JAEA-Review
2025-033
DOI:10.11484/jaea-review-2025-033

JAEA-Review

動画像からの特徴量抽出結果に基づいた高速3次元炉内環境モデリング（委託研究）

ー令和5年度　英知を結集した原子力科学技術・人材育成推進事業ー

High-speed 3D Modeling for Nuclear Reactor Environment Based on Feature Extraction Results from Video Images
(Contract Research)
– FY2023 Nuclear Energy Science & Technology and Human Resource Development Project –

福島廃炉安全工学研究所　廃炉環境国際共同研究センター
札幌大学

Collaborative Laboratories for Advanced Decommissioning Science,
Fukushima Research and Engineering Institute
Sapporo University

November 2025
Japan Atomic Energy Agency
日本原子力研究開発機構

国立研究開発法人日本原子力研究開発機構　研究開発推進部　科学技術情報課
〒 319-1112　茨城県那珂郡東海村大字村松 4 番地 49
E-mail: ird-support@jaea.go.jp

JAEA-Review 2025-033

動画像からの特徴量抽出結果に基づいた高速 3 次元炉内環境モデリング
（委託研究）
－令和 5 年度 英知を結集した原子力科学技術・人材育成推進事業－

日本原子力研究開発機構　福島廃炉安全工学研究所
廃炉環境国際共同研究センター

札幌大学

（2025 年 6 月 9 日受理）

日本原子力研究開発機構（JAEA）廃炉環境国際共同研究センター（CLADS）では、令和 5 年度 英知を結集した原子力科学技術・人材育成推進事業（以下、「本事業」という。）を実施している。

本事業は、東京電力ホールディングス株式会社福島第一原子力発電所（1F）の廃炉等をはじめとした原子力分野の課題解決に貢献するため、国内外の英知を結集し、様々な分野の知見や経験を従前の機関や分野の壁を越えて緊密に融合・連携させた基礎的・基盤的研究及び人材育成を推進することを目的としている。

平成 30 年度の新規採択課題から実施主体を文部科学省から JAEA に移行することで、JAEA とアカデミアとの連携を強化し、廃炉に資する中長期的な研究開発・人材育成をより安定的かつ継続的に実施する体制を構築した。

本研究は、令和 5 年度に採択された研究課題のうち「動画像からの特徴量抽出結果に基づいた高速 3 次元炉内環境モデリング」の令和 5 年度分の研究成果について取りまとめたものである。

本研究は、1F 廃炉に向けて、原子炉格納容器及び原子炉建屋内を調査する際に撮影した動画像を入力し、指定された時間、動画像から抽出された特徴量に応じて、周辺情報を補強した上で情報量が大きい立体復元手法を選択し、作業空間を 3 次元モデリングする研究開発を行う。

令和 5 年度は、写真測量、深層学習に基づいた立体復元手法による解析から、良好な立体復元を得るための有効な撮影条件を抽出する手法及び少ないデータから指定された時間までに立体復元結果を生成できるように特徴量を抽出する手法の検証を行った。さらに、動画像から抽出された点群データをセグメンテーションに適用し、インスタンスラベルが付された部品に分類した。

本報告書は、日本原子力研究開発機構の英知事業における委託業務として、札幌大学が実施した成果を取りまとめたものである。

廃炉環境国際共同研究センター：〒979-1151 福島県双葉郡富岡町大字本岡字王塚 790-1

JAEA-Review 2025-033

High-speed 3D Modeling for Nuclear Reactor Environment Based on Feature Extraction Results from Video Images
(Contract Research)
－ FY2023 Nuclear Energy Science & Technology and Human Resource Development Project －

Collaborative Laboratories for Advanced Decommissioning Science,
Fukushima Research and Engineering Institute
Japan Atomic Energy Agency
Tomioka-machi, Futaba-gun, Fukushima-ken

Sapporo University

(Received June 9, 2025)

The Collaborative Laboratories for Advanced Decommissioning Science (CLADS), Japan Atomic Energy Agency (JAEA), had been conducting the Nuclear Energy Science & Technology and Human Resource Development Project (hereafter referred to “the Project”) in FY2023.

The Project aims to contribute to solving problems in the nuclear energy field represented by the decommissioning of the Fukushima Daiichi Nuclear Power Station (1F), Tokyo Electric Power Company Holdings, Inc. (TEPCO). For this purpose, intelligence was collected from all over the world, and basic research and human resource development were promoted by closely integrating/collaborating knowledge and experiences in various fields beyond the barrier of conventional organizations and research fields.

The sponsor of the Project was moved from the Ministry of Education, Culture, Sports, Science and Technology to JAEA since the newly adopted proposals in FY2018. On this occasion, JAEA constructed a new research system where JAEA-academia collaboration is reinforced and medium-to-long term research/development and human resource development contributing to the decommissioning are stably and consecutively implemented.

Among the adopted proposals in FY2023, this report summarizes the research results of the “High-speed 3D modeling for nuclear reactor environment based on feature extraction results from video images” conducted in FY2023.

The present study aims to develop a 3D model for a workspace that maximizes the amount of information based on the features extracted from video, which is taken when surveying the primary containment vessel and inside the reactor building as part of the decommissioning of 1F, considering within a specified time.

In FY2023, we verified extracting effective shooting conditions for obtaining 3D reconstruction based on photogrammetry and the method extracting feature values that can generate 3D restoration results from a small amount of data within a specified time based on deep learning. In addition, we applied point cloud data extracted from video to segmentation and classified it into parts with instance labels.

Keywords: 3D Modeling, Structure from Motion, Multi-view Stereo, Image Processing, Deep Learning, Segmentation, Generative Adversarial Networks

This work was performed by Sapporo University under contract with Japan Atomic Energy Agency.

目次

Contents

This is a blank page.

1. 英知を結集した原子力科学技術・人材育成推進事業の概要

文部科学省では、「東京電力（株）福島第一原子力発電所の廃止措置等研究開発の加速プラン（平成 26 年 6 月文部科学省）」等を踏まえ、平成 27 年度から「英知を結集した原子力科学技術・人材育成推進事業」（以下、「本事業」という。）を立ち上げ、「戦略的原子力共同研究プログラム」、「廃炉加速化研究プログラム」及び「廃止措置研究・人材育成等強化プログラム」を推進している。

具体的には、国内外の英知を結集し、国内の原子力分野のみならず様々な分野の知見や経験を、機関や分野の壁を越え、国際共同研究も含めて緊密に融合・連携させることにより、原子力の課題解決に資する基礎的・基盤的研究や産学が連携した人材育成の取組を推進している。

一方、日本原子力研究開発機構（以下、「JAEA」という。）では、平成 27 年に廃炉国際共同研究センター（以下、「CLADS」という。現：廃炉環境国際共同研究センター）を組織し、「東京電力ホールディングス（株）福島第一原子力発電所の廃止措置等に向けた中長期ロードマップ」等を踏まえ、東京電力ホールディングス株式会社福島第一原子力発電所廃炉（以下、「1F 廃炉」という。）に係る研究開発を進めている。

また、平成 29 年 4 月に CLADS の中核拠点である「国際共同研究棟」の運用を開始したことを踏まえ、今後は CLADS を中核に、廃炉の現場ニーズを踏まえた国内外の大学、研究機関等との基礎的・基盤的な研究開発及び人材育成の取組を推進することにより、廃炉研究拠点の形成を目指すことが期待されている。

このため、本事業では平成 30 年度の新規採択課題から実施主体を文部科学省から JAEA に移行することで、JAEA とアカデミアとの連携を強化し、廃炉に資する中長期的な研究開発・人材育成をより安定的かつ継続的に実施する体制を構築することとし、従来のプログラムを、①共通基盤型原子力研究プログラム、②課題解決型廃炉研究プログラム、③国際協力型廃炉研究プログラム、④研究人材育成型廃炉研究プログラム（令和元年度より新設）に再編した。

2. 平成 30 年度 採択課題

平成 30 年度採択課題については以下のとおりである。

課題数：19 課題

共通基盤型原子力研究プログラム	11 課題（若手研究 6 課題、一般研究 5 課題）
課題解決型廃炉研究プログラム	6 課題
国際協力型廃炉研究プログラム	2 課題（日英）

平成 30 年度 採択課題一覧

共通基盤型原子力研究プログラム

【若手研究】

課題名	研究代表者	所属機関
被災地探査や原子力発電所建屋内情報収集のための半自律ロボットを用いたセマンティックサーベイマップ生成システムの開発	河野 仁	東京工芸大学
汚染土壌の減容を目的とした重液分離による放射性微粒子回収法の高度化	山﨑 信哉	筑波大学
ラドンを代表としたアルファ核種の吸入による内部被ばくの横断的生体影響評価	片岡 隆浩	岡山大学
炉心溶融物の粘性及び表面張力同時測定技術の開発	大石 佑治	大阪大学
iPS 細胞由来組織細胞における放射線依存的突然変異計測系の確立	島田 幹男	東京工業大学
レーザー共鳴イオン化を用いた同位体存在度の低いストロンチウム 90 の迅速分析技術開発	岩田 圭弘	東京大学

共通基盤型原子力研究プログラム

【一般研究】

課題名	研究代表者	所属機関
放射性核種の長期安定化を指向した使用済みゼオライト焼結固化技術の開発	新井 剛	芝浦工業大学
燃料デブリ取り出しを容易にするゲル状充填材の開発	牟田 浩明	大阪大学
レーザー蛍光法を用いた燃料デブリ変質相の同定	斉藤 拓巳	東京大学
過酷炉心放射線環境における線量測定装置の開発	岡本 保	木更津工業高等専門学校
レーザー加工により発生する微粒子の解析と核種同定手法の開発	長谷川 秀一	東京大学

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
合金相を含む燃料デブリの安定性評価のための基盤研究	桐島 陽	東北大学
ガンマ線画像スペクトル分光法による高放射線場環境の画像化による定量的放射能分布解析法	谷森 達	京都大学
燃料デブリ取出し時における放射性核種飛散防止技術の開発	鈴木 俊一	東京大学
アルファダストの検出を目指した超高位置分解能イメージング装置の開発	黒澤 俊介	東北大学
ナノ粒子を用いた透明遮へい材の開発研究	渡邉 隆行	九州大学
先端計測技術の融合で実現する高耐放射線燃料デブリセンサーの研究開発	萩原 雅之	高エネルギー加速器研究機構

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
放射性微粒子の基礎物性解明による廃炉作業リスク低減への貢献	五十嵐 康人	茨城大学
放射線耐性の高い薄型 SiC 中性子検出器の開発	三澤 毅	京都大学

3. 令和元年度 採択課題

令和元年度採択課題については以下のとおりである。

課題数：19 課題

共通基盤型原子力研究プログラム　7 課題（若手研究 2 課題、一般研究 5 課題）
課題解決型廃炉研究プログラム　4 課題
国際協力型廃炉研究プログラム　4 課題（日英 2 課題、日露 2 課題）
研究人材育成型廃炉研究プログラム 4 課題

令和元年度 採択課題一覧

共通基盤型原子力研究プログラム

【若手研究】

課題名	研究代表者	所属機関
ウラニル錯体化学に基づくテーラーメイド型新規海水ウラン吸着材開発	鷹尾 康一朗	東京工業大学
動作不能からの復帰を可能とする多連結移動ロボットの半自律遠隔操作技術の確立	田中 基康	電気通信大学

共通基盤型原子力研究プログラム

【一般研究】

課題名	研究代表者	所属機関
一次元光ファイバ放射線センサを用いた原子炉建屋内放射線源分布計測	瓜谷 章	名古屋大学
低線量・低線量率放射線被ばくによる臓器別酸化ストレス状態の検討	鈴木 正敏	東北大学
単一微粒子質量分析法に基づくアルファ微粒子オンラインモニタリングに向けた基礎検討	豊嶋 厚史	大阪大学
幹細胞動態により放射線発がんを特徴付ける新たな評価系の構築	飯塚 大輔	量子科学技術研究開発機構
耐放射線性ダイヤモンド半導体撮像素子の開発	梅沢 仁 (～R2.3.31) 大曲 新矢 (R2.4.1～)	産業技術総合研究所

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
Multi-Physics モデリングによる福島 2・3 号機ペデスタル燃料デブリ深さ方向の性状同定	山路 哲史	早稲田大学
燃料デブリ取出しに伴い発生する廃棄物のフッ化技術を用いた分別方法の研究開発	渡邉 大輔	日立ＧＥニュークリア・エナジー
アパタイトセラミックスによる ALPS 沈殿系廃棄物の安定固化技術の開発	竹下 健二 (～R3.6.30) 塚原 剛彦 (R3.7.1～)	東京工業大学
拡張型スーパードラゴン多関節ロボットアームによる圧力容器内燃料デブリ調査への挑戦	高橋 秀治	東京工業大学

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
高い流動性および陰イオン核種保持性を有するアルカリ刺激材料の探索と様々な放射性廃棄物の安全で効果的な固化	佐藤 努	北海道大学
再臨界前の中性子線増に即応可能な耐放射線FPGAシステムの開発	渡邊 実	静岡大学 (～R3.3.31) 岡山大学 (R3.4.1～)

国際協力型廃炉研究プログラム（日露共同研究）

課題名	研究代表者	所属機関
燃料デブリ取出し臨界安全技術の高度化	小原 徹	東京工業大学
微生物生態系による原子炉内物体の腐食・変質に関する評価研究	金井 昭夫	慶應義塾大学

研究人材育成型廃炉研究プログラム

課題名	研究代表者	所属機関
燃料デブリ取り出し時における炉内状況把握のための遠隔技術に関する研究人材育成	淺間 一	東京大学
化学計測技術とインフォマティックスを融合したデブリ性状把握手法の開発とタイアップ型人材育成	高貝 慶隆	福島大学
放射線・化学・生物的作用の複合効果による燃料デブリ劣化機構の解明	大貫 敏彦 (～R2.3.31) 竹下 健二 (～R3.6.30) 塚原 剛彦 (R3.7.1～)	東京工業大学
燃料デブリ分析のための超微量分析技術の開発	永井 康介	東北大学

4. 令和 2 年度 採択課題

令和 2 年度は、2 つのプログラムにおいて、研究課題の採択を決定した。
公募の概要は以下のとおりである。

公募期間：令和 2 年 3 月 17 日～令和 2 年 5 月 14 日（課題解決型）
令和 2 年 5 月 13 日～令和 2 年 7 月 15 日（国際協力型）

課題数：10 課題
課題解決型廃炉研究プログラム　8 課題（若手研究 2 課題、一般研究 6 課題）
国際協力型廃炉研究プログラム　2 課題（日英）

令和 2 年度 採択課題一覧

課題解決型廃炉研究プログラム
【若手研究】

課題名	研究代表者	所属機関
燃料デブリにおける特性の経年変化と環境劣化割れの調査	楊 会龍 (～R4.7.31) 村上 健太 (R4.8.1～)	東京大学
健全性崩壊をもたらす微生物による視認不可腐食の分子生物・電気化学的診断及び抑制技術の開発	岡本 章玄	物質・材料研究機構

課題解決型廃炉研究プログラム

【一般研究】

課題名	研究代表者	所属機関
遮蔽不要な臨界近接監視システム用ダイヤモンド中性子検出器の要素技術開発	田中 真伸	高エネルギー加速器研究機構
α／β／γ 線ラジオリシス影響下における格納容器系統内広域防食の実現:ナノバブルを用いた新規防食技術の開発	渡邉 豊	東北大学
β、γ、X 線同時解析による迅速・高感度放射性核種分析法の開発	篠原 宏文	日本分析センター
合理的な処分のための実機環境を考慮した汚染鉄筋コンクリート長期状態変化の定量評価	丸山 一平	東京大学
溶脱による変質を考慮した汚染コンクリート廃棄物の合理的処理・処分の検討	小崎 完	北海道大学
マイクロ波重畳 LIBS によるデブリ組成計測の高度化と同位体の直接計測への挑戦	池田 裕二	アイラボ

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
革新的水質浄化剤の開発による環境問題低減化技術の開拓	浅尾 直樹	信州大学
無人航走体を用いた燃料デブリサンプルリターン技術の研究開発	鎌田 創	海上・港湾・航空技術研究所

5. 令和 3 年度 採択課題

令和 3 年度は、2 つのプログラムにおいて、研究課題の採択を決定した。
公募の概要は以下のとおりである。

公募期間：令和 3 年 3 月 16 日～令和 3 年 5 月 17 日（課題解決型）
令和 3 年 4 月 13 日～令和 3 年 7 月 1 日（国際協力型 日英共同研究）
令和 3 年 7 月 12 日～令和 3 年 8 月 18 日（国際協力型 日露共同研究）

課題数：12 課題
課題解決型廃炉研究プログラム 8 課題
国際協力型廃炉研究プログラム 2 課題（日英）、2 課題（日露）

令和 3 年度 採択課題一覧

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
建屋応答モニタリングと損傷イメージング技術を活用したハイブリッド型の原子炉建屋長期健全性評価法の開発研究	前田 匡樹	東北大学
燃料デブリ周辺物質の分析結果に基づく模擬デブリの合成による実機デブリ形成メカニズムの解明と事故進展解析結果の検証によるデブリ特性データベースの高度化	宇埜 正美	福井大学
ジオポリマー等による PCV 下部の止水・補修及び安定化に関する研究	鈴木 俊一	東京大学
世界初の同位体分析装置による少量燃料デブリの性状把握分析手法の確立	坂本 哲夫	工学院大学
アルファ微粒子の実測に向けた単一微粒子質量分析法の高度化	豊嶋 厚史	大阪大学
連携計測による線源探査ロボットシステムの開発研究	人見 啓太朗	東北大学
中赤外レーザー分光によるトリチウム水連続モニタリング手法の開発	安原 亮	自然科学研究機構

課題名	研究代表者	所属機関
福島原子力発電所事故由来の難固定核種の新規ハイブリッド固化への挑戦と合理的な処分概念の構築・安全評価	中瀬 正彦	東京工業大学

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
福島第一原子力発電所の廃止措置における放射性エアロゾル制御及び除染に関する研究	Erkan Nejdet (～R4.1.31) 三輪 修一郎 (R4.2.1～)	東京大学
燃料デブリ取り出しのための機械式マニピュレータのナビゲーションおよび制御	淺間 一	東京大学

国際協力型廃炉研究プログラム（日露共同研究）

課題名	研究代表者	所属機関
福島第一発電所 2、3 号機の事故進展シナリオに基づく FP・デブリ挙動の不確かさ低減と炉内汚染状況・デブリ性状の把握	小林 能直	東京工業大学
非接触測定法を用いた燃料デブリ臨界解析技術の高度化	小原 徹	東京工業大学

6. 令和 4 年度 採択課題

令和 4 年度は、2 つのプログラムにおいて、研究課題の採択を決定した。
公募の概要は以下のとおりである。

公募期間：令和 4 年 3 月 1 日～令和 4 年 5 月 6 日（課題解決型）
令和 4 年 4 月 7 日～令和 4 年 6 月 16 日（国際協力型 日英共同研究）

課題数：8 課題
課題解決型廃炉研究プログラム　6 課題
国際協力型廃炉研究プログラム　2 課題（日英）

令和 4 年度 採択課題一覧

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
革新的アルファダスト撮像装置と高線量率場モニタの実用化とその応用	黒澤 俊介	東北大学
3 次元線量拡散予測法の確立と γ 線透過率差を利用した構造体内調査法の開発	谷森 達	京都大学
α汚染可視化ハンドフットクロスモニタの要素技術開発	樋口 幹雄	北海道大学
高放射線耐性の低照度用太陽電池を利用した放射線場マッピング観測システム開発	奥野 泰希	京都大学 (～R5.3.31) 理化学研究所 (R5.4.1～)
障害物等による劣悪環境下でも通信可能なパッシブ無線通信方式の開発	新井 宏之	横浜国立大学
無線 UWB とカメラ画像分析を組合せたリアルタイム 3D 位置測位・組込システムの開発・評価	松下 光次郎	岐阜大学

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
耐放射線プロセッサを用いた組み込みシステムの開発	渡邊 実	岡山大学
マイクロ・ナノテクノロジーを利用したアルファ微粒子の溶解・凝集分散に及ぼすナノ界面現象の探求	塚原 剛彦	東京工業大学

7. 令和 5 年度 採択課題

令和 5 年度は、2 つのプログラムにおいて、研究課題の採択を決定した。
公募の概要は以下のとおりである。

公募期間：令和 5 年 3 月 1 日～令和 5 年 4 月 14 日（課題解決型）
令和 5 年 4 月 12 日～令和 5 年 6 月 15 日（国際協力型 日英共同研究）

課題数：9 課題
課題解決型廃炉研究プログラム 7 課題
国際協力型廃炉研究プログラム 2 課題（日英）

これらの提案について、外部有識者から構成される審査委員会において、書面審査及び面接審査、日英共同研究については二国間の合同審査を実施し、採択候補課題を選定した。
その後、PD（プログラムディレクター）・PO（プログラムオフィサー）会議及びステアリングコミッティでの審議を経て、採択課題を決定した。

令和 5 年度 採択課題一覧

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
遮蔽不要な耐放射線性ダイヤモンド中性子計測システムのプロトタイプ開発	金子 純一	北海道大学
簡易非破壊測定に向けた革新的な n・γ シンチレーション検出システムの開発	鎌田 圭	東北大学
ペデスタル部鉄筋コンクリート損傷挙動の把握に向けた構成材料の物理・化学的変質に関する研究	五十嵐 豪	名古屋大学
動画像からの特徴量抽出結果に基づいた高速 3 次元炉内環境モデリング	中村 啓太	札幌大学
放射性コンクリート廃棄物の減容を考慮した合理的処理・処分方法の検討	小崎 完	北海道大学

課題名	研究代表者	所属機関
高バックグラウンド放射線環境における配管内探査技術の開発	鳥居 建男	福井大学
PCV 気相漏洩位置及び漏洩量推定のための遠隔光計測技術の研究開発	椎名 達雄	千葉大学

国際協力型廃炉研究プログラム（日英共同研究）

課題名	研究代表者	所属機関
革新的分光画像解析による燃料デブリの可視化への挑戦と LIBS による検証	牟田 浩明	大阪大学
燃料デブリ除去に向けた様々な特性をもつメタカオリンベースのジオポリマーの設計と特性評価	Yogarajah Elakneswaran	北海道大学

本報告書は、以下の課題の令和 5 年度分の研究成果について取りまとめたものである。

課題解決型廃炉研究プログラム

課題名	研究代表者	所属機関
動画像からの特徴量抽出結果に基づいた高速 3 次元炉内環境モデリング	中村 啓太	札幌大学

研究成果を取りまとめた成果報告書を付録として添付する。

付録

成果報告書

This is a blank page.

令和 5 年度

日本原子力研究開発機構

英知を結集した原子力科学技術・人材育成推進事業

動画像からの特徴量抽出結果に基づいた 高速 3 次元炉内環境モデリング

（契約番号　R05I106）

成果報告書

令和 6 年 3 月

学校法人札幌大学

本報告書は、国立研究開発法人日本原子力研究開発機構の「英知を結集した原子力科学技術･人材育成推進事業」による委託業務として、学校法人札幌大学が実施した「動画像からの特徴量抽出結果に基づいた高速３次元炉内環境モデリング」の令和５年度分の研究成果を取りまとめたものである。

目次

執筆者リスト

研究代表者
学校法人札幌大学　准教授　中村　啓太

再委託先
公立大学法人岩手県立大学　教授　間所　洋和

連携先
国立研究開発法人日本原子力研究開発機構
廃炉環境国際共同研究センター
遠隔技術ディビジョン
空間情報応用制御研究グループ マネージャー　羽成　敏秀

表一覧

図一覧

略語一覧

PCV	: Primary Containment Vessel	(原子炉格納容器)
1F	: 東京電力ホールディングス株式会社福島第一原子力発電所	
SfM	: Structure from Motion	(複数画像からカメラの位置姿勢と対象の座標を取得する技術)
CLADS	: Collaborative Laboratories for Advanced Decommissioning Science	(廃炉環境国際共同研究センター)
NDF	: Nuclear Damage Compensation and Decommissioning Facilitation Corporation	(原子力損害賠償・廃炉等支援機構)
JAEA	: Japan Atomic Energy Agency	(国立研究開発法人日本原子力研究開発機構)
AI	: Artificial Intelligence	(人工知能)
MVS	: Multi-View Stereo	(多視点ステレオ)
3DCG	: 3D Computer Graphics	(3 次元コンピュータグラフィックス)
API	: Application Programming Interface	(ソフトウェアやプログラムの間をつなぐインターフェース)
ICP	: Iterative Closest Point	(2 点群を整合させるアルゴリズムの 1 つ)
ROS	: Robot Operating System	(ロボット用のソフトウェアプラットフォーム)
東電 HD	: 東京電力ホールディングス株式会社	
HP	: HomePage	(ホームページ)
ROV	: Remotely Operated Vehicle	(遠隔操作型の無人潜水機)
IQA	: Image Quality Assessment	(画像の品質を高精度に計測する方法の 1 つ)
PIQE	: Perception based Image Quality Evaluator	(統計的な特徴量を使用して画質を評価する手法の 1 つ)
SIFT	: Scale Invariant Feature Transform	(スケール不変特徴量変換)
ORB	: Oriented fast and Rotated BRIEF	(画像における特徴点、特徴量を抽出するアルゴリズムの 1 つ)
CLAHE	: Contrast Limited Adaptive Histogram Equalization	(適用的ヒストグラム平坦化)
GPU	: Graphics Processing Unit	(画像処理装置)
SAM	: Segment Anything Model	(セグメンテーションのための基盤モデル)
BB	: Bounding Box	(バウンディングボックス)
GT	: Ground Truth	(AI モデルの出力の学習やテストに使用される実際のデータ)
ConvNets	: Convolutional Neural Networks	(畳込みニューラルネットワーク)
MLP	: Multi-Layer Perceptron	(多層パーセプトロン)
IoU	: Intersection over Union	(物体検出における評価指標の 1 つ)

CUDA	: Compute Unified Device Architecture	（NVIDIA が開発・提供している GPU 向けの汎用並列コンピューティングプラットフォーム）
NeRF	: Neural Radiance Fields	（さまざまな角度から撮影した複数の写真から自由視点画像を生成する技術）
LiDAR	: Light Detection And Ranging	（光検出と測距）
PSNR	: Peak Signal-to-Noise Ratio	（ピーク信号対雑音比）
札幌大学	: 学校法人札幌大学	
岩手県立大学	: 公立大学法人岩手県立大学	

概略

東京電力ホールディングス株式会社福島第一原子力発電所（以下、「1F」と略す。）の廃炉に向けて、原子炉格納容器（以下、「PCV」と略す。）および原子炉建屋内を調査する際に撮影した動画像を入力とし、指定された時間、動画像から抽出された特徴量に応じて、周辺情報を補強したうえで情報量が大きい立体復元手法を選択し、作業空間を 3 次元モデリングする研究開発を行う。研究課題は、(1)シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係、(2)動画像からの迅速な 3 次元モデリング手法の研究開発、(3)深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング、という 3 項目で構成される。各研究課題を通じて、動画像からの迅速な 3 次元モデリング手法を導入したシステム構築を目指し、炉内および建屋内の差分を検出しつつ現場状況を把握することに貢献する。

以下に、3 カ年計画の 1 年目である令和 5 年度の業務実績を述べる。

(1) シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係

実際の PCV を模した暗所環境を構築し、光の反射、レンズの焦点距離、撮影時のカメラ振動を考慮したシミュレーションを行い、写真測量用の人工画像の生成を行った。そして、生成した人工画像を入力した写真測量を行うことで、写真測量対象環境に適した光源の強さ、撮影時の焦点距離があることを確認した。また、撮影時のカメラ振動によって、獲得する写真測量結果の変化も同時に調査した。さらに、アーム付き移動ロボットおよびアクションカメラを導入することで、さまざまな姿勢で撮影できるカメラシステムの検討を行った。

(2) 動画像からの迅速な 3 次元モデリング手法の研究開発（連携先：JAEA）

1F 調査画像の特性分析を行い、画像品質評価手法の導入により画像補正による画質スコアの改善と画像特徴点のマッチング数の向上に関連性があることを示した。また、深層学習を適用した画像特徴点を抽出する手法の調査・分析を行い、深層学習の適用により従来手法と比べて良好な結果が得られることを確認した。そして、深層学習の適用によって抽出された画像特徴点をもとにした立体復元結果の復元精度の評価および画像特徴点抽出数・マッチング数の増加に伴う計算時間の増加を今後の課題として抽出した。さらに、イメージスティッチングの適用性検証の結果、カメラ軌道および投影モデルの選択がパノラマ画像合成の性能に与える影響を確認し、球面投影型モデル、アフィン投影型モデル以外の合成方法についても実験的に検証していくこととした。

(3) 深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング（再委託先：岩手県立大学）

動画像データから Structure from Motion（SfM）を用いて抽出された点群データを、パノプティックパーツセグメンテーションにより、インスタンスラベルが付された部品に分類した。セグメンテーションに使用するバックボーンを、性能、速度、メモリ使用量、拡張性、実装の容易さなど、複数の要素を考慮しつつ選定した。ハードウェアの性能や制約条件に合わせた最適化を施し、処理速度と精度のトレードオフ関係を最適化した。

(4) 研究推進

研究代表者の下で各研究項目間ならびに廃炉環境国際共同研究センター（CLADS）等との連携を密にして研究を進めた。また、研究実施計画を推進するための打合せや会議等を開催した。

1. はじめに

1F 1-3 号機の PCV 内部には大量の燃料デブリが存在しており、廃炉に向けて燃料デブリを取り出す準備が進められている。原子力損害賠償・廃炉等支援機構（NDF）の「東京電力ホールディングス（株）福島第一原子力発電所の廃炉のための技術戦略プラン 2022」によると、令和 5 年度後半に 2 号機の試験的取り出しに着手し、その後段階的な取り出し規模の拡大等の一連の作業を進める予定となっている。燃料デブリの取り出し戦略として、取り出し方法の検討において直接的な映像調査が必要とされている。作業計画の策定やオペレータが作業空間を適切に認知するために、作業空間の情報を 3 次元かつ俯瞰的に提示できることが望ましい。

そこで本研究では、調査により得られた動画像から作業空間の 3 次元モデルを生成することを提案する。指定された時間、動画像から抽出された特徴量に応じて、より情報量が大きい立体復元手法を選択し、作業空間の 3 次元モデリングを行う。図 1.1-1 に本申請研究で提案する 3 次元モデリング手法の概念図を示す。

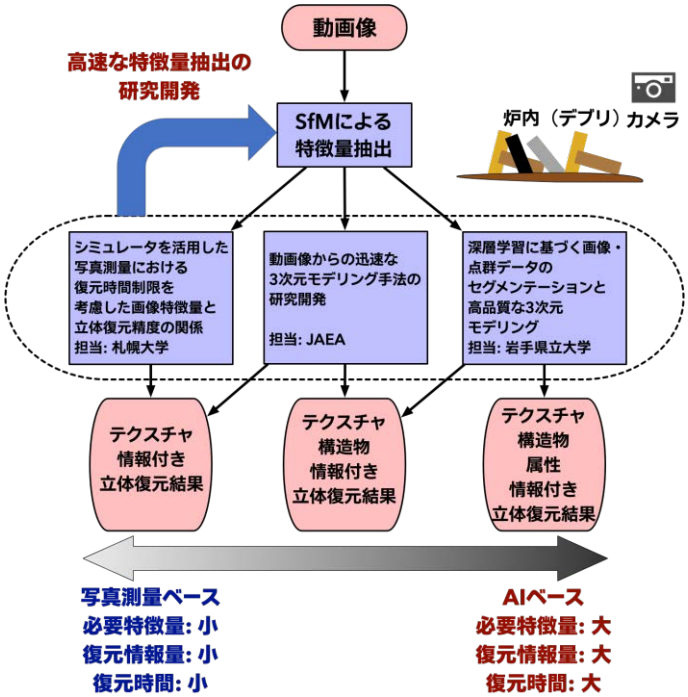


図 1.1-1　本申請研究が提案する 3 次元モデリング手法の概念図

本研究は、以下の実施項目から構成される。

(1) シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係

(2) 動画像からの迅速な 3 次元モデリング手法の研究開発

(3) 深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング

(1) については札幌大学において実施、(2) については JAEA において実施、(3) については岩手県立大学で実施する。

また、プログラムでは、上記の研究開発を通じて各研究機関で適宜相互連携しながら、動画像からの迅速な 3 次元モデリング手法を写真測量、シミュレーションおよび人工知能（AI）技術の側面から研究開発する。最終的に各研究結果を統合し『指定した時間、抽出された特徴量に応じて、より情報量が大きい立体復元結果を自動的に生成するプロトタイプシステム』を構築することを目指す。

2. 業務計画

2.1 全体計画

1F の廃炉に向けて、指定された時間、動画像から抽出された特徴量に応じて、周辺情報を補強したうえで情報量が大きい立体復元手法を選択し、作業空間を 3 次元モデリングする研究開発を行う。研究課題は、(1)シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係、(2)動画像からの迅速な 3 次元モデリング手法の研究開発、(3)深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング、という 3 項目で構成される。年度別全体計画を図 2.1-1 に示す。

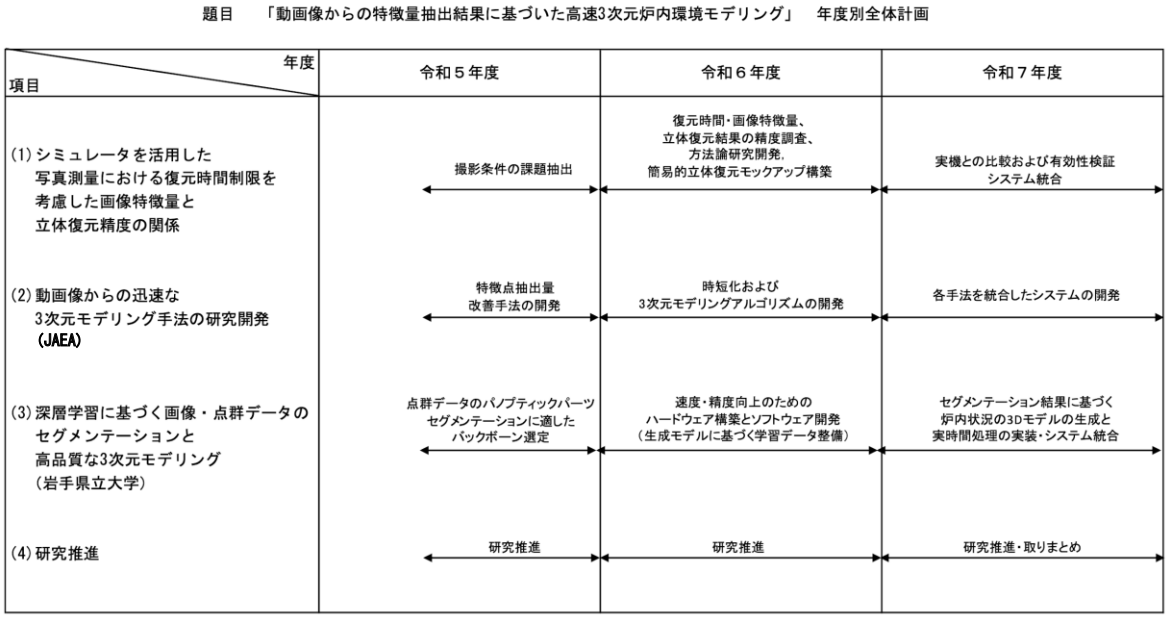


図 2.1-1　年度別全体計画

2.2　実施体制

実施体制を図 2.2-1 に示す。

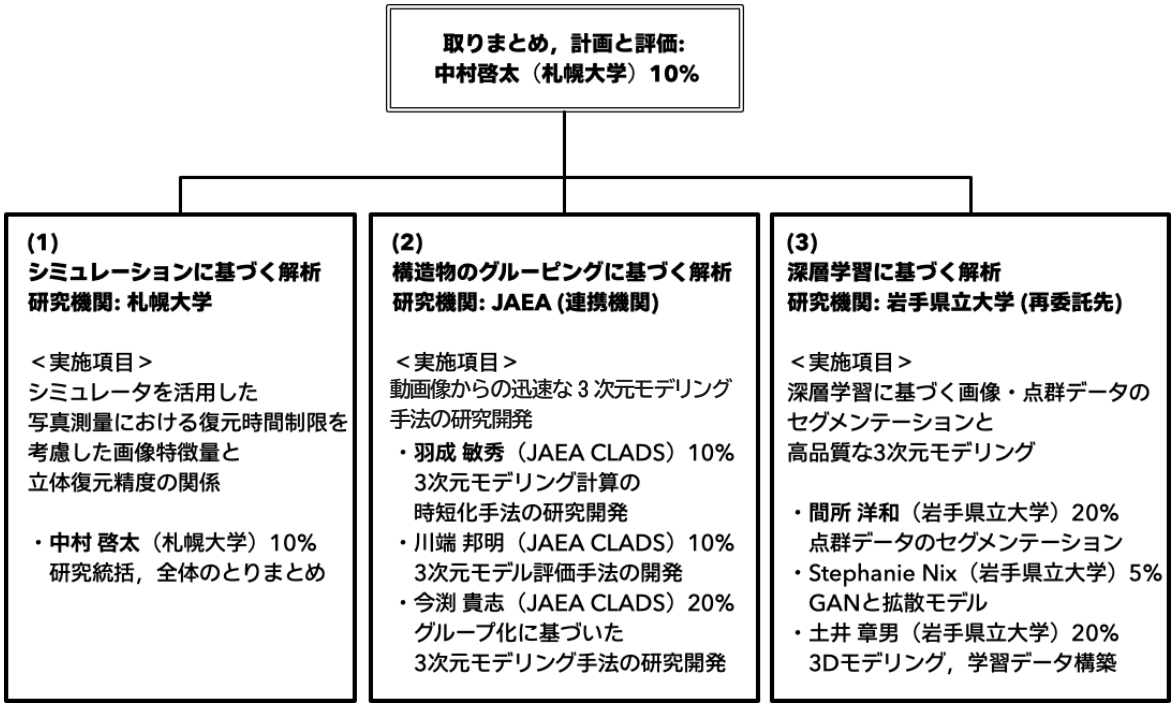


図 2.2-1　実施体制図

2.3　令和 5 年度の成果の目標および業務の実施方法

2.3.1　シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係（札幌大学）

シミュレータを活用し、廃炉作業のような画像取得に制限がある環境を考慮した写真測量による画像からの立体復元における撮像条件の課題を抽出する。特に、画像特徴量に影響を与える光の反射、レンズの焦点距離、撮影時のカメラ振動に注目し、これらの要素が変化した際の写真測量による画像間特徴量および立体復元結果の精度の関係を明らかにする。また、シミュレーションで得られた結果と実際のカメラで撮影した画像による立体復元結果を比較するため、さまざまな姿勢で撮影できるカメラシステムの構築を行う。

2.3.2　動画像からの迅速な 3 次元モデリング手法の研究開発（連携先：JAEA）

最新の PCV 調査の動画像データを対象として、画像の特性分析および深層学習を用いた手法等を複合的に適用した画像特徴点抽出手法等の開発を実施する。これらの手法を組み合わせて、構造物のグループ化に基づく 3 次元モデリングアルゴリズムによる環境情報の提示、立体復元計算の時短化アルゴリズムによる高速化等の令和 6 年度に必要な研究開発の課題について抽出を行う。併せて、イメージスティッチングによる高解像度画像の生成手法および画像間のオプティカルフローに基づく遠隔操作機器の移動・停止の自動記録生成手法の適用性検証を実施する。

2.3.3　深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング（再委託先：岩手県立大学）

動画像データから SfM を用いて抽出された点群データを、パノプティックパーツセグメンテーションにより、インスタンスラベルが付された部品に分類する。セグメンテーションに使用するバックボーンを性能、速度、メモリ使用量、拡張性、実装の容易さなど、複数の要素を考慮しつつ選定する。ハードウェアの性能や制約条件に合わせた最適化を施し、処理速度と精度のトレードオフ関係を最適化する。

2.3.4　研究推進

研究代表者の下で各研究項目間ならびに CLADS 等との連携を密にして研究を進める。また、研究実施計画を推進するための打合せや会議等を開催する。

3. 実施内容および成果

3.1 シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係（札幌大学）

(1) 令和 5 年度概要

令和 5 年度では、シミュレータを活用し、廃炉作業のような画像取得に制限がある環境を考慮した写真測量による画像からの立体復元における撮像条件の課題を抽出した。特に、画像特徴量に影響を与える光の反射、レンズの焦点距離、撮影時のカメラ振動に注目し、これらの要素が変化した際の写真測量による画像間特徴量および立体復元結果の精度の関係を明らかにした。また、シミュレーションで得られた結果と実際のカメラで撮影した画像による立体復元結果を比較するため、さまざまな姿勢で撮影できるカメラシステムの構築を行った。(2)以降にそれぞれの実施内容および結果について述べる。

(2) 令和 5 年度実施内容および成果

撮像条件の変化によって、SfM[3.1-1]と Multi-View Stereo (MVS) [3.1-2]による写真測量手法で獲得される立体復元結果は容易に変化する。つまり、撮影機会が少ない廃炉活動において、撮影条件が SfM-MVS による立体復元にどのような影響を与えるかを事前に調査し、立体復元結果を予測することが非常に重要である。廃炉活動を行うために PCV 内部のような暗所において、立体復元性能に影響を及ぼすと考えられる照明の強さ（消費電力）および焦点距離の違いによって獲得される立体復元結果を調査することは重要であり、実現場でどのように画像を取得するかを計画することにつながる。しかしながら、調査を行うためのモックアップの設置や調整には、環境の整備や各種カメラ・ 照明などの機材の手配など、高いコストを要する。そこで、3DCG ソフトである Blender[3.1-3]を使用して、シミュレーションで構築した仮想環境から立体復元検証用画像データを生成し、立体復元に有効な撮影条件を検証する。この Blender には Application Programming Interface(API)が用意されており、プログラミング言語 Python で指定したカメラおよびライティング設定で撮影した画像を自動的に生成できる。Blender で構築した立体復元対象となる仮想環境を図 3.1-1 に示す。PCV を参考にし、直径 5.5 m、高さ 1.5 m の円柱の内部に復元対象物である 1 辺 0.3 m の立方体を等間隔で設置した。円柱の表面には錆のテクスチャ、各立方体には木目のテクスチャをそれぞれ貼り付けている。

図 3.1-1 Blender 作成した仮想環境（上面図）

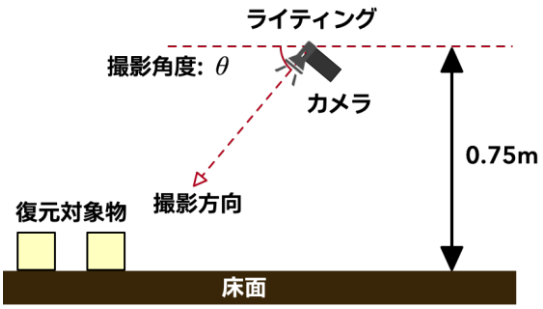


図 3.1-2 カメラ、ライティング設定

さらに、1F の廃炉プロジェクトで開発されたモジュール化されたレール構造[3.1-4]に応用することを考慮して、図 3.1-2 に示すように、床面から高さ 0.75 m の位置に仮想カメラおよびライティングを設置し、カメラの撮影角度 θ を変更できるように設定した。廃炉環境に導入することを考え、撮影中はこの撮影角度を固定とし、後にさまざまな撮影角度で撮影した画像を組み合わせることで、実現場における立体復元戦略を計画できるように設定した。そして、仮想レールの上面を移動することで、床面にある復元対象物を撮影するシミュレーションを行う。撮影シミュレーションの軌道を図 3.1-3 に示す。また、図 3.1-4 に示すように、ライティングの向きはカメラの撮影方向と同一として扱い、仮想カメラと同様に動くように設定し、仮想環境を照らしながら撮影を行うものとする。

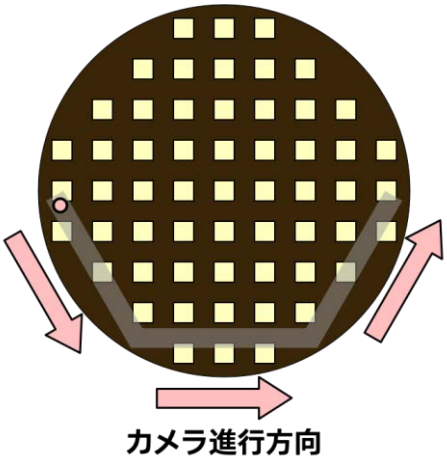


図 3.1-3　カメラ軌道（上面図）

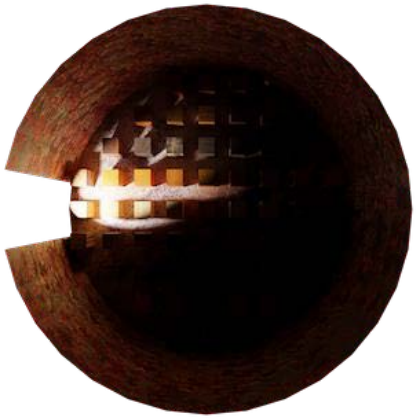


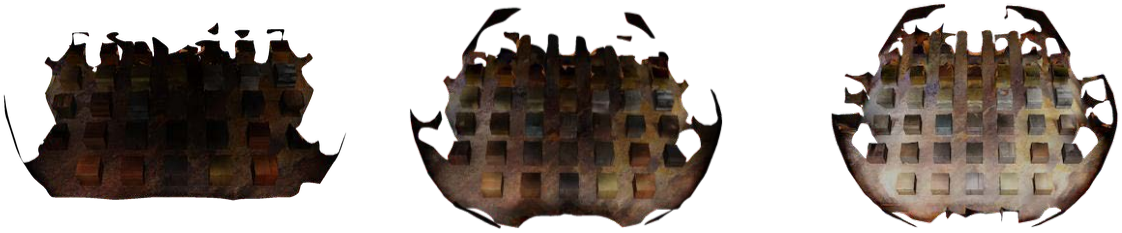
図 3.1-4　シミュレーションの 1 コマ（上面図）

この実験では、カメラの焦点距離変更をカメラの視野角変更として扱う。そして、アクションカメラを参考にし、視野角を 67 度、85 度、103 度、121 度の 4 種類を用意し、被写界深度として、床面にピントが合うようにした。本研究では、カメラの撮影角度 θ を 45 度で撮影を行う。そして、照明の強さを 10 W、100 W、1000 W の 3 種類を用意し、視野角・照明の強さを変化させて、合計で 12 パターンの撮影シミュレーションを行い、検証用の画像を生成する。撮影シミュレーションから等間隔で合計 41 枚の画像を生成し、SfM-MVS に適用する。アクションカメラを参考に、生成される画像の解像度を 1920 px×1080 px に設定した。図 3.1-5 に各ライティングにおけるカメラの視野角 85 度で撮影した同フレームの画像を示す。

(a) 照明の強さ 10 W　(b) 照明の強さ 100 W　(c) 照明の強さ 1000 W

図 3.1-5　視野角 85 度における同じフレームで生成した画像例

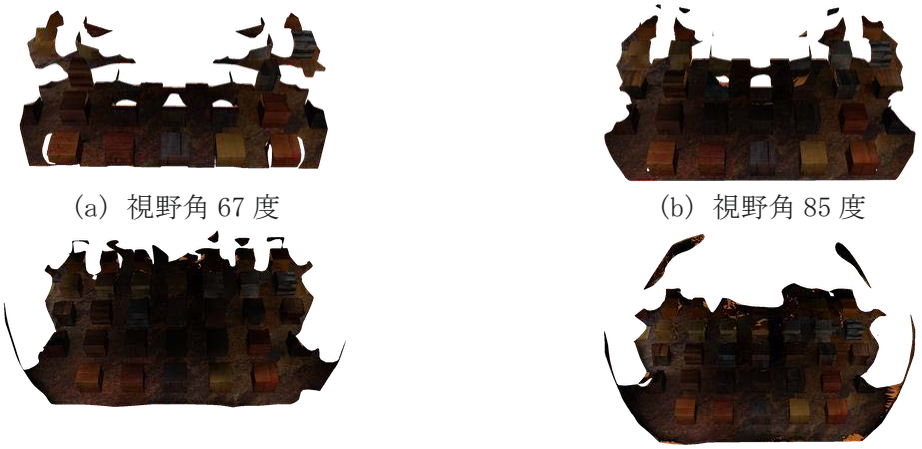
生成した画像群を SfM-MVS に適用して立体復元を行う。本検証では、立体復元手法に関する研究[3.1-5]で採用されている COLMAP[3.1-6]と OpenMVS[3.1-7]の組み合わせを採用した。図 3.1-6 に各ライティングにおけるカメラの視野角 103 度で撮影した画像による立体復元結果を示す。



(a) 照明の強さ 10 W　　(b) 照明の強さ 100 W　　(c) 照明の強さ 1000 W

図 3.1-6　各ライティングにおけるカメラの視野角 103 度で撮影した画像による立体復元結果例

この結果から、照明の強さが小さいとき、テクスチャを十分に照らすことができないため、3 つの結果の中で欠落がより多い立体復元結果になることがわかる。そして、照明の強さが大きいとき、暗闇環境の中でも全体的にテクスチャを十分に照らすことができるため、3 つの結果の中で復元対象物である立方体のテクスチャをより多く復元できている立体復元結果になることを確認できる。ただし、照明の強さが大きすぎる場合、画像の白飛びが発生し、画像特徴量を得られないため、撮影されている対象物の立体復元が獲得されない可能性もあるため、復元対象物とライティングの位置関係に応じて、照明の強さを適切に設定する必要があることがシミュレーションを用いて確認できる。また、図 3.1-7 に各視野角における照明の強さが 10 W で撮影した画像による立体復元結果を示す。



(a) 視野角 67 度　　(b) 視野角 85 度

(c) 視野角 103 度　　(d) 視野角 121 度

図 3.1-7　各視野角における照明の強さが 10 W で撮影した画像による立体復元結果例

この結果から、復元対象が撮影された範囲の立体復元結果を獲得できることを確認できる。撮影軌道が同じであるため、カメラの視野角が広い場合、対象物全体が撮影されることになり、より広く復元結果を得られることを視覚的に確認できる。一方で、カメラの視野角が狭い場合、環境全体ではなく、復元対象物をより大きく撮影するため、復元された対象物付近の床面に欠落が多いことも視覚的に確認できる。ただし、照明の強さが 100 W かつ視野角が 121 度の場合、画像間マッチングができなかったため、復元結果を獲得できなかった。そして、表 3.1-1 に各撮影パターンで獲得した高密度点群数を示す。

表 3.1-1　各撮影パターンにおける獲得された高密度点群数

		視野角			
		67 度	85 度	103 度	121 度
照明の強さ	10 W	2,765,716	2,497,194	1,848,413	1,162,014
	100 W	2,887,366	2,785,029	2,368,345	
	1000 W	2,747,547	2,791,632	2,637,095	2,455,409

この結果から視野角が広くなるにつれ、獲得される高密度点群数が少なくなることがわかる。また、今回対象とした環境は、照らしながら撮影を行うシミュレーションであるため、照明の強さが小さい、かつ視野角が広すぎる場合、対象物が映っている部分が少ないため、画像特徴量を得ることができないため。獲得される高密度点群がより少なくなることがわかる。

写真測量で獲得した立体復元結果を定量的に評価するために、Blender で作成した環境を点群化したデータをリファレンスとし、写真測量で獲得した立体復元結果との比較を行った。比較には、リファレンスデータと写真測量で得られた各高密度点群間で Iterative Closest Point (ICP) アルゴリズム[3.1-8]によるマッチング処理を行い、参考文献[3.1-9]に示してあるように、2 データ間の差異およびずれ量に基づいて立体復元結果の精度を算出する。精度を定義するにあたって、リファレンスデータを構成する各点に対し、写真測量で獲得した高密度点群の中で最も距離が近い点を求め、その 2 点の距離が最大許容量の範囲内で納まっている点群数の比率を適合率として評価する。本検証では、最大許容量 5 mm に対する適合率を採用する。

表 3.1-2 に各撮影パターンにおける適合率を示す。

表 3.1-2　各撮影パターンにおける適合率（最大許容量：5 mm）

		視野角			
		67 度	85 度	103 度	121 度
照明の強さ	10 W	96.85%	98.37%	97.17%	93.14%
	100 W	99.52%	99.22%	97.15%	
	1000 W	99.10%	99.20%	96.29%	91.63%

この結果から、立体復元結果を得られた 11 パターンにおいて、立体復元結果の適合率が 90%以上であるため、復元結果の精度が高いことがわかる。ただし、視野角が広すぎる場合、他の視野角の場合よりも適合率が小さく、獲得される高密度点群数が少なくなる。よって、欠落箇所をより少なくし、より精度を高い立体復元を行いたい場合、視野角を狭すぎず、広すぎないように適切に設定する必要が明らかになった。

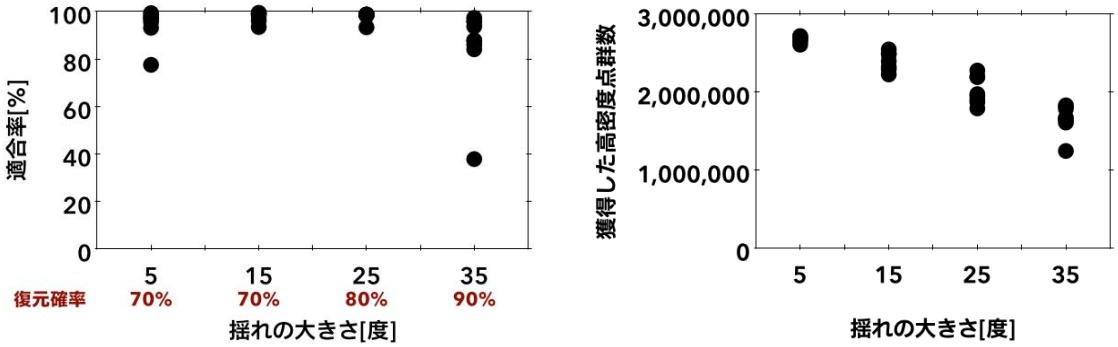
次に、実際の撮影を考えた場合、撮影中のカメラ撮影角度およびライティングの向きが固定ではなく変動する。つまり、カメラおよびライティングの向きが揺らぐと考えられる。そこで、乱数を用いてこの揺らぎを擬似的に実装し、揺らぎの大きさが写真測量で得られる復元結果に与える影響を調査する必要がある。そこで、画像を生成する際に、図 3.1-2 に示しているカメラおよびライティング向きに対して正規分布に従う乱数を加えることで、人工的な揺らぎを実装した。今回の調査では、乱数の範囲を以下に示す 4 パターンとして、揺らぎの大きさに対する影響を調査した。

1. ±5 度 （つまり、θ が 40 度から 50 度となる）
2. ±15 度（つまり、θ が 30 度から 60 度となる）
3. ±25 度（つまり、θ が 20 度から 70 度となる）
4. ±35 度（つまり、θ が 10 度から 80 度となる）

また、今回の調査は、図 3.1-1 に示す仮想環境した環境と以下に示す 4 パターンの撮影条件で、乱数で実装した揺らぎを考慮した暗所環境を照らしながら撮影するシミュレーションを行い、 人工画像を生成し、前述した写真測量手法に適用する。本調査では 10 試行を行った。

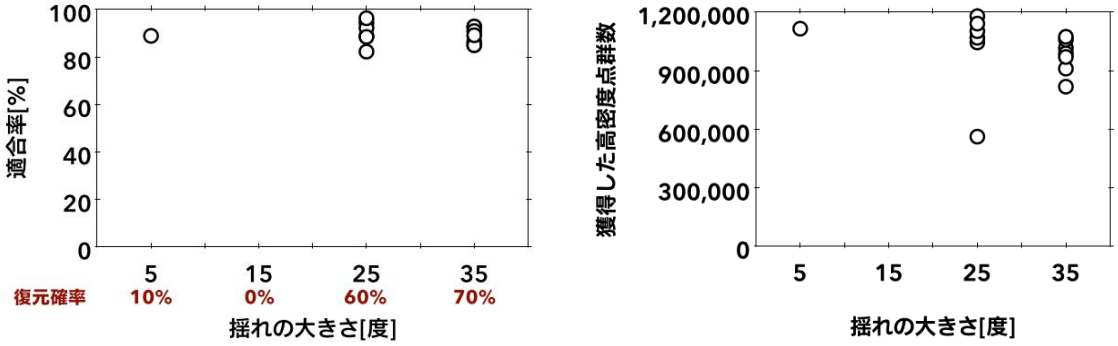
① 照明の強さ 10 W、視野角 67 度
② 照明の強さ 10 W、視野角 121 度
③ 照明の強さ 1000 W、視野角 67 度
④ 照明の強さ 1000 W、視野角 121 度

図 3.1-8～図 3.1-11 にパターン①～④に対して、揺れの大きさに対する復元成功確率、復元結果の適合率および獲得した高密度点群数の関係を示す。



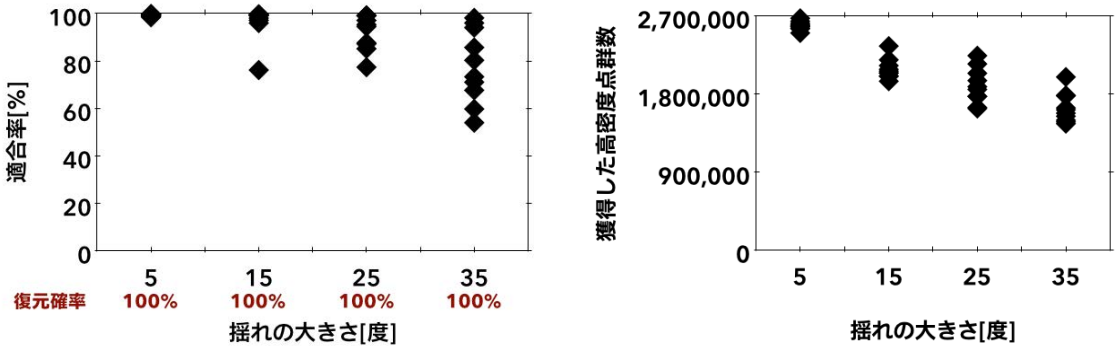
(a) 復元成功確率・適合率　　(b) 高密度点群数

図 3.1-8　照明の強さ 10 W、視野角 67 度における揺れの大きさに対する変化



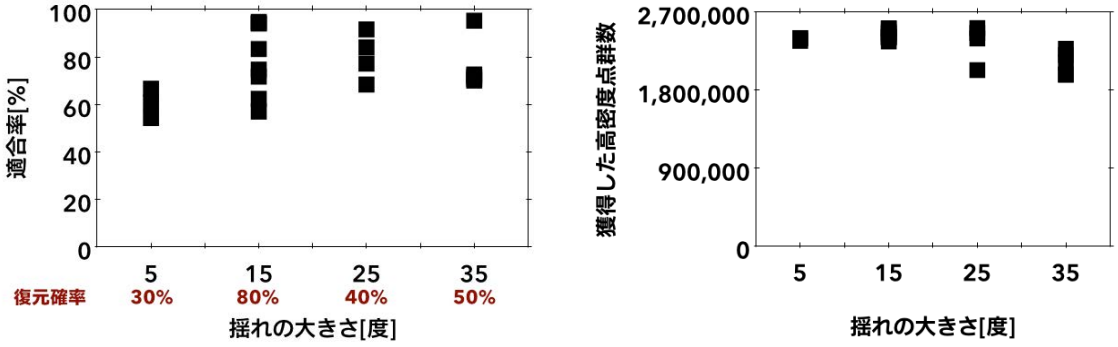
(a) 復元成功確率・適合率　　(b) 高密度点群数

図 3.1-9　照明の強さ 10 W、視野角 121 度における揺れの大きさに対する変化



(a) 復元成功確率・適合率　　(b) 高密度点群数

図 3.1-10　照明の強さ 1000 W、視野角 67 度における揺れの大きさに対する変化



(a) 復元成功確率・適合率　　(b) 高密度点群数

図 3.1-11　照明の強さ 1000 W、視野角 121 度における揺れの大きさに対する変化

パターン①「照明の強さ 10 W、視野角 67 度」の場合、高確率で復元でき、復元成功できた場合、高適合率であることがわかる。一方で、揺らぎが大きくなるほど、獲得する高密度点群数が少なくなることを確認できる。パターン②「照明の強さ 10 W、視野角 121 度」の場合、揺れの大きさが小さい場合、ほぼ復元できないことがわかる。一方で、揺らぎが大きくなるほど、獲得する高密度点群数にばらつきが大きくなることを確認できる。パターン③「照明の強さ 1000 W、視野角 67 度」の場合、揺れの大きさに関係なく復元できることが確認できるが、揺らぎが大きくなるほど、復元結果の適合率のばらつきが大きくなることがわかる。 一方で、揺らぎが大きくなるほど、獲得する高密度点群数のばらつきが大きくなることを確認できる。パターン④「照明の強さ 1000 W、視野角 121 度」の場合、揺れの大きさに対して、復元成功確率にばらつきがあることを確認できる。また、揺れの大きさに対して、復元結果の適合率のばらつきも大きくなることがわかる。一方で、揺らぎが大きくなるほど、獲得する高密度点群数が少なくなることを確認できる。

これらの結果から復元対象環境に応じて、適切な照明の強さ、カメラ視野角が存在することが明らかになった。また、設定した照明の強さ、カメラ視野角によって、許容できる揺れの大きさも明らかになった。復元対象環境に応じて、これらの値は変化する。ゆえに、あらかじめシミュレーションを用いて、これらのパラメータを検討することは、撮影回数が制限される廃炉作業によって重要であるといえる。

さらに、シミュレーションで得られた結果と実際のカメラで撮影した画像による立体復元結果を比較するため、さまざまな姿勢で撮影できるカメラシステムの構築を行った。図 3.1-12 に示すように、小型モバイルマニピュレーターを搭載している TurtleBot3 MM Gen3[3.1-10]にアクションカメラを搭載することで、現実に即した移動しながらさまざまな姿勢で撮影できるシステムを構築した。この TurtleBot3 MM Gen3 は Robot Operating System (ROS) 1 や ROS2[3.1-11]で制御できるため、シミュレーションと同様な方法で画像および動画の取得が可能になる。

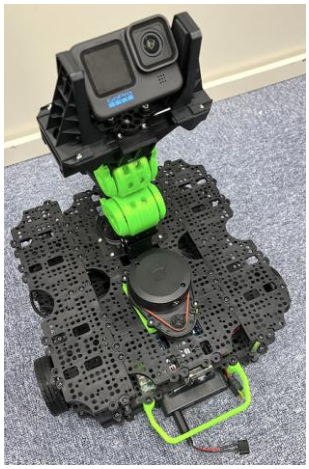


図 3.1-12　さまざまな姿勢で撮影できるカメラシステムの検討

(3) まとめ

シミュレータを活用し、写真測量で重要な画像特徴量に影響を与える光の反射、レンズの焦点距離、撮影時のカメラ振動に注目し、これらの要素が変化した際の写真測量による画像間特徴量および立体復元結果の精度の関係を数値計算実験によって明らかにした。また、シミュレーションで得られた結果と実際のカメラで撮影した画像による立体復元結果を比較するため、さまざまな姿勢で撮影できるカメラシステムの構築を行った。

参考文献

[3.1-1] Ullman, S., The interpretation of structure from motion、Proceedings of the Royal Society of London, Series B, Biological Sciences, 203, 1153, 1979, pp.405-426. https://doi.org/10.1098/rspb.1979.0006

[3.1-2] Seitz, S.M. et al., A comparison and evaluation of multi-view stereo reconstruction algorithms, 2006 IEEE computer society conference on computer vision and pattern recognition, 1, IEEE, 2006. https://doi.org/10.1109/CVPR.2006.19

[3.1-3] Blender, https://www.blender.org/ （参照：2024年4月1日）.

[3.1-4] Yokomura, R. et al., Rail DRAGON: Long-Reach Bendable Modularized Rail Structure for Constant Observation Inside PCV, IEEE Robotics and Automation Letters, 9, 4, 2024, pp.3275-3282. https://doi.org/10.1109/LRA.2024.3366022

[3.1-5] Kataria, R. et al., Improving structure from motion with reliable resectioning, 2020 International Conference on 3D vision, IEEE, 2020. https://doi.org/10.1109/3DV50981.2020.00014

[3.1-6] Schönberger, J.L. et al., Structure-from-motion revisited, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. https://doi.org/10.1109/CVPR.2016.445

[3.1-7] Cernea, D., OpenMVS: Multi-View Stereo Reconstruction Library, https://cdcseacave.github.io/openMVS （参照：2024年4月1日）.

[3.1-8] Besl, P.J. et al., Method for registration of 3-D shapes, Sensor fusion IV: control paradigms and data structures, 1611, 1992. https://doi.org/10.1117/12.57955

[3.1-9] Nakamura, K. et al., 3D reconstruction considering calculation time reduction for linear trajectory shooting and accuracy verification with simulator, Artificial Life and Robotics, 28, 2, 2023, pp.352-360. https://doi.org/10.1007/s10015-022-00835-x

[3.1-10] TurtleBot3 MM Gen3, https://e-shop.robotis.co.jp/product.php?id=491 （参照：2024年4月1日）.

[3.1-11] ROS (Robot Operating System), http://wiki.ros.org/ （参照：2024年4月1日）.

3.2　動画像からの迅速な 3 次元モデリング手法の研究開発（連携先：JAEA）

（1）令和 5 年度概要

令和 5 年度は、最新の PCV 内部調査の動画像データを対象として画像の特性分析および深層学習を用いた手法等を複合的に適用した画像特徴点抽出手法等の開発を行い、これらの手法を組み合わせて構造物のグループ化に基づく 3 次元モデリングアルゴリズムによる環境情報の提示、立体復元計算の時短化アルゴリズムによる高速化等の次年度に必要な研究開発の課題について抽出を行った。また、イメージスティッチングによる高解像度画像の生成手法および画像間のオプティカルフローに基づく遠隔操作機器の移動・停止の自動記録生成手法の適用性検証も併せて実施した。以下でそれぞれの実施内容および結果について述べる。

（2）令和 5 年度実施内容および成果

まず、PCV 内部調査の動画像データを対象とした画像の特性分析について述べる。画像の特性分析では東京電力ホールディングス（以下、「東電 HD」と略す。）の HomePage（HP）上で公開されている動画・写真ライブラリーの中から、動画を選定することとした。

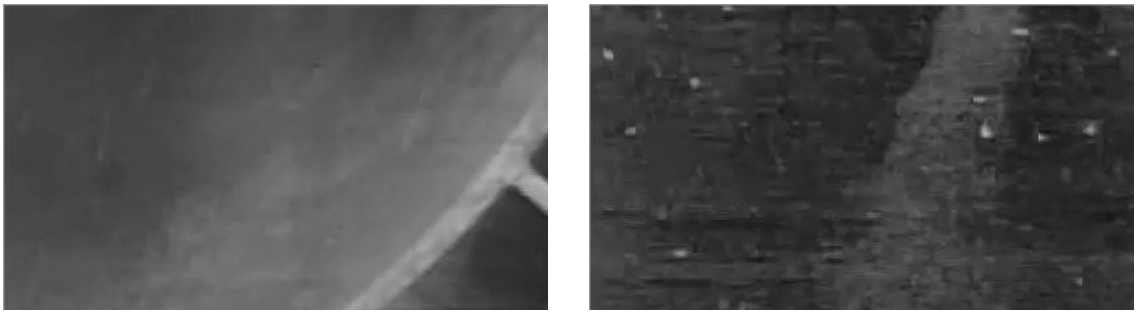
令和 5 年度は、最新の PCV 内部調査の動画像データを対象として画像の特性分析および深層学習を用いた手法等を複合的に適用した画像特徴点抽出手法等の開発を行い、これらの手法を組み合わせて構造物のグループ化に基づく 3 次元モデリングアルゴリズムによる環境情報の提示、立体復元計算の時短化アルゴリズムによる高速化等の令和 6 年度に必要な研究開発の課題について抽出を行った。また、イメージスティッチングによる高解像度画像の生成手法および画像間のオプティカルフローに基づく遠隔操作機器の移動・停止の自動記録生成手法の適用性検証も併せて実施した。

PCV 内部調査の動画像データを対象とした画像の特性分析では、東電 HD の HP 上で公開されている動画・写真ライブラリーの中から動画を選定することとした。PCV 内部において動画像取得を行う場合、種々の制約条件により照明不足による低コントラストや放射線ノイズ、構造物のぼやけが生じやすい。そこで、照明不足による低コントラストや放射線ノイズが発生しているシーンを多数含む、令和 4 年 5 月に行われた小型 Remotely Operated Vehicle（ROV）による 1 号機 PCV 内部調査[3.2-1]時の動画を対象とした。動画から抽出した画像の一例を図 3.2-1 に示す。図より、画像は全体的に暗くコントラストが低いこと、図 3.2-1(b)の後方監視カメラでは放射線ノイズが発生していることも確認できる。このように PCV 内部調査で得られる画像は種々の制約条件により、品質が高くないことが多い。そのため、3 次元モデルの生成に適した画像の品質について考えることは迅速に 3 次元モデリングを実行するうえで重要な要素となる。そこで、PCV 内部調査で得られた画像の品質を評価し、画像の定量的な品質評価と 3 次元モデリングへの影響における関係を調べるための検証を実施した。

(a)：水中監視カメラ　　(b)：後方監視カメラ

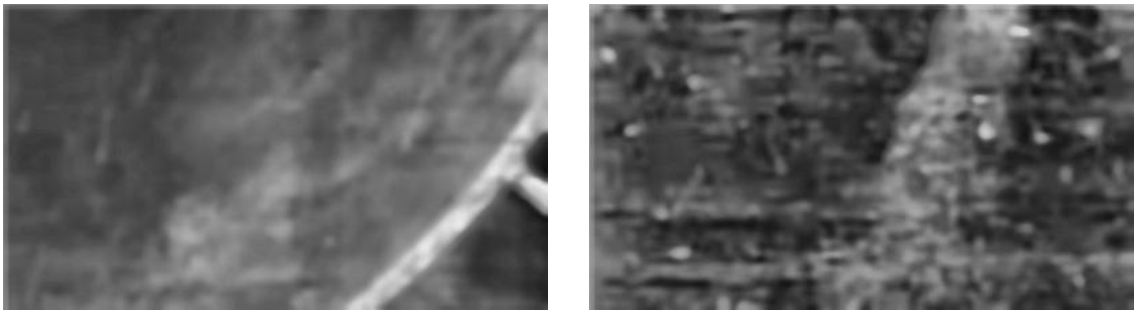
図 3.2-1　PCV 内部調査画像（出典：東電 HD）

画像品質を評価するために、動画像の品質評価手法（以下、「IQA」と略す。）の指標を導入した。これにより、主観的にしか判断できなかった画像品質の定量的な評価が可能となった。PCV内部調査で得られる動画像は制約条件下で取得されるため、比較できる基準画像が存在しないことから、基準画像を必要としない品質評価手法（No-Reference IQA：NR-IQA）を採用することとした。NR-IQAとしてPerception based Image Quality Evaluator（PIQE）[3.2-2]を用いて品質評価を行った。PIQEは学習済みモデルを必要とせず、画像中の局所特徴量からスコアを算出する。PIQEのスコアは値が0～100で表現され、小さいほど画像品質が高いことを示す。このスコアを指標としてPCV内部調査で得られた画像に対して定量的な品質評価を行った。ここで、画像からの特徴点の抽出に用いられる特徴量抽出手法（例えば、SIFT[3.2-3]やORB[3.2-4]等）は一般的にグレイスケール画像を入力としている。そのため、本報告では評価対象の画像をグレイスケール画像に変換し、入力画像とした。画像の変換とPIQEのスコアの計算には、Python3とコンピュータビジョンのオープンソースライブラリOpenCV4[3.2-5]を用いて実装した。各手法の実行環境はApple MacBook Pro（macOS 12.7.4、CPU:Core i9、8 core、RAM:64 GB、DDR4）を用いた。変換したグレイスケール画像とPIQEのスコアを図3.2-2に示す。また、PCV内部調査では、制約条件により十分な品質の画像を取得できない場合もあるため、画像補正によりどの程度品質が改善するかを把握することは重要である。そこで、評価対象の画像に対する補正が画像品質へ与える影響も検証した。画像補正は、以前行った実験[3.2-6]によって、画像品質の改善に対して汎用的に効果が期待できるコントラスト調整を適用することとした。本報告では、画像中の局所領域ごとにコントラストの調整が可能でコントラストの過剰な増加を抑制できる適用的ヒストグラム平坦化（以下、「CLAHE」と略す。）[3.2-7]を採用して画像補正を実施した。図3.2-3に画像補正を行った結果を示す。図3.2-2および図3.2-3より、水中監視カメラ画像および後方監視カメラ画像ともにCLAHEの適用によってPIQEのスコアが小さくなっており、画像品質が改善していることを示している。また、暗所部分が鮮明化されていることが目視でも確認できる。



PIQE score = 62.48　　PIQE score = 64.87
(a) 水中監視カメラ　　(b) 後方監視カメラ

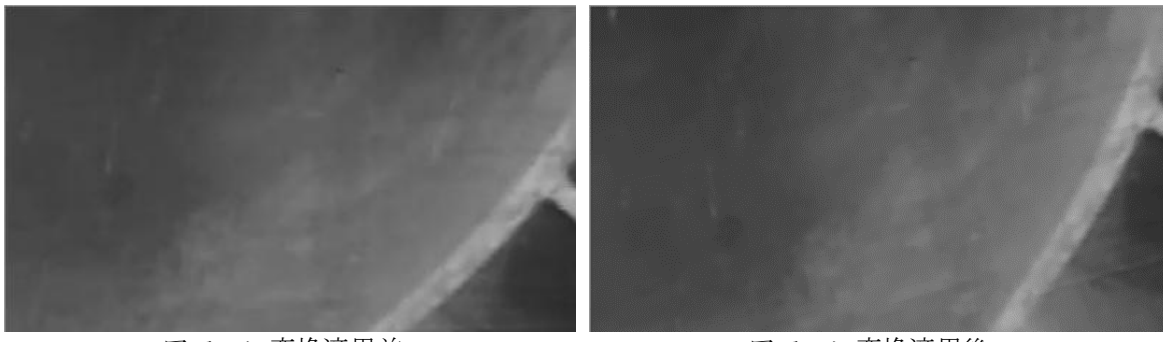
図3.2-2　グレイスケール画像



PIQE score = 41.03　　PIQE score = 57.81
(a) 水中監視カメラ　　(b) 後方監視カメラ

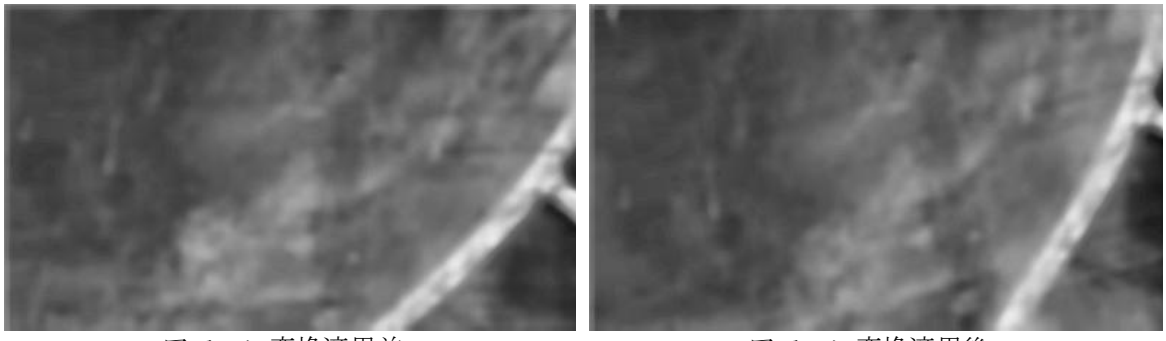
図3.2-3　補正画像

画像品質は、画像からの特徴点の抽出およびマッチングに影響を与えるため、 SfM のような画像からの 3 次元モデリングの復元精度を左右する。そこで、画像品質のスコアを基準として画像補正が特徴点の抽出およびマッチングに与える影響を評価するために、選定した原画像（図 3.2-2）と補正画像（図 3.2-3）に対してアフィン変換を行い、画像中の対応点の座標位置が計算可能な画像のペアを作成して検証を実施した。本報告では、アフィン変換として画像の水平方向に 10 px シフトし、原画像の中心座標を原点として反時計回りに 15 度回転させた。また、アフィン変換によって生じる画像中の欠損部分を取り除くために外縁部のトリミングも行い、画像の解像度を 400 px×224 px とした（図 3.2-4、図 3.2-5）。



アフィン変換適用前　　アフィン変換適用後

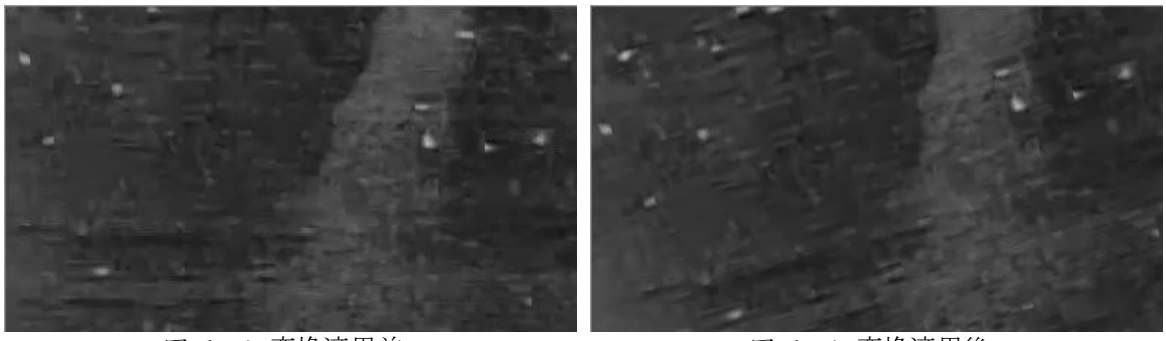
(a) 原画像



アフィン変換適用前　　アフィン変換適用後

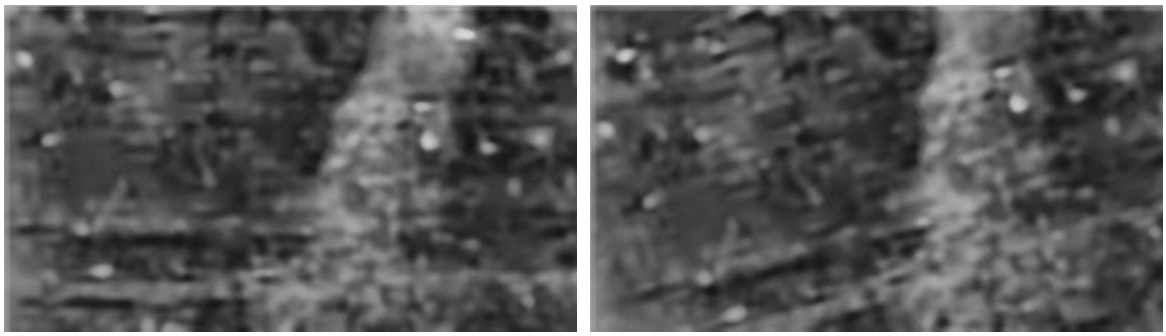
(b) 補正画像

図 3.2-4　水中監視カメラ画像へのアフィン変換の適用



アフィン変換適用前　　アフィン変換適用後

(a) 原画像



アフィン変換適用前　　アフィン変換適用後

(b) 補正画像

図 3.2-5　後方監視カメラ画像へのアフィン変換の適用

図 3.2-4 および図 3.2-5 のペア画像は、アフィン変換によって作成されているため、特徴点マッチングにおける対応点同士の座標情報を求めることが可能であり、精度評価が実施できる。抽出された特徴点のマッチングは、2 次元ノルムを使った総当りマッチングによって対応点を探索した。ここで、誤マッチングを抑えつつ良好なマッチング結果を得るために用いられている Lowe's ratio test[3.2-3]の閾値をパラメータとしてマッチングの精度評価を行うこととした。画像中の特徴点抽出には特徴量記述子として SIFT を用いた。Lowe's ratio test のしきい値を 0.75 としたケースの結果を表 3.2-1、表 3.2-2 に示す。

表 3.2-1　水中監視カメラ画像からの特徴点の抽出およびマッチング結果

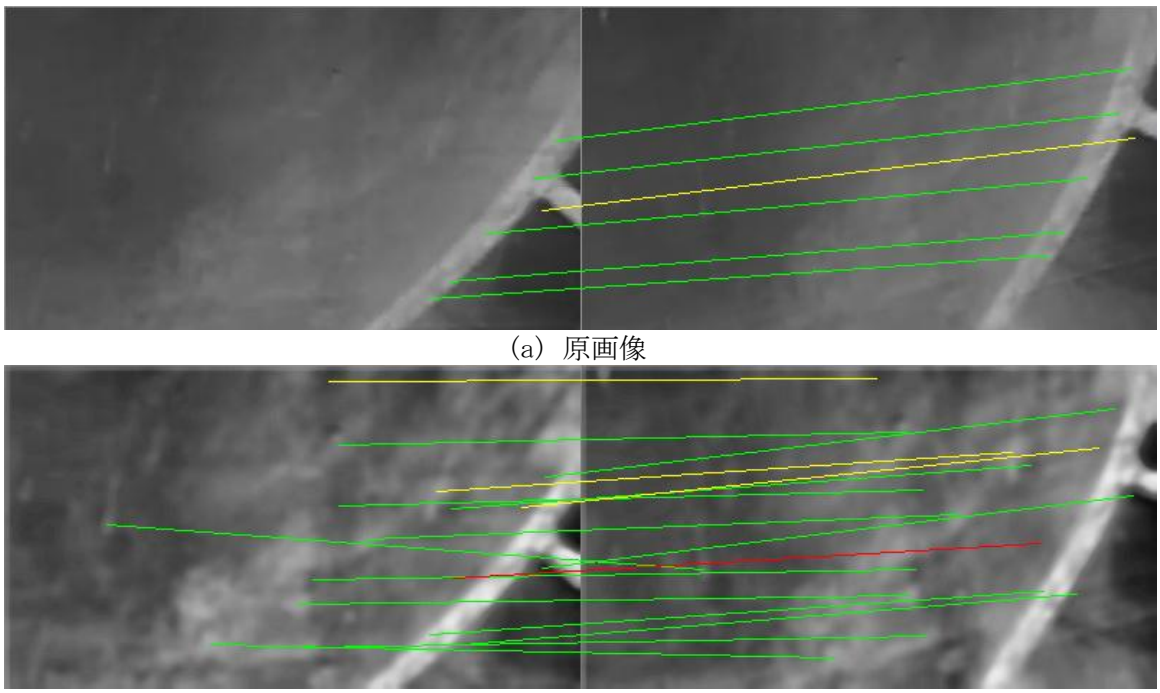
	PIQE スコア 1	特徴点抽出数 1	PIQE スコア 2	特徴点抽出数 2	マッチング数
原画像	62.48	9	69.17	9	6
補正画像	41.03	54	45.02	44	17

表 3.2-2　後方監視カメラ画像からの特徴点の抽出およびマッチング結果

	PIQE スコア 1	特徴点抽出数 1	PIQE スコア 2	特徴点抽出数 2	マッチング数
原画像	64.87	62	56.98	62	28
補正画像	57.81	365	55.76	358	174

表 3.2-1、表 3.2-2 の添え字 1、2 は、アフィン変換適用前と適用後をそれぞれ表している。この表から画像補正の適用によって特徴点の抽出数が大幅に改善され、マッチング数もそれに

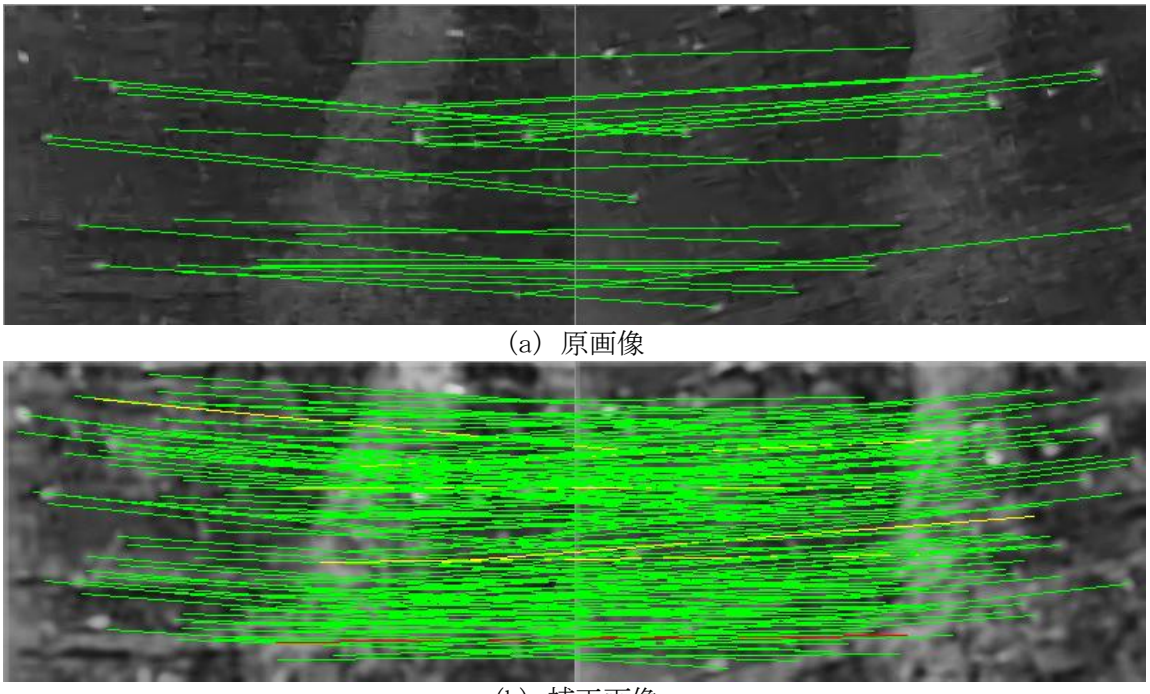
伴って増加していることが確認できる。さらに、図 3.2-6、図 3.2-7 に表 3.2-1、表 3.2-2 に対応した特徴点のマッチング結果を示す。図中では、対応付けられた特徴点同士を線で結んでいる。緑色の線はアフィン変換から計算された正しい座標位置とのずれが 1 px 未満、黄色の線は 1 px 以上 5 px 未満、赤色の線は 5 px 以上となっている。図より、補正画像は原画像と比較して特徴点のマッチング数が増加していることも確認できる。



(a) 原画像

(b) 補正画像

図 3.2-6　水中監視カメラ画像の特徴点マッチング結果

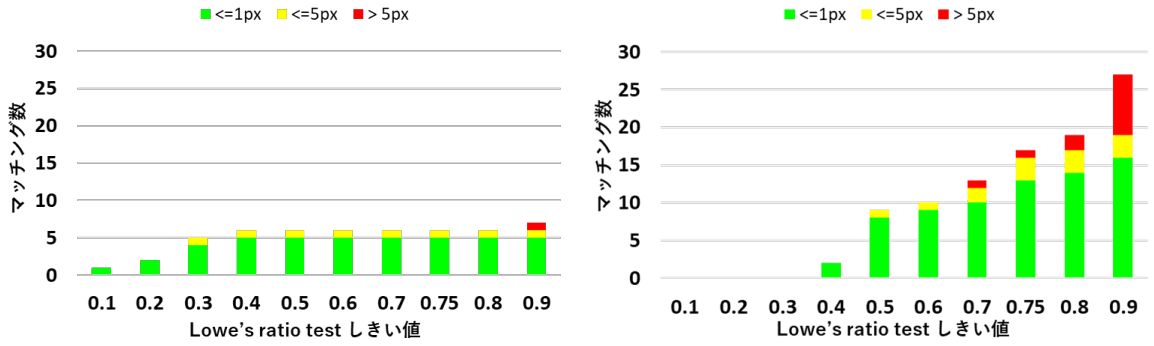


(a) 原画像

(b) 補正画像

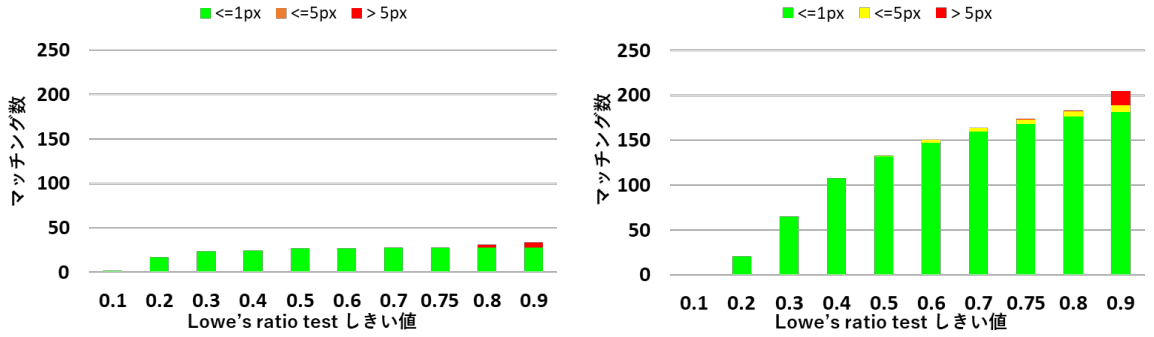
図 3.2-7　後方監視カメラ画像の特徴点マッチング結果

図 3.2-8、図 3.2-9 に Lowe's ratio test のしきい値をパラメータとしたマッチング数の変化を示す。図中のヒストグラムの色は図 3.2-6 と同様にアフィン変換から計算された正しい座標位置とのずれが 1 px 未満を緑色、1 px 以上 5 px 未満を黄色、5 px 以上を赤色としている。原画像の傾向としてしきい値の増加に伴い、ずれが 1 px 未満のマッチング数が徐々に収束し、その後に誤マッチングの割合が増加していくことが確認できる。原画像の結果と比較して、補正画像はマッチング数が大幅に増加しており、CLAHE の適用によってマッチング精度が向上していることを示唆している（図 3.2-8、図 3.2-9）。表 3.2-1、表 3.2-2 および図 3.2-8、図 3.2-9 のマッチング結果から原画像と補正画像を比較すると、PIQE スコアの改善に伴ってマッチング数が増加していることが確認できる。そのため、画像品質評価のスコアを判断基準として、3 次元モデルの生成に適した画像であるかを事前に判断が可能となる。それによって、低品質の画像に対しては、適切な画像補正を実施することで 3 次元モデリングの復元精度を向上させることが可能となる。



（a）原画像　　（b）補正画像

図 3.2-8　水中監視カメラ画像の特徴点マッチング数のしきい値依存性



（a）原画像　　（b）補正画像

図 3.2-9　後方監視カメラ画像の特徴点マッチング数のしきい値依存性

次に、SfM および MVS による立体復元計算の特徴点抽出において、従来手法や深層学習を適用した手法を用いた場合の復元結果についての分析および複数手法を複合的に適用する方法の開発を実施した。

令和 6 年度に実施予定のグループ化に基づく 3 次元モデリング手法の開発では、立体復元で生成した点群に基づいてモデリングを行うことから、精緻な 3 次元モデリングを実施するためには立体復元の段階で高密度な点を生成できることが望ましい。前述の通り SfM および MVS においては、画像特徴点抽出および画像間の対応点マッチングの結果が復元する点群の密度に大きく影響する。つまり、高密度な点の生成には良好な特徴点を得ることが必要となる。しかし、PCV 内部調査映像には、放射線に起因する撮像ノイズや水中における濁りや浮遊物等の影響が

あり、良好な特徴点抽出が難しいことが想定される。従来の SfM では、画像特徴点抽出手法として SIFT[3.2-3]が用いられてきたが、アーチファクト（ノイズ等を原因とした実際には存在しない虚像）の影響を受けやすく[3.2-8]、照明の大きな変化[3.2-9]に弱いという課題が指摘されており、良好な特徴点抽出および対応点マッチングが難しいことが予想される。ここで、近年の深層学習による特徴点抽出手法は、学習方法の工夫により上述の課題を克服しつつあり[3.2-8]、PCV 内部調査動画像に対しても良好な結果を得られることが期待できる。

そこで、令和 5 年度は、東電 HD が公開している PCV 内部調査映像から抽出した動画像を対象に従来手法および深層学習を適用した手法を用いて、特徴点抽出数や対応点マッチング数、復元結果および処理時間の違いを分析することで、令和 6 年度の開発に向けた課題抽出を行った。図 3.2-10 は、分析対象として 1 号機の PCV 内部調査映像から選定して切り出した 3 つの動画像を示している[3.2-10][3.2-11][3.2-12]。

(a)動画像 1（気中カメラ、放射線ノイズ）

(b) 動画像 2（水中カメラ、濁りと浮遊物）

(c) 動画像 3（水中カメラ、濁りと気泡）

図 3.2-10　対象の PCV 内部調査動画像の一例（出典：東電 HD）

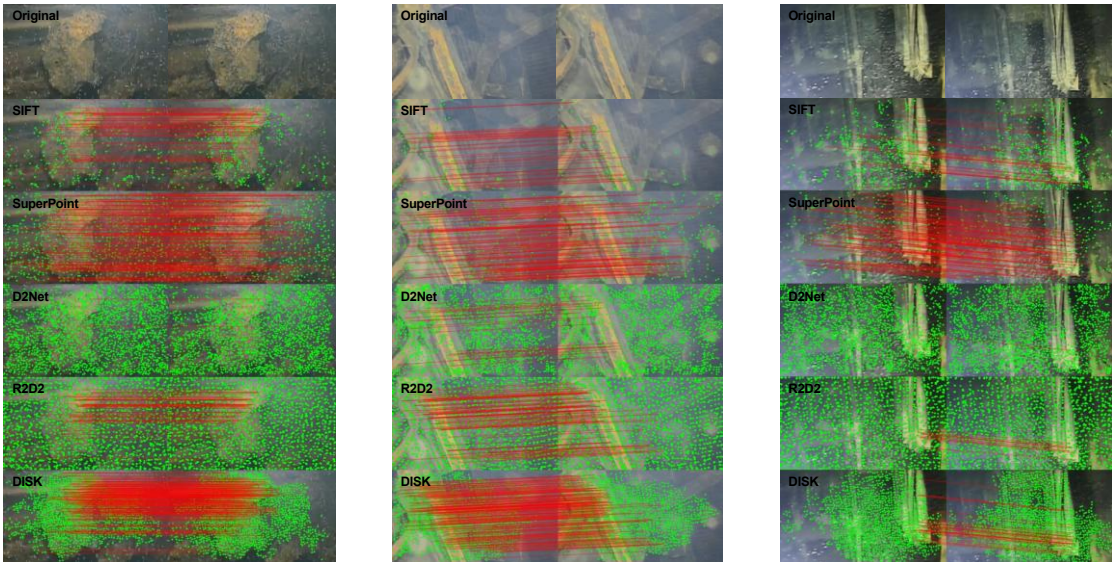
これらの動画は、国際廃炉研究開発機構が開発した水中 ROV によって撮影されたものである[3.2-10]。動画像 1 は画像枚数が 176 枚であり、気中カメラによって天井面の構造物を映したものであり、放射線の影響による撮像ノイズが確認できる。動画像 2 は、画像枚数が 220 枚であり、水中カメラによって構造物を映したものであり、水中の濁りによって右奥のグレーチングがぼやけていることに加えて浮遊物が確認できる。動画像 3 は、画像枚数が 110 枚であり、水中カメラによって構造物を映したものであり、水中の濁りおよび気泡によって奥に映る構造物がぼやけた映像である。取得された映像には撮影日時等の情報が表示されているが、立体復元計算に悪影響を与えることからトリミングにより除去を行い、画像のサイズは 560 px×315 px となった。

ここで、立体復元計算には画像に基づいた立体復元を用いた局所位置探索のツールボックスである hloc[3.2-9]を使用して分析を行った。代表的な深層学習を適用した特徴点抽出手法として、SuperPoint[3.2-13]、D2-Net[3.2-14]、R2D2[3.2-15]、DISK[3.2-16]の 4 手法および従来手法である SIFT の計 5 手法によって結果比較を行うこととした。対応点マッチングでは、最近傍法を用いた ratio-test 法[3.2-3]を用いるが、SuperPoint のみ記述子の形式が違うことから SuperPoint に基づいて提案された SuperGlue[3.2-17]を用いることとした。立体復元計算に使用した計算機として、OS が Ubuntu 20.04、CPU が Intel Xeon Gold 5317 (3.00 GHz、12 Core) x2、RAM が DDR4 3200 512 GB、GPU が NVIDIA RTX A6000 48 GB の計算機を用いた。

動画像 1 から 3 に対して特徴点抽出および対応点マッチングを行い、その結果に基づいて立体復元計算による密な点群の生成を行った。表 3.2-3 および図 3.2-11 は、5 手法を用いた場合の抽出した特徴点数、対応点マッチングの対応組数および処理に要した時間の一例を示している。図 3.2-11 の結果は、緑色の点が特徴点、赤い線がマッチング結果をそれぞれ示している。

表 3.2-3　特徴点抽出および対応点マッチング結果の一例

手法名	動画像 1			動画像 2			動画像 3		
	特徴点	対応組	時間[s]	特徴点	対応組	時間[s]	特徴点	対応組	時間[s]
SIFT	726	55	0.12	238	38	0.12	1,128	28	0.13
SuperPoint	1,292	172	0.09	386	136	0.08	682	141	0.09
D2-Net	3,638	2	0.11	1,594	21	0.09	3,407	1	0.11
R2D2	6,679	46	0.11	6,274	83	0.09	6,443	10	0.11
DISK	3,074	183	0.03	3,007	175	0.05	3,087	31	0.03



(a) 動画像 1　　(b) 動画像 2　　(c) 動画像 3

図 3.2-11　特徴点抽出および対応点マッチング可視化結果の一例

表 3.2-3 に示す通り、動画像 1 と 2 において、従来手法（SIFT）と比較して深層学習を適用した手法（SuperPoint、D2-Net、R2D2、DISK）において多くの特徴点が抽出できていることを確認した。また、SuperPoint および DISK の 2 手法では、対応点マッチング数の増加が確認できた。動画像 1 では、全ての手法において放射線ノイズを特徴点として捉えており、ratio-test 法の対応点マッチング結果では当該領域を除外できているが、SuperPoint および SuperGlue では、放射線ノイズの領域を除外しつつ背面の構造物においても良好なマッチングが行えている。動画像 2 では、SuperPoint および DISK において、水中の濁りによりぼけた背面のグレーチング等の構造物に対しても特徴点を抽出でき、良好なマッチングが行えている。これらに対して、動画像 3 では、SuperPoint が SIFT の特徴点抽出数を下回ったが、これは SIFT で気泡を特徴点として多く抽出していることが原因と考えられる。 SuperPoint の結果では、特徴点抽出数自体は少ないが、気泡以外の特徴点を網羅的に抽出できており、SuperGlue による対応点マッチングの結果についても、従来の ratio-test 法による対応組数より多いことが確認できた。D2-Net、R2D2 および DISK においては、特徴点抽出数は多いが、対応点マッチングにおいて撮像ノイズや水中の濁り、浮遊物、気泡が影響していることで良好な対応点マッチング数は少ない結果となった。処理時間は、全ての深層学習を適用した手法において、従来手法と比較して処理が早いことを確認した。処理速度および対応点マッチングの結果から、SuperPoint および SuperGlue による特徴点抽出および対応点マッチングが種々の影響がある動画像に対して安定して良好であることがわかった。

この結果を踏まえて、SuperPoint および SuperGlue と従来手法である SIFT および ratio-test 法の組み合わせによって得た、対応点マッチング結果を用いた立体復元結果の比較を行った。表 3.2-4 は立体復元計算によって復元された点の数および処理時間、図 3.2-12 は MVS を用いた密な立体復元結果をそれぞれ示している。なお、立体復元結果は比較が行えるように手作業で大凡の位置座標とスケールを合わせた。

表 3.2-4　立体復元の結果

手法名	動画像 1		動画像 2		動画像 3	
	復元点	処理時間[s]	復元点	処理時間[s]	復元点	処理時間[s]
SIFT	210,325	1,325.56	13,756	119.31	22,692	231.52
SuperPoint	473,693	2,124.33	286,030	2,060.69	140,388	1,151.70

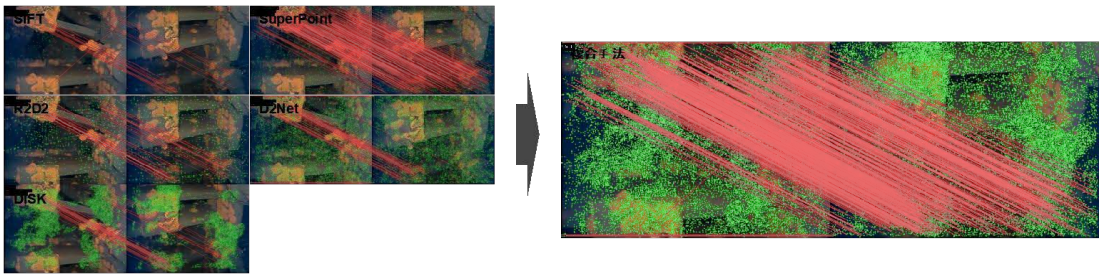
(a) 動画像 1　(b) 動画像 2　(c) 動画像 3

図 3.2-12　SIFT および SuperPoint による密な立体復元結果

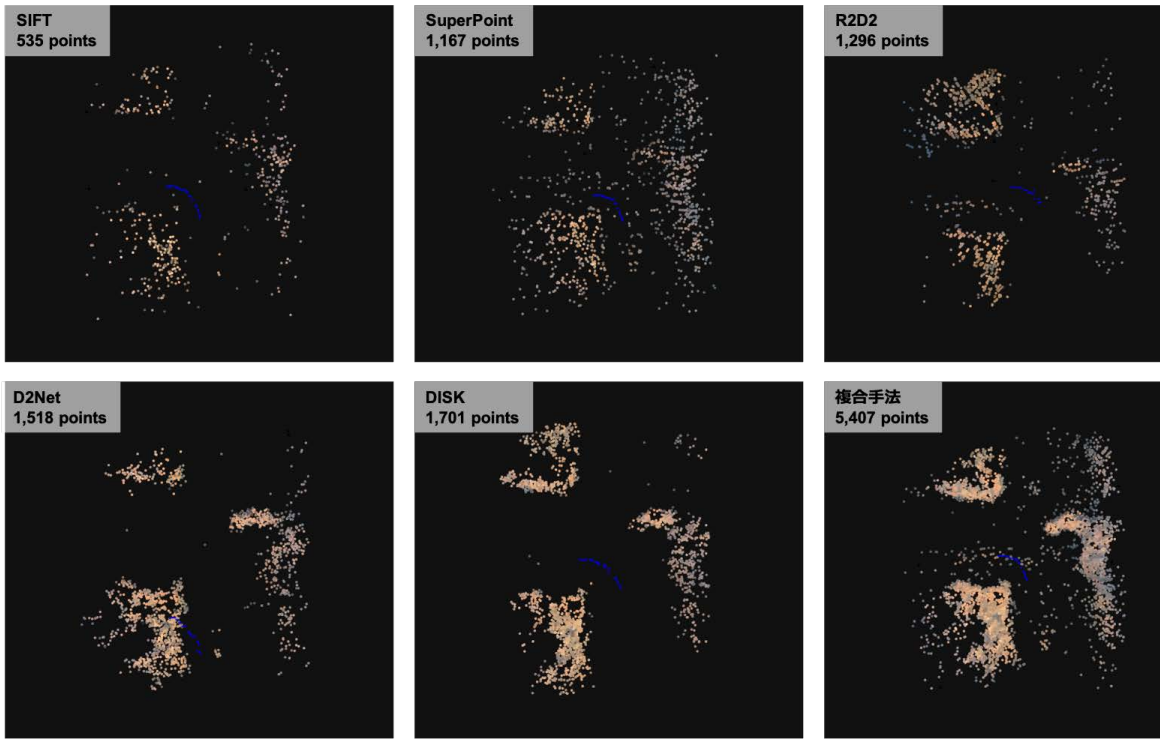
表 3.2-4 に示す通り、全ての動画像の立体復元結果において、深層学習を適用した手法が従来手法より多くの点を復元できることを確認した。図 3.2-12(a)～(c)に示す通り、深層学習を適用した復元結果では、構造物が高い密度で復元できていることがわかる。また、図 3.2-12(b)および図 3.2-12(c)に示すように、水中の濁りや浮遊物、気泡の影響によって特徴点が抽出できていなかった領域が、深層学習を適用した手法により特徴点を抽出できているため、復元されていることが確認できる。それぞれの手法における処理時間を比較すると、全ての動画像において従来手法による立体復元結果の処理時間が短いことがわかる。これは、抽出された特徴点数が多いほど、それ以降の三角測量等の立体復元計算にコストがかかるためである。今回の分析では、対象となる動画像の枚数が少なかったが、数時間におよぶ映像から抽出した動画像を対象とした場合、処理時間は顕著に増加することが想定されることから、計算処理の効率化を考えていく必要がある。また、各立体復元結果において構造物の間にアーチファクトとなる点が復元されるため構造物の視認性が悪いといえる。今後、構造物のグループ化アルゴリズムの開発によって、構造物の復元点の強調またはアーチファクトの除去を行う手法について研究開発や立体復元の精度について評価を行う必要がある。

ここまでの深層学習を適用した特徴点抽出手法を用いた立体復元結果について分析から、複数の手法を複合的に適用することで、さらに密度の高い立体復元が可能であることが考えられる。例えば、図 3.2-11 の特徴点抽出結果に示すように SuperPoint は、ぼやけた構造物を含め

て特徴点を網羅的に抽出しているのに対して、DISK ではテクスチャがはっきりとした構造物を重点的に特徴点の抽出を行っている。このように、手法ごとに異なる特徴点抽出特性を持つため、特徴点抽出および対応点マッチングの結果を複合的に用いることでより網羅的な特徴点抽出が可能となる。そこで、5 手法の特徴点抽出結果および対応点マッチング結果を統合して重複点除去を行う手法（以下、「複合手法」と略す。）を開発し、立体復元計算を行った。図 3.2-13 は、東電 HD が公開している PCV 内部調査映像[3.2-18]から抽出した動画像（画像枚数 66 枚）を例に、複合手法の特徴点抽出および対応点マッチング結果を統合した可視化結果および立体復元結果を示す。



(a) 統合した特徴点抽出結果および対応点マッチング結果



(b) 立体復元結果

図 3.2-13　手法を複合的に適用した特徴点および立体復元結果

立体復元の結果、5 手法を用いて復元された点の数が、535～1,701 点であるのに対して、複合手法では、5,407 点と多くの点が復元できた。図 3.2-13(b)に示す通り、複合手法では、他の手法を単体で適用した場合より、構造物表面および全体的に多くの点が復元できていることが確認できる。 図 3.2-14 は、5 手法および複合手法を用いて SfM による疎な立体復元計算までを行った場合の復元点と処理時間をグラフ化したものである。立体復元における復元点および

処理時間には相関がみられ、特徴点抽出数と立体復元工程の処理時間はトレードオフの関係性にあることが明らかとなった。取得した調査映像に基づいて、次の作業計画に立体復元結果を確認するような活用場面を考えた場合、時短アルゴリズムを開発することが必要といえる。今後、密な立体復元についても同様に立体復元結果と処理時間について分析を実施する予定である。

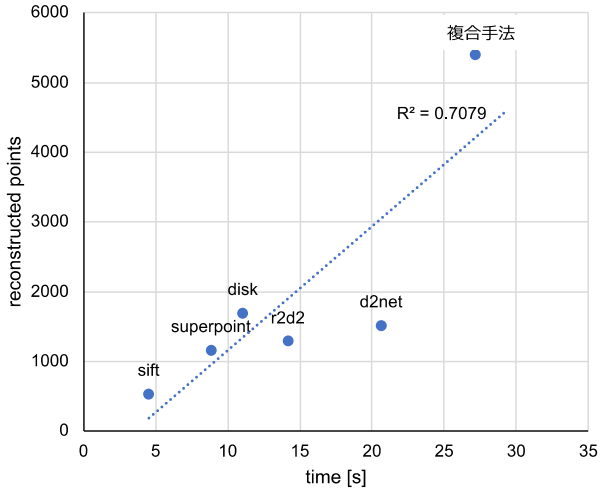


図 3.2-14　復元結果と処理時間の関係について

以上の結果から、立体復元における特徴点抽出において深層学習を適用した手法は、従来手法と比較して、より広範囲の領域を密度高く立体復元できることがわかった。さらに、複数の特徴点抽出手法を複合的に適用することで、各手法の特性を統合して密度の高い立体復元が可能であることがわかった。課題として、立体復元の精度を明らかにすること、アーチファクトとなる復元点の判別および除去する手法の開発が必要であること、特徴点抽出数の増加に伴う立体復元計算の処理コストの増加の 3 点を抽出した。

さらに、3 次元モデリングの代替手法として、イメージスティッチングによる高解像度画像の生成手法の適用性検証を行った。イメージスティッチングとは、複数の画像を相互に繋ぎ合わせて広い範囲をカバーする画像を合成する手法（stich、スティッチとは「縫う」を意味する。）の総称である。図 3.2-15 は、3 枚の画像に対してイメージスティッチング処理を適用した例を示している。

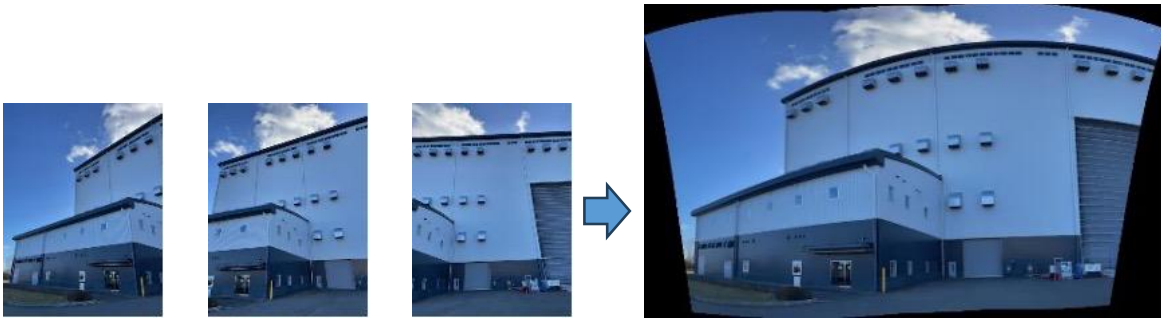


図 3.2-15　イメージスティッチングの例

令和 5 年度は、廃炉作業環境におけるイメージスティッチングの適用性を検証するために、1F において取得された動画から抽出された画像を対象にして実験を行った。イメージスティッチングの代表的な方法として画像をフィッティングするモデルを仮定して縫い合わせる手法が一般的である。その中の代表的な方法には、球面投影型モデルに基づいた合成方法とアフィン

投影型モデルに基づいた合成方法があり、それぞれのモデルを用いてイメージスティッチング計算を行った結果を比較することとした。

ここで、イメージスティッチングを行った計算機環境としては、計算機の OS に Ubuntu20.0.4、計算プログラムは C 言語と OpenCV4.2.0 を用いて実装し、GCC9.4.0 によってコンパイルして実行した。プログラムには、OpenCV4.2.0 の Stitcher クラスを用いて計算している。ここで、図 3.2-16 は東電 HD が公開している 3 号機 PCV 内部調査によって得られた動画[3.2-18] 170724_02j.mp4 の冒頭の約 8 秒間部分の画像 267 枚（30fps）を 5 枚おきに選択して、イメージングスティッチングを適用した結果について示している。

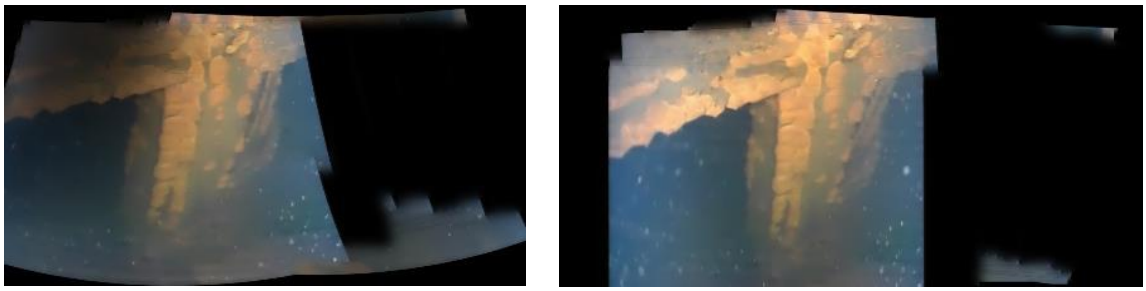


図 3.2-16　3 号機 PCV 内部調査画像への球面投影型モデル（左）
およびアフィン投影型モデル（右）の適用結果

図 3.2-17 は、東電 HD が公開している 1 号機 PCV 内部調査によって得られた動画[3.2-1] 220523_02j.mp4 の一部（4 分割の右上）をトリミングした約 6 秒間部分の画像 171 枚（30fps）を 4 枚おきに選択して、イメージングスティッチングを適用した結果について示している。

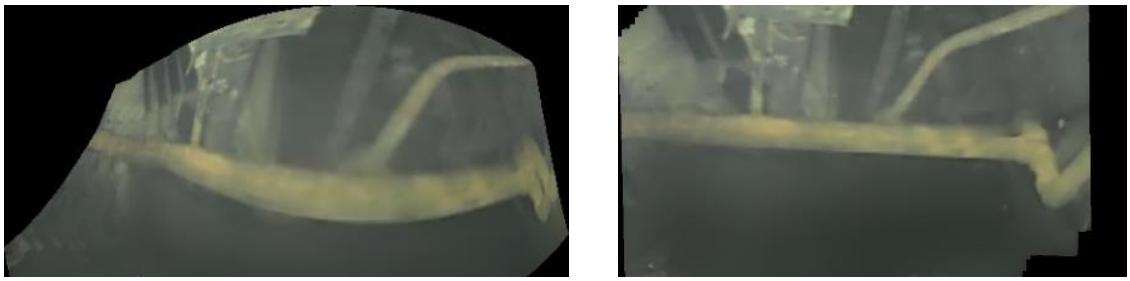


図 3.2-17　1 号機 PCV 内部調査画像への球面投影型モデル（左）
およびアフィン投影型モデル（右）の適用結果

今回対象にした動画は、短時間かつ移動距離がそれほど大きくないケースであるためどちらのモデルを用いた場合でもイメージスティッチング処理は成功している。図 3.2-16 の場合はどちらも同様な結果が得られているが、計算原理通り、球面投影モデルの場合は曲面に展開されている印象で、アフィン投影モデルは平面に投影されている印象が見て取れる。

また、図 3.2-17 の場合は球面投影型モデルを用いた結果は、動画での視認の印象に対して歪んだものになった。これらの結果より、ロボット等が移動しながら撮影を行うという条件では、撮像対象までの距離とカメラ-対象物の相対位置関係がイメージスティッチングによる 2 次元画像合成の性能に影響を与えることがわかった。今後は、今回適用した 2 つのモデル以外の合成方法についても実験的に検証して研究していく。また、移動量の少ない画像を計算対象としないことにより、計算処理の効率化についても考えていく必要がある。

(3)まとめ

最新の PCV 内部調査の動画像データを対象として、画像の特性分析および深層学習を用いた手法等を複合的に適用した画像特徴点抽出手法等の開発を実施した。これらの手法を組み合わせて、令和 6 年度に必要な研究開発の課題について抽出を行った。併せて、イメージスティッチングによる高解像度画像の生成手法および画像間のオプティカルフローに基づく遠隔操作機器の移動・停止の自動記録生成手法の適用性検証を実施した。

具体的には、1F PCV 内部調査画像の特性分析を行い、画像品質評価手法の導入により画像補正による画質スコアの改善と画像特徴点のマッチング数の向上に関連性があることを示した。また、深層学習を適用した画像特徴点を抽出する手法の調査・分析を行い、深層学習の適用により従来手法と比べて良好な結果が得られることを確認した。そして、課題として立体復元の精度を明らかにすること、アーチファクトとなる復元点の判別および除去する手法の開発が必要であること、特徴点抽出数の増加に伴う立体復元計算の処理コストの増加の 3 点を抽出した。さらに、イメージスティッチングの適用性検証の結果、カメラ軌道および投影モデルの選択がパノラマ画像合成の性能に与える影響を確認し、球面投影型モデル、アフィン投影型モデル以外の合成方法についても実験的に検証していくこととした。

参考文献

[3.2-1] 東電 HD 動画・写真ライブラリー, 写真集:福島第一原子力発電所 1 号機原子炉格納容器内部調査（水中 ROV-A2）の実施状況（5 月 19 日の作業状況）, https://photo.tepco.co.jp/date/2022/202205-j/220523-01j.html （参照：2024 年 3 月 31 日）.

[3.2-2] Venkatanath, N. et al., Blind image quality evaluation using perception based features, 2015 twenty first national conference on communications (NCC), IEEE, 2015. https://doi.org/10.1109/NCC.2015.7084843

[3.2-3] Lowe. D.G., Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60, 2004, pp.91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

[3.2-4] Rublee, E. et al., ORB: An efficient alternative to SIFT or SURF, 2011 International conference on computer vision, IEEE, 2011. https://doi.org/10.1109/iccv.2011.6126544

[3.2-5] OpenCV -Open Computer Vision Library, https://opencv.org/ （参照：2024 年 3 月 31 日）.

[3.2-6] 羽成ら, 廃炉作業における原子炉内環境モデリングに向けた画像品質評価, 第 24 回計測自動制御学会システムインテグレーション部門講演会, 3D3-03, 2023, pp.3237-3238. https://jglobal.jst.go.jp/detail?JGLOBAL_ID=202402263484738897 （参照：2024 年 3 月 31 日）.

[3.2-7] OpenCV document, Histograms - 2: Histogram Equalization, https://docs.opencv.org/4.1.1/d5/daf/tutorial_py_histogram_equalization.html （参照：2024 年 3 月 31 日）.

[3.2-8] Liu, C. et al., A review of keypoints’ detection and feature description in image registration, Scientific programming, 2021, 2021, pp.1-25. https://doi.org/10.1155/2021/8509164

[3.2-9] Sarlin, P.-E. et al., From coarse to fine: Robust hierarchical localization at large scale, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019. https://doi.org/10.1109/CVPR.2019.01300

[3.2-10] 東電 HD 動画・写真ライブラリー, 写真集:福島第一原子力発電所 1 号機 PCV 内部調査（ROV-A2）の実施状況（3 月 14～16 日の作業状況）, https://photo.tepco.co.jp/date/2022/202203-j/220324-01j.html （参照：2024 年 3 月 31 日）.

[3.2-11] 東電 HD 動画・写真ライブラリー, 写真集:福島第一原子力発電所 1 号機原子炉格納容器内部調査の実施状況（2 月 10 日時点）, https://photo.tepco.co.jp/date/2022/202202-j/220210-01j.html （参照：2024 年 3 月 31 日）.

[3.2-12] 東電 HD 動画・写真ライブラリー, 写真集:福島第一原子力発電所 1 号機原子炉格納容器内部調査（ROV-A2）の実施完了（2023 年 3 月 28 日～30 日の作業状況）, https://photo.tepco.co.jp/date/2023/202304-j/230404-01j.html （参照：2024 年 3 月 31 日）.

[3.2-13] DeTone, D. et al., SuperPoint: Self-supervised interest point detection and description, Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2018. https://doi.org/10.1109/CVPRW.2018.00060

[3.2-14] Dusmanu, M. et al., D2-Net: A Trainable CNN for Joint Detection and Description of Local Features, IEEE Conference on Computer Vision and Pattern Recognition, 2019. https://doi.org/10.1109/CVPR.2019.00828

[3.2-15] Revaud, J. et al., R2D2: Reliable and repeatable detector and descriptor, Advances in neural information processing systems, 32, 2019. https://proceedings.neurips.cc/paper_files/paper/2019/hash/3198dfd0aef271d22f7bcddd6f12f5cb-Abstract.html （参照：2024 年 3 月 31 日）.

[3.2-16] Tyszkiewicz, M. et al., DISK: Learning local features with policy gradient、Advances in Neural Information Processing Systems, 33, 2020, pp.14254-14265, https://proceedings.neurips.cc/paper_files/paper/2020/hash/a42a596fc71e17828440030074d15e74-Abstract.html （参照：2024 年 3 月 31 日）.

[3.2-17] Sarlin, P.-E. et al., SuperGlue: Learning feature matching with graph neural networks, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020. https://doi.org/10.1109/CVPR42600.2020.00499

[3.2-18] 東電 HD 動画・写真ライブラリー, 写真集:福島第一原子力発電所 3 号機　PCV 内部調査進捗　～22 日調査速報～, https://photo.tepco.co.jp/date/2017/201707-j/170722-01j.html （参照：2024 年 3 月 31 日）.

3.3　深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング
（再委託先：岩手県立大学）

(1)　令和 5 年度概要

岩手県立大学では、図 3.3-1 に示す研究計画の全体構成に沿って、パノプティックパーツセグメンテーションによるインスタンス分類を中心に、関連研究のサーベイ、要素技術やモデルの選定、ソフトウェアによる実装、GPU の導入とサーバの設定を含む実行環境の構築、各モデルの性能検証、精度の比較評価を、セグメンテーション用バックボーンの選定ならびにハードウェアの性能や制約条件に合わせた最適化と並行して、研究開発に取り組んだ。

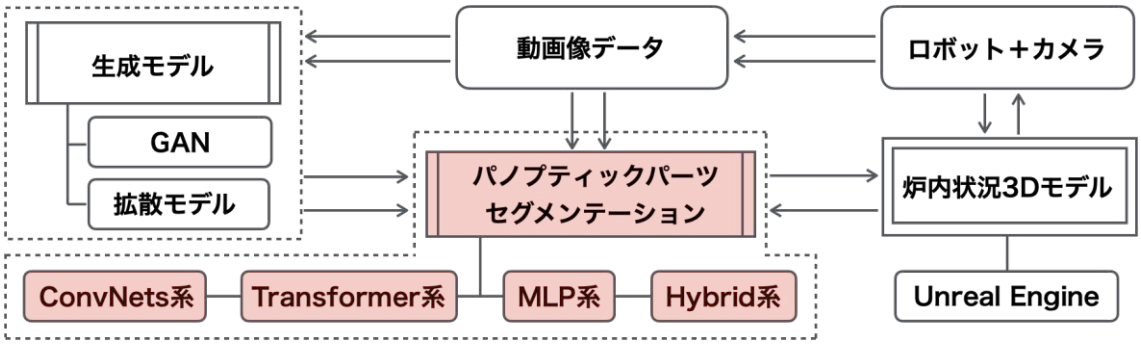


図 3.3-1 研究計画の全体構成と当該年度の研究実施対象（薄赤色部分）

(2)　令和 5 年度実施内容および成果

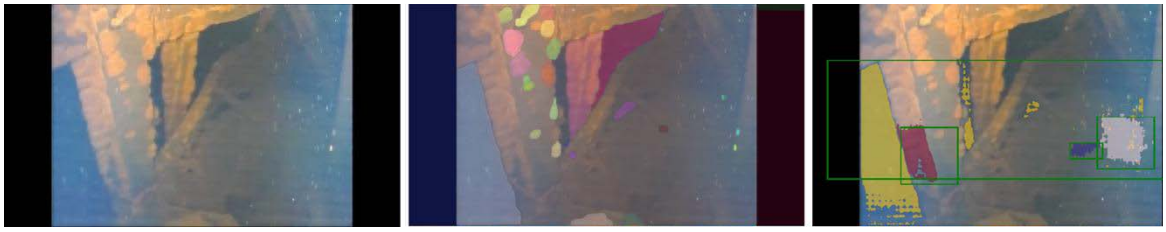
①　パノプティックパーツセグメンテーションによるインスタンス分類

本研究では、動画像データから SfM を用いて抽出された点群データを、パノプティックパーツセグメンテーションにより、インスタンスラベルが付された部品に分類した。以下にその詳細結果を示す。

1）Vanilla SAM によるセグメンテーション結果

Segment Anything Model（SAM）[3.3-1]は、Meta 社が提案したセグメンテーションのための基盤モデル[3.3-2]として、あらゆる物体のセグメンテーションを可能にするフレームワークである。令和 5 年に同社の公式ブログおよび GitHub と arXiv で公開された後、その特性と能力により、画像認識やセグメンテーションの幅広い分野における新たなパラダイムを開拓した。画像内の各ピクセルがどの物体に属するかを識別するタスクとして位置付けられるセグメンテーションにおいて、SAM の登場により、画像内の物体の形状や位置を詳細に理解することが実現された。従来のセグメンテーションモデル[3.3-3]は、特定のクラスに属するオブジェクトとスタッフをセグメンテーションの対象としていたが、SAM はその限界を打破し、あらゆる物体のセグメンテーションをファインチューニング[3.3-4]なしで実行することにより、セグメンテーションの応用範囲を大幅に拡張した。SAM の根幹を成す重要な特性は、大量のラベル付きデータを必要とせず、少量のラベル付きデータのみを用いて学習するメカニズムである。この学習方法により、ラベル付きデータの取得に伴うコストの大幅な削減を実現した。また、文脈内学習と類推により、SAM は新たなクラスの物体に対しても迅速に適応でき、その汎用性を一層高めている他、同時にその性能を評価するために、SA-1B と名付けられた新たなベンチマークも提案した。SA-1B は、従来のセグメンテーションのベンチマークが包含していない多様な物体とシーンを含んでいるため、SAM の性能をより広範で公正な基準で評価することが基盤モデルとして実現された。

図 3.3-2 に Vanilla SAM によるインスタンスセグメンテーション結果を示す。左パネルに示す 1F の公開 ROV 映像から切り出した任意の静止画像を対象とした。中央パネルに画像の全領域の全画素を対象としたセグメンテーション結果を示す。透明度が低く照明の照射範囲の限られた水中画像のため、画像の端付近の物体はセグメンテーション対象から除外され、中央から左側付近の上部にインスタンスが集中している。右パネルには、矩形で指定できるバウンディングボックス（Bounding Box：BB）を設定し、BB 内部のみをセグメンテーション対象とした結果を示す。



（左：原画像、中央：全領域に対する結果、右：BB 内のみの結果）

図 3.3-2　Vanilla SAM によるインスタンスセグメンテーション結果（出典：東電 HD）

Vannilla SAM には、ポジティブポイントとネガティブポイントの指定により、ユーザの意図を最小限のインタラクションで指示する機能が装備されている。溶融し水棺状態の物体に対して、画素単位のアノテーションを Ground Truth（GT）として施すことは現実的ではないため、本研究では当該ポイント指定によるセグメンテーションの効果を検証した。

図 3.3-3 にポジティブポイントのみ、ポジティブとネガティブの両方のポイント、BB とネガティブポイントの組み合わせの 3 種類のセグメンテーション結果を示す。左パネルに示すポジティブポイントのみのセグメンテーション結果では、画像中の緑色星マークがマウスクリックにより指定したポジティブポイントであり、画像の左側の対象物に対して、同一の性質を有すると推測される領域がセグメンテーションされている。中央パネルの結果は、画像中央付近をネガティブポイントとして指定した結果である。赤色の星マークが、ネガティブポイントに該当する。セグメンテーションされない領域に差異が発生しなかったことから、右パネルに示す実験結果では、緑色の矩形枠で示す BB 内部に対して、左パネルおよび中央パネルのセグメンテーションでは、ポジティブと指定された領域の一部をネガティブに指定した。その結果、柱状の構造物の左右において、異なるセグメンテーション結果が得られた。

（左：ポジティブポイントのみ、中央：ポジティブとネガティブの両方のポイント、
右：BB とネガティブポイントの組み合わせ）

図 3.3-3　Vanilla SAM によるポイント指定のインスタンスセグメンテーション結果
（出典：東電 HD）

2）FastSAM によるセグメンテーション結果

SAM の課題として、計算負荷の高さが挙げられる。GPU を潤沢に使えない環境や即時に結果を必要とする場合には、SAM の計算負荷がボトルネックとして顕在化する。大量の ROV 映像に対する実時間処理を考慮して、本実験では FastSAM について評価した。Transformer 系アーキテクチャを用いている Vanilla SAM に対して、FastSAM はこの問題の解決を目的として、セグメントタスクを全インスタンスのセグメンテーションとプロンプトによる選択の 2 ステップに分解するアプローチを採用している。具体的には、Convolutional Neural Networks（ConvNets）系列の物体検出器として名高い You Only Look Once（YOLO）ファミリ[3.3-5]の YOLOv8-seg を用いて画像内の全てのインスタンスをセグメンテーションした後、ポイントプロンプト、ボックスプロンプト、テキストプロンプトによる選択を行うことで、対象物体を特定する手法の FastSAM は、Vanilla SAM と同等の性能を維持しながら 50 倍程度の高速化が実現されている。さまざまなベンチマーク課題において FastSAM と Vanilla SAM を比較評価した結果、FastSAM が高速でありながら Vanilla SAM とほぼ同等の性能を達成できることを実証している。さらに異常検出や顕著性の高いオブジェクト分割など、実世界のアプリケーションにおいても FastSAM の有用性が示されている。

FastSAM によるセグメンテーション結果を図 3.3-4 に示す。FastSAM では、BB とインスタンスセグメンテーションが同時に実行される。抽出されるインスタンスの粒度を支配するパラメータを変更したのが左パネルと中央パネルの結果である。中央パネルと同一温度で、異なるシーンに適用した結果が右パネルである。注目度の高いパーツが抽出されているが、他の部品との区分に関しては、FastSAM のみでは解析できない。また、高速化を優先する関係上、バックボーンが ConvNets 系統のため、アテンションに関する概念と枠組みを持たない制約から、固定的な範囲と関連性でのフィードバックによる表現に限定される。

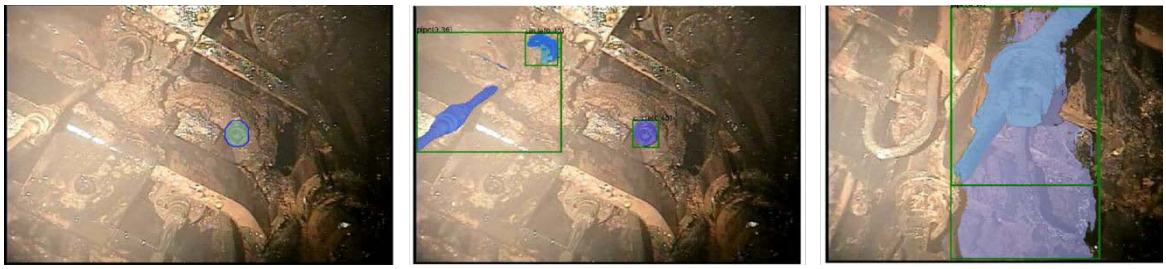


図 3.3-4　FastSAM によるインスタンスセグメンテーション結果（透明度の高いシーン）
（出典：東電 HD）

続いて、徐々に透明度が低くなる映像に対して、FastSAM を適用した。図 3.3-5 に示す結果では、正方形に近い部品のみならず、長棒状の部品も抽出しており、セグメンテーションの機能的な動作が確認できた。しかしながら、透明度がさらに低い画像に当該モデルを適応した図 3.3-6 に示す実験結果では、BB が画像全体に設定されている。この結果は、水をカテゴリとして抽出したと推測できるが、対局的に水はオブジェクトでなくスタッフに属するため、インスタンスセグメンテーションとセマンティックセグメンテーションが混在したパノプティックセグメンテーションの結果と位置付けられる。一方、水の BB とオーバラップする関係でインスタンスが抽出されている。温度を変化させた場合に抽出対象も変化しているが、中央からやや右側に位置するナットに関しては、いずれの結果においても抽出されている。ただし、左パネルの結果では、ナットが 2 領域のインスタンスに分割されている。

図 3.3-5　FastSAM によるインスタンスセグメンテーション結果（透明度が低下したシーン）
（出典：東電 HD）

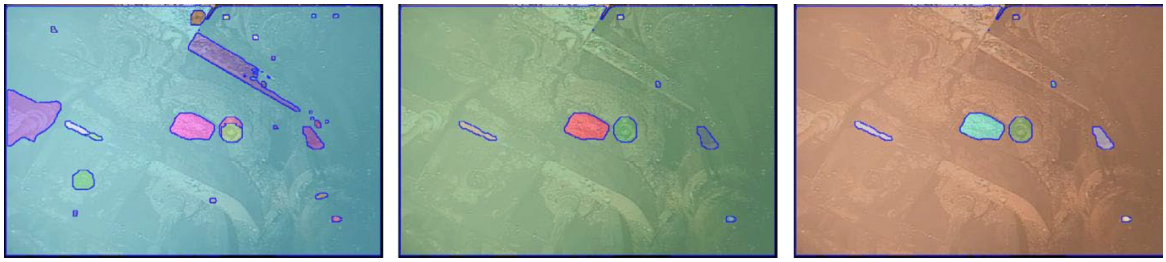


図 3.3-6　FastSAM によるインスタンスセグメンテーション結果（透明度が低いシーン）
（出典：東電 HD）

3）Semantic-SAM によるセグメンテーション結果

任意の粒度で物体や物体の一部をセグメント化し認識することができるSemantic-SAM[3.3-6]は、画期的な画像セグメンテーションモデルとして登場した。Semantic-SAM は、2 種類の主要な特性から構成されている。1 番目として、異なる粒度の 3 種類のデータセットを統合し、物体とその部分の分類を分離させることで、豊富な意味的情報をモデルに取り込むことを実現している。2 番目として、学習時に複数選択学習手法を導入したことにより、クリック等の単一のインタラクションで、複数レベルのマスクを生成するとともに複数の正解マスクに適用する多粒度対応能力である。Semantic-SAM では、Vanilla SAM において提案された SA-1B データセットに加えて、一般的なセグメンテーション、部分セグメンテーションデータセットを統合して学習させることにより、モデルがセマンティック認識能力と多粒度対応能力を有していることが確認できた。さらに、SA-1B データセットを他のセグメンテーションタスク、例えばパノプティックセグメンテーションや部分セグメンテーションと組み合わせることで、Vanilla SAM と比較して性能の向上を実現している。さらに、Semantic-SAM は、従来のセグメンテーションモデルが抱えていた粒度の固定化に伴う問題を解決し、さまざまな観点から物体を認識する他、必要な粒度でのセグメンテーションを実現した。

Semantic-SAM による実験結果を図 3.3-7 に示す。各結果の左パネルが原画像、中央パネルがセグメンテーション結果、右パネルがラベルの付された結果である。中央パネルと右パネルにおいて、インスタンスの粒度が異なるのは、ラベルを割り振るだけの領域が形成されていない微小インスタンスが、右パネルの結果では省略されているからである。ラベルは SAM による事前学習により形成された範囲内で割り振られていることから、1F のデータセットに基づくファインチューニングにより、文脈学習を超える再ラベリングの可能性を検討することが今後の新たな課題である。一方、再ラベリングに関連するアノテーションの労力と計算負荷とのトレードオフが、連鎖発生する新たな課題として現出する。ファインチューニング後の後処理としては、言語モデルを用いて RAG（Retrieval Augmented Generation）[3.3-7]の枠組みの導入により、再学習せずに再ラベリングが実現できると考えられる。

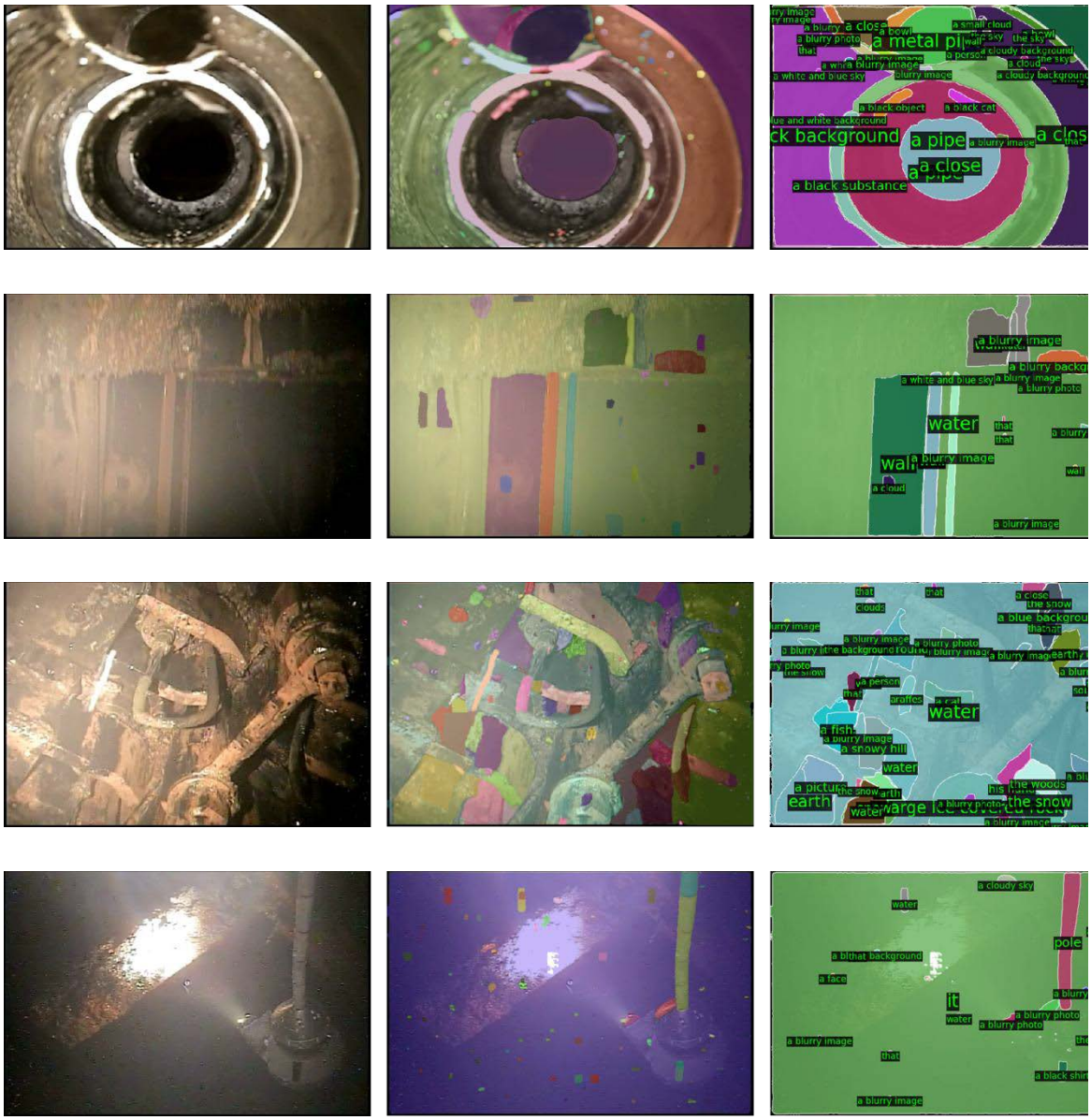


図 3.3-7　SemanticSAM によるラベル付きインスタンスセグメンテーション結果
（出典：東電 HD）

② セグメンテーション用バックボーンの選定【令和 5 年度】

令和 5 年度の本研究では、セグメンテーションに使用するバックボーンを性能、速度、メモリ使用量、拡張性、実装の容易さなど、複数の要素を考慮しつつ選定した。以下にその詳細結果を示す。

深層学習のネットワークモデルは、図 3.3-8 に示す通り、大別して、バックボーン、ネック、ヘッドからの 3 モジュールから構成される。この中で、バックボーンは、深層学習モデルの中心的な部位であり、主に特徴を抽出するために使用される[3.3-8]。バックボーンは、画像、テキスト、音声などの入力データを受け取り、特徴マップを生成するためのニューラルネットワークアーキテクチャを指す。バックボーンに対して、ヘッドはバックボーンが出力した特徴マップを処理し、タスク固有の出力を生成するために使用される。ヘッドは、通常はバックボーンの出力を平坦化し、全結合層や畳み込み層を追加して、最終的な出力を生成するためのニューラルネットワークアーキテクチャを指す。また、ネックはバックボーン

とヘッドの間にある部分であり、特徴マップを処理して情報を圧縮するために使用される。加えてネックは、バックボーンからの特徴マップを受け取り、畳み込み層やプーリング層を使用して特徴を抽出し、次元を削減する。ネックにおけるモデルパラメータ削減は、計算効率の向上に寄与する。バックボーン、ヘッド、ネックの各パートは、深層学習モデルの中で密接に連携しており、それぞれが重要な役割を果たしている。特にバックボーンは、特徴の抽出に特化し、ネックは特徴を処理し、ヘッドは特定のタスクに応じた出力を生成することに特化している。

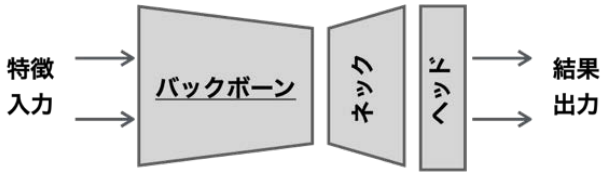


図 3.3-8　深層学習モデルの基本構造（バックボーン、ネック、ヘッド）

深層学習におけるバックボーンは、主に 4 種類の構造が存在する。古典的なバックボーンが、Convolutional Neural Network（ConvNets）系[3.3-9]である。ConvNets の歴史は古く、その基本構造として、昭和 55 年に Fukushima が Neocognitron[3.3-10]を提案している。ConvNets 系バックボーンの構造は、畳み込み層とプーリング層を組み合わせた階層構造を採用しており、主に画像や時系列データなどに対して効果的であり、なおかつ局所的な相関関係を捉える能力を備えている。

次に挙げられるのが Transformer 系[3.3-11]の構造であり、Self-Attention メカニズムを基盤とした全結合層の積み重ねから構成されている。Transformer 系バックボーンは元来、平成 29 年に自然言語処理向けに開発されたが、令和 2 年にパッチに分割された画像に適用した Vision Transformer（ViT）[3.3-12]が、視覚情報処理分野に新たなパラダイムを与えた。Transformer 系の特徴は、長距離の依存関係を捉えることが可能であり、順序付けられたデータに適していることである。

3 番目に挙げられるのが Multi-Layer Perceptron（MLP）系であり、単純に全結合層を積み重ねた構造から形成されている。MLP の歴史は最も古く、その名が示す通り Rosenblatt が昭和 32 年に提案したパーセプトロン[3.3-13]まで遡る。ただし、その当時は線形分離問題すら解けない状態であり、誤差逆伝播法[3.3-14]の登場を待つまで冬の時代が続き、ConvNets へと時代は切り替わったが、現在の MLP は、大規模化により、全結合のネットワーク構造が見直され、再び注目を集めている。MLP 系の特徴は、構造が単純でありながら、計算資源が比較的少なくて済み、さらに非線形変換に強いことである。

4 番目として、これらの 3 種類を組み合わせたものが、Hybrid 系の構造である。Hybrid 系の最大の特徴は、異なる性質を持つ構造を組み合わせることで、より高い表現力を実現できる点にある。ConvNet 系は長年にわたり、画像認識などの視覚タスクにおいて主流を占めていたが、長距離依存性を捉えにくいという課題が存在した。一方、Transformer 系は長距離依存性を捉える能力を有するものの、言語モデルに由来するため、2 次元構造を直接扱うのが難しいという課題が残されていた。また、MLP 系は構造が単純過ぎて表現力に乏しいという課題を抱えている。そのため、Hybrid 系が登場し、異なるアプローチの長所を組み合わせることで高い表現力を実現しようとするアプローチが試みられた。本実験では、表 3.3-1 に示す 3 系統のバックボーンを比較した。なお、Hybrid 系のバックボーンに関しては、これらを組み合わせた。

表 3.3-1　各系統のバックボーンにおける代表モデル

系	モデル名	原著論文	出版年
ConvNet	PointNet	[3.3-15]	平成 29 年
	PointNet++	[3.3-16]	平成 29 年
	MinkowskiNet	[3.3-17]	令和元年
	PointNeXt	[3.3-18]	令和 4 年
Tranformer	PointFormer	[3.3-19]	令和 3 年
	Point Transformer	[3.3-20]	令和 3 年
	Point Transformer V2	[3.3-21]	令和 4 年
	StratifiedFormer	[3.3-22]	令和 4 年
	Swin3D	[3.3-23]	令和 4 年
	PointGPT	[3.3-24]	令和 5 年
	OctFormer	[3.3-25]	令和 5 年
MLP	PointMLP	[3.3-26]	令和 4 年

ベンチマーク実験結果では、PointNet++と PointTransformer は、原著論文と同様の平均 Intersection over Union (IoU) が得られた。これらの手法は、ポイントクラウドデータを効果的に処理し、高い精度で物体検出やセグメンテーションなどのタスクを達成することが示唆される。一方、PointNeXt は、検証に用いた GPU サーバにおいて、最大性能を引き出すプログラムが正常に動作しないという問題が発生した。これは、NVIDIA 社製の計算プラットフォームであり、深層学習分野では事実上の標準としての地位を築いている Compute Unified Device Architecture (CUDA) とそのハードウェアの制約や環境の影響によるものであり、後述する本研究事業において構築した GPU サーバであれば、PointNeXt の動作することを確認した。また、CUDA11 に対応した RTX A6000 が 3 台搭載された Workstation も存在し、当該環境では PointNeXt が動作することを確認した。これらの GPU は、より高い処理能力を持ち、PointNeXt が正常に動作する可能性があり、適切なハードウェア環境で使用することにより PointNeXt の効果的な利用が期待される。これらを踏まえて本ベンチマークでは、PointNeXt の使用が優位との結論に達したが、環境制約を排除できれば Hybrid 系バックボーンへの移行に結び付くと考えているものの、令和 6 年度での課題として繰り越す結論に至った。

③ ハードウェアの性能や制約条件に合わせた最適化【令和 5 年度】

令和 5 年度の本研究では、ハードウェアの性能や制約条件に合わせた最適化を施し、処理速度と精度のトレードオフ関係を最適化した。以下にその詳細結果を示す。ここでは、構築したハードウェアの基本性能を示した後、1F の ROV 映像から切り出した連続画像を対象として、廃炉作業における画像の高精度化と 3 次元化の予備実験を踏まえて、生成拡散事前学習モデルを利用して劣化画像のブラインド画像復元タスク（超解像タスク）を行うフレームワークである Diffusion Based Blind Image Restoration (DiffBIR) [3.3-27]と生成拡散事前学習モデルを利用して劣化画像のブラインド画像復元タスクもしくは超解像タスクを行うフレームワークであり、複雑なシーンの視点を合成するための最先端の結果を達成する方法の Neural Radiance Fields (NeRF) [3.3-28]を用いた評価結果を示す。

1）構築した GPU サーバ環境

構築した GPU サーバ環境を図 3.3-9 に示す。本システムは、NVIDIA 社製の H100 を内蔵する GPU サーバと A100 を 2 枚内蔵する GPU サーバから構成される。サーバラックの下側には、停電に備えて、UPS を 2 台装備している。また、1F の映像を大規模に格納し、高速に各 GPU へとデータを受け渡しするために、SSD サーバをサーバラックの上部に備えている。GPU に関しては、NVIDIA 社の公式資料に基づく性能情報によると、H100 と A100 は、それぞれ同社の最新世代の GPU アーキテクチャである Hopper 世代と Ampere 世代に基づいた高性能のフラグシップ GPU である。H100 は 4 nm プロセスを採用し、144 個の TensorCore と 80 GB の HBM3e メモリを搭載しており、4 Tensor PFLOPS のテンソル性能と優れた電力効率を実現している。一方、A100 は 7 nm プロセスで 108 個の TensorCore、40/80 GB の HBM2e メモリを搭載し、1.9 Tensor PFLOPS のテンソル性能を持ち、訓練と推論の両方で高性能を発揮する汎用的な製品で、AI 応用のみならず、クラウドサービスでも広く採用されている。

図 3.3-9　構築した GPU サーバ環境（左：正面、中央：側面、右上：H100、右下：A100×2）

両者のプロセスとアーキテクチャの違いは、電力効率にも影響を与えており、H100 は A100 に比べてワット当たりの演算性能が高い。カタログ上の処理速度差は、AI 推論の場合、H100 は最大 4.5 倍、モデル訓練でも最大 2 倍の高速化が謳われているが、実際のベンチマークでは、1.2～1.4 倍程度に留まっており、両者の価格差を鑑みた場合には、H100 が A100 の 2 倍程度であることから、メモリ空間が同一であることを踏まえて、本研究では、両方の GPU の導入に至った。ただし、令和 5 年末の段階で A100 の生産が打ち切られたことから、令和 6 年度以降の GPU 増強には、H100 を追加する計画であるが、価格性能比では、H100 が A100 に劣るため、Ampere 世代もしくは Ada 世代を並行して導入する計画である。

2）DiffBIR によるベンチマーク結果

マスク画像復元のための革新的なアプローチが導入された DiffBIR は、事前に学習されたテキストから画像への拡散モデルを復元プロセスに統合することで注目された。当該フレームワークは、実世界のシナリオに対処するために設計された 2 段階のパイプラインを介して操作される。最初の段階では、著者らは多様な劣化に対応できる復元モジュールを事前に訓練することに焦点を当てている。この段階では、モデルの汎化能力を向上させ、さまざまな実世界シナリオで効果を確保されており、また、その後潜在的な拡散モデルの生成能力を活用して現実的な画像復元を実現した。具体的には、インジェクション型調整サブネット（LAControlNet）の導入が微調整を促進し、事前に訓練された拡散モデル[3.3-29]に基づくコンポーネントがその生成力を保持する。さらに、復元品質と忠実度のバランスを実現する

ための操作可能なモジュールを導入された。これは、推論中のノイズ除去プロセスに潜在的な画像ガイダンスを組み込むことで達成され、利用者が望む結果を達成する柔軟性を提供した。原著論文において例証されている広範な実験的評価結果は、DiffBIR が各種既存モデルに対して優越していることに加えて、特に盲目的な画像超解像度および盲目的な顔の復元タスクにおいて、その有効性が示された。これらの評価は、合成および実世界のデータセットの両方で行われ、提案されたフレームワークの堅牢性と効果を示唆した。これらの特徴から、DiffBIR は、盲目的な画像復元の領域で有望なアプローチを提供し、拡散モデルからの先進的な技術を活用して実世界の課題に効果的に対処した。

DiffBIR による画像復元結果を図 3.3-10 に示す。左パネルが原画像、右パネルが 50 ステップ目の復元結果である。Intel 社製 13 世代の i9-13900k を使用した場合、0.5～1.0 秒の範囲で処理できた。なお、当該 CPU を搭載した計算機環境の RAM は 5 世代の DDR メモリの 64 GB であった。この結果から、H100 もしくは A100 を使う計算負荷ではなかったため、NVIDIA 製の RTX 4090 を使用して画像復元を施したところ、0.01 秒以内に処理が完了したことから、実時間処理の対象として問題のない範囲であることを確認した。なお、RTX 4090 のメモリ容量は、24 GB である。

図 3.3-10　DiffBIR による復元結果（左：入力画像、右：出力画像）
（出典：東電 HD）

鮮明度の低い画像に、DiffBIR を適用した結果を図 3.3-11 に示す。画像の右上および右下にスーパーインポーズされている文字と数値を含めて、画質が改善している。ただし、古典的な機械学習法と異なり、深層学習モデルでは、End-To-End のネットワークにより内部表現の一部としてこのような前処理が施されることから、定性的に画像の復元が確認できたとしても、セグメンテーションの精度として定義されている評価指標での改善に結び付かないことが多々あり、逆に値が低下する現象も内在している。

図 3.3-11　DiffBIR による低品質画像の復元結果（左：入力画像、右：出力画像）
（出典：東電 HD）

(3) NeRF によるベンチマーク結果

本研究では、画像の画質復元に加えて、3 次元空間の情報が 2 次元空間へと縮退している通常のカメラの映像から、深層学習に基づく 3 次元復元を実現することにより、ステレオカメラ、深度カメラ、レンジセンサ、Light Detection And Ranging（LiDAR）を使わないアプローチの実現と現場実装を目指している。しかしながら、3 次元復元では、ポイントクラウドとしての特徴入力に加えて、GPU による演算が必須と位置付けられている NeRF をベンチマーク対象とした。NeRF の処理過程では、カメラ位置を推定する COLMAP が含まれているため、令和 6 年度以降の取り組み課題の一部としても、その基礎特性を評価した。

画像生成手法として Mildenhall らにより令和 2 年に発表された NeRF の核心概念は、3 次元空間内の各点における放射輝度をモデル化することであり、このフレームワークにより、高品質でリアルな画像の合成が実現されている。NeRF の原著論文では、3 次元シーンを各点における放射輝度を近似するニューラルネットワークによってモデル化する手法として定義されており、任意の視点からのリアルな画像を合成できる。NeRF は、従来の手法と異なり、3 次元シーンを直接モデル化することによって、高品質な画像生成を切り拓いた。一方、従来モデルでは、画像生成において 3 次元空間の構造を明示的にモデル化することが難しく、多くの場合、複雑な手法や大量のデータが必要であったが、NeRF では各点の放射輝度をニューラルネットワークで表現することで、3 次元空間内の複雑な構造を柔軟に捕捉した。さらに、NeRF は視点合成（view synthesis）というタスクに特化しており、与えられた 3 次元シーンから任意の視点における画像を生成することができる。これは、従来の手法では困難であった視点間の一貫性を保持しつつ、リアルな画像を合成することができるという点で画期的である。このように、NeRF は 3 次元シーンの表現と画像生成において革新的な手法であり、その性能は実世界の多様なシーンにおいても高い評価を受けている。

NeRF に入力した画像の一部を図 3.3-12 に示す。これらの画像に対して COLMAP を施した。ここで、COLMAP は、NeRF において重要な役割を果たす構造から動画を復元するための構造化画像技術である。COLMAP は、入力された複数の画像から特徴点を検出しマッチングを行い、対応する特徴点の 3 次元座標と各カメラの位置・方向を同時に推定することでスパースな 3 次元点群を復元する。この特徴点検出とマッチングでは、スケール不変特徴点検出器などの高度なコンピュータビジョンアルゴリズムが用いられ、ロバストな点対応が得られる。さらに、行列からの分解や最適化手法を用いて、点群とカメラ位置が同時に精密に推定され、その後、このスパースな点群を密なデプスマップに変換し、3 次元メッシュに統合する。ここでは、ステレオ視差や多視点ステレオ法を用いて、点群からデプスマップを密に推定し、ボク

セル解像度で表現される 3 次元メッシュに変換される。最終処理としては、入力画像からテクスチャを投影することで、テクスチャ付きの高解像度な 3 次元モデルを生成する。テクスチャリングでは、画像からの射影と視点ブレンディングの手法が用いられ、一貫性のあるテクスチャが付与される。NeRF では、この COLMAP から得られる高精度な幾何構造とカメラ位置の推定値が、ニューラルネットワークの学習における 3 次元シーン表現の初期値として利用される。なお、COLMAP の出力は、ニューラルネットワークの座標変換やレイサンプリングに使われ、ニューラルレンダリングによる微調整が行われるフレームワークである。一方、COLMAP は、NeRF における幾何的な基礎を提供する重要な前処理であるものの、後続する 3 次元生成の結果に影響を与えるため、近年では COLMAP を使用しない処理（NoPe-NeRF[3.3-30]、BARF[3.3-31]、NeRF--[3.3-32]など）が検討されており、本研究においても、令和 6 年度以降の実装課題に位置付けている。

図 3.3-12　NeRF の各モデルへの入力画像
（出典：東電 HD）

NeRF には、さまざまなモデルが既に多数提案されている。本実験では、NeRF のフレームワークとして Nerfstudio[3.3-33]を使用した。Nerfstudio では、Vanilla NeRF において課題とされていた膨大な計算負荷に対して、高速性と低メモリの両方を実現した Nerfacto が提供されている。Nerfacto では、空間分割データ構造の導入し、シーンを階層的な空間分割データ構造として Octree 等で表現することで、密度の疎な領域を効率的にスキップしている。このメカニズムにより、レンダリングの計算量が大幅に削減されるとともに入力の 3 次元座標を直接ネットワークに渡すのではなく、座標の特徴量を事前に畳み込みにより符号化する。このエンコーディングにより、ネットワークの入力チャンネル数が削減され、メモリ使用量と計算コストが抑制されている。

本実験では、Vanilla NeRF より 1 桁以上高速に学習でき、シーンの詳細な表現能力は若干劣る側面があるものの、10B レイ級の大規模シーンも扱える Nerfacto を用いた。Nerfacto による 3 次元復元結果を図 3.3-13 に示す。左右のパネルに視点を変えた画像を表示している。Nerfacto の短所である中心付近以外の疎な特徴に対しては、解像度が低下していることが確認できる。なお、計算時間は 48 GB のメモリ空間を備える NVIDIA 社製 GPU の A6000 で、数分間程度であった。一方、精度の高い NeRF の派生系としてサーベイ論文等で評価されている Mip-NeRF[3.3-34]を使用した場合は、処理時間が 60 時間まで増加した。なお、NeRF がピクセルごとに単一の光線をサンプリングするのに対し、Mip-NeRF は連続的なスケールでシーンを表現し、光線の代わりにアンチエイリアス化された円錐台を効率的にレンダリングすることで実現されている。エラー率の低減のアドバンテージを有するものの、シーンを連続的なスケールで表現するための追加の計算と処理が必要で、なおかつ Mip-NeRF は Vanilla NeRF がピクセルごとに単一の光線をサンプリングするのに対し、アンチエイリアス化された円錐台をレンダリングする。一方、詳細な表現を可能とするが、レンダリング過程に追加の制約を引き起こす。

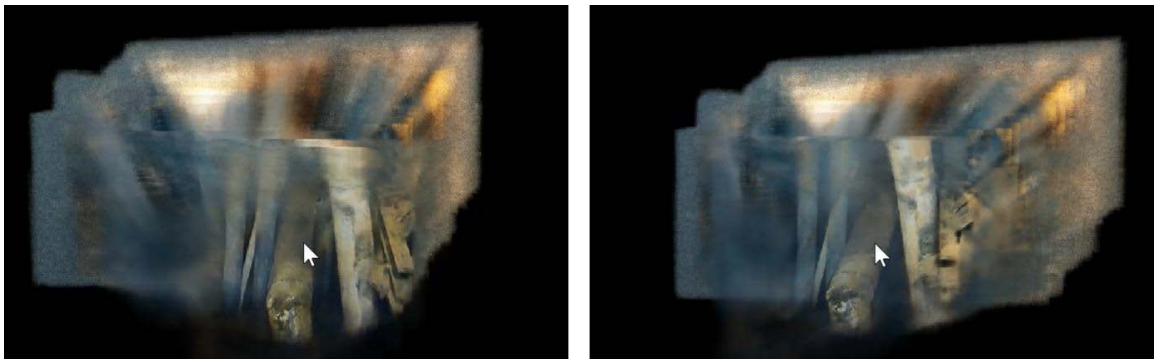


図 3.3-13　Nerfacto による復元結果
（出典：東電 HD）

1F の ROV 映像は、水中シーンから構成されるため、本研究では、Mip-NeRF ではなく NeRF in the wild[3.3-35]の一部として提案されている SeaThru-NeRF[3.3-36]の処理性能をベンチマークの一端として評価した。水中や霧がかった場合における物体の見え方への影響については、未だ解決されていない課題が残されている。従来の NeRF やその派生モデルでは、このような散乱媒体中のシーンは考慮されていなかった。しかしながら、NeRF はボリュームレンダリングに基づいているため、適切にモデル化されれば媒体の影響を内在的に考慮する能力を持っている。本研究では、散乱媒体中の NeRF に対する新しいレンダリングモデルを開発した。このモデルは、水中透過の画像生成モデルに基づいており、シーン情報と媒体パラメータの両方を学習可能な適切なアーキテクチャを提案している。シミュレーションと実環境のデータセットを用いて、我々の手法の有効性を実証した。水中での新規視点から生成された画像は、非常に現実的な光追跡が再現されていた。さらに興味深いことに、カメラとシーンの間の散乱媒体を除去し、媒体によって大きく遮蔽されていた遠方の物体の外観と奥行きを再構成することができた。SeaThru-NeRF による実験結果を図 3.3-14 に示す。9 種類の視点からの画像を表示しているが、これらはいずれも今回の映像のカメラにはない視点である。

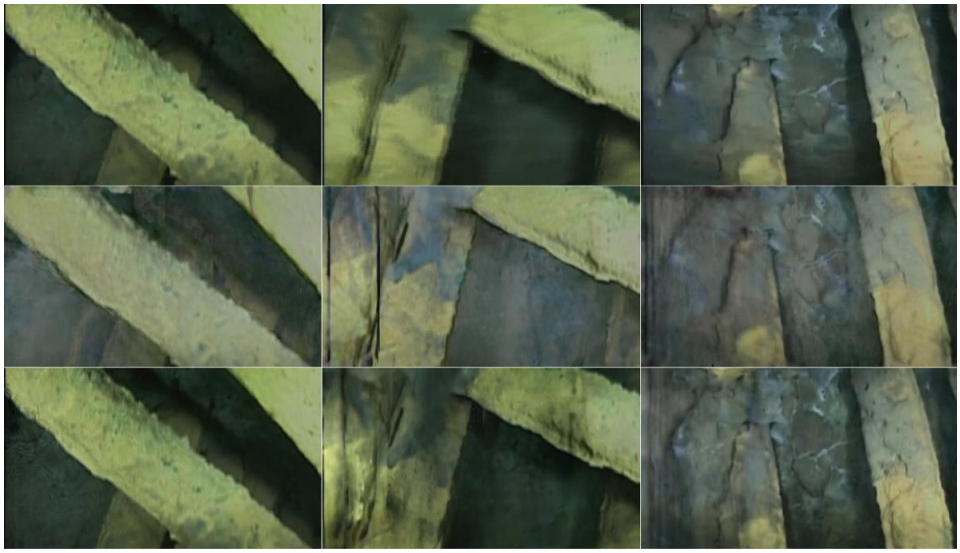


図 3.3-14　SeaThru-NeRF による復元結果
（出典：東電 HD）

定量的な評価として、3 種類の画像品質測定法を用いた結果を表 3.3-2 に示す。画像圧縮など非可逆圧縮を使ったコーデックの再現性の品質尺度である Peak Signal-to-Noise Ratio（PSNR）は画像の品質を予測する方法で、デジタルテレビや映画の画像など、他の種類のデジタル画像とビデオの知覚品質を予測するための方法の SSIM（Structural Similarity Index Measure）[3.3-37]、AlexNet や VGG などの学習済み画像分類ネットワークの畳み込み層が出力する特徴量をもとにした基準の Learned Perceptual Image Patch Similarity（LPIPS）[3.3-38]の 3 種類の評価指標に対して、Nerfacto と SeaThru-NeRF の結果を比較した。小数第 2 位の有効数字における各指標の平均値は、PSNR（Nerfacto）が 18.59、PSNR（SeaThru）が 32.80、SSIM（Nerfacto）が 0.66、SSIM（SeaThru）が 0.92、LPIPS（Nerfacto）が 0.35、LPIPS（SeaThru）が 0.15 であった。個々のフレームでは各値の変動が生じているが、いずれの手法においても、SeaThru-NeRF による結果の優位性が示された。ただし、予備検証の位置付けでの本実験では、評価対象としたデータ数が限られていることから、現在入手している 600 GB の全映像での評価を、本年度の取り組みにおいて構築した GPU サーバによる学習環境に計算負荷に応じて段階的に適用する計画である。

表 3.3-2　3 種類の画像品質測定法を用いた比較結果

ファイル名	PSNR(nerfacto)	PSNR(seathru)	SSIM(nerfacto)	SSIM(seathru)	LPIPS(nerfacto)	LPIPS(seathru)
frame_00010	18.7442	35.2687	0.7258	0.917	0.3674	0.2023
frame_00020	21.3456	30.5311	0.7235	0.9185	0.3258	0.1575
frame_00030	21.6116	31.0471	0.7254	0.9208	0.3207	0.1546
frame_00040	21.7218	30.5922	0.728	0.9136	0.3052	0.1444
frame_00050	19.4407	35.7952	0.6571	0.9516	0.3219	0.0934
frame_00060	19.7209	34.3669	0.6612	0.9475	0.3142	0.0917
frame_00070	18.0244	36.4718	0.6332	0.9432	0.3655	0.116
frame_00097	17.1821	35.8294	0.6263	0.9584	0.298	0.08
frame_00107	17.0033	35.2383	0.6191	0.9593	0.3064	0.0836
frame_00117	17.6334	25.1911	0.6536	0.8578	0.4665	0.3416
frame_00204	17.2468	33.3747	0.6543	0.9508	0.4728	0.0701
frame_00214	18.7905	34.8203	0.7073	0.9437	0.2987	0.1055
frame_00224	22.1615	25.6835	0.7596	0.8824	0.3815	0.2439
frame_00255	16.6556	19.59	0.5732	0.6684	0.6684	0.5626
frame_00267	16.8942	35.0403	0.6125	0.9598	0.3053	0.0734
frame_00277	17.1032	35.6151	0.6224	0.9627	0.3014	0.0646
frame_00287	17.6851	35.436	0.646	0.9609	0.2918	0.0732
frame_00297	17.0232	36.2291	0.6171	0.9603	0.3072	0.0741
frame_00306	17.1774	37.1738	0.6273	0.9642	0.3028	0.0714
平均	18.58765789	**32.80497895**	0.6617315789	**0.9232052632**	0.3537631579	**0.1475736842**

(4) まとめ

動画像データから SfM を用いて抽出された点群データをパノプティックパーツセグメンテーションにより、インスタンスラベルが付された部品に分類した。セグメンテーションに使用するバックボーンを性能、速度、メモリ使用量、拡張性、実装の容易さなど、複数の要素を考慮しつつ選定した。ハードウェアの性能や制約条件に合わせた最適化を施し、処理速度と精度のトレードオフ関係を最適化した。

東電 HD のウェブサイトで公開されている ROV が撮影した 1F の映像から、COLMAP と NeRF を用いて点群データを抽出した。PointNet に由来する代表的なモデルによりインスタンスセグメンテーションを実施し、分類精度の基礎データを取得した。バックボーンについては、ConvNets 系、Transformer 系、MLP 系の 3 種類を比較した結果、Transformer 系と MLP 系のハイブリッドモデルが優位であることが示唆された。

基盤モデルの SAM を用いた評価実験では、1F の映像に対するインスタンスセグメンテーションを試みた。これらのモデルは従来のアプローチを超える性能を発揮し、作業現場の状況をリアルタイムでインスタンスレベルの詳細さで認識できる可能性を示した。特に、Semantic-SAM は、物体とその部分の分類を行う点で優れていた。ハードウェアの性能に関しては、NVIDIA 社製の H100 を 1 枚搭載した GPU サーバと A100 を 2 枚搭載した GPU サーバを構築した。性能比較実験から後者が前者の 2 倍のメモリに対して、速度は前者が後者の 1.2 倍程度の結果が得られた。

以上から 3 カ年計画の 1 年目である令和 5 年度の業務項目を実施し、所期の目標を達成した。今後は、構築した GPU サーバ環境を活用しつつ、実データに対する学習を重ね、リアルタイム対応を実現していく必要がある。さらに、現場への実装に向けて、ロバスト性やセキュリティ対策なども継続して検討する計画である。

参考文献

[3.3-1] Kirillov, A. et al., Segment anything, Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023. https://doi.org/10.1109/ICCV51070.2023.00371

[3.3-2] Zhou, C. et al., A comprehensive survey on pretrained foundation models: A history from BERT to chatGPT, arXiv preprint arXiv:2302.09419, 2023. https://doi.org/10.48550/arXiv.2302.09419

[3.3-3] Kirillov, A. et al., Panoptic segmentation, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019. https://doi.org/10.1109/CVPR.2019.00963

[3.3-4] Radenović, F. et al., Fine-tuning CNN image retrieval with no human annotation, IEEE transactions on pattern analysis and machine intelligence,41, 7, 2018, pp.1655-1668. https://doi.org/10.1109/TPAMI.2018.2846566

[3.3-5] Terven, J. et al., A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS, Machine Learning and Knowledge Extraction, 5, 4, 2023, pp.1680-1716. https://doi.org/10.3390/make5040083

[3.3-6] Li, F. et al., Semantic-SAM: Segment and recognize anything at any granularity, arXiv preprint arXiv:2307.04767, 2023. https://doi.org/10.48550/arXiv.2307.04767

[3.3-7] Lewis, P. et al., Retrieval-augmented generation for knowledge-intensive nlp tasks, Advances in Neural Information Processing Systems, 33, 2020, pp.9459-9474.

[3.3-8] Elharrouss, O. et al., Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches, arXiv preprint arXiv:2206.08016, 2022. https://doi.org/10.48550/arXiv.2206.08016

[3.3-9] Lee, Y., Handwritten digit recognition using k nearest-neighbor, radial-basis function, and backpropagation neural networks, Neural computation, 3, 3, 1991, pp.440-449. https://doi.org/10.1162/neco.1991.3.3.440

[3.3-10] Fukushima, K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, Biological cybernetics, 36, 4, 1980, pp.193-202. https://doi.org/10.1007/BF00344251

[3.3-11] Vaswani, A. et al., Attention is all you need, Advances in neural information processing systems, 30, 2017, https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html（参照：2024年3月31日）.

[3.3-12] Dosovitskiy, A. et al., An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint arXiv:2010.11929, 2020. https://doi.org/10.48550/arXiv.2010.11929

[3.3-13] Rosenblatt, F., The perceptron, a perceiving and recognizing automaton Project Para, Cornell Aeronautical Laboratory, 1957.

[3.3-14] Rumelhart, D.E. et al., Learning representations by back-propagating errors, nature 323, 6088, 1986, pp.533-536. https://doi.org/10.1038/323533a0

[3.3-15] Qi, C.R. et al., PointNet: Deep learning on point sets for 3D classification and segmentation, Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. https://doi.org/10.1109/CVPR.2017.16

[3.3-16] Qi, C.R. et al., Pointnet++: Deep hierarchical feature learning on point sets in a metric space, Advances in neural information processing systems, 30, 2017, https://proceedings.neurips.cc/paper_files/paper/2020/hash/a42a596fc71e17828440030074d15e74-Abstract.html（参照：2024年3月31日）.

[3.3-17] Choy, C. et al., 4D spatio-temporal convNets: Minkowski convolutional neural networks, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019. https://doi.org/10.1109/CVPR.2019.00319

[3.3-18] Qian, G. et al., PointNeXt: Revisiting pointNet++ with improved training and scaling strategies, Advances in Neural Information Processing Systems, 35, pp.23192-23204, 2022, https://proceedings.neurips.cc/paper_files/paper/2022/hash/9318763d049edf9a1f2779b2a59911d3-Abstract-Conference.html（参照：2024年3月31日）.

[3.3-19] Pan, X. et al., 3D object detection with pointformer, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021. https://doi.org/10.1109/CVPR46437.2021.00738

[3.3-20] Zhao, H. et al., Point transformer, Proceedings of the IEEE/CVF international conference on computer vision, 2021. https://doi.org/10.1109/ICCV48922.2021.01595

[3.3-21] Wu, X. et al., Point transformer V2: Grouped vector attention and partition-based pooling, Advances in Neural Information Processing Systems, 35, 2022, pp.33330-33342, https://proceedings.neurips.cc/paper_files/paper/2022/hash/d78ece6613953f46501b958b7bb4582f-Abstract-Conference.html（参照：2024年3月31日）.

[3.3-22] Wang, Q. et al., Window normalization: enhancing point cloud understanding by unifying inconsistent point densities, arXiv preprint arXiv:2212.02287, 2022. https://doi.org/10.48550/arXiv.2212.02287

[3.3-23] Lai, X. et al., Stratified transformer for 3D point cloud segmentation, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022. https://doi.org/10.1109/CVPR52688.2022.00831

[3.3-24] Chen, G. et al., PointGPT: Auto-regressively generative pre-training from point clouds, Advances in Neural Information Processing Systems, 36, 2023, https://proceedings.neurips.cc/paper_files/paper/2023/hash/5ed5c3c846f684a54975ad7a2525199f-Abstract-Conference.html（参照：2024年3月31日）.

[3.3-25] Wang, P., OctFormer: Octree-based transformers for 3D point clouds, ACM Transactions on Graphics (TOG), 42, 4, 2023, pp.1-11. https://doi.org/10.1145/3592131

[3.3-26] Ma, X. et al., Rethinking network design and local geometry in point cloud: A simple residual MLP framework, arXiv preprint arXiv:2202.07123, 2022. https://doi.org/10.48550/arXiv.2202.07123

[3.3-27] Lin, X. et al., DiffBIR: Towards blind image restoration with generative diffusion prior, arXiv preprint arXiv:2308.15070, 2023. https://doi.org/10.48550/arXiv.2308.15070

[3.3-28] Mildenhall, B. et al., NeRF: Representing scenes as neural radiance fields for view synthesis, Communications of the ACM, 65, 1, 2021, pp.99-106.https://doi.org/10.1145/3503250

[3.3-29] Rombach, R. et al., High-resolution image synthesis with latent diffusion models, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022. https://doi.org/10.1109/CVPR52688.2022.01042

[3.3-30] Bian, W. et al., NoPe-NeRF: Optimising neural radiance field with no pose prior, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023. https://doi.org/10.1109/CVPR52729.2023.00405

[3.3-31] Lin, C.-H. et al., BARF: Bundle-adjusting neural radiance fields, Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021. https://doi.org/10.1109/ICCV48922.2021.00569

[3.3-32] Wang, Z. et al., NeRF--: Neural radiance fields without known camera parameters, arXiv preprint arXiv:2102.07064, 2021. https://doi.org/10.48550/arXiv.2102.07064

[3.3-33] Tancik, M. et al., Nerfstudio: A modular framework for neural radiance field development, ACM SIGGRAPH 2023 Conference Proceedings, 2023. https://doi.org/10.1145/3588432.3591516

[3.3-34] Barron, J.T. et al., Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields, Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021. https://doi.org/10.1109/ICCV48922.2021.00580

[3.3-35] Martin-Brualla, R. et al., NeRF in the wild: Neural radiance fields for unconstrained photo collections, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021. https://doi.org/10.1109/CVPR46437.2021.00713
[3.3-36] Levy, D. et al., SeaThru-NeRF: Neural radiance fields in scattering media, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023. https://doi.org/10.1109/CVPR52729.2023.00014
[3.3-37] Wang, Z. et al., Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing, 13, 4, 2004, pp.600-612. https://doi.org/10.1109/TIP.2003.819861
[3.3-38] Zhang, R. et al., The unreasonable effectiveness of deep features as a perceptual metric, Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. https://doi.org/10.1109/CVPR.2018.00068

3.4 研究推進

毎月 1 回程度のプロジェクト会議を実施するとともに、研究項目間での進捗報告および議論を行った。

第 1 回プロジェクト会議：令和 5 年 9 月 7 日（木）開催、WEB 会議
第 2 回プロジェクト会議：令和 5 年 10 月 24 日（火）開催、WEB 会議
第 3 回プロジェクト会議：令和 5 年 11 月 21 日（火）開催、WEB 会議
英知事業ワークショップ：令和 5 年 12 月 5 日（火）開催、富岡町文化交流センター 学びの森
第 4 回プロジェクト会議：令和 5 年 12 月 19 日（火）開催、WEB 会議
中間フォロー：令和 6 年 1 月 12 日（金）開催、WEB 会議

最後に、令和 5 年度の対外発表（国内会議発表、国際会議発表および論文投稿）について、以下に示す。

【国内学会発表】

[1] 中村啓太，羽成敏秀，間所洋和，今渕貴志，川端邦明，Nix Stephanie，土井章男，動画像からの特徴量抽出結果に基づいた高速 3 次元炉内環境モデリングに向けた研究,第 24 回計測自動制御学会システムインテグレーション部門講演会，3D3-02，2023.12.
[2] 羽成敏秀，今渕貴志，中村啓太，川端邦明，廃炉作業における原子炉内環境モデリングに向けた画像品質評価，第 24 回計測自動制御学会システムインテグレーション部門講演会，3D3-03，2023.12.
[3] 土井章男，高志毅，高橋弘毅，加藤徹，山下圏，NeRF 技術を用いた画像からの可視化と 3D モデリング，第 52 回テレイマージョン技術研究会，2024.2.

【論文発表】

[1] K. Nakamura, T. Hanari, T. Matsumoto, K. Kawabata, H. Yashiro, A Study on the Effects of Photogrammetry by the Camera Angle of View Using Computer Simulation, Journal of Robotics and Mechatronics, vol.36, no.1, 2024.2, pp.115-124.

4. 結言

本研究では、1F の廃炉に向けて、PCV および原子炉建屋内を調査する際に撮影した動画像を入力とし、指定された時間、動画像から抽出された特徴量に応じて、周辺情報を補強したうえで情報量が大きい立体復元手法を選択し、作業空間を 3 次元モデリングする研究開発を行うことを目的とする。以下に、3 カ年計画の 1 年目である令和 5 年度の業務実績を述べる。

シミュレータを活用した写真測量における復元時間制限を考慮した画像特徴量と立体復元精度の関係（札幌大学）

実際の PCV を模した暗所環境を構築し、光の反射、レンズの焦点距離、撮影時のカメラ振動を考慮したシミュレーションを行い、写真測量用の人工画像の生成を行った。そして、生成した人工画像を入力した写真測量を行うことで、写真測量対象環境に適した光源の強さ、撮影時の焦点距離があることを確認した。また、撮影時のカメラ振動によって、獲得する写真測量結果の変化も同時に調査した。さらに、アーム付き移動ロボットおよびアクションカメラを導入することで、さまざまな姿勢で撮影できるカメラシステムの検討を行った。

動画像からの迅速な 3 次元モデリング手法の研究開発（連携先：JAEA）

1F 調査画像の特性分析を行い、画像品質評価手法の導入により画像補正による画質スコアの改善と画像特徴点のマッチング数の向上に関連性があることを示した。また、深層学習を適用した画像特徴点を抽出する手法の調査・分析を行い、深層学習の適用により従来手法と比べて良好な結果が得られることを確認した。そして、深層学習の適用によって抽出された画像特徴点をもとにした、立体復元結果の復元精度の評価および画像特徴点抽出数・マッチング数の増加に伴う計算時間の増加を今後の課題として抽出した。さらに、イメージスティッチングの適用性検証の結果、カメラ軌道および投影モデルの選択がパノラマ画像合成の性能に与える影響を確認し、球面投影型モデル、アフィン投影型モデル以外の合成方法についても実験的に検証していくこととした。

深層学習に基づく画像・点群データのセグメンテーションと高品質な 3 次元モデリング（再委託先：岩手県立大学）

動画像データから SfM を用いて抽出された点群データを、パノプティックパーツセグメンテーションにより、インスタンスラベルが付された部品に分類した。セグメンテーションに使用するバックボーンを、性能、速度、メモリ使用量、拡張性、実装の容易さなど、複数の要素を考慮しつつ選定した。ハードウェアの性能や制約条件に合わせた最適化を施し、処理速度と精度のトレードオフ関係を最適化した。

研究推進

研究代表者の下で各研究項目間ならびに CLADS 等との連携を密にして研究を進めた。また、研究実施計画を推進するための打合せや会議等を開催した。

本研究のアウトカムとしては、自動的に生成された炉内環境の復元結果から、作業前における立体的な炉内状況および未探索箇所の把握が可能になり、オペレータの空間認識を支援できると考えている。また、蓄積される教師データによる全体の精度向上によって、より高精度の 3 次元モデルの提供が可能になると想定している。また、撮影された動画像群から高速に特徴点群を抽出し、生成した点群および要求された時間内に応じた、復元情報量が大きくなるように自動的に炉内環境の立体復元結果データを獲得できるように、各研究機関で適宜相互連携しながら研究を行う。

This is a blank page.