

# ドローン映像からの3次元復元の リアルタイム化のためのキーフレーム選択と形状復元

橋本 瑞樹<sup>1,a)</sup> 松永 竜太郎<sup>1,b)</sup> 西山 和範<sup>1,c)</sup> 金澤 靖<sup>1,d)</sup>

## 概要：

東日本大震災の津波による行方不明者や遺失物の捜索は、現在も閑上浜などで続けられており、このような捜索では砂浜の地表面の形状や漂着物などの把握が重要となる。本研究では、ドローンを用いて、これらの捜索活動を効率化することを目的とし、本稿では、主な捜索対象となる砂浜シーンに対し、ドローンによる撮影の特徴を利用したキーフレーム選択と形状復元の並列化とGPUを用いた高速化により、リアルタイムに3次元復元するだけでなく、リアルタイムモザイク画像生成を同時に行う方法を提案する。提案法では、オプティカルフローによる運動推定を用いてキーフレームを選択し、砂浜の平面性を利用した射影変換ベースの復元と局所的なループクロージングを利用して、リアルタイムかつ高精度な形状復元を行うと共に、フロー情報を用いてモザイク画像生成を行う。実映像を用いた実験により、提案法の有効性を示す。

## Key-frame selection and Reconstruction for Real-time 3-D Reconstruction from UAV camera

MIZUKI HASHIMOTO<sup>1,a)</sup> RYOTARO MATSUNAGA<sup>1,b)</sup> KAZUNORI NISHIYAMA<sup>1,c)</sup>  
YASUSHI KANAZAWA<sup>1,d)</sup>

## Abstract:

After the Great East Japan Earthquake, the search activity for missing people and items is still performed at Yuriage beach. In the beach, we must search not only the underground of the beach but also the surface of that. In addition, it is important to select a candidate search locations because the target range of the sands is very huge. The purpose of this research is to improve the search activity by using a video from a drone camera. In this paper, we propose the key-frame selection and reconstruction method for real-time 3-D reconstruction using a drone. Using the property of the sands and drone movings over that, we use a motion estimation by optical flow for selecting key-frames and adopt a homography-based method for real-time 3-D reconstruction. We also adopt local loop-closing for improving accuracy. In addition, using the flow information between key-frames, we can generate a mosaiced image simultaneously. We show the effectiveness of our method by real video experiments.

## 1. はじめに

東日本大震災において発生した津波による行方不明者や遺失物の捜索活動が閑上浜などにおいて現在も続けられて

いる [17] (図 1)。この捜索活動は、図 1 (a) のように地中レーダを用いて地中の埋没物を探索し、同図 (b) のように反応があった部分を人手で掘り起こすという時間と手間のかかる方法で行われており、その効率化が求められている。このために、人工衛星や航空機に搭載されたレーダを用いた捜索場所候補の絞り込みが行われているが、上空からのレーダ画像には地中の物体と同様に地上物体の反射の影響を受けるため、レーダ画像上ではその区別が難しい。それらを区別するために、実際の表層の情報を取得することが

<sup>1</sup> 豊橋技術科学大学 情報・知能工学系,  
Department of Computer Science and Engineering,  
Toyohashi University of Technology, Japan.

a) m-hashimoto@img.cs.tut.ac.jp

b) r-matsunaga@img.cs.tut.ac.jp

c) nishiyama@img.cs.tut.ac.jp

d) kanazawa@cs.tut.ac.jp

重要となるが、砂浜の表層は天候によって日々変化する。そこで、捜索日当日に出来る限り最新の情報を取得することも必要となる。加えて、日々変化する砂浜の表層においては、漂着物も重要な捜索対象であり、その漂着物の位置把握や認識も重要となる。

本研究では、広範囲をリアルタイムに観測可能なドローンを用い、上空から撮影された映像から、地表面の形状や漂着物などをリアルタイムに3次元復元するだけでなく、同時にモザイク画像を生成して、広範囲な砂浜の漂着物のチェックを行うことを可能とすることで、捜索活動の効率化を図ることを目的とする。本稿では、広範囲な砂浜を撮影しながら、リアルタイムに3次元復元を行うために、ドローンでの撮影に特化したキーフレーム選択と砂浜のようなほぼ平面状の形状に適した復元法を提案する。

砂浜における撮影においてはドローンは砂浜とほぼ平行に、ほぼ一定の速度で、網羅的に撮影することが多く、ホバリングや移動方向の転換など、3次元復元には適していないフレームも多い。そこで提案法では、このような冗長なフレームに対し、オプティカルフローによる運動推定により取り除くだけでなく、移動量を一定に保つようにキーフレームの選択を行う。また砂浜がほぼ平面であることを用い、射影変換ベースの形状復元を採用する。これらを並列処理とGPUを用いた高速化により、リアルタイムに処理を行う。

## 2. 関連研究

良く知られている通り、移動するカメラで撮影された画像から3次元形状を復元する手法は Structure from Motion (SFM)、ロボット分野では SLAM (Simultaneous Localization and Mapping) と呼ばれ、極めて多くの研究が行われている。

SFMの手法として最も著名なものに Bundler[15] が挙げられる。この Bundler は、あるシーンを撮影した大量の未整列な画像群から3次元復元を行う手法であり、まず各画像から SIFT[9] による特徴抽出と対応付けを行う。そして画像間の対応により推定した基礎行列から、既知あるいは推定した焦点距離情報を用いて5点アルゴリズムによるカメラ間の相対運動の推定と三角測量による対応点の復元を行い、最後にバンドル調整[16]により、高精度な形状復元を実現している。しかし、この Bundler のみでは画像間で対応の取れた疎な3次元点群しか得られないため、多くの場合、後処理として PMVS2[4] を行い、復元した3次元点群を補完し、比較的密な形状復元を得ている。この Bundler を用いた復元では基礎行列をベースに移動量推定を行うため、砂浜のようなほぼ平面状のシーンに対しては、基礎行列が退化するため、結果として復元結果に歪みが生じることも多い。また、Bundler ではバンドル調整の処理時間が無視できず、リアルタイムな3次元復元処理は難しい。



図1 閑上浜での捜索活動の様子。

ARの分野で著名なものとして PTAM (Parallel Tracking and Mapping) [8] や DTAM (Dense Tracking and Mapping in Real-Time) [13] が挙げられる。PTAM はマーカーレス AR のための単眼カメラによる SLAM であり、シーン中の特徴点をトラッキングすることにより自己位置を推定する。この PTAM の実装ではトラッキングとマッピングの処理を分離して並列に処理し、局所的なバンドル調整を導入することにより、リアルタイムに高精度な SLAM を可能にした。DTAM は、その名の通り、リアルタイムに密な点群の復元と自己位置推定を行う手法であり、輝度誤差と正則化項によるエネルギーによる最適化により、画像間の密なマッピングを求めている。疎な点群ベースの PTAM に比べ、密なマッチングを行うために、復元形状および自己位置推定とも高精度な手法となっている。しかしながら、PTAM や DTAM は主に AR が目的であるため、比較的小さく、起伏の大きなシーンを対象としており、地形などの広いシーンでは難しい。

SLAM において代表的なものとして LSD-SLAM[2] や ORB-SLAM [12] が挙げられる。LSD-SLAM は非特徴点ベースの単眼 SLAM であり、輝度誤差最小化によるトラッキングとカメラ姿勢に対する最適化により、比較的密な点群のリアルタイムな復元を実現しており、工場などの大規模なシーンにも適用可能となっている。しかし、画像内のエッジ部分に存在する点群を用いているため、砂浜のようなテクスチャが均質なシーンでは精度が低下する。ORB-SLAM[12] は特徴点ベースの SLAM であり、ORB 特徴量を用いた高速な処理に加え、バンドル調整とカメラ姿勢の最適化、ループクロージングにより高精度な SLAM を行っている。

これらの手法はいずれも一般のシーン、つまり比較的起伏の多いシーンで、テクスチャも比較的リッチなシーンを対象としているため、砂浜のようなほぼ平面でかつテクスチャもほぼ一定であるようなシーンに対しては、カメラの移動量の計算の精度が下がり、復元した形状に歪みを残す結果となる。

## 3. 提案手法

提案法は、ドローンにより撮影された砂浜の映像から復元することを目指しており、リアルタイム化のために次の2つの処理を行う。

- ドローンの移動に特化したキーフレームの選択
- 平面状形状に対応するための射影変換ベースの形状復元

キーフレーム選択では、ドローンで撮影された映像からオプティカルフローを求めて運動推定を行い、ホバリングや純粋回転に近いなどのような、復元に適さないフレームを検出すると共に、一定量の移動を検出し、それらをキーフレームとして選択する。形状復元では、対応点を検出後、砂浜がほぼ平面であることを考慮して、射影変換ベースの3次元復元を行う。対応する画像間の大域的な平面点群に対する射影変換を求め、その平面上の点だけでなく、平面上にない点も復元を行う。これにより、砂浜のような平面状のシーンに対しても高精度に復元が可能となる。そして、これらの処理に対するリアルタイム処理を実現するために、並列化およびGPUによる処理を用いる。

### 3.1 キーフレーム選択

ドローンにより撮影された映像には静止中のほぼ変化のないフレームや純粋回転のような復元に適さないフレームが多く含まれる。そのようなフレームを使用することは処理時間や復元精度の低下の問題があるため、リアルタイム処理においては復元に適切なフレームのみを使用する必要がある。提案法ではLucas-Kanade法[10]によるオプティカルフローを用いた運動推定により入力映像から冗長なフレームを除去し、復元に使用するキーフレームを選択する。同時にオプティカルフローを用いて特徴点のトラッキングを行うことで対応付けを行う。また、リアルタイム化のためにオプティカルフローの計算はGPUを用いて高速化する。キーフレーム選択の流れは以下のとおりである。

- (1) 先頭フレームの画像のレンズ歪みを補正して、キーフレームとして登録する。
- (2) 以下を入力の流れの終わりまで繰り返す。
  - (a) 画像のレンズ歪みを補正し、オプティカルフローによる運動推定と特徴点のトラッキングを行う。
  - (b) 選択基準に当てはまったフレームをキーフレームとして登録し、特徴点のトラッキング結果を対応点として保存する。
  - (c) 新たなキーフレームが登録された場合はGFTT[14]により特徴点の検出を行う。

以下、詳細に説明する。

#### 3.1.1 オプティカルフローによる運動推定

冗長なフレームおよび復元に適さないフレームを除去するために、オプティカルフローによる運動推定を行う。ここでは連続したフレーム間でオプティカルフローを計算し、フローベクトルの大きさと方向のヒストグラムを作成し、その特徴を利用する。図2に作成されたヒストグラムの例を示す。ドローンの並進中は方向、大きさのヒストグラム双方に強いピークが生じるのに対し、回転中の場合は

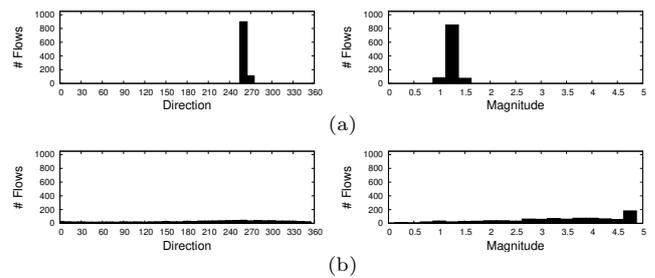


図2 ヒストグラムの例(左:フローベクトルの方向, 右:大きさ). (a) 並進, (b) 回転.

双方とも平坦なヒストグラムとなる。これらの特徴を利用し、ヒストグラムのピークの位置と大きさに対してしきい値を用いて判別することで運動の種類を推定する。またオプティカルフローの検出を行う位置は、映像のアスペクト比に影響されないよう、画像中央の正方領域に内接する円内に等間隔で設定する。

#### 3.1.2 キーフレームの選択基準

この運動推定の結果を利用して、キーフレームの選択を行う。ここでは以下の基準に当てはまったフレームをキーフレームとする。

- 並進フレームであること。
- 前キーフレームからの重なり率がしきい値を下回っていること。

これらにより回転中や静止中のフレームを除去しつつ、画像の重なりを基準に等間隔でキーフレームを選択することができる。重なり率は並進フレームのフローベクトルの平均値を積算して、画像の移動量から計算する。

### 3.2 3次元形状の復元

前述のキーフレーム選択によって得られた連続した2枚のキーフレームを用いて形状の復元を行う。一般の基礎行列あるいは基本行列ベースの復元では、対象とする砂浜のようなほぼ平面のシーンはしばしば退化してしまい、復元形状に歪みを生じることも多い。そこで提案法では平面が多いシーンに対しても退化せず、高精度な復元が可能な射影変換ベースの復元を行う[6], [7]。また、隣接して復元された形状を、順次合成することにより広い領域の形状復元を行う。形状復元の大まかな流れは以下の通りとなる。

- (1) 新しいキーフレームが登録されるまで待機する。
  - (2) SURF特徴量[1]により前キーフレームと対応付けを行う。
  - (3) 射影変換ベースの形状復元を行う。
  - (4) 形状合成、局所的なループレゾリングを行う。
- 対応付けに用いるSURF特徴量の計算には高速化のため、GPUを使用する実装を用いる。

#### 3.2.1 射影変換ベースの形状復元

射影変換ベースの復元では、射影変換の関係を満たす平面状の点を平面上に復元し、満たさない立体部分の点を射

影変換からのずれを利用して復元する。

まず、エッジ極線拘束条件による RANSAC[3] により、画像間全体の対応から誤対応を除去する。次に、得られた対応組から、射影変換を拘束条件とした RANSAC により、平面上の点の対応を抽出する。この RANSAC は提案法の精度に関わるため厳しい条件で行い、多くのサンプルを高速で処理するため、GPU を用いて処理を行う。次に、平面上の点の対応に対し、FNS 法 [5] を用いて射影変換行列  $H$  を推定し、その行列を分解して平面のパラメータ（法線ベクトル  $\mathbf{n}$ 、原点からの距離  $d$ ）とカメラの相対的な運動（並進  $\mathbf{t}$ 、回転  $\mathbf{R}$ ）を計算する。そして、得られた平面とカメラの運動パラメータを用いて 3 次元位置の計算を行う。

ここでは、焦点距離  $f$  は校正済みで既知とし、対応点  $(x, y)$ ,  $(x', y')$  をそれぞれ同次ベクトル  $\mathbf{x} = (x/f, y/f, 1)^\top$ ,  $\mathbf{x}' = (x'/f, y'/f, 1)^\top$  で表す。

まず平面上の点については、最適補正 [5] を用いて対応点の位置を平面上で視線が交わるよう補正し、次の式で 3 次元位置  $\mathbf{X}$  を計算する。

$$\mathbf{X} = \frac{d}{(\mathbf{n}, \mathbf{x})} \mathbf{x} \quad (1)$$

ここで、 $(\mathbf{a}, \mathbf{b})$  はベクトル  $\mathbf{a}$  と  $\mathbf{b}$  の内積とする。

平面に乗っていない点については、パララックスを用いた 3 次元位置の計算を行う [11]。このとき、図 3 に示すように、第 1 画像上の点  $\mathbf{x}$  は射影変換行列  $H$  により、第 2 画像上で

$$\tilde{\mathbf{x}}' = \mathcal{Z}[H\mathbf{x}] \quad (2)$$

となる。ただし、 $\mathcal{Z}$  は第 3 要素を 1 に正規化する演算とする。しかしながら、平面上にない実際の対応点の位置  $\mathbf{x}'$  とは異なった位置となり、この  $\mathbf{x}'$  と  $\tilde{\mathbf{x}}'$  の差をパララックスとよぶ。提案法では、これを利用して次のように 3 次元位置を計算する。

(1) 平面上の仮想的な 3 次元位置  $\tilde{\mathbf{X}}$  を次式で計算する。

$$\tilde{\mathbf{X}} = \frac{d}{(\mathbf{n}, \mathbf{x})} \mathbf{x} \quad (3)$$

(2) 平面からの距離  $\Delta\mathbf{x}$  を次式で計算する。

$$\Delta\mathbf{x} = \frac{1}{2}(\mathcal{Z}\mathbf{x} + \mathcal{Z}'\mathbf{R}\mathbf{x}' + \mathbf{x}_p) \quad (4)$$

ただし、次のようにおいた。

$$\begin{aligned} \mathbf{x}_p &= \tilde{\mathbf{X}}_z \mathbf{R}(\mathbf{x}' - \tilde{\mathbf{x}}'), \\ Z &= \frac{(\mathbf{x}_p, \mathbf{x})\|\mathbf{x}'\|^2 - (\mathbf{x}, \mathbf{R}\mathbf{x}')(\mathbf{x}_p, \mathbf{R}\mathbf{x}')}{\|\mathbf{x}\|^2\|\mathbf{x}'\|^2 - (\mathbf{x}, \mathbf{R}\mathbf{x}')^2}, \\ Z' &= \frac{(\mathbf{x}, \mathbf{R}\mathbf{x}')(\mathbf{x}_p, \mathbf{x}) - (\mathbf{x}_p, \mathbf{R}\mathbf{x}')\|\mathbf{x}\|^2}{\|\mathbf{x}\|^2\|\mathbf{x}'\|^2 - (\mathbf{x}, \mathbf{R}\mathbf{x}')^2} \end{aligned} \quad (5)$$

(3) 3 次元位置  $\mathbf{X}$  を次のように計算する。

$$\mathbf{X} = \tilde{\mathbf{X}} + \Delta\mathbf{x} \quad (6)$$

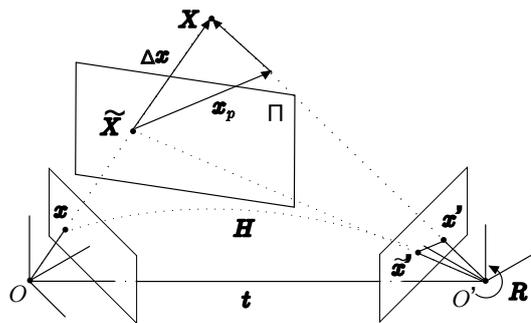


図 3 シーン中の平面と対応点の関係。

この復元の際、最適な補正を行うことも考えられるが、ここではリアルタイム処理のために、簡略化した上記手順を用いる。

また復元された 3 次元形状には、誤対応などの原因により、誤差の大きな外れ値が含まれる可能性がある。そのような点は形状合成の精度低下の原因となるのに加えて結果の視認性も悪くなる。そこで、3 次元位置が他の点から大きく離れている点を外れ値として除去する。その手順を以下に示す。

(1) 復元された各 3 次元点  $\mathbf{X}_i$ ,  $i = 1, \dots, N$  に対して、その  $K$  個の近傍点の集合  $\mathcal{X}_i = \{\mathbf{X}_\alpha\}$ ,  $\alpha = 1, \dots, K$  を求め、 $\mathbf{X}_i$  からの距離の平均  $\bar{d}_i$  を計算する。

$$\bar{d}_i = \frac{1}{K} \sum_{\mathbf{X}_\alpha \in \mathcal{X}_i} \|\mathbf{X}_\alpha - \mathbf{X}_i\| \quad (7)$$

(2) 近傍との距離の平均  $\bar{d}_i$ ,  $i = 1, \dots, N$  の標準偏差  $\sigma$  を次のように計算する。

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N \bar{d}_i^2 - \bar{d}^2}, \quad \bar{d} = \frac{1}{N} \sum_{i=1}^N \bar{d}_i \quad (8)$$

(3)  $\bar{d}_i > \alpha\sigma$  となる点を外れ値として除去する。

ここで、近傍の個数  $K$  と定数  $\alpha$  はどの程度の外れ値を除去するかのパラメータとなる。提案法では  $K = 7$ ,  $\alpha = 3$  を用いた。

### 3.2.2 形状の合成

2 つのキーフレーム間で復元された形状を合成して、より大きな形状を求める。この形状の合成では、既知の対応点の情報を用いて、既に合成済みの大きな形状と新たに復元された部分形状の相対的なスケールを推定し、それらが一致するよう補正を行い、形状を合成する。

ここでは、まず対応候補から 3 次元相似変換を拘束条件とする RANSAC により、誤対応を除去する。その後、得られた対応すべてを用いて、3 次元相似変換を計算し直し、合成する。

### 3.2.3 局所的なループクロージング

前述の形状復元と合成では連続する 2 つのキーフレーム間の対応点のみに基づいて行われるため、合成における誤差の累積され、結果として大きな復元形状になればなるほ



図 4 生成されたモザイク画像の例

ど、より大きな誤差が生じてしまう。これを防止するためには、一般的には観測が閉じていること、すなわちループクロージングを用いた最適化を行うが、大域的なループクロージングは極めて計算コストが高く、リアルタイム処理には適さない。そこで提案法では、局所的なループクロージングを検出し、それを用いて補正を行うことで誤差の蓄積を緩和する。

まず形状合成の前に、最新のキーフレームと空間的に隣接したキーフレーム間との対応付けを行う。合成を行うキーフレームと共通の特徴点を  $N$  点以上持つキーフレームの中で最も遠いカメラまでの距離を  $d_m$  とし、 $\alpha d_m$  以下のキーフレームとの間で対応付けを行う。そして対応付けられた点を同一の点として、それらの中点を新たな位置として点の統合を行うことで、局所的なループのクロージング処理を行う。これにより、計算速度を落とすことなく、より広範囲の対応を利用して合成が行え、誤差の蓄積も軽減される。

### 3.3 モザイク画像の生成

提案法では、3次元形状のリアルタイム復元だけでなく、地表面の目視搜索を効率化するために、地表面の状況把握のため、簡易的なモザイク画像の生成も同時に行う。

このモザイク画像生成では、ドローンの動きはほぼ平面上の並進と仮定し、対応点からキーフレーム間の並進量を計算することで、簡易的にモザイク画像を生成する。生成されたモザイク画像の例を図4に示す。

## 4. 実験

提案法の有効性を確認するため、前述の提案法を実装したシステムを用いて実画像による実験を行った。実験は従来法との比較実験、提案法の精度評価の2つを行った。事前にカメラキャリブレーションを行って得られたカメラパラメータを利用した。実験環境を表1に示す。

表 1 実験の環境.

項目	仕様
CPU	Intel Core i7 7700HQ @ 2.60GHz
RAM	16GB
GPU	NVIDIA GeForce GTX 1060 6GB

表 2 実験に用いた動画の長さ.

シーン	フレーム数	動画時間 [s]
1	6579	219
2	10966	365
3	2542	84
4	5395	180

### 4.1 従来法との比較実験

まず従来法との比較を行った。比較対象として、Bundler[15] + PMVS2[4] と LSD-SLAM[2] を用いた。また、事前に撮影した動画を入力としてオフラインで実行した。動画の撮影時の解像度は  $3840 \times 2160$  であるが、提案法および Bundler には  $1280 \times 720$  にリサイズした動画を使用した。また Bundler の入力には提案法によって選択されたキーフレームを使用した。LSD-SLAM の入力には  $640 \times 360$  にリサイズした動画を使用した。ここで、提案法のフレーム選択の重なり率のしきい値は 0.8 とした。

使用した映像は主に平面部分を撮影した平面シーン2つ(シーン1, 2)、車や堤防などの立体的な部分を含んだ立体部分を含むシーン2つ(シーン3, 4)の計4シーンである。各映像の情報を表2に示す。

平面シーン1に対し、提案法によって選択されたキーフレームの一部を図5に、それぞれの手法で復元した形状を図6に示す。また立体部分を含むシーン4に対して提案法によって選択されたキーフレームの一部を図7に、それぞれの手法で復元した形状を図8に示す。

見てわかる通り、いずれのシーンにおいても、提案法が少ない歪みで形状復元を行えていることが確認できた。平面シーンでは、Bundler や LSD-SLAM は形状が大きく歪んでいたり復元に失敗しているが、提案法では歪みの少ない形状を復元できている。提案法は支配的平面を基準とした平面ベースの復元を行っているため、このようなほぼ平面のシーンでも安定した復元が行えることがわかる。

立体部分を含むシーンでは、提案法は立体部分を含めて精度良く復元を行うことができ、このような立体部分においても提案法が形状復元を行えることを確認した。Bundler では PMVS2 により密な点群が得られているが、シーンによっては平面部分に歪みが生じた。LSD-SLAM では誤差の蓄積により堤防がずれて復元されてしまっている。

各手法の処理時間と復元された点数の比較を表3に示す。提案法はほぼリアルタイムな処理を行うことができていくことがわかる。テクスチャが少ない部分などで対応数が少なくなり疎な点群となっているが、砂浜におかれたごみの山などを確認することは可能で、シーンの把握は十分

可能であると考え、一方 Bundler では PMVS2 により密な点群が得られているが、処理時間が長くオフライン処理しか行うことはできない。LSD-SLAM はリアルタイムな処理が可能で、その性質上非常に密な点群が得られたが、砂浜のような均質なシーンでは精度が低く点のばらつきが大きいため、詳細な形状の確認が難しかった。

## 4.2 実画像を用いた精度評価実験

次に提案法の形状復元の精度を確認するため、別の実映像を用いた精度の評価を行った。精度評価では長さが既知の物体を撮影した映像を用いて復元を行い、復元形状のスケールを合わせて長さの誤差を評価した。図 9 に復元された形状と計測箇所を示す。

表 4 は 1 の長さにスケールを合わせて評価を行った結果である。誤差は 10[cm] から 30[cm] 程度、割合としては 1[%] 以下となった。誤差の要因としては形状合成のスケールの推定誤差が蓄積したものと考えられる。しかし、リアルタイム処理であることを考慮すると、本研究の目的に対して十分な精度が得られていると考えられる。

## 5. まとめ

ドローン映像からの 3 次元復元をリアルタイムに行うためのキーフレーム選択と形状復元の方法を提案した。提案法では、オプティカルフローを用いた運動推定によるキーフレーム選択とパララックスを用いた平面に適した形状復元による 3 次元復元を並列化と GPU を用いた高速化によりリアルタイムに行う。実画像を用いた実験により、砂浜のような平面的なシーンにおいても従来法と比較して高精度な復元をほぼリアルタイムに行えることを示した。

今後の課題として、実利用を考慮したトラッキングの失敗からの自動的な復帰や視認性を向上させるためにパッチを貼ることで密な点群を得ること、ローカルバンドル調整による高精度化などが挙げられる。

## 謝辞

本研究の一部は、日本学術振興会科学研究費補助金 基盤研究 (C)(No. 15K00232) および基盤研究 (B)(No. 15H02997) による。

## 参考文献

[1] Bay, H., Tuytelaars, T. and Van Gool, L.: SURF: Speeded Up Robust Features, *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I* (Leonardis, A., Bischof, H. and Pinz, A., eds.), Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 404–417 (2006).

[2] Engel, J., Schöps, T. and Cremers, D.: LSD-SLAM: Large-Scale Direct Monocular SLAM, *European Conference on Computer Vision (ECCV)* (2014).

[3] Fischler, M. A. and Bolles, R. C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM*, Vol. 24, No. 6, pp. 381–395 (1981).

[4] Furukawa, Y. and Ponce, J.: Accurate, Dense, and Robust Multi-View Stereopsis, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2007).

[5] Kanatani, K. and Niitsuma, H.: Optimal Two-View Planar Scene Triangulation, *IPSS Transactions on Computer Vision and Applications*, Vol. 3, pp. 67–79 (2011).

[6] Kanazawa, Y. and Kawakami, H.: Detection of planar regions with uncalibrated stereo using distributions of feature points, *The 15th British Machine Vision Conference (BMVC2004)*, London, U.K., pp. 247–256 (2004).

[7] Kanazawa, Y., Sakamoto, T. and Kawakami, H.: Robust 3-D reconstruction using one or more homographies with uncalibrated stereo, *Sixth Asian Conference on Computer Vision (ACCV2004)*, Jeju Island, Korea, pp. 503–508 (2004).

[8] Klein, G. and Murray, D.: Parallel Tracking and Mapping for Small AR Workspaces, *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*, Nara, Japan (2007).

[9] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints, *Int. J. Comput. Vision*, Vol. 60, No. 2, pp. 91–110 (2004).

[10] Lucas, B. and Kanade, T.: An iterative image registration technique with an application to stereo vision, *7th Int. Joint Conference on Artificial Intelligence*, Vancouver, British Columbia, pp. 674–679 (1981).

[11] Matsunaga, R., Hashimoto, M., Kanazawa, Y. and Sonoda, J.: Accurate 3-D reconstruction of sands from UAV image sequence, *2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)*, pp. 1–6 (2016).

[12] Mur-Artal, R., Montiel, J. M. M. and Tardes, J. D.: ORB-SLAM: A Versatile and Accurate Monocular SLAM System, *IEEE Transactions on Robotics*, Vol. 31, No. 5, pp. 1147–1163 (2015).

[13] Newcombe, R. A., Lovegrove, S. J. and Davison, A. J.: DTAM: Dense Tracking and Mapping in Real-time, *Proceedings of the 2011 International Conference on Computer Vision, ICCV '11*, Washington, DC, USA, IEEE Computer Society, pp. 2320–2327 (2011).

[14] Shi, J. and Tomasi, C.: Good features to track, *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593–600 (1994).

[15] Snavely, N., Seitz, S. M. and Szeliski, R.: Photo Tourism: Exploring Photo Collections in 3D, *SIGGRAPH Conf. Proc.* (2006).

[16] Triggs, B., McLauchlan, P., Hartley, R. and Fitzgibbon, A.: Bundle Adjustment – A Modern Synthesis, *VISION ALGORITHMS: THEORY AND PRACTICE, LNCS*, Springer Verlag, pp. 298–375 (2000).

[17] 東日本大震災捜索活動協議会実行委員会：東日本大震災行方不明者捜索活動協議会，<http://54renzock53.wixsite.com/mpsa>.

表 3 復元に要した時間と復元点数.

シーン	提案法		Bundler+PMVS2		LSD-SLAM	
	時間 [s]	点数	時間 [s]	点数	時間 [s]	点数
1	220	104,837	2,344	14,052	219	14,357,646
2	375	436,491	26,672	1,593,913	365	29,726,982
3	84	41,685	834	288,016	84	4,820,475
4	181	56,240	932	557,291	180	6,286,926

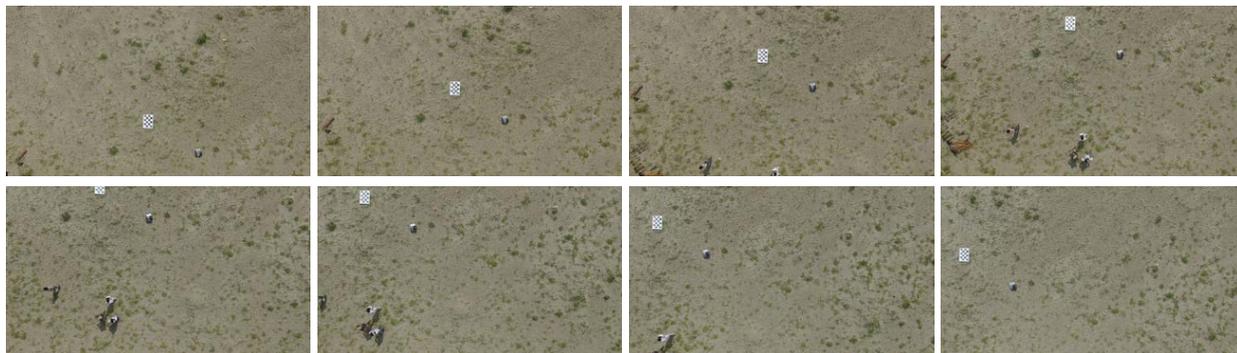
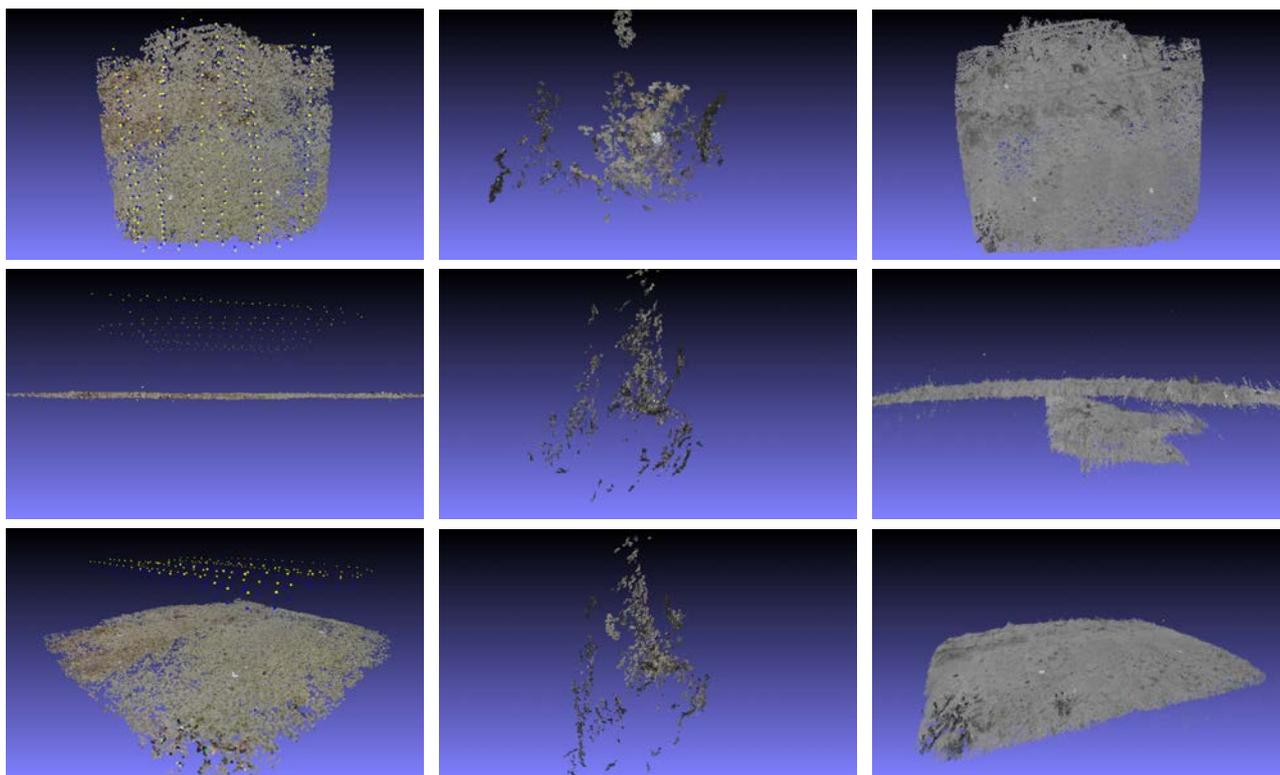


図 5 提案法により平面シーンで選択されたキーフレームの一部.



提案法

Bundler + PMVS2

LSD-SLAM

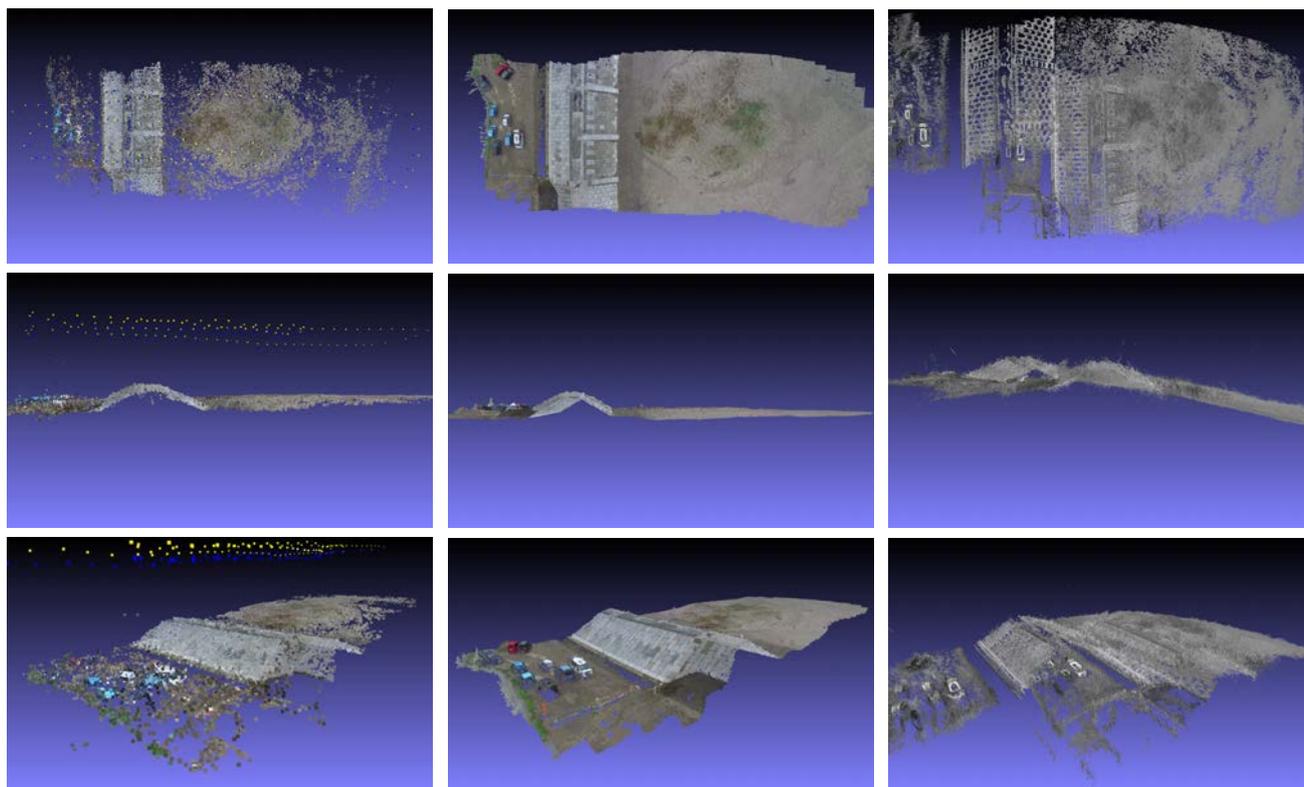
図 6 平面シーンの復元結果の比較. 上から, それぞれ, 真上, 真横, 斜めから見た形状.

表 4 復元精度の評価結果.

番号	実測値	復元形状	誤差	誤差 [%]
1	570	570.0	—	—
2	3335	3365.02	30.02	0.90
3	3335	3317.29	17.71	0.53
4	2770	2760.52	9.48	0.34



図 7 提案法により立体部分を含むシーンで選択されたキーフレームの一部.



提案法

Bundler + PMVS2

LSD-SLAM

図 8 立体部分を含むシーンの復元結果の比較. 上から, それぞれ真上, 真横, 斜めから見た形状.

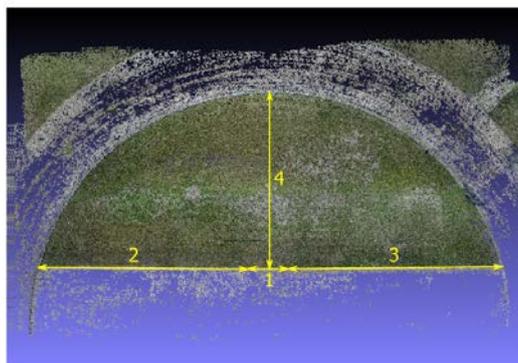


図 9 精度評価の復元形状と計測箇所.