回転型ステレオカメラを用いた屋外シーンの密な三次元再構成

吉田 武史*・深尾 隆 則*

Dense 3D Reconstruction of Outdoor Scenes Using a Rotational Stereo Camera

Takeshi YOSHIDA* and Takanori FUKAO*

This paper proposes a method to construct a three-dimensional (3D) model of an outdoor scene containing a large amount of information by combining local stereo images which are acquired from multiway. It is difficult to reconstruct outdoor scenes accurately using a single view of a stereo camera because the distance to a target is very far compared with the stereo camera's baseline length. Therefore the proposed method makes the 3D model accurately by reducing the uncertainties of a 3D point from various angles to the region which includes the correct point. Multi-view stereo images are captured using a rotational stereo camera that swings back and forth and can capture not only the upper surface of an object but also the side surface of the object. As an experiment, a blimp robot with a rotational stereo camera is used to capture aerial stereo images of the ground. The proposed method achieves the dense and accurate reconstruction of the outdoor 3D model.

Key Words: stereo camera, 3D reconstruction, UAV

1. はじめに

阪神淡路大震災や東日本大震災をはじめとする自然災害や大事故などの大規模災害現場においては、初期の救援活動が重要とされており、安全かつ迅速な被災地の情報収集が必要とされるが、現在はそのようなシステムは確立されていない。災害現場では道路の寸断や交通渋滞が発生する可能性が考えられ、行政機関やレスキュー隊に提示する情報を地上から収集することは困難であると考えられる。したがって、道路の被災状況による影響を受けずに災害現場の情報収集には航空機を用いて上空から行なうことが有効である。そこで、本論文では、航空機の中で比較的低空で安全に飛行可能である点から自律飛行船ロボットを用いている¹⁾。情報収集には本研究で開発した回転型ステレオカメラを用いる。回転型ステレオカメラは上空から撮影する際に建築物をさまざまな方向から撮影することができる利点をもつ。たとえばステレオカメラの視線方向を鉛直下向きに固定した場合、建築物の屋根や道路に対しては十分な情報を得ることができるが、建築物の側面の情報は非常に少ないものとなる。そこでステレオカメラの光軸方向を進行方向から逆の方向まで前後に回転させることで建築物側面の情報も取得できるようにし、より正確な三次元モデルの構築を行なう。広範囲の画像を撮影する際に

一般的に用いられる全方位カメラと比べると、回転型ステレオカメラは取得する画像群の中で重複するデータが少なくなり、1つの方向に対して得られる画像の解像度が高くなるという点もまた優れている。また、収集した情報の提示方法としては、回転型ステレオカメラによって対象を撮影し、災害現場の三次元モデルを生成、これを最終的にはGIS（地理情報システム）などの地図上に重ねることを目標としている。三次元モデルを用いることで従来の二次元画像と異なり、注目部分の拡大縮小だけでなく、視点切替が容易に行なえるようになる。

ステレオカメラは距離計測に一般的に用いられるレーザレンジファインダと比べると色情報と位置情報を同時に取得できる点に利点がある。またレーザレンジファインダはそれ自体が高価であるだけでなく、色情報を付加するためにカメラが別に必要となるため、ステレオカメラのみでシステムを構築できればシステム全体の低コスト化が図れる。しかしステレオカメラは対象までの距離が長くなればなるほど、その距離計測の精度は低下し、飛行船による上空からの計測では非常にその精度は悪くなってしまう。またステレオカメラを長距離の距離計測に用いるケースは少なく、そのような状況における確立した三次元モデル生成法はいまだ提案されていない。

そこで本論文では回転型ステレオカメラによる複数方向からのステレオ画像を基に長距離の計測に対して精度のよい三次元モデル構成法を提案する。具体的には、各視点からの距離計測における不確かさ領域の共通部分を真の点を含む領域として限定する。これにより実現される自律飛行船ロボットによる三次元モデル生成システムは、地震などの災害時の情

* 神戸大学大学院工学研究科 神戸市灘区六甲台町 1-1

* Graduate School of Engineering, Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe

(Received April 2, 2012)

(Revised September 5, 2012)

報収集だけでなく、三次元の地理情報の記録や、三次元モデルを得るには時間とコストがかかる文化遺産の記録など、さまざまな用途に用いることができる。

2. 複数画像による三次元再構成とその問題点

本論文で用いるステレオカメラは二眼ステレオであるが、さまざまな角度からのステレオ画像を統合するという点から多眼ステレオとしても利用する。そこで、多眼ステレオに関する関連研究とその問題点と、二眼ステレオの問題点について述べる。

2.1 多眼ステレオ

多眼ステレオのアルゴリズムはSeitzらによって4つのカテゴリーに分類されている²⁾。これら4つの分類とはコスト関数を計算した三次元ボリュームから面を抽出する手法、コスト関数を最小化することで反復して面を展開する手法、三次元シーンの距離画像を統合するといった画像空間を用いる手法、画像間で検出してマッチングを行なった特徴点を再構成し、面を当てはめる手法である。本論文で用いる手法はさまざまな角度から撮影した二眼ステレオによる距離画像を統合するという点から3番目のカテゴリーに分類することができる。Furukawaらによるとこの手法はさまざまな距離画像を統合する点に関しては複雑ではあるが、屋内環境における小さい物体や屋外シーンのどちらの対象に対しても同一のアルゴリズムによって扱うことができる点からフレキシブルな手法であるとしている³⁾。以降、三次元シーンの距離画像統合に分類される手法のいくつかを示す。

Goeseleらはある参照画像に対してk近傍法を用いて選ばれた局所的な画像群に対して、SSSD (Sum of SSD)⁴⁾の評価をNCC (Normalized Cross-Correlation) に置き換えたマルチベースラインステレオを使うことで正確な距離画像の生成を行なった⁵⁾。加えて彼らはレーザレンジファインダに用いられる距離画像の統合手法を用いて、これらの距離画像から三次元モデルを構築した。Strechaらは奥行き推定と可視性の推定をEMアルゴリズムに関連する隠れマルコフモデルの生成モデルとしてモデル化する手法を提案した⁶⁾。彼らは多視点の画像を用いた距離画像の獲得とその外れ値の検出のみを対象としており、距離画像の統合に関しては言及していない。BradleyらはGoeseleらの手法を拡張した。彼らは生成した距離画像に対してウィンドウの大きさが可変となるマッチングを用いることでワイドベースラインステレオにおける対応点探索の正確さを高めている⁷⁾。加えて彼らはVisual hullとDisparity ordering constraintを拘束条件として用いることで外れ値を除去した。しかし彼らの手法はその拘束条件によってカメラと対象までの距離を固定していることからその適用環境が限定される。

これらの手法は密な三次元モデルを再構成することができるが、屋内の限られた領域における小さな物体を主な対象としている。そのため、これらのアルゴリズムを屋外シーンの再構成に用いるには、物体が存在する境界領域 (バウンディン

グボックス)を設定することが困難であることから、適用することが難しい。屋外シーンの再構成に関する関連研究としてはFurukawaらによって提案されたPMVS (Patch-based multi-view stereo)がある³⁾。

PMVSは特徴点による画像間の対応付け、三次元モデルのパッチの拡張、外れ値の除去の3つのステップで行なわれる。初めにPMVSはHarris corner detectorとDifference of Gaussianの2種類の特徴点を検出し、エッジポラ拘束を用いて特徴点の対応付けを行なう。つぎに対応付けられた特徴点から物体表面を近似する小さな接平面であるパッチを求める。つぎにパッチと隣り合う領域に特徴点がなかった場合、パッチを接平面方向に拡張していく。そのように拡張したパッチに対して、このパッチの位置および向きから各カメラから観測可能かどうかを判定し、カメラ間で不一致がある場合は外れ値であるとしてパッチを除去する。

PMVSによって得られる三次元モデルは1画素以下の精度で対応付けが行ないやすい特徴点を起点としているため、三次元モデルを構成する点の三次元位置はおおよそ正確である。一方、このモデルは特徴点のみから得られる三次元モデルのパッチを拡張した擬似的に密なモデルであるため、画像中の対象となる部分のパターンが一様で特徴点が少ない場合、もしくは同じ領域を撮影している画像枚数が少なく、複数枚数の画像におけるその領域の特徴点が少ない場合は求められるパッチが少なくなるため、パッチを拡張しきれず欠損が残ってしまう問題がある。

2.2 ステレオカメラの三次元計測における不確かさ

Fig. 1で示すような、同じ内部パラメータをもつ2台のカメラを、光軸方向にそろえ、また視点の高さもそろえて配置することで構成するステレオを平行ステレオという。平行ステレオは左右画像間で対応する点 $p_r(x_r, y)$, $p_l(x_l, y)$ を求めることで幾何学的関係から対応点の三次元位置 $P(X, Y, Z)$ を(1)式により計算することができる。ここで f はカメラの焦点距離、 b はカメラ間のベースライン、 d は対応点の視差とする。

$$X = \frac{b(x_l + x_r)}{2d}, \quad Y = \frac{by}{d}, \quad Z = \frac{bf}{d} \quad (1)$$

(1)式は画像の構成要素である画素が大きさをもたない点として計算している。しかし実際には画素はCCD 1画素分の

Fig. 1 Parallel stereo used to obtain 3D view

Fig. 2 Uncertainty in the parallel stereo

Table 1 Example of uncertainty

	dX	dY	dZ
2 m	1.02 mm	1.02 mm	4.41 mm
20 m	10.23 mm	10.23 mm	441.19 mm
Ratio	10	10	100

大きさをもっている．そのため，一対の対応付けされた画素の組から求められる点の三次元位置は Fig. 2 のような範囲で不確かさをもつ．またこの不確かさの範囲は (2) 式から求めることができる．ここで δ は 1 画素の大きさとなる．

この不確かさの計測点が 2 m の場合と 20 m の場合における dX と dY と dZ ，これらの比を Table 1 に示す．同表から本論文において想定しているカメラから計測点までの距離 20 m では dZ の値は無視できないほど大きく，また距離 2 m のときと比べて，その変化量が著しく大きいことがわかる．

$$\begin{aligned}
 dX &= \frac{b(x_l + x_r + 2\delta)}{2d} - \frac{b(x_l + x_r)}{2d} = \frac{b\delta}{d} \\
 dY &= \frac{by}{d-1} - \frac{by-\delta}{d-1} = \frac{b\delta}{d-1} \\
 dZ &= \frac{bf}{d-\delta} - \frac{bf}{d+\delta} = \frac{2b\delta f}{d^2 - \delta^2}
 \end{aligned} \quad (2)$$

3. 回転型ステレオカメラシステム

本論文で用いる回転型ステレオカメラを Fig. 3 に示す．これは 2 台のカメラの光軸をそろえ，平行となるように 1 m の長さの回転軸の両端に配置している．この軸を回転させることで前後 60° の範囲で視点の光軸方向を変更することができる．ステレオカメラの回転制御は，飛行船搭載コンピュータから角度指令値を受信したモータコントローラが回転軸に取り付けられているロータリーエンコーダとの角度と指令値の偏差に応じて，DC モータを指定角度まで回転することで行なっている．また各カメラの内部パラメータとカメラ間の外部パラメータは，円を格子状に配置したキャリブレーションパターンを用いたステレオキャリブレーションを行なうことで求めている．本論文で用いる回転型ステレオカメラはベースラインが 1 m であるため，キャリブレーションを行なうにあたり必要な大きさを維持しつつ，作成と実験場への携行の容易さから A1 の大きさのキャリブレーションボードを用い

(a) Overview of the rotational stereo camera

(b) The mounted rotational stereo camera on the blimp robot

Fig. 3 Rotational stereo camera unit

Fig. 4 Difference in visible area

ている．

上空からステレオ画像を撮影する際，1 つの方向のみから撮影すると，Fig. 4 (a) のようにカメラの視線方向に垂直に近い面は高い解像度で撮影ができる反面，カメラの視線方向に平行に近い面は見える部分は少ないため解像度は低くなる．そのためこれらの画像から再構成される三次元モデルは対象の上面は密であるが，対象の側面は粗なモデルとなる．一方，ステレオカメラを前後に回転させて撮影したステレオ画像は Fig. 4 (b) のように上面の情報に加えて，対象の側面についても十分な情報を得ることができる．これらの画像によって再構成される三次元モデルは上面のみでなく側面も密な三次元モデルとなる．飛行船ロボットが低空飛行可能であるといった特徴から建築物側面の見える効果が高いこと，また効率的な情報収集のため直線軌道を中心とした巡回となることから，重複するデータは少なくなり，カメラの視点方向に対して垂直に近くなる面を撮影する機会が多くなるなど，カメラの回転による効果は高いと考えられる．

4. 密な三次元再構成

一対のステレオ画像から得られる距離画像から構築した三次元モデルの点群は距離画像中の画素数に相当する量の点が再構成されるため，点の数が多いため密なモデルとなる．しかし

2.2 節で述べたように、一対の距離画像から得られる三次元点の位置は、距離が大きくなるにつれて誤差が大きくなる。一方、2.1 節で述べたように PVMS によって得られる三次元点の位置は、その基となる特徴点ベースのマッチングが 1 画素以下の精度で画像間のマッチングを行なうことができるため非常に正確である。しかし、画像中に特徴点となる部分は多くはなく、特徴点ベースによって得られる三次元モデルは非常に粗なモデルとなる。そこで PVMS は特徴点から得た三次元モデルのパッチを拡張することで三次元モデルを密にしている。しかし特徴点が少ない部分に関しては拡張しきれず欠損が残ってしまう問題がある。そこで本論文では **Fig. 5** に示すようにさまざまな角度から撮影したステレオ画像から得られる三次元計測における不確かさの共通部分を求めることで、より確からしい真の点を含む領域として限定していく。これにより視差画像を三次元再構成したモデルの点群の量を保ちつつ、三次元点の位置の精度の高いモデルが構築できる。このような二眼ステレオカメラの不確かさに注目し精度を高める研究はいくつかなされている^{8),9)}。しかしこれらの手法は二眼ステレオ単体での使用を考えられており、三次元点の位置の精度の向上に限度がある。またこれらの手法を多眼ステレオに拡張するには不確かさの共通部分の表現が非常に複雑になる。そこで本論文では二眼ステレオの複数視点での使用のために CG や CAD に用いられている CSG を二眼ステレオの不確かさのモデルとして応用する。CSG は 1 つのモデルが連立不等式の集合であるため、複数視点による不確かさの共通部分を簡単に表現することができる。また本手法はある共通する対象を撮影する視点間の角度が大きければ、大きいほど精度が増すという特徴をもつ。このことから本手法は回転型ステレオカメラシステムの対象をさまざまな角度から撮影できるという点を有効に活用できる。

4.1 ステレオ画像間の外部パラメータ導出および対応点探索

本論文ではステレオ画像間の並進ベクトル、および回転行列は Structure-from-motion と Bundle adjustment を用いることで推定している¹⁰⁾。この処理に関して本論文では Snively らによる SfM ソフトウェアである Bundler を用い

ている。

Bundler が行なう処理として、初めに入力画像ごとの SIFT 特徴量による特徴点を求め、ANN による最近傍探索を行なうことで画像間での特徴点の対応付けを行なう。つぎに 8 点アルゴリズムにより基本行列の推定を行ない、RANSAC を用いることでこの特徴点の対応付けの外れ値を除去する。このように求まった特徴点の対応付けから画像の接続関係を示す Image connectivity graph を作成する。

Image connectivity graph から初期復元を行なうペアを選択し、5 点アルゴリズムを用いることでカメラ間の外部パラメータを推定して、カメラ間の相対並進ベクトルと相対回転行列を求める。このとき初期ペアの内の片方のカメラ座標が全体のワールド座標となる。ここから Bundle adjustment を用いて、この外部パラメータの推定誤差を最小化する。その後、初期ペアの画像と共通して写っている物の多い画像を 1 枚ずつ追加し、その都度 DLT 法によりカメラの並進ベクトルと回転行列を求め、Bundle adjustment による全カメラ分の推定誤差最小化を行なっていく。Bundler はカメラ間の並進ベクトルと回転行列と同時に特徴点の三次元位置も求めるが、本論文ではカメラ間の並進ベクトルと回転行列のみを用いている。

また撮影した角度の異なるステレオ画像間ではテンプレートマッチングを用いて共通する点を求めることは、その見えの違いの大きさから難しい。そこで本論文ではテンプレートマッチングを用いることなくステレオ画像間の共通部分を求めるために三重線形拘束による Epipolar transfer を用いている¹¹⁾。

Fig. 6 に示す 3 つの画像の間で Image3 と Image1 との基礎行列を F_{31} 、Image3 と Image2 との基礎行列を F_{32} とする。今、Image1 における点 x_1 と Image2 における点 x_2 の対応付けがされているとき、Image3 において x_1 と x_2 と対応付けられる点 x_3 の位置を求める。

基礎行列 F_{31} が既知であるとき、エピポーラ拘束より点 x_1 に対応付けられる点 x_3 はエピポーラ線 $F_{31}x_1$ 上に存在する。また同様に基礎行列 F_{32} が既知であるとき、エピポーラ拘束より点 x_2 に対応付けられる点 x_3 はエピポーラ線 $F_{32}x_2$ 上

Fig. 5 Intersection of uncertainties

Fig. 6 Epipolar transfer

に存在する．これら 2 つのエピポーラ線の交点は (3) 式により求められ，この交点が Image3 の対応点 $\mathbf{x}_3$ となる．同次座標において 2 つの直線の交点はこの 2 つの直線の外積から求まる．

$$\mathbf{x}_3 = (F_{31}\mathbf{x}_1) \times (F_{32}\mathbf{x}_2) \quad (3)$$

4.2 CSG による Visual Hull の表現

画像間の共通部分を用いて三次元形状を求める手法に Shape from Silhouette と呼ばれる手法がある¹²⁾．Shape from Silhouette は各画像において対象物体のシルエット領域とカメラ座標中心からなる視錐体 (Visual cone) を定義し，これらの共通部分を求める．この共通部分は視体積 (Visual hull) と呼ばれ，Visual hull の各画像の投影像はその各画像のシルエットと一致する．

Shape from Silhouette の実装は画素を単位として二次元平面シルエット輪郭を離散化する方法と，三次元空間を voxel で離散化する方法の 2 つに分けることができる．しかし本論文は Visual hull を用いて三次元物体を構成するのではなく，Visual hull 内のある点を求めることを目的としている．そのため本論文ではこれらの手法は用いず，Visual hull の共通部分を表現するために CG や CAD で使われるソリッドモデル表現法である CSG (Constructive solid geometry) を用いる．このソリッドモデルはモデルの体積，重心などのマスプロパティを求めることができる点を特徴としているため，本論文の用途に適している．CSG はあらかじめ定義された基本立体 (プリミティブ) を組み合わせることで複雑な形状を表わすことができる表現法である．このプリミティブを組み合わせることで新しい立体を定義するには，プリミティブの間で和や差，共通部分である積といった形状の論理演算を行なう．プリミティブの論理演算の例を Fig. 7 に示す．

プリミティブは連立不等式の集合として定式化される．つまり，この連立不等式をすべて満たす点はプリミティブの内部もしくは境界上の点となり，満たさない点はプリミティブ外部の点となる．このことから CSG で表現された立体の共通部分を求めるためには，各プリミティブで定義された境界条件をすべて満たすことができる点集合を求めればよいということになる．

4.3 真の点を含む領域の導出

平行ステレオカメラによる三次元計測の真の点を含む領域は Fig. 2 に示すように六面体であり，この六面体は空間を平面により切り分けた 2 つの空間の一方の積集合とみなすことができる．1 つの平面により切り分けられる空間の領域の点集合を半空間と呼び，(4)，(5) 式で表わすことができる．こ

Fig. 7 Boolean operation of primitives

こで $\mathbf{a}$ は $|\mathbf{a}| = 1$ となる空間を切り取る平面の法線ベクトル， $\mathbf{x}_0$ は平面を通る点とする．(4) 式は空間を切り取る平面の法線ベクトルと同じ方向の半空間を表わし，(5) 式は空間を切り取る平面の法線ベクトルと逆の方向の半空間を表わす．

$$P(x, y, z) = \{(x, y, z) = \mathbf{x} \in \mathbb{R} \mid \mathbf{a}^\top (\mathbf{x} - \mathbf{x}_0) \geq 0\} \quad (4)$$

$$P(x, y, z) = \{(x, y, z) = \mathbf{x} \in \mathbb{R} \mid \mathbf{a}^\top (\mathbf{x} - \mathbf{x}_0) \leq 0\} \quad (5)$$

ここで，Fig. 8 に示すように k 枚目の視点の画素による真の点を含む領域を構成する点を $p_1, p_2, \dots, p_8$ ，これらの点がある世界座標系へ座標変換した後の点を $x_1, x_2, \dots, x_8$ と定義する．また $x_1, x_2, \dots, x_8$ からできる平面をそれぞれ $P_1, P_2, \dots, P_6$ とする． k 枚目のカメラ座標から世界座標系へ変換する回転行列，並進ベクトルをそれぞれ $\mathbf{R}, \mathbf{T}$ とするとこれらの関係は (6) 式で表わすことができる．

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \end{bmatrix} = \mathbf{R} \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \end{bmatrix} + \mathbf{T} \quad (6)$$

これらの点から (4)，(5) 式より真の点を含む領域を構成する 6 つの半空間 $S_{k1} \sim S_{k6}$ は (7)～(12) 式で求まる．ここで半空間 $S_{k1} \sim S_{k6}$ はそれぞれ三角形 $x_4x_1x_3$ ，三角形 $x_2x_1x_6$ ，三角形 $x_2x_6x_3$ ，三角形 $x_8x_4x_7$ ，三角形 $x_8x_5x_4$ ，三角形 $x_8x_7x_5$ の面から求めている．

$$S_{k1}(x, y, z) = \quad (7)$$

$$\left\{ (x, y, z) = \mathbf{x} \in \mathbb{R} \mid \left(\frac{(\mathbf{x}_1 - \mathbf{x}_4) \times (\mathbf{x}_3 - \mathbf{x}_4)}{\|(\mathbf{x}_1 - \mathbf{x}_4) \times (\mathbf{x}_3 - \mathbf{x}_4)\|} \right)^\top (\mathbf{x} - \mathbf{x}_4) \geq 0 \right\}$$

$$S_{k2}(x, y, z) = \quad (8)$$

$$\left\{ (x, y, z) = \mathbf{x} \in \mathbb{R} \mid \left(\frac{(\mathbf{x}_1 - \mathbf{x}_2) \times (\mathbf{x}_6 - \mathbf{x}_2)}{\|(\mathbf{x}_1 - \mathbf{x}_2) \times (\mathbf{x}_6 - \mathbf{x}_2)\|} \right)^\top (\mathbf{x} - \mathbf{x}_2) \geq 0 \right\}$$

$$S_{k3}(x, y, z) = \quad (9)$$

$$\left\{ (x, y, z) = \mathbf{x} \in \mathbb{R} \mid \right.$$

Fig. 8 Region including true point by six half spaces

$$S_{k4}(x, y, z) = \left\{ \left(\frac{(\mathbf{x}_6 - \mathbf{x}_2) \times (\mathbf{x}_3 - \mathbf{x}_2)}{|(\mathbf{x}_6 - \mathbf{x}_2) \times (\mathbf{x}_3 - \mathbf{x}_2)|} \right)^\top (\mathbf{x} - \mathbf{x}_2) \geq 0 \right\} \quad (10)$$

$$S_{k5}(x, y, z) = \left\{ (x, y, z) = \mathbf{x} \in \mathbb{R} \left| \left(\frac{(\mathbf{x}_4 - \mathbf{x}_8) \times (\mathbf{x}_7 - \mathbf{x}_8)}{|(\mathbf{x}_4 - \mathbf{x}_8) \times (\mathbf{x}_7 - \mathbf{x}_8)|} \right)^\top (\mathbf{x} - \mathbf{x}_8) \geq 0 \right. \right\} \quad (11)$$

$$S_{k6}(x, y, z) = \left\{ (x, y, z) = \mathbf{x} \in \mathbb{R} \left| \left(\frac{(\mathbf{x}_5 - \mathbf{x}_8) \times (\mathbf{x}_4 - \mathbf{x}_8)}{|(\mathbf{x}_5 - \mathbf{x}_8) \times (\mathbf{x}_4 - \mathbf{x}_8)|} \right)^\top (\mathbf{x} - \mathbf{x}_8) \geq 0 \right. \right\} \quad (12)$$

さらに (7)~(12) 式より k 枚目のある画素における真の点を含む領域は (13) 式で求まる。

$$S_k(x, y, z) = \bigcap_{j=1}^6 S_{kj} = S_{k1} \cap S_{k2} \cap S_{k3} \cap S_{k4} \cap S_{k5} \cap S_{k6} \quad (13)$$

したがって全 n 枚における共通する画素の真の点を含む領域の積集合は (14) 式で求まる。

$$V(x, y, z) = \bigcap_{i=1}^n \bigcap_{j=1}^6 S_{ij}(x, y, z) \quad (14)$$

求まった Visual hull から具体的な 1 点を求めるが、重心をその点として計算する。Visual hull は複数の六面体が重なった部分の積集合であるため、その形状は非常に複雑であると予想される。そのため重心の計算には、無作為に標本点を選んで関数値を評価し、その標本に基づいて積分を計算するモンテカルロ法を用いる¹³⁾。

Fig. 9 に示すように、ある Visual hull の体積 V を求める。この形状を体積が既知の値 V_0 で囲む。つぎに囲んだ形状の中で一様な確率密度で多数の点をランダムに発生させる。発生させた点が Visual hull の内部の点かどうかを 1 つずつ判定する。Visual hull は (14) 式で定義されるが、同式は境界条件の積集合であるため、この点の内外判定は容易である。発生させた点のうち、Visual hull 内部に入った点を求める。ここでは発生させた点の個数を N 、そのうち Visual hull 内部に

Fig. 9 Calculation of volume with Monte Carlo method

入った点の個数を n とする。このとき発生させた点 $\mathbf{p}_i$ が外部の点であるとき $f_i = 0$ 、 $\mathbf{p}_i$ が内部の点であるとき $f_i = \mathbf{p}_i$ となり、Visual hull の重心 $\mathbf{g}$ は (15) 式で求められる。

$$\mathbf{g} = \frac{1}{n} \sum_{i=1}^N \begin{cases} f_i = \mathbf{p}_i & \text{if } \mathbf{p} \in V(x, y, z) \\ f_i = 0 & \text{if } \mathbf{p} \notin V(x, y, z) \end{cases} \quad (15)$$

5. 実験結果

5.1 シミュレーション実験

提案手法の精度および点群の量を確認するために、**Fig. 10** (a) に示すコンピュータ上の仮想空間に配置した箱状の物体上空を **Fig. 10** (b) に示す軌道で回転型ステレオカメラが移動しながら撮影した画像を用いて処理を行なう。このとき、箱状物体の高さは 3 m、カメラが移動する高度は約 20 m とする。またカメラのスペックは **Table 2** に示したものと同一性能とする。撮影した画像の例を **Fig. 11** に示す。ここでは画像間の対応点探索の成功率を高めるために、物体および地面にテクスチャパターンを貼付けている。本実験ではモンテカルロ法のサンプル数は 2000 としている。

Fig. 10 Layout and flight path in simulation

Table 2 Specification of the camera, Point Gray Research GRAS-20S4C-C

Image sensor model	Sony ICX274 1/1.8"
Resolution	1624 × 1224
Pixel size	4.40 × 4.40 μm
Max frame rate	30 fps
Interface	IEEE1394b

Fig. 11 Example of the captured image

Fig. 12 3D model reconstructed with proposed method

Fig. 13 3D model reconstructed with PMVS

Table 3 Example of uncertainty

	Proposed method	PMVS
RMSE of the measuring	25.8 mm	11.9 mm
Count of points	538195	9979

Fig. 10 (a) に示す物体を撮影した画像から提案手法を用いて得られた三次元モデルを Fig. 12 に示す。また比較のため PMVS を用いて得られた三次元モデルを Fig. 13 に示す。ここで Fig. 12 と Fig. 13 中の物体の下側が空白になっているのは、三次元再構成を行なうために用いたどの画像からも死角になっているためである。

Fig. 10 (a) と Fig. 12 と Fig. 13 を比較した場合、Fig. 13 の側面には欠損が多いことがわかる。これは移動しながら撮影を行なっているため、撮影した全画像中で側面が写っている割合が少なく、PMVS の起点となる特徴点が十分に得られなかった結果であると考えられる。一方、提案手法はテンプレートマッチングによる左右の画像の対応付けによる距離画像を基準としており、対応点探索の間違いがなければ 1 組のステレオ画像全体分の点の三次元距離が求まる。そのため特徴点基準の PMVS よりも多い点群数が得られ、密な三次元再構成ができていることがわかる。

つぎに三次元モデルの精度を調べるために、得られた三次元モデルの上面部分について地面からの高さの真値との二乗平均平方根誤差 (RMSE) を算出し、それぞれの手法との比較を行なった。また得られた三次元モデルの上面部分の点群の点数を算出した。これらの結果を Table 3 に示す。同表より本論文の提案手法は PMVS と比較して、RMSE は約 2 倍であるが、Table 1 に示した二眼ステレオでの 20 m の場

Fig. 14 RC blimp for information gathering

(a) JAXA Taiki Aerospace Research Field

(b) Reconstructed target

Fig. 15 Experimental field

(a) View1

(b) View2

(c) View3

(d) View4

Fig. 16 Examples of captured images

合の誤差 441 mm と比較して非常に小さくなっている。また点群の点数は非常に多く、密なモデルとなっていることがわかる。

5.2 屋外実験

本論文で用いる飛行船を Fig. 14 に示す。全長約 12m の RC 飛行船をベースに各種機器を搭載し、自律制御が行なえるように改造したものである¹⁾。

実験として Fig. 15 (a) に示す JAXA 大樹航空宇宙実験場にて飛行船ロボットに搭載した回転型ステレオカメラを用いて上空から実験場内の建築物 (Fig. 15 (b)) に対して、角度を変えて撮影した 6 組のステレオ画像群に対して提案手法を用いることで屋外シーンの三次元再構成を行なった。本実験ではシミュレーション実験同様、モンテカルロ法のサンプル数は 2000 としている。上空から撮影した画像の一部を Fig. 16 に示す。これらの画像から回転型ステレオカメラは建築物のさまざまな面を撮影できていることがわかる。

比較のため PMVS を用いて得られた三次元モデルを Fig. 17 に示す。これらの画像から PMVS を用いて三次元

(a) View1

(b) View2

Fig. 17 3D model reconstructed with PMVS

(a) View1

(b) View2

Fig. 18 3D model reconstructed with proposed method

再構成したモデルは建築物の屋根については非常に正確に三次元再構成できており、全体として外れ値が非常に少ない一方、多くの欠損をもっていることがわかる。

つぎに本論文の提案手法を用いて同様のステレオ画像に対して得られた三次元モデルを **Fig. 18** に示す。同図より本手法で三次元再構成を行なったモデルは PMVS を用いて三次

元再構成を行なったモデルに対して、全体的に点が密であることがわかる。加えて、このモデルは PMVS による三次元モデルに比べて、壁面の再構成に成功しており、密な三次元再構成に成功していることがわかる。

6. おわりに

本論文ではさまざまな角度から撮影したステレオ画像から得られる点の三次元座標の不確かさをモデル化し、その不確かさの共通部分を求めることで真の点を含む領域として限定していくことによって、より正確で密な屋外シーンの三次元モデルを構築する手法を提案した。さらに実験として回転型ステレオカメラを飛行船ロボットに搭載し、上空から実験場にある建築物を撮影した。これらの画像に対して、提案手法を用いることで屋外シーンのより正確で密な三次元モデルの構築を確認できた。

謝辞 本研究は(財)セコム科学技術振興財団の助成により実施した。ここに謝意を表します。

参考文献

- 1) H. Saiki, T. Fukao, T. Urakubo and T. Kohno: Hovering control of outdoor blimp robots based on path following, *Journal of Robotics and Mechatronics*, **23**-2, 2124/2129 (2011)
- 2) S.M. Seitz, B. Curless, J. Diebel, D. Scharstein and R. Szeliski: A comparison and evaluation of multi-view stereo reconstruction algorithms, *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 519/528 (2006)
- 3) Y. Furukawa and J. Ponce: Accurate, dense, and robust multiview stereopsis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**-8, 1362/1376 (2010)
- 4) M. Okutomi and T. Kanade: A multiple-baseline stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**-4, 63/69 (1993)
- 5) M. Goesele, B. Curless and S.M. Seitz: Multi-view stereo revisited, *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2402/2409 (2006)
- 6) C. Strecha, R. Fransens and L.V. Gool: Combined depth and outlier estimation in multi-view stereo, *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2394/2401 (2006)
- 7) D. Bradley, T. Boubekeur and W. Heidrich: Accurate multi-view reconstruction using robust binocular stereo and surface meshing, *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1/8 (2008)
- 8) S.D. Blostein and T.S. Huang: Error analysis in stereo determination of 3-d point positions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **9**-6, 752/765 (1987)
- 9) L. Matthies and A. Shafer: Error modeling in stereo navigation, *International Journal of Robotics and Automation*, **3**-3, 239/248 (1987)
- 10) N. Snavely, S.M. Seitz and R. Szeliski: Modeling the world from Internet photo collections, *International Journal of Computer Vision*, **80**-2, 189/210 (2008)
- 11) R. Hartley and A. Zisserman: *Multiple View Geometry in Computer Vision*, Cambridge University Press, second

edition (2003)

- 12) G. Slabaugh, B. Culbertson, T. Malzbender and R. Shafer:
A survey of methods for volumetric scene reconstruction
from photographs, *Proceedings of the 2001 Eurographics
Conference on Volume Graphics*, 81/101 (2001)
- 13) W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T.
Vetterling: *Numerical Recipes in C*, Cambridge Univer-
sity Press (1992)

=====

[著 者 紹 介]

吉 田 武 史 (学生会員)

2007 年神戸大学工学部機械工学科卒業。2008 年同大学大学院工学研究科機械工学専攻博士前期課程修了。2010 年同大学大学院博士後期課程に進学。ステレオカメラを用いた三次元再構成などのコンピュータビジョンに関する研究に従事。情報処理学会、日本ロボット学会の学生会員。

深 尾 隆 則 (正会員)

1992 年京都大学工学部航空工学科卒業, 94 年同大学大学院工学研究科応用システム科学専攻修士課程修了。96 年同大学博士後期課程中退, 同年助手となり, 同大学大学院情報学研究科助手を経て, 2004 年神戸大学工学部機械工学科助教授, 2007 年同大学大学院工学研究科機械工学専攻准教授。この間, 2001~2003 年米国カーネギーメロン大学ロボティクス研究所客員研究員。ロボティクスなどに関する研究に従事。博士 (情報学)。システム制御情報学会, 日本機械学会, 自動車技術会, 情報処理学会, IEEE などの会員。

=====