える。

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

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 Quaternion \mathbf{q}_f で評価している。この Quaternion は i_f 、 j_f 、 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。

歪んだレンジデータにおける制約

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に歪んでおり、結果として正しくない形状が得られる。ただし、このことは距離画像全体を見た場合であって、局的には正確な情報を含んでいる。以下の制約では、この局的に正確な距離情報を用いた制約条件を示す。

われわれの FLRS では、計測原理に time-of-flight を用いており、距離画像面のみで、ラスタスキャンのようにレーザーを発している。1 回の計測に要する時間はわかっているため、このことから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。

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

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。

実は、この制約は非常に重要である。画像のみからカメラの運動を推定する”Structure from Motion”や”Shape from Motion”といった技術 [13] は、スケールに関する曖昧性が残されてしまう。一方、この制約では絶対的な距離データを用いているため、カメラパラメータの推定のみならず、スケールの曖昧性の除去にも有用である。

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

$$F_C = \sum_{p=1}^P \left(\left(x_{fp} - i_{fp}(\vec{s}_p - \vec{t}_{fp}) \right)^2 + \left(y_{fp} - j_{fp}(\vec{s}_p - \vec{t}_{fp}) \right)^2 + \left(z_{fp} - k_{fp}(\vec{s}_p - \vec{t}_{fp}) \right)^2 \right)$$

が得られる。

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

$$F = w_A F_A + w_B F_B + w_C F_C$$

を最小化することによって、カメラの運動パラメータに関してのリファインメントをおこなう。

この全コスト関数の最小化には、共役勾配法をもちいる。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]。

また、前章で述べたように、最小化計算の初期解として、Perspective-Factorization による解をもちいた。これは、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ためである。実際 Perspective-Factorization によって、非常に近い近似解を採用することができた。

4. 実験結果

4.1 FLRS システム

われわれの FLRS システムは、スキャナ、コントローラ、PC の 3 つによって構成され、この 3 つを気球によって吊るす。

スキャナ部分は、レーザーレンジファインダを含み、特に気球に吊るすことを考えて製作した。Figure 3 にその内部の様子を示す。この部分は、スポット照射型のレーザー測距機と 2 つのミラーとからなる。レーザー光源には、Z+F 社の LARA25200 を使用した。2 つのミラーは、レーザーが向かう方向を制御しており、ポリゴンミラーが距離画像における横方向のスキャンを司っており、スウィングミラーが縦方向のスキャンを司っている。詳細を Table 1 に示しておいた。

コントローラ部分は、信号処理装置とインターフェースとからなる。信号処理装置では PC からの信号を受け取り、2 つのミラーの制御をおこなう。レーザー測距機からのレンジデータと、エンコーダからのミラー角度のデータとをもちいて、1 枚のレンジデータに変換し、PC に蓄積される。

Table 1: The spec of the FLRS

| | |
|-----------------------------------|-------------------|
| Light source | 23mW, 780nm |
| Scanning length | 2 - 25.2 m |
| Range resolution | 0.38mm |
| Vertical resolution | 160 pixel |
| Horizontal resolution | 1800 pixel |
| Rotational speed (polygon mirror) | 2400 rpm |
| Swing range (swing mirror) | -15 - +15 deg |
| Scanning time | 1 sec/frame |
| Weight | 21.2kg |
| Size (depth × width × height) | 27 × 53 × 30 (cm) |

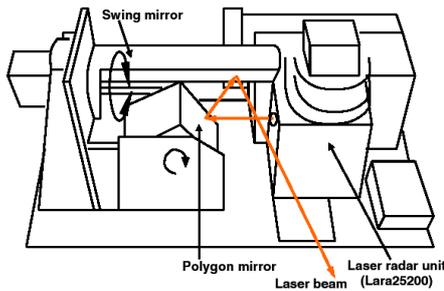
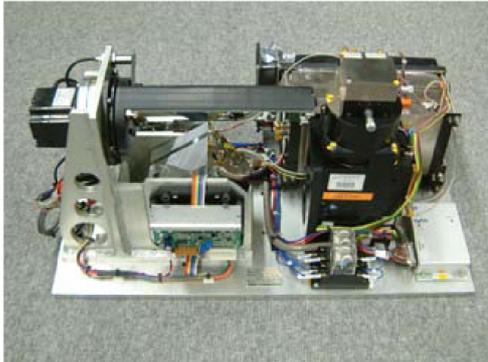


Figure 3: The interior of scanner unit

4.2 バイヨン寺院における実験

画像列のトラッキング

われわれの FLRS システムでは、1 回の計測で 72 フレーム撮影している (30fps)。各画像において特徴点を抽出するのに、スケール変化に対してロバストな Lowe オペレータ [5] を使用した。画像間で特徴点を追跡するため、特徴点まわりでのウィンドウマッチングをおこなうことで、全フレームわたり対応点をつなげた。以上の手続きを経て、72 フレームから 100 点程度の特徴点を追跡することができた。

カメラ運動の平滑化

ここでは、カメラパラメータのリファインメントの効果を検証するため、カメラの併進成分の軌跡を示す (Figure 4)。Figure 4 で上のグラフでは、Perspective Factorization で得られたカメラの併進成分の軌跡、同図の下グラフでは、リファインメント処理後の同軌跡を示す。

Factorization のみでのカメラパラメータ推定によって、ある程度の軌跡が描かれているが、局所的に見ると、凹凸があり、滑らかなカメラの運動とはいえない。一方、リファインメ

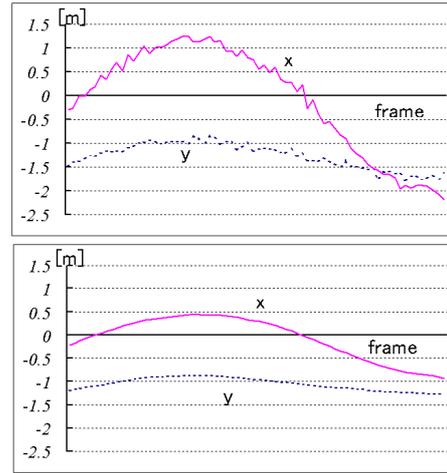


Figure 4: The trace of the camera translation.

ントを施したあとでは、軌跡が滑らかになっており、気球の運動として妥当な様子を示している。このことより、前章で述べたコスト関数の最小化によるリファインメントによって、カメラの動きを滑らかにする効果が得られていることがわかる。

復元形状の評価

本論文で提案している形状復元手法の精度について評価するため、FLRS から得られたデータを復元した形状と、他のレンジセンサから得られた形状とを比較した。この”他のレンジセンサ”とは、地上に固定して置かれた Cyrax2500 であり、このセンサから得られたデータを正しい解とする。

ICP アルゴリズム [1][14] によって両者のデータをアラインし、重なり合った部分について考察をおこなう。アラインメントの結果を Figure 5 に示す。上の図は、Cyrax2500 によるデータと、FLRS から得られたそのままのデータとを合わせた結果である。尊顔あたりのデータは合っているが、それより上の部分、特に楕円で示した部分で、あまりレンジデータとフィットしていない。一方、下の図では、Cyrax2500 によるデータと、本手法によって修正したデータとの合わせである。データ全体に渡って両者がフィットしており、上の図において楕円で示しておいた部分もきちんと合っていることがわかる。

Figure 6 は、本手法による効果を視覚化したものである。ここでは、FLRS から得られた形状がどれだけ Cyrax2500 による形状に近づいたかを示してある。左側の図が、Cyrax2500 によるデータと、FLRS から得られたそのままのデータとを合わせたものである。このとき、両データの間で、点間距離が 6.0cm 以下だったときには明るい色で表示し、それ以上であったときには暗い色で表示した。同様に、右側の図は、Cyrax2500 によるデータと本手法によって修正したデータとを合わせたものである。この結果より、本手法によって、FLRS から得られたデータが良好に修復されていることがわかる。

5 まとめと今後の課題

本稿では、FLRS による計測で問題となる計測形状データの修復法について述べた。気球に同時に搭載したビデオカメラから、センサの動きを精度良く推定することで、データ修復をおこなった。まず、Perspective-Factorization をおこなうことで、カメラ運動の推定をおこなった。つぎに、3 つの条件を課

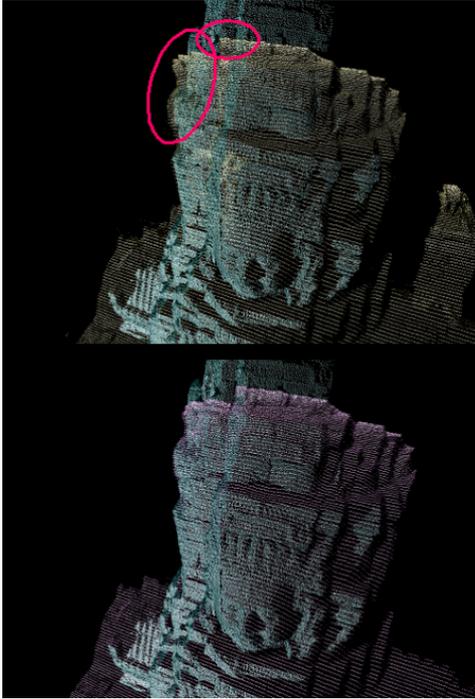


Figure 5: Range data before and after the recovery process.

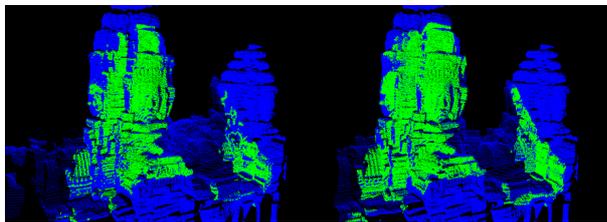


Figure 6: The comparison between the Cyrax 2500's (the correct data) and the original distorted data (left), and that between the Cyrax 2500's and the recovered data.

した非線型最適化問題を解くことで、カメラ運動推定をリファインメントした。その結果、歪んだデータを精度良く修復できることが示された。また、本手法は気球に特化したもののみならず、滑らかに運動するセンサに対して、一般的に適用することができる。

一方、技術的な課題もまだ多く残されている。例えば、現システムでは、特徴点は全フレームに渡ってトラックされていなければならない、オクルージョンに対応できていない。したがって、気球の運動が激しい場合にはカメラの運動を推定することが難しい。今後は、このようなオクルージョンに対応できるよう改良する予定である。

References

- [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] M.Han and T.Kanade, "Perspective factorization methods for Euclidean reconstruction," *CMU-RI-TR-99-22*, 1999.
- [4] Y.Hirota, T.Masuda, R.Kurazume, K.Hasegawa and K.Ikeuchi "Flying range finder and its data registration algorithm," *Proc. of ICRA*, pp.3155-3160, 2004.
- [5] D.G.Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, Vol.60, No.2, pp.91-110, 2004.
- [6] D.Miyazaki, T.Oishi, T.Nishikawa, R.Sagawa, K.Nishino, T.Tomomatsu, Y.Yakase and K.Ikeuchi "The great buddha project: Modelling cultural heritage through observation" *Proc. of VSMM*, pp.138-145, 2000.
- [7] 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.
- [8] 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.
- [9] W.H.Press, B.P.Flannery, S.A.Teukolsky and W.T.Vetterling, "Numerical Recipes in C," *Cambridge Univ. Press*, 1988.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] J.Weng, T.S.Huang and N.Ahuja, "Motion and structure from two perspective views: algorithms, error analysis and error estimation," *IEEE Trans. on PAMI*, Vol.11, No.5, pp.451-476, 1989.
- [14] Z.Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, Vol.13, pp.119-152, 1994.
- [15] W.Zhao, D.Nister and S.Hsu, "Alignment of Continuous Video onto 3D Point Clouds," *Proc. of CVPR*, Vol.2, pp.964-971, 2004.