

# カメラの移動撮影に基づく 簡便なオンライン3次元計測システムの検討

山尾 創輔<sup>1</sup> 三浦 衛<sup>1</sup> 酒井 修二<sup>1</sup> 伊藤 康一<sup>1</sup> 青木 孝文<sup>1</sup>

概要：本稿では、カメラで静止画像を撮影しながら、高精度かつ密な3次元点群を逐次的に計測する簡便なオンライン3次元計測システムを提案する。これまでに提案されている密な3次元計測手法は、一般的に入力画像が全て揃ってから計測処理を行う。これらの手法は、画像が追加されるたびに最初から処理をやり直す必要があるため、撮影しながら3次元形状を計測することができない。一方で、提案システムは、撮影のたびに、(i) Structure from Motion によるカメラパラメータ推定と、(ii) 位相限定相関法に基づく密な対応付けを実行する。これらの処理は数秒しかかからないので、非同期かつ並列に処理を実行することで、撮影するたびに物体の3次元形状を逐次的に計測することが可能となる。性能評価実験を通して、従来手法に匹敵する精度でオンライン3次元計測が可能であることを示す。

## 1. はじめに

近年、3次元プリンタの高性能化・低価格化が急速に進み、消費者向けの3次元プリンティングが注目されている。しかしながら、非専門家でも3次元プリンティングを利用できる環境を実現するには、解決すべき技術的な課題がいくつかある。特に重要なものとして、非専門家にとっては、3次元プリンタに入力する3次元データの作成が困難であるという問題がある。実世界に存在する物体の3次元データを作成するためには、物体の3次元モデルを計測する必要がある。一般に、対象物体の3次元計測では、レーザスキャンやパターン光投影に基づく計測システムが用いられる。しかし、このようなシステムは、対象物体に依存する特別な計測装置が必要となるため、システムの導入費用が高価になったり、応用範囲が限定されてしまうという問題がある。そのため、日常生活の場面における実用的な3次元プリンティング環境を実現するには、誰でも手軽に使える3次元モデリングシステムが必要となる。

コンピュータビジョンの分野では、簡便で手軽な3次元モデリングへのアプローチの一つとして、複数枚の画像からの3次元形状復元に関する研究が古くから盛んである[1-3]。近年では、複数枚の静止画像から高精度かつ密な3次元形状を復元する手法が提案されており[4]、複数枚の静止画像をクラウド上にアップロードすることで3次

元モデルを生成できる Web アプリケーション [5] や、商用の3次元モデリングシステム [6] も存在する。しかし、これらのシステムは、オフラインの処理で構成されているため、入力画像をすべて用意するまで3次元モデリング処理を始めることができない。満足する3次元モデルが得られなければ、入力画像を撮り直したり、新しい画像を追加したりすることで、3次元復元処理を最初からやり直す必要があるため、3次元モデルの再計測に多くの時間を要してしまう。さらに、満足する3次元モデルを生成するための画像撮影には、撮影位置や計測処理に関する技術的な知識が必要となる。

このような理由から、非専門家でも手軽に3次元モデリングできるシステムは、画像の撮影と並行して3次元復元できるようなオンラインのシステムであることが望ましい。オンラインの3次元計測システムを用いることで、利用者は、3次元計測結果をリアルタイムに確認し、満足する3次元モデルを生成するために必要な撮影位置を把握することができる。これまでに提案されたオンラインの3次元計測システムは、動画像を入力とするシステムが一般的である。このような、動画像の撮影に基づくオンラインの密な3次元計測システムとして、[7-9] が提案されている。しかし、動画像の撮影に基づく3次元計測システムは、動画像を処理するための高性能な計算資源が必要となったり、モーションブラーを抑えて高品質な動画像を撮影するのが難しかったりするため、利用の負担が大きい。また、静止画像に比べると、動画像の解像度や画質が低いいため、動画像からでは高品質な3次元モデリングが難しいという問題

<sup>1</sup> 東北大学 大学院情報科学研究科  
Graduate School of Information Sciences, Tohoku University,  
Sendai-shi, Miyagi, 980-8579, Japan

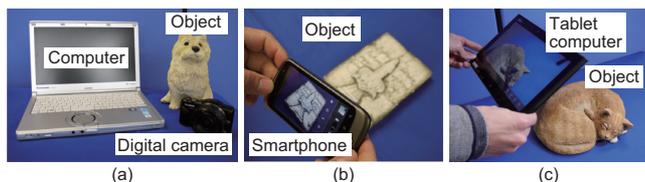


図 1 提案する 3 次元計測システムの構成例: (a) デジタルカメラと計算機, (b) スマートフォン, (c) タブレット端末

もある。

一方, 半澤らは, 静止画像の撮影に基づくオンラインの 3 次元復元システムを提案している [10]. 半澤らは, カメラによる静止画像の撮影と並行して物体の 3 次元形状を復元し, 3 次元復元の状況をリアルタイムに可視化することで, 誰でも手軽に使える 3 次元復元システムを実現している. しかし, 半澤らのシステムでは, 疎な 3 次元点群として物体の形状を復元する. 物体の高品質な 3 次元形状を取得するためには, 疎な 3 次元点群ではなく, 密な 3 次元点群を復元するシステムであることが望ましい. しかしながら, 静止画像から密な 3 次元点群をオンラインで計測するシステムはこれまでに提案されていない.

本稿では, Structure from Motion (SfM) と位相限定相関法 (Phase-Only Correlation: POC) に基づく画像対応付け手法と組み合わせることで, 静止画像から密な 3 次元点群をオンラインで計測するシステムを提案する. 提案システムは, 新しい画像が撮影されると, SfM によってカメラの撮影位置と姿勢を推定する. そして, 撮影位置と姿勢が判明したカメラの画像を用いて, POC に基づく画像対応付けによって密な 3 次元点群を復元する. 性能評価実験を通して, 従来の高精度な 3 次元計測手法に匹敵する精度で, 密な 3 次元点群をオンラインで計測できることを示す.

## 2. システムの概要

提案システムは, 図 1 のように, デジタルカメラと計算機から構成される. 図 1 (a) の場合, Wi-Fi 通信機能を搭載したカメラまたはメモリーカードを利用し, 撮影した画像を計算機へ自動的に送信することで, 計測の利便性を向上させることが可能である. また, 図 1 (b), (c) の場合, スマートフォンやタブレット端末のように, カメラと汎用プロセッサの両方を搭載した機器を用いることで, より小型で携行性の高いシステムの構築が可能である.

提案システムを用いて安定に 3 次元形状を計測するためには, ぼけやハレーションによる入力画像の画質の低下を抑える必要がある. 近年の消費者向けデジタルカメラは, 焦点距離やシャッタースピード, 絞りを自動で調節することで, 焦点の不整合や手ぶれ, ハレーションによる画質の低下を抑制する. そのため, 利用者は, 撮影に関する専門的な知識を持たなくても, 3 次元計測に適した静止画像を撮影することができる.

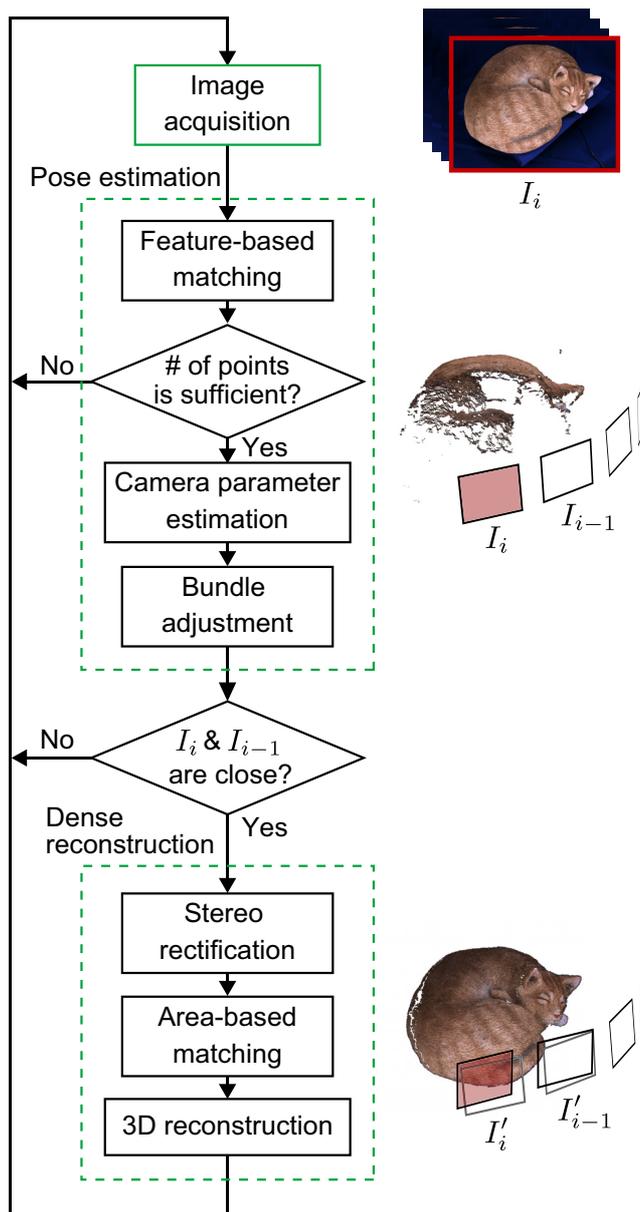


図 2 静止画像から密な 3 次元点群を逐次的に計測するオンラインシステムの処理フロー:  $i$  番目に撮影した画像  $I_i$  と, その直前に撮影した画像  $I_{i-1}$  から,  $i$  番目のカメラの位置と姿勢を推定する.  $I_i$  と  $I_{i-1}$  が十分に近接する場合は, これらの画像をステレオ平行化し, 領域ベースの画像対応付け手法を用いて密な 3 次元点群を復元する

図 2 に示すように, 提案システムの処理は, 大きく分けて, (i) 静止画像の取得, (ii) カメラの位置・姿勢推定, (iii) 密な 3 次元復元の 3 つのステップから構成される. まず, 任意の視点から計測対象を撮影した画像  $I_i$  を取得する. 次に, 画像  $I_i$  と, その直前に撮影した画像  $I_{i-1}$  の間で, 特徴ベースの画像対応付けを行う. そして, SfM [11, 12] により, 画像  $I_i$  と  $I_{i-1}$  の対応点から,  $I_i$  を撮影したカメラの位置と姿勢を推定する.  $I_i$  と  $I_{i-1}$  を撮影したカメラの位置が十分に近いならば, 画像  $I_i$  と  $I_{i-1}$  をステレオ平行化し, 平行化されたステレオ画像ペアの間で, POC [13] に基づく密な画像対応付けを行う. そして, 得られた対応点

から物体の3次元点群を復元し、物体全体の3次元点群に追加する。提案システムは、画像ごとに上述の処理を非同期かつ並列に行うことで、カメラの移動撮影と並行して物体全体の密な3次元点群を更新し、リアルタイムに計測の状況を可視化する。

### 3. カメラの位置・姿勢推定

本章では、提案システムにおけるカメラの位置・姿勢推定について説明する。カメラの位置・姿勢推定は、(i) 特徴ベースの画像対応付け、(ii) カメラパラメータの推定、(iii) バンドル調整による最適化から構成される。以下では、これらの具体的な処理を説明する。

#### (i) 特徴ベースの画像対応付け

新たに撮影した画像  $I_i$  と、その直前に撮影した画像  $I_{i-1}$  の対応関係を求める。これらの画像の間には、焦点距離の変化や移動撮影の影響により、拡大縮小や回転、非線形な変形が含まれる。そのため、提案システムでは、特徴ベースの画像対応付け手法を用いて画像間の対応点を求める。

特徴点検出および特徴量記述には、Speeded Up Robust Features (SURF) [14] を利用する。SURF は、画像間の幾何学的変形や輝度変化に対してロバストである Scale-Invariant Feature Transform (SIFT) [15] 特徴量を近似することで、SIFT に匹敵する性能で高速な特徴量記述を行う。カメラパラメータの推定に必要な数の対応点が得られなければ、画像  $I_i$  を破棄し、もう一度  $I_i$  の撮影を待機する。

#### (ii) カメラパラメータの推定

特徴ベースの画像対応付けにより取得した対応点ペアから、画像  $I_i$  のカメラパラメータを推定する。まず、画像  $I_i$  を撮影したカメラの内部パラメータ  $A_i$  を求める。内部パラメータ行列  $A$  は次式で定義される。

$$A = \begin{pmatrix} f \frac{w}{D} & 0 & \frac{w}{2} \\ 0 & f \frac{h}{D} & \frac{h}{2} \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

ここで、 $f$  はカメラの焦点距離、 $w$ 、 $h$  はそれぞれ画像の幅と高さ、 $D$  はカメラの撮像素子の幅である。焦点距離  $f$  と画像の解像度 ( $w, h$ ) は、Exif (Exchangeable image file format) [16] 情報から取得できる。また、撮像素子の幅  $D$  は、カメラの仕様を参照することで知ることができる。

次に、 $I_i$  を撮影したカメラの外部パラメータを推定する。 $i = 2$  の場合、5点アルゴリズム [17] を用いて、対応点から  $I_i$  の外部パラメータを推定する。一方で、 $i > 2$  の場合は、Kneip らの手法 [18] を用いて、対応点における3次元点と2次元点の幾何学的な関係から  $I_i$  の外部パラメータを推定する。このとき、5点アルゴリズムと Kneip らの手法と、RANSAC (RANdom SAMple Consensus) [19] を組み合わせることで、誤対応点に対してロバストにパラメータ

を推定する。推定された  $I_i$  のカメラパラメータを用いて、 $I_i$  と  $I_{i-1}$  の対応点から、計測対象の疎な3次元点群を復元する。このとき、 $I_i$  と  $I_{i-1}$  のカメラ位置とのなす角が小さすぎたり大きすぎたりするような3次元点や、再投影誤差が閾値よりも大きい3次元点を誤対応点として除去する。また、後段の最適化処理の計算コストを抑えるために、近傍点との距離が十分に小さい3次元点については、それらの平均の座標を持つ3次元点として統合する。さらに、カメラの位置・姿勢推定の状況を確認できるように、カメラの位置・姿勢と3次元点群の様子をリアルタイムに可視化する。

#### (iii) バンドル調整による最適化

求めたカメラパラメータと3次元点群の精度が後続の処理の安定性に大きく影響するため、バンドル調整に基づく再投影誤差の最小化により、これらのパラメータを最適化する。本稿では、バンドル調整の対象範囲に応じて、大域バンドル調整と局所バンドル調整の2種類を考える。

大域バンドル調整は、全ての画像のカメラパラメータと3次元点群を最適化する。推定されたカメラパラメータの集合を  $P = \{p_i\}$  ( $1 \leq i \leq K$ )、復元された3次元点群の座標の集合を  $Q = \{q_j\}$  ( $1 \leq j \leq L$ ) とする。ここで、 $K$  は対象となるカメラの数、 $L$  は対象となる3次元点群  $Q$  の点数である。大域バンドル調整は、次式で定義されるコスト関数  $E_g(P, Q)$  を最小化する。

$$E_g(P, Q) = \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^L \|m_{i,j} - m_{\text{rep}}(p_i, q_j)\|^2 \quad (2)$$

ここで、 $m_{i,j}$  は、画像  $I_i$  における  $q_j$  の画像座標である。また、 $m_{\text{rep}}(p_i, q_j)$  は、 $p_i$  を用いて  $q_j$  を投影したときの画像座標である。大域バンドル調整は、カメラの数や3次元点数が多くなるにつれて計算量が急激に増加するため、一定のフレーム間隔で繰り返し行う。

局所バンドル調整は、画像  $I_i$  のカメラパラメータと、 $I_i$  から観測される3次元点群を最適化する。画像  $I_i$  のカメラパラメータを  $p_i$ 、 $I_i$  から観測される3次元点群の集合を  $Q' = \{q'_j\}$  ( $1 \leq j \leq L'$ ) とする。ここで、 $L'$  は対象となる3次元点群  $Q'$  の点数である。局所バンドル調整は、次式で定義されるコスト関数  $E_l(p_i, Q')$  を最小化する。次式で定義されるコスト関数  $E_l(p_i, Q')$  の最小化を行う。

$$E_l(p_i, Q') = \frac{1}{2} \sum_{j=1}^{L'} \|m_j - m_{\text{rep}}(p_i, q'_j)\|^2 \quad (3)$$

局所バンドル調整は、大域バンドル調整に比べて計算量が少ないため、カメラパラメータを推定するたびに適用する。バンドル調整で最適化するコスト関数は非線形関数であるため、非線形最小二乗アルゴリズムの1つである Levenberg-Marquardt (LM) 法を用いる。このとき、スパース行列を利用したバンドル調整 [20] を用いることで大幅な高速化が可能である。

#### 4. 密な 3 次元復元

本章では、提案システムにおける密な 3 次元復元について説明する。密な 3 次元復元は、(i) ステレオ平行化、(ii) 領域ベースの密な画像対応付け、(ii) 密な 3 次元点群の復元から構成される、ただし、3 次元計測の品質の低下を抑えるために、画像  $I_i$  と  $I_{i-1}$  のカメラの位置が十分に近接する場合に密な 3 次元復元を行う。提案システムでは、3 次元点群の重心に対する  $I_i$  と  $I_{i-1}$  のカメラの視差角が閾値よりも小さい場合に、カメラが近接していると判定する。以下では、ステレオ平行化と領域ベースの画像対応付けの具体的な処理を説明する。

##### (i) ステレオ平行化

ステレオ画像ペア  $I_i, I_{i-1}$  をステレオ平行化することで、平行化後のステレオ画像ペア  $I'_i, I'_{i-1}$  を求める。ステレオ平行化は、平行ステレオカメラで撮影したような画像となるように、画像ペアを変形する処理である [11]。平行化後のステレオ画像では、画像ペアの間の幾何学的変形が画像の水平座標軸方向のみに縮退する。これにより、ステレオ画像ペアの間の射影変形を軽減し、領域ベースの画像対応付けの安定化を図ることができる。また、ステレオ画像ペアの間の対応点探索が一次元の走査となるため、画像対応付けの計算コストを抑えることができる。平行化後のステレオ画像ペア  $I'_i, I'_{i-1}$  は、カメラパラメータから計算される透視投影行列を用いて、ステレオ画像  $I_i, I_{i-1}$  を変形することで得られる。

##### (ii) 領域ベースの密な画像対応付け

領域ベースの画像対応付け手法を用いて、平行化後のステレオ画像ペア  $I'_i, I'_{i-1}$  の間の密な対応点を求める。領域ベースの画像対応付け手法は、特徴ベースの画像対応付け手法とは異なり、 $I'_{i-1}$  上の任意の位置に配置された参照点と対応する  $I'_i$  上の点を求めることができる。 $I'_{i-1}$  上に多くの参照点を配置することで、ステレオ画像ペアの間の密な対応点を得られるため、高品質な 3 次元形状を計測することができる。

提案手法では、領域ベースの画像対応付け手法として、POC に基づく画像対応付け手法 [13] を用いる。領域ベースの画像対応付け手法の中でも、POC に基づく画像対応付けは、画像間の輝度変化にロバストであることが知られている。この特徴は、照明条件の変化やカメラのオートゲインにより生じる画像間の輝度変化に対して有効である。また、POC に基づく画像対応付けでは、相関ピークの解析的なモデルに基づいて高精度に対応付けを行うため、高精度な 3 次元形状の計測が可能である。

##### (iii) 密な 3 次元点群の復元

三角測量の原理に基づいて、ステレオ画像ペア  $I_i, I_{i-1}$  の対応点を 3 次元復元する。そして、 $i$  番目より前に復元した 3 次元点群に、 $I_i$  と  $I_{i-1}$  から復元した密な 3 次元点

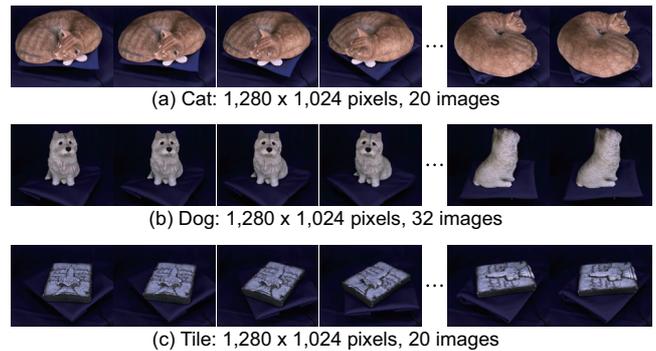


図 3 性能評価実験で用いるデータセット: (a) 猫の置物, (b) 犬の置物, (c) デコレーションタイル



図 4 3 次元レーザスキャナで計測した真値の 3 次元メッシュモデル: (a) 猫の置物, (b) 犬の置物, (c) デコレーションタイル

群を追加する。ここで、大域バンドル調整によって、 $i$  番目より前に推定したカメラパラメータが更新されている可能性があるため、物体全体の 3 次元点群の座標を計算し直す。このとき、座標の再計算に要する処理時間を抑えるために、ステレオ平行化後の画像ペア  $I'_i, I'_{i-1}$  の対応点の座標をステレオ平行化前の画像ペア  $I_i, I_{i-1}$  上の座標に変換して保持しておき、3 次元座標の再計算に用いる。また、計測の状況を確認できるように、復元した密な 3 次元点群をリアルタイムに可視化する。最後に、物体全体の密な 3 次元点群を復元し終わったら、3 次元点群の品質を高めるために、近傍点との平均距離が大きい点をアウトライヤとして除去する。

#### 5. 性能評価実験

本章では、カメラの移動撮影によって用意したデータセットを用いて、提案システムの 3 次元計測結果の品質と 3 次元計測に要する処理時間を評価する。評価では、従来の高精度かつ密な 3 次元計測システムとの比較を行う。従来システムでは、(i) Scale Invariant Feature Transform (SIFT) に基づく SfM を用いたカメラの位置・姿勢推定と、(ii) Patch-based Multi-View Stereo (PMVS) [4] を用いた密な 3 次元復元を行う。

##### 5.1 実験環境

計測対象は、図 3 に示すような、猫の置物 (W30cm × D30cm × H10cm)、犬の置物 (W20cm × D15cm × H20cm)、デコレーションタイルの (W20cm×D10cm×H5cm) の 3 種類である。使用したカ

メラは Point Gray 社製の Flea 3 (FL3-U3-13E4C-C) であり, 使用したレンズは, SPACECOM 社製の JHF12M-MP である. 撮影した画像は,  $1,280 \times 1,024$  ピクセルのカラー画像である. 撮影した画像の枚数は, 猫の置物, 犬の置物, デコレーションタイルのそれぞれで 20, 30, 20 枚である. 従来システムがオフラインの処理から構成されるため, これらのデータセットをあらかじめ撮影して用意しておく. カメラから対象物体までの距離は約 1 m である. 従来システムにおいて, SfM によるカメラパラメータの推定が不安定になることがあるため, カメラの内部パラメータ (焦点距離および画像中心) は, 事前に撮影したチェッカーパターンを用いて, Zhang らのキャリブレーション手法 [21] によって求めておく. 3次元計測誤差を定量的に評価するために, コニカミノルタ社の3次元レーザスキャナ VIVID 910 を用いて, 計測対象の真値の3次元メッシュモデル (図 3) を計測しておく.

撮影した静止画像から, 従来システムと提案システムを用いて対象物体の密な3次元点群を計測する. 提案システムに対しては, 実際の撮影時間の影響を加味するために, 2秒間隔で順番に画像を入力する. 画像はブルーバックで撮影されており, 青色領域を抽出することで背景領域のマスクを生成し, 背景領域に対する処理を省略する. そして, 計測した3次元点群と真値の3次元メッシュモデルとの間で ICP (Iterative Closest Point) による位置合わせを行い, 3次元計測誤差を評価する. このとき, 計測した3次元点群にはスケールの不定性が存在するため, ICP による位置合わせにおいて, 回転と並進だけでなく, スケールも推定する.

従来システムの実装では, SfM によるカメラの位置・姿勢推定に Visual SFM [22] を, 密な3次元復元に PMVS2 [4] を用いる. 入力する画像は, ブルーバック領域をマスクしておく. PMVS2 における画像ピラミッドの階層数を 0 とする. また, NCC (Normalized Cross-Correlation) に基づく画像対応付けのウィンドウサイズを  $7 \times 7$  ピクセルに, NCC の相関値の閾値を 0.7 に, 画像を分割するセルのサイズを  $2 \times 2$  に設定する. その他のパラメータは, Furukawa らの手法 [4] で設定されている値と同じものを用いる.

提案システムは C++ で実装する. SURF に基づく対応付けは, OpenCV [23] を用いて実装し, Fast-Hessian Detector の閾値を 100 とする. SfM におけるアウトライヤ除去では, 再投影誤差が 1.5 ピクセル以上の3次元点をアウトライヤと判定する. また, 隣接するカメラペアとの視差角が 0.5 度以下あるいは 120 度以上となる3次元点についてもアウトライヤと判定する. 大域バンドル調整と局所バンドル調整は, Sparse Bundle Adjustment (SBA) を用いて実装する. 大域バンドル調整は, 再投影誤差の RMS (Root Mean Square) が 0.5 ピクセルよりも大きいときに実行する. 密な3次元復元は, 3次元点群に対するステレ

表 1 3次元計測点数

Data set	Conventional	Proposed
Cat	284,289	572,876
Dog	143,446	239,107
Tile	81,418	243,640

表 2 3次元計測誤差の RMS [mm]

Data set	Conventional	Proposed
Cat	0.59	0.71
Dog	0.90	1.13
Tile	0.85	0.54

オ画像ペアの視差角が 5 度以下であるときに実行する. 画像に配置する参照点は, 1 ピクセル間隔で格子状に配置する. POC に基づく画像対応付けのウィンドウサイズを  $64 \times 64$  ピクセルに, ピーク値の閾値を 0.6 に設定する. カメラの位置・姿勢と3次元点群の可視化, および, 統計的なフィルタによるアウトライヤ除去は, Point Cloud Library (PCL) [24] を用いて実装する. 統計的なフィルタによるアウトライヤ除去では, 近傍 30 点とのユークリッド距離の RMS が 1.5 以上である点を除去する.

従来システムと提案システムの実行環境は, Windows 7 Professional, Intel® Core™ i7-990X (3.47GHz), RAM 24 GB である.

## 5.2 3次元計測の品質の評価

表 1, 2 に, 従来システムと提案システムにおける, 3次元計測点群の点数と3次元計測誤差の RMS を示す. また, 図 5 に, 3次元計測結果と3次元計測誤差のカラーマップを示す. 表 1 から, 提案システムは, 従来システムよりも密な3次元点群を計測できることが確認できる. また, 表 2 から, 提案システムは, 従来システムに匹敵する精度で3次元点群を計測できることが確認できる. 従来システムが高精度な3次元計測システムの一つであることから, 提案システムの3次元計測精度は十分に高いと言える.

提案システムでは, 計測処理の前半で復元された3次元点群と後半で復元された3次元点群とが重なる領域で, 3次元計測精度が低下する傾向がある. これは, カメラパラメータ推定の誤差が蓄積することで, 異なるステレオ画像ペアから復元された3次元点群が完全には重ならないことに起因する. この問題に対して, ループクローズング処理によりロバストなカメラの位置・姿勢推定を行ったり, 最終的に得られた3次元点群に ICP を適用したりすることで, 3次元計測精度の向上が期待できる.

また, 図 6 に, Poisson Surface Reconstruction [25] を用いて, 提案システムの3次元計測結果から生成した3次元メッシュモデルを示す. 図 6 から, 提案システムは, 従来システムに匹敵する品質で3次元モデリングが可能であ

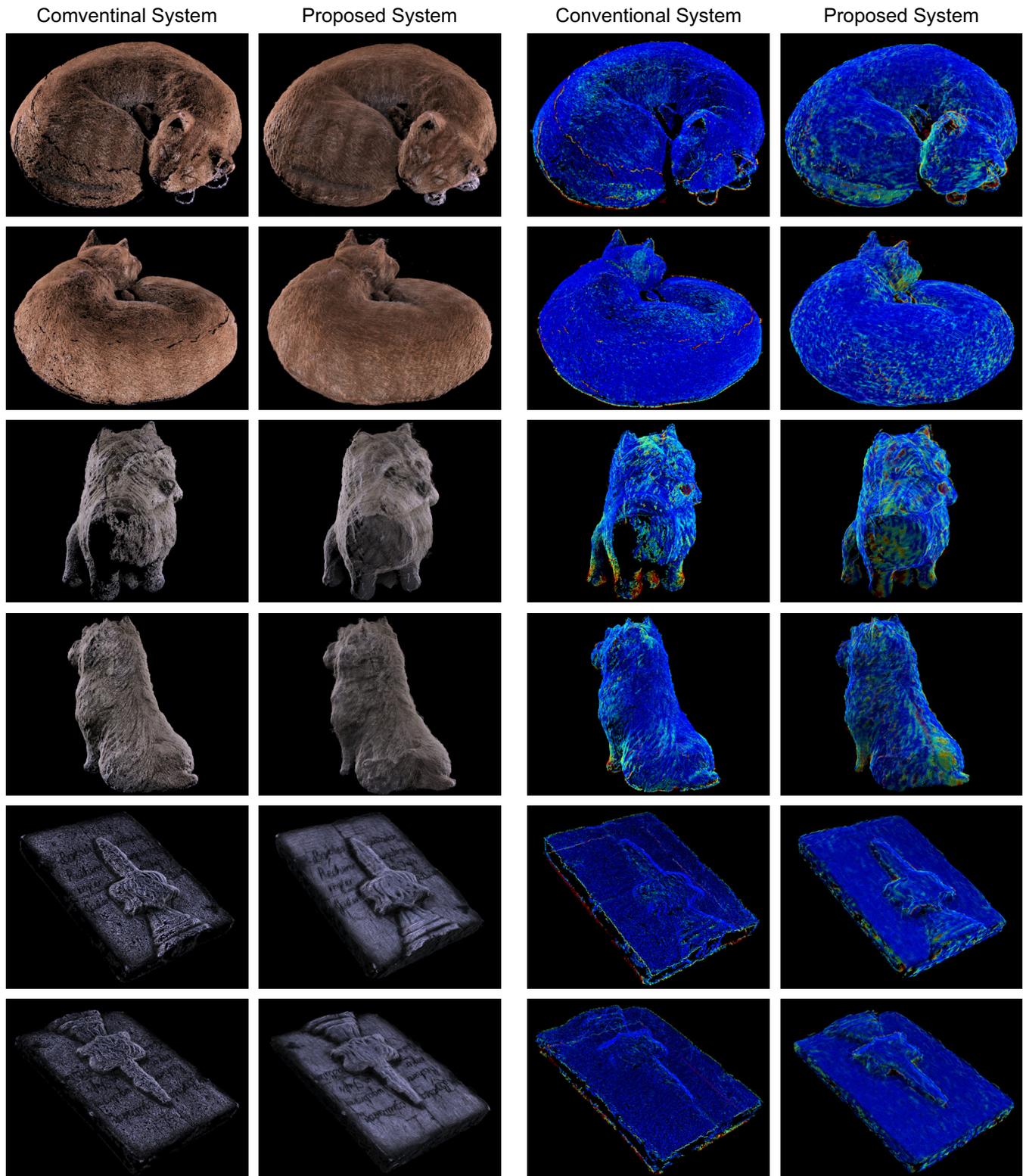


図 5 3次元計測点群 (1-2 列目) と 3次元計測誤差のカラーマップ (3-4 列目): 青 (0 mm), 赤 (3 mm)

表 3 物体全体の 3 次元計測に要する処理時間 [sec.]

Data set	Conventional	Proposed
Cat	620	66
Dog	435	96
Tile	285	58

ることが確認できる。

### 5.3 3次元計測に要する処理時間の評価

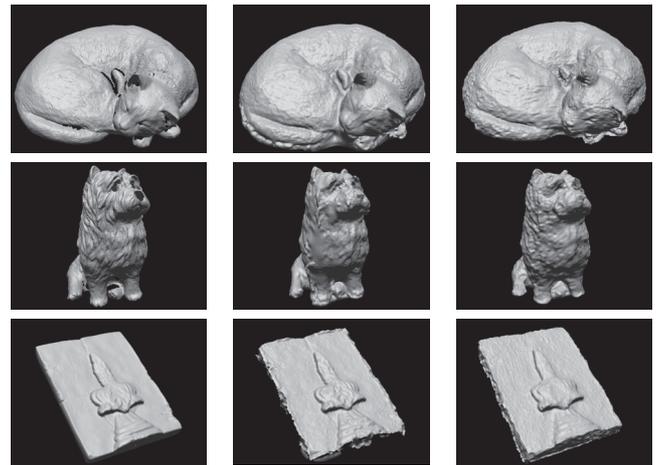
表 3 に、従来システムと提案システムにおける、物体全体の 3 次元計測に要する処理時間を示す。表 3 から、従来システムは、物体全体の 3 次元計測を終えるまでに数分の処理時間を要することが確認できる。一方、提案システムは、従来システムに比べて 5 ~ 10 倍速く物体全体を 3 次元計測できる。

図 7 に、提案システムにおける、画像 1 枚あたりの処理時間の内訳を示す。図 7 から、提案システムは、画像 1 枚あたり 10 秒未満の処理で 3 次元点群を更新できることがわかる。内訳を見ると、提案システムでは、画像の読み込みや、特徴ベースの画像対応付け、密な 3 次元復元に要する処理時間が、入力画像の枚数によらずにほぼ一定であることが確認できる。しかし、カメラの位置・姿勢推定については、入力画像の枚数が増えるにつれて、処理時間も増加する傾向がある。これは、入力画像の枚数が増えることで、バンドル調整の対象となるパラメータ数が指数的に増加するためである。この問題を解消するためには、マルチコアによる高速なバンドル調整 [26] を用いて、カメラの位置・姿勢推定に要する時間を削減するのが有効である。

以上の評価により、提案システムは、従来の高精度な 3 次元復元システムに匹敵する精度で、より密な 3 次元点群をオンラインで計測できることを示した。

### 5.4 汎用デジタルカメラによる 3 次元計測

図 8 に、汎用デジタルカメラ (Panasonic LUMIX DMC-GF1) で撮影した画像からの 3 次元計測結果を示す。撮影した画像は、1,280 × 960 ピクセルのカラー画像である。カメラの内部パラメータは、画像の Exif 情報から算出したものを初期値として推定する。その他の設定は、従来システムと提案システムの両方において、前述の実験と同様である。従来手法については、猫の置物やウサギのぬいぐるみのように、カメラの視点数が限られているために、3 次元点群が疎になっていることが確認できる。木製の立方体については、Visual SFM によるカメラの位置・姿勢推定に失敗してしまうため、側面の形状を計測できなかった。一方で、提案システムは、どの物体についても密な 3 次元点群を計測できている。



(a) Ground truth (b) Conventional system (c) Proposed system

図 6 真値の 3 次元メッシュモデル (左) と、3 次元計測結果から生成した 3 次元メッシュモデル (中央と右)

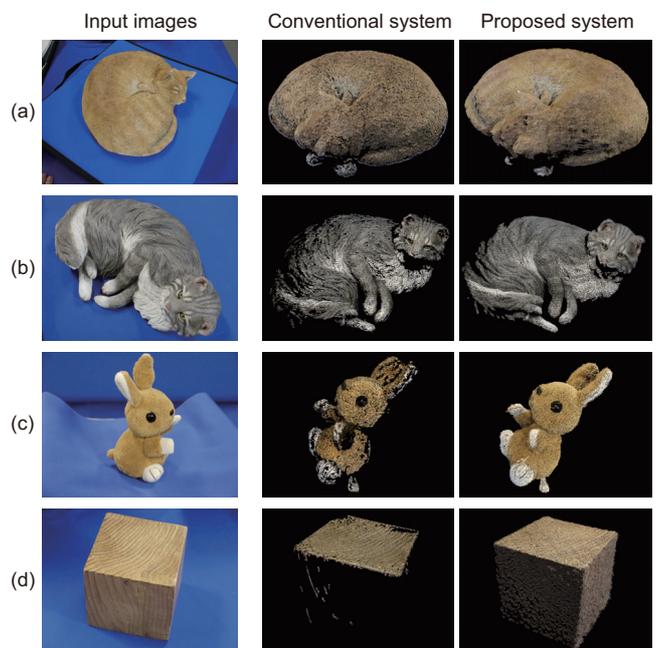


図 8 汎用デジタルカメラで撮影した静止画像からの 3 次元計測結果: (a) 猫の置物 (入力画像 15 枚), (b) 猫の置物 (入力画像 24 枚), (c) ウサギのぬいぐるみ (入力画像 15 枚), (d) 木製の立方体 (入力画像 16 枚)

## 6. まとめ

本稿では、カメラで移動撮影した静止画像から密な 3 次元点群を計測するオンラインシステムを提案した。性能評価を通して、提案システムは、従来の高精度な 3 次元計測システムに匹敵する精度で、より密な 3 次元点群をリアルタイムに計測できることを示した。今後は、計測状況をもとに、利用者に対して計測に適切な撮影方法をアドバイスするようなインタラクティブなシステムの開発や、大規模な物体を対象とした 3 次元計測を考えている。

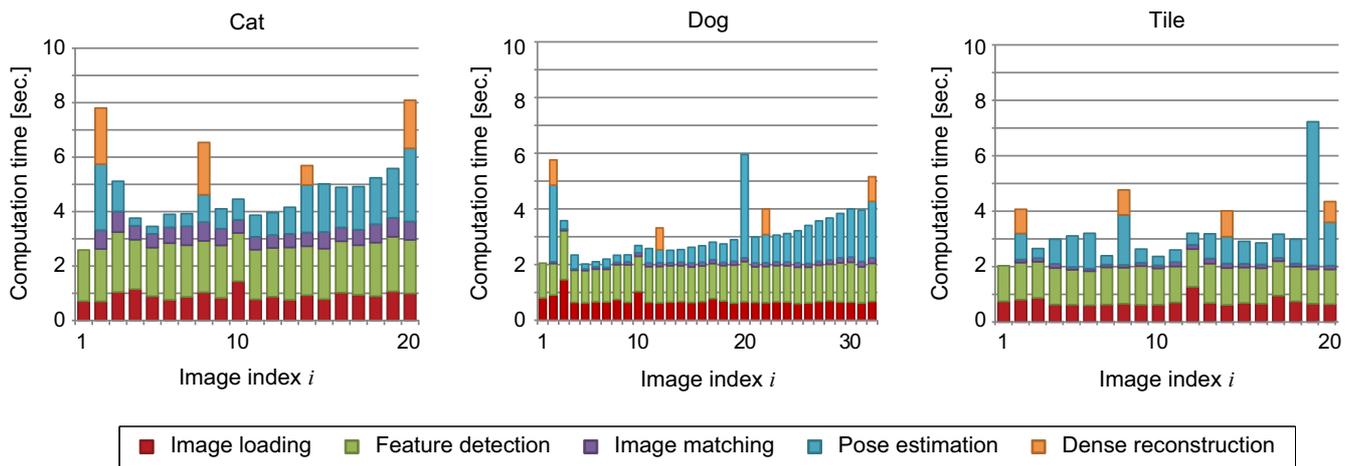


図 7 提案システムにおける画像一枚あたりの処理時間の内訳

参考文献

[1] Seitz, S. M., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R.: A comparison and evaluation of multi-views stereo reconstruction algorithms, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 519–528 (2006).

[2] Strecha, C., von Hansen, W., Gool, L. V., Fua, P. and Thoennessen, U.: On benchmarking camera calibration and multi-view stereo for high resolution imagery, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 1–8 (2008).

[3] Agarwal, S., Snavely, N., Simon, I., Seitz, S. M. and Szeliski, R.: Building Rome in a day, *Proc. Int'l Conf. Computer Vision*, pp. 72–79 (2009).

[4] Furukawa, Y. and Ponce, J.: Accurate, dense, and robust multiview stereopsis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 32, No. 8, pp. 1362–1376 (2010).

[5] Autodesk 123D Catch — 3D model from photos. <http://www.123dapp.com/catch>.

[6] Agisoft PhotoScan — [www.agisoft.ru](http://www.agisoft.ru/products/photoscan). <http://www.agisoft.ru/products/photoscan>.

[7] Newcombe, R.: DTAM: Dense tracking and mapping in real-time, *Proc. Int'l Conf. Computer Vision* (2011).

[8] Wendel, A., Maurer, M., Graber, G., Pock, T. and Bischof, H.: Dense reconstruction on-the-fly, pp. 1450–1457 (2012).

[9] Tanskanen, P., Kolev, K., Meier, L., Camposeco, F., Saurer, O. and Pollefeys, M.: Live metric 3D reconstruction on mobile phones, *Proc. Int'l Conf. Computer Vision*, pp. 65–72 (2013).

[10] 半澤悠樹, 鳥居秋彦, 奥富正敏: オンライン撮影に適した実用的な SfM システム, *電子情報通信学会論文誌 D*, Vol. J96-D, No. 8, pp. 1753–1763 (2013).

[11] Szeliski, R.: *Computer Vision: Algorithms and Applications*, Springer-Verlag New York Inc. (2010).

[12] Hartley, R. and Zisserman, A.: *Multiple View Geometry*, Cambridge University Press (2004).

[13] Takita, K., Muquit, M. A., Aoki, T. and Higuchi, T.: A sub-pixel correspondence search for computer vision applications, *IEICE Trans. Fundamentals*, Vol. E87-A, No. 8, pp. 1913–1923 (2004).

[14] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L.: Speeded-up robust features (SURF), *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346–359 (2008).

[15] Lowe, D. G.: Distinctive image features from scale-invariant keypoints, *Int'l J. Comput. Vision*, Vol. 60, No. 2, pp. 91–110 (2004).

[16] Japan Electronics and Information Technology Industries Association: Exchangeable image file format for digital still cameras. <http://www.jeita.or.jp/>.

[17] Nistér, D.: An Efficient Solution to the Five-Point Relative Pose Problem, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 756–770 (2004).

[18] Kneip, L., Scaramuzza, D. and Siegwart, R.: A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation, *IEEE*, pp. 2969–2976 (2011).

[19] Fischler, M. A. and Bolles, R. C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. ACM*, Vol. 24, No. 6, pp. 381–395 (1981).

[20] Lourakis, M. I. A. and Argyros, A. A.: SBA: A software package for generic sparse bundle adjustment, *ACM Trans. Math. Software*, Vol. 36, No. 1, pp. 1–30 (2009).

[21] Zhang, Z.: Flexible camera calibration by viewing a plane from unknown orientations, *Proc. Int'l Conf. Computer Vision*, Vol. 1, pp. 666–673 (1999).

[22] Wu, C.: VisualSFM: A Visual Structure from Motion System, <http://homes.cs.washington.edu/~ccwu/vsfm/>.

[23] Open Computer Vision Library. <http://sourceforge.net/projects/opencvlibrary/>.

[24] Point Cloud Library. <http://pointclouds.org/>.

[25] Kazhdan, M., Bolitho, M. and Hoppe, H.: Poisson surface reconstruction, *Proc. Symp. Geometry Processing*, pp. 61–70 (2006).

[26] Wu, C., Agarwal, S., Curless, B. and Seitz, S.: Multicore Bundle Adjustment, *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pp. 3057–3064 (2011).