海面等領域推定と単眼深度推定を活用した港湾施設側面の点検用画像作成方法の提案

藤井 純一郎† 石井 明† 宮本 冬馬† 山岡 淳† 里村 大樹‡
†八千代エンジニヤリング株式会社 ‡国土交通省 国土技術政策総合研究所
E-mail: jn-fujii@yachiyo-eng.co.jp

1 はじめに

港湾施設は高度経済成長期にその多くが整備され，今後は建設後50年が経過して老朽化している施設が増加することが見込まれている[1]．港湾施設の点検は，これまで主に技術者による目視点検により実施されており，より効率的かつ的確な維持管理手法が求められている．これに対し，国土技術政策総合研究所では UAV 空撮画像から AI 等を活用して港湾施設の変状を抽出すること等により点検診断を支援する「UAV を活用した港湾の施設の点検診断システム」（UAV 点検診断システム）の開発を行っている[2]．現状の UAV 点検診断システムは，施設上面部の点検用画像（オルソ画像・DSM 画像）を入力として変状抽出を行うが，施設の劣化度判定等を行うためには施設側面部でも同様の変状抽出が必要である．しかし，施設側面の点検用画像を作成する際に海面や遠景が映りこむことで，SfM 解析において特徴点マッチングに失敗する場合がある．

本研究は，港湾施設側面の画像を対象に，SfM 解析時の特徴点マッチングを改善することを目的として，海面等領域推定と単眼深度推定に基づき点検対象外領域をマスキングする手法を提案し，港湾の実画像で検証実験を行ったものである．

図-1 UAV 点検診断システム概要

2 UAV 点検診断システムの概要と課題

2.1 システム概要

UAV 点検診断システムは，港湾の施設の点検診断ガイドライン[3]に基づき，港湾施設の変状を自動で抽出し，劣化度および性能低下度の評価を支援するシステムである．UAV で撮影した港湾施設の画像から SfM 解析により三次元データ化を行い，そこで得られた点検用画像からひび割れ・段差・錆汁・鉄筋露出等の変状を検出する機能を有する（図- 1）．変状は位置・大きさを自動計測し，港湾技術者による劣化度判定の基礎資料として用いられる．なお，点検用画像作成時に SfM 解析を行うのは，変状の位置精度を担保するため，測量結果に基づく誤差を検証したオルソ画像・DSM 画像を作成するためである．

2.2 施設側面部に適用する際の問題点

現行の UAV 点検診断システムは施設上面部を対象としている．それに対し，筆者らは予備実験として施設側面部への適用を試行して以下の課題を把握した．

施設側面の画像の場合，港湾施設の背後に海面や空・山・構造物・地平線など遠景が映りこむ．このような画像を用いて SfM 解析を行うと，海面や遠景の特徴点マッチングに失敗し（図- 2），結果的に不自然な三次元形状が作成される場合があった（図- 3）．

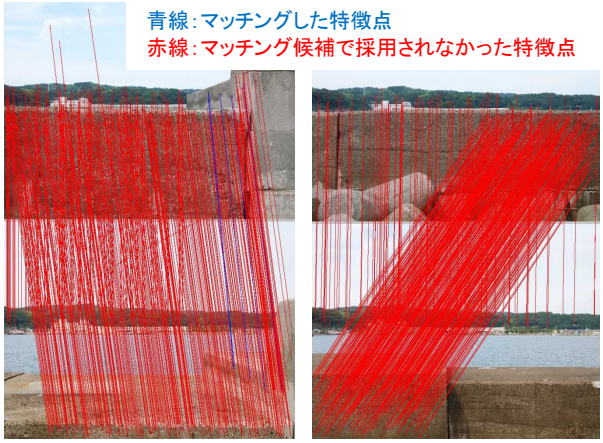

図-2 特徴点マッチングの失敗例

図-3 不自然な三次元化例

3 先行研究

現行の UAV 点検診断システムにおいても，施設上面部の点検用画像を作成する際に特徴点マッチングの阻害要因となる海面や空（海面等）の領域推定・除去を実装している[4]．具体的には Deeplab v3+[5] を用いて海面等領域の Semantic Segmentation を行うモデルを事前に学習しておき，そのモデルを用いて SfM 解析の前処理として海面等領域を推論・マスキングすることで特徴点マッチングの改善を図っている．ただし本研究で対象とする施設側面の画像では，海面だけでなく遠景が映りこむため，これらの領域の推定・マスキングが課題として残されている．

遠景の領域の推定のためには，被写体との距離を認識する方法があり，撮影時にハードウェアで距離を記録する深度センシングカメラ・ステレオカメラやレーザー照射により点群を取得する LiDAR がすでに商用化されている．ただし，専用のハードウェアを UAV に追加で搭載する必要があるため，全国の港湾に汎用的に適用するには課題がある．

一方で，撮影後にソフトウェアで二次元の画像から画素ごとの距離を推定する単眼深度推定の研究が進んでいる．初期の単眼深度推定の研究では，深度情報のラベル付きデータ（GroundTruth）を用いた教師あり学習の事例があるが，教師データを収集・作成することが課題であった[6][7][8]．それに対し Zhou ら[9]は動画を入力として深度を推定する DepthCNN とカメラ姿勢を推定する PoseCNN を学習することにより，深度情報のラベルなしでの単眼深度推定を実現した．ただし，適用シーンやドメインごとにバリエーションを確保した動画を，その都度収集し学習を行う必要がある．

近年は大量のデータで事前学習することにより，用途の異なるシーンやドメインに対してロバストな単眼深度推定を実現する手法が提案されている．Ranftl ら[10]は，測定方法や環境が異なるデータセットを同時に学習することを可能とする損失関数を提案し，複数のデータセットで事前学習済みのモデルを公開しており，zero-shot 深度推定を実現した．さらに Yang ら[11]は，ラベル付きデータとラベルなしデータを組み合わせた半教師あり学習手法を提案した．複数のラベル付きデータセット（約 150 万枚）とラベルなしデータセット（約 6200 万枚）を学習することで，汎化性能が大幅に向上したことが報告されている．

4 提案手法

本研究では，UAV 点検診断システムを港湾施設側面に適用する際の問題点を解消するため，点検対象外領域である海面等・遠景の領域を海面等領域推定と単眼深度推定によりマスキングする手法を提案する（図- 4）．本手法により SfM 解析時の特徴点マッチングが改善され，変状抽出用の点検用画像作成の省力化および精度向上が期待できる．

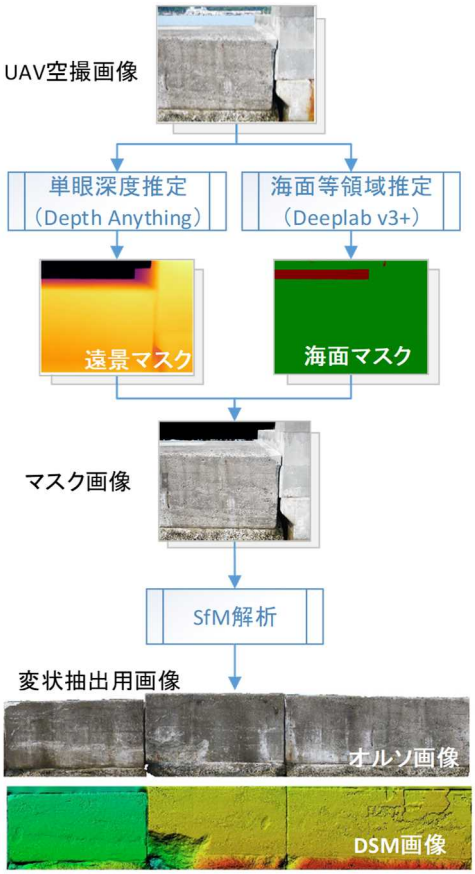

図-4 提案手法のフロー

5 検証実験

5.1 実験方法

提案手法の妥当性を確認するため，図-4 のフローに従って，港湾施設側面の実画像を用いて変状抽出用画像の作成を行った．改善効果を検証するため，海面等領域推定・単眼深度推定の有無の異なる条件で比較を行った．

5.1.1 使用データ

港湾施設側面の画像，および SfM 解析に用いる測量結果が整備されていた港湾施設のデータを用いた．各港の画像枚数・仕様は下表の通りである．

表-1 使用データ

対象	画像枚数	解像度
A 港	291 枚	5280 x 3956
B 港	188 枚	5472 x 3648

5.1.2 使用モデル・ソフト

(1)単眼深度推定モデル(遠景マスク)

本研究では全国の港湾施設側面に適用することを視野に入れているため，先行研究でロバストに zero-shot 深度推定が可能とされる Yang らによって提案された DepthAnything[11]を採用した．実装は公式サイトで公開されている事前学習モデルを用い，Encoder は ViT-S により深度推定を実施した．

実験に先立ち，港湾施設側面の画像で深度推定を行い，定性的に問題がないことを確認し，併せて遠景としてマスクする領域の閾値を検討する予備実験を行った(図- 5)．深度推定結果からマスクする領域は，Depth＜1.0 のピクセルとした．

なお，この閾値は本研究では予備実験でアドホックに決めたものであるが，UAV 点検診断システムで用いる UAV 空撮画像は港湾施設から一定距離で撮影するため，将来的には論理的に決めることができると考えている．

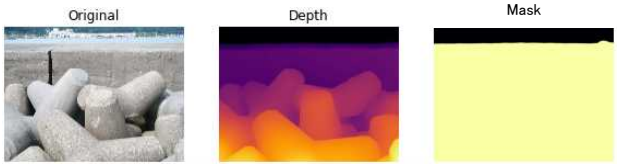

図-5 単眼深度推定結果の例
(Depth2 は Depth＜1.0 で 2 値化したもの)

(2)海面等領域推定モデル(海面マスク)

海面等領域推定は里村ら[4]による先行研究で精度検証がされており，UAV 点検診断システムにも実装済みの既存のモデルを踏襲した．

本モデルは Deeplab v3+[5]を用いて，国土技術政策総合研究所で独自に作成した全国の港湾の海面等領域のアノテーションデータ 645 枚を学習したものである．画像は学習/検証に分割し，Deeplab v3+に合わせてタイル分割して用いたため，学習に使った画像枚数は約 10 万枚であった．精度評価は mIoU=85.3%と報告されている．

実験に先立ち，港湾施設側面の画像で推論を実施し，側面の画像に対しても問題なく海面等領域推定が行えることを定性的に確認した(図- 6)．

図-6 海面等領域推定結果の例
(Mask の赤領域が海面等領域)

(3)SfM 解析ソフト

SfM 解析および変状抽出用画像の作成は，商用ソフトの Metashape ver2.0 を用いて実施した．後述する評価指標も Metashape で算定したものを採用している．

5.1.3 実験ケース

海面・遠景のマスクを行わないで変状抽出用画像を作成するケースをベースラインとして，海面・遠景のマスクの有無の別により表-2 のケース間で評価・考察を行った．

表-2 実験ケース

ケース	遠景マスク	海面マスク
original(baseline)	-	-
back	○	-
sea	-	○
back+sea(proposed)	○	○

5.1.4 評価方法

提案手法で期待される変状抽出用画像作成の省力化および精度向上を定量化する指標として，以下を評価した．

（1）省力化指標：タイポイント数

省力化の直接的な指標としては，特徴点マッチング失敗による作業のやり直しや特徴点の手動修正などにかかる作業時間の短縮が考えられるが，これら作業が必ずしも発生するわけではなく，かつ，その作業時間は対象港湾施設ごとの画像特性や作業者の熟練度に依存するため定量評価が難しい．そのため，本実験ではそれらの作業増のリスクに直結する特徴点マッチングの成功数を評価指標とする．

SfM 解析の過程では，各画像内の特徴点（キーポイント）を抽出し，次にそれらを画像間でマッチングを図る（タイポイント）．提案手法が有効に働くと，遠景や海面等の領域の特徴点が減る分，マッチングの成功率が上がり，タイポイントが増えることが期待される．本実験で用いた SfM 解析ソフト Metashape では，キーポイント・タイポイントの上限を設定することができるため，各ケースで同じ数のキーポイントを発生させ，特徴点マッチングに成功したタイポイントの数を評価指標として採用した．

Metashape の解析設定は表-3 の通りである．

表-3 SfM 解析設定

項目	設定値
精度	中
キーポイント上限	40,000
タイポイント上限	2,000

（2）精度指標：カメラ・標定点・検証点座標誤差

精度向上の指標として，カメラ・標定点・検証点の座標の誤差を評価する．

SfM 解析では上記のタイポイント検出（特徴点マッチング）の後に，カメラの内部パラメータ・外部パラメータと特徴点座標を同時に最適化計算により求める．その際，UAV 点検診断システムでは SfM 解析結果の精度を担保するため，地上測量などにより座標値が既知の点を標定点・検証点として手動でタイポイントを作成し，絶対位置の精度検証を行う．標定点座標は SfM 解析時の制約条件として最適化を行い，その結果得られる検証点座標の推定値と実測値の誤差を算出する．SfM 解析の最終的な精度指標としては検証点座標誤差が用いられるが，本研究では最適化計算の過程で得られるカメラ外部パラメータの一つであるカメラ座標，および標定点座標も併せて検証することとした．

特徴点マッチングが改善することにより，タイポイントの数が増えた場合，最適化計算の誤差が小さくなる可能性がある．ただし誤差はタイポイントの数のみに依存するわけではなく，数が少なくても質の高いタイポイントが得られていれば誤差は小さくなる．

なお，本研究で用いた港湾施設の画像データは，Z 座標（高度）として標高値でなく，対地高度と思われる値が記録されており誤差が正しく評価できなかったため，XY 座標の誤差で評価を行った．

5.2 結果と考察

実験結果を表-4 に示す．省力化指標と精度指標について，結果の考察を加える．

表-4 実験結果（**太字**は各評価指標の最良の結果，下線は対象港湾別の最良の結果）

ケース	対象	タイポイント数（↑）	カメラ座標誤差[cm]（↓）			標定点座標誤差[cm]（↓）			検証点座標誤差[cm]（↓）		
			平均	最大	標準偏差	平均	最大	標準偏差	平均	最大	標準偏差
original	A港	55,091	12.498	24.881	5.035	2.672	3.830	1.242	2.617	5.336	1.691
	B港	60,527	156.785	175.182	13.197	0.264	1.066	0.337	0.068	0.682	0.272
	平均	57,809	84.642	100.032	9.116	**1.468**	2.448	0.790	1.342	3.009	0.981
back	A港	61,118	12.509	23.960	4.975	2.689	3.852	1.240	2.657	5.238	1.674
	B港	60,385	156.782	175.188	13.200	0.259	1.022	0.325	0.068	0.686	0.280
	平均	60,752	84.645	**99.574**	9.088	1.474	**2.437**	0.783	1.362	2.962	**0.977**
sea	A港	57,285	11.879	27.467	4.934	3.062	4.163	1.034	2.433	5.125	1.798
	B港	61,813	156.816	175.245	13.184	0.256	1.031	0.331	0.076	0.719	0.264
	平均	59,549	**84.347**	101.356	**9.059**	1.659	2.597	**0.683**	**1.255**	**2.922**	1.031
back+sea	A港	61,683	12.786	26.603	5.175	2.820	3.999	1.222	2.584	5.312	1.746
	B港	61,876	156.822	175.248	13.181	0.259	1.023	0.326	0.075	0.711	0.263
	平均	**61,780**	84.804	100.926	9.178	1.539	2.511	0.774	1.329	3.011	1.004

5.2.1 省力化指標

省力化に関する評価指標であるタイポイント数は，ベースラインの original に比べて点検対象外領域をマスクしたケースの方が増える結果となった．唯一 B 港の back が original に比べてわずかにタイポイントが減っているが，その差は全体のタイポイント数に対して微小である．これは B 港では遠景が映った画像枚数が限られたため，偶発的に当該画像の遠景部分の特徴点がマッチングしていたことによるものと考えた．

全体としては A 港，B 港とも，遠景と海面等の両方にマスクをかけた back+sea が最良の結果を示した．back と sea を比べると，A 港と B 港で傾向が異なる結果となった．これは画像に占める遠景の領域と海面等の領域の多寡によるものと考えた．従って両ケースの優劣は一概には決められず，対象港湾や撮影条件により変わると想定している．また，B 港は A 港に比べて全体的に差が僅少であり，これは元画像を確認したところ遠景や海面等の領域が少なかったことによるものと考えた．

back+sea ケースのタイポイントが最も多くなったことから，想定通り点検対象外領域をマスクすることで特徴点マッチングが成功する割合が増加したことが示された．これにより 2.2 で示した問題が生じる可能性が下がるため，提案手法は省力化に寄与することが示唆された．

5.2.2 精度指標

精度向上に関する評価指標は各ケース間で大きな差は現れず，マスクの有無によってベースラインから精度が向上する明確な結果は得られなかった．最良のケースも指標毎・対象港湾毎に異なる結果となった．一方で，どのケースでも極端に精度が低下するものはなく，B 港のカメラ座標を除いて誤差は各ケースで mm 単位の違いしか出ていない．標定点・検証点の座標は基準点からの TS 観測により計測しているが，基準点座標は RTK-GNSS により座標を計測している．また，カメラ座標も RTK-GNSS により計測している．RTK-GNSS は元々数 cm 程度の誤差があるため，mm 単位の誤差は無視できるレベルと考えられる．なお，B 港のカメラ座標の誤差の値が大きく出ているのは，撮影時に画像に記録されている GPS の誤差によるものと考えている．

以上より，提案手法により点検対象外領域をマスクしてタイポイントが増えても，SfM 精度に大きな影響はないことが示唆される．

6 おわりに

本研究で得られた知見，および今後の課題を以下にまとめる．

6.1 まとめ

本研究は，港湾施設側面の画像を対象に，SfM 解析時の特徴点マッチングを改善することを目的として，海面等領域推定と単眼深度推定に基づき点検対象外領域をマスキングする手法を提案し，港湾の実画像で実験を行った．特徴点マッチング改善により期待される効果として，省力化と精度向上の 2 つの観点から定量評価を行い，以下の知見が得られた．

- 点検対象外となる遠景および海面をマスクすることで特徴点マッチングが改善することを明らかにした
- 改善の度合いは画像内の遠景・海面マスクの占める領域に依存するが，両方をマスクすることで最良の特徴点マッチングが得られた
- 検証点等の座標精度についてはマスクの有無で大きく変わらなかった
- 提案手法は特徴点マッチングを改善することで SfM 解析作業の省力化が期待でき，精度低下も見られないことから，港湾施設側面の点検用画像作成に適用可能である

6.2 今後の課題

本研究では，国土技術政策総合研究所が過去に港湾施設上面部の点検を目的として撮影した画像から，副次的に施設側面部も映っている画像を用いて実験を行った．本来は施設側面の点検目的で撮影を行った画像で検証を行うことが望ましい．特に対象港湾により GPS の誤差が大きく異なったことから，撮影位置の取得条件を厳密に定義して画像を収集すべきである．また，標定点・検証点についても同様の理由により施設上面部のみに設置していたが，施設側面部の点検用画像を作成する場合は側面部，特に施設上面部から離れた水面近くに標定点・検証点を設置すべきである．

次に，提案手法による点検用画像作成の省力化は，本研究では間接的な指標としてタイポイント数で評価を行ったが，より直接的な処理時間や作業時間といった評価は今後の課題である．

また本研究では，海面等領域推定と単眼深度推定により点検対象外領域を正確にマスキングできることを前提として，提案手法の妥当性を示した．しかし，港湾施設の実画像では必ずしも海面や遠景を精度よく推定できるとは限らない．特に海面等領域推定においては，水面近くの濡れたコンクリート面や透明度の高い海面付近の消波ブロックなど，水面上か水面下か人目でも判断が難しい場合がある（図- 7）．このような画像では，海面除去を行ってもタイポイントが増えない事例が確認されている．従って提案手法の検証と並行して，海面等領域推定の精度向上には継続して取り組む必要がある．

図-7 海面等領域推定が難しい例

最後に，提案手法の実用化に向けては，本研究では2港を対象として実験を行ったが，対象港湾を増やして検証を行うことが望ましい．その上で，UAV 点検診断システムに提案手法を組み込み，施設側面の点検にシステムを対応させることが課題である．

参考文献

[1] 日本港湾協会: 数字で見る港湾 2023, 2023.

[2] 里村大樹，辻澤伊吹，山本康太: UAV・AI を活用した港湾等のインフラ維持管理に関する点検診断システム開発（その1），国土技術政策総合研究所資料, No.1135, 2020.

[3] 国土交通省港湾局: 港湾の施設の点検診断ガイドライン, 2021.

[4] 里村大樹: AI による海面等の推定と錆汁・鉄筋露出の検出―港湾の施設の点検診断効率化を目指して―, AI・データサイエンス論文集, vol.3, No.J2, pp. 360-371, 2022.

[5] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam: Encoder-decoder with atrous separable convolution for semantic image segmentation, In European Conference on Computer Vision, 2018

[6] D. Hoiem, A. A. Efros, and M. Hebert: Automatic photo pop-up, ACM Transactions on Graphics, vol. 24, no.3, 2005.

[7] A. Saxena, M. Sun, and A. Y. Ng: Make3D：Learning 3D scene structure from a single still image, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no.5, 2009.

[8] D. Eigen, C. Puhrsch, and R. Fergus: Depth map prediction from a single image using a multi-scale deep network, In Advances in Neural Information Processing Systems, 2014.

[9] T. Zhou, M. Brown, N. Snavely, and D. G. Lowe: Unsupervised learning of depth and ego-motion from video, In Computer Vision and Pattern Recognition, 2017.

[10] R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, and V. Koltun: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[11] L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng, and H. Zhao: Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data, arXiv:2401.10891, 2024.