

エネルギー科学研究科
エネルギー社会・環境科学専攻修士論文
環境再構成モデルによる
題目： レンダリング画像を利用した
リローカリゼーション手法の開発

指導教員： 下田 宏 教授

氏名： 木村 太郎

提出年月日： 平成29年2月10日(金)

論文要旨

題目：環境再構成モデルによるレンダリング画像を利用したリローカリゼーション手法の開発

下田研究室, 木村太郎

要旨：

放射性汚染物質の慎重な取り扱いや膨大な時間を必要とする原子力発電プラントの解体作業を支援するために、拡張現実感（Augmented Reality：AR）を利用する研究が行われている。情報を直感的に提示できるARを用いることにより、効率的かつ安全に作業を行えるようになる可能性がある。ARを利用するためには、カメラの姿勢をリアルタイムで計測するトラッキング技術が必要となる。トラッキングには、カメラ姿勢の推定に失敗した場合に、現在のカメラ画像のみを用いて処理を再開させるリローカリゼーションと呼ばれる技術が必要である。しかし、従来のリローカリゼーション手法には、使用可能な領域が事前に撮影した領域に制限されるという問題がある。実際の解体作業支援を想定すると、使用可能な範囲の制限が無く、カメラを自由に移動させることができるリローカリゼーション手法を実現する必要がある。

そこで、本研究では撮影画像から環境の色と三次元構造を再構成した環境再構成モデルによるレンダリング画像を利用することにより、使用可能な範囲が制限されるという問題を緩和するリローカリゼーション手法を開発することを目的とする。

提案手法では、環境再構成モデルを用いて、事前に撮影していない領域で撮影した画像もレンダリングにより生成し、生成したレンダリング画像とその画像を生成した時のカメラ姿勢を組にしてキーフレームデータベース（キーフレームDB）に格納する。そして、入力の実画像と類似したレンダリング画像をキーフレームDBから抽出し、そのレンダリング画像と組のカメラ姿勢を現在のカメラ姿勢の推定値とする。しかし、環境再構成モデルを用いて生成したレンダリング画像を使用することにより、実画像には写っているがレンダリング画像には現れない場合があるなどの問題がある。そのため、このような実画像とレンダリング画像の間の差異に起因する問題を解決する必要がある。そこで、本提案手法は類似画像検索処理に周辺領域や信頼度を導入することにより、実画像とレンダリング画像の間の差異による悪影響の緩和を試みた。

提案手法を、原子炉廃止措置研究開発センターの充填ポンプ室内を対象として評価した結果、環境再構成モデルを用いて生成したレンダリング画像により、Randomized Fernのような既存手法の持つ、使用可能な領域が事前に撮影した領域に制限されるという問題を緩和できていることが示された。また、提案手法は周辺領域や信頼度を用いることにより、正確に実画像と類似したレンダリング画像を抽出しリローカリゼーションを成功させることができることが示された。

今後の課題として、想定する作業領域が広がるとレンダリング画像の数も増えるため、レンダリング画像を保存するキーフレームDBの構造を改良することにより類似画像検索の処理時間を削減することが必要である。また、環境が変化した場合に、自動的に環境再構成モデルやキーフレームDBを更新する手法や、多少の環境変化に対してもロバストな類似画像検索処理を導入するなどが必要となる。

目次

第 1 章 序論	1
第 2 章 研究の背景と目的	3
2.1 研究の背景	3
2.1.1 既存のトラッキング手法とその問題点	3
2.1.2 既存のリローカリゼーション手法とその問題点	9
2.2 研究の目的と意義	16
第 3 章 環境再構成モデルを用いたリローカリゼーション手法の提案	18
3.1 提案手法の概要	18
3.2 キーフレームデータベースの構築	20
3.2.1 環境再構成モデルの作成	22
3.2.2 レンダリング画像の生成と特徴量の計算	23
3.2.3 レンダリング間隔の設定	27
3.3 類似画像検索とカメラ姿勢の推定	29
3.3.1 類似画像検索の概要	29
3.3.2 画像特徴の抽出と無効領域の処理	30
3.3.3 相違度の計算	32
3.3.4 カメラ姿勢の推定	34
3.4 提案手法の予想される利点と欠点	36
第 4 章 提案手法の評価	37
4.1 評価の概要	37
4.2 評価の方法	38
4.2.1 評価に用いる指標	38
4.2.2 評価用データセットの作成	39
4.2.3 Iterative Closest Point(ICP) アルゴリズムによる推定値の補正	43
4.3 提案手法を評価するための環境	44

4.3.1	評価に用いたソフトウェアおよびハードウェア	44
4.3.2	リローカリゼーションの成否判定の閾値の設定	45
4.3.3	作成した評価用データセット	46
4.3.4	レンダリング画像の生成	55
4.3.5	本研究で使用する Randomized Fern の実装	60
4.3.6	提案手法のパラメータの決定	61
4.4	評価結果と考察	65
4.4.1	中心ピクセル数と周辺領域のサイズの影響の評価結果と考察 . .	65
4.4.2	既存手法との比較	68
4.4.3	類似画像検索に Randomized Fern を用いた場合の評価	70
4.4.4	サンプリング方式の比較	72
第 5 章 結論		79
謝 辞		81
参 考 文 献		82

目次

2.1	AR を用いたプラント作業支援の例 [8][9]	3
2.2	人工マーカの例	4
2.3	自然特徴の例	5
2.4	自然特徴を用いたカメラ姿勢の推定	6
2.5	再投影誤差の最小化による現在のカメラ姿勢の推定	7
2.6	オクルージョンの例	8
2.7	トラッキングとリローカリゼーションの関係	9
2.8	キーフレームの保存	11
2.9	類似画像検索によるカメラ姿勢の推定	11
2.10	Randomized Fern における相違度の計算	13
2.11	キーフレームベース手法を用いた場合の復帰可能範囲の限界	13
2.12	環境再構成モデルの欠損領域	14
3.1	提案手法で扱うデータと処理の流れ	18
3.2	トラッキング対象とする環境の撮影	19
3.3	環境再構成モデルを用いたレンダリング画像の生成	20
3.4	キーフレームデータベースの構築処理の流れ	21
3.5	RGB 画像とデプス画像及び三次元環境再構成モデル	22
3.6	カメラ姿勢のサンプリングとレンダリング画像の生成	23
3.7	RGB-D カメラの RGB カメラとデプスセンサ	24
3.8	カメラの透視投影モデル	24
3.9	実画像とレンダリング画像	25
3.10	レンダリング画像の無効領域	26
3.11	レンダリング画像における周辺領域の領域内平均の計算処理の流れ	27
3.12	類似画像検索処理の流れ	30
3.13	周辺領域を用いた画像間の相違度の計算	31
3.14	実デプス画像の無効領域	32
3.15	実画像における周辺領域の領域内平均の計算処理の流れ	33

3.16	周辺領域による有効ピクセルの利用	34
3.17	画像間の相違度の計算の流れ	35
4.1	初期姿勢の違いによる各データセットの座標系の違い	40
4.2	環境内での三脚の設置の様子	41
4.3	データセット用画像の撮影時の様子	41
4.4	ICP を用いた環境再構成モデルの座標系の統一	42
4.5	ICP を用いたカメラ姿勢の推定値の補正	43
4.6	Microsoft Xbox One Kinect の外観	44
4.7	充填ポンプ室の様子	47
4.8	Area1 の撮影領域	48
4.9	Area2 の撮影領域	48
4.10	Area3 の撮影領域	48
4.11	Area1 の撮影時のカメラ軌道	49
4.12	Area2 の撮影時のカメラ軌道	50
4.13	Area3 の撮影時のカメラ軌道	51
4.14	モデル 1	52
4.15	モデル 2	53
4.16	モデル 3	54
4.17	カメラ姿勢の 6 パラメータ	55
4.18	カメラ姿勢のサンプリングの範囲	55
4.19	Area1 で等間隔サンプリングされたカメラ姿勢と入力用データセットの カメラ軌道	57
4.20	Area2 で等間隔サンプリングされたカメラ姿勢と入力用データセットの カメラ軌道	57
4.21	Area3 で等間隔サンプリングされたカメラ姿勢と入力用データセットの カメラ軌道	58
4.22	Area3 において深度値標準偏差サンプリングにより得られたカメラ姿勢 と入力用データセットのカメラ軌道	59
4.23	一般的なガウシアンフィルタと本研究で用いるガウシアンフィルタ	61
4.24	τD_{real} を変化させた場合の提案手法の復帰率の変化	62
4.25	τD_{ren} を変化させた場合の提案手法の復帰率の変化	63
4.26	τC_{ren} を変化させた場合の提案手法の復帰率の変化	63

4.27 τC_{ren} と τD_{ren} を変化させた場合の提案手法の復帰率の変化	64
4.28 τC_{ren} と τD_{real} を変化させた場合の提案手法の復帰率の変化	64
4.29 τD_{real} と τD_{ren} を変化させた場合の提案手法の復帰率の変化	65
4.30 等間隔サンプリングを用いた場合の提案手法の復帰率	66
4.31 等間隔サンプリングを用いた場合の処理時間	67
4.32 提案手法と Randomized Fern が類似画像として抽出した画像の例 (Area3 を処理対象とした場合)	70
4.33 提案手法と Randomized Fern が類似画像として抽出した画像の例 (Area2 を処理対象とした場合)	71
4.34 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似 画像として抽出された画像の例 (Area1 を処理対象とした場合)	73
4.35 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似 画像として抽出された画像の例 (Area2 を処理対象とした場合)	73
4.36 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似 画像として抽出された画像の例 (Area3 を処理対象とした場合)	74
4.37 深度値標準偏差サンプリングを用いた場合の提案手法の復帰率	75
4.38 深度値標準偏差サンプリングを用いた場合の処理時間	76
4.39 等間隔サンプリングと深度値標準偏差サンプリングの復帰率の比較	77
4.40 等間隔サンプリングと深度値標準偏差サンプリングの処理時間の比較	78

表 目 次

4.1	評価に使用した PC の仕様と開発環境	44
4.2	開発に使用したソフトウェアおよびライブラリ	45
4.3	Xbox One Kinect のセンサの仕様 ^[52]	45
4.4	Xbox One Kinect の寸法・重量 ^[52]	45
4.5	リローカリゼーションの成否判定の閾値	45
4.6	DB 用データセットの撮影領域とフレーム数	47
4.7	入力用データセットの撮影領域とフレーム数	49
4.8	各撮影領域のサンプリング時のカメラ姿勢のパラメータの範囲	56
4.9	各サンプリング手法の場合のレンダリング画像数	58
4.10	等間隔サンプリング時に全フレームのレンダリングにかかる時間	58
4.11	本研究の信頼度の閾値の設定	64
4.12	Randomized Fern の復帰率と処理時間	69
4.13	中心ピクセル数が 500 の場合の提案手法の最大の復帰率と処理時間	69
4.14	類似画像検索に Randomized Fern を用いた場合と提案手法の復帰率と処理時間	71

第 1 章 序論

2011年3月11日の東日本大震災により発生した福島第1原子力発電プラントの事故以降、国内の原子力エネルギーに対する認識は大きく変化した。原子力発電プラント運用の安全性への信頼は失われ、全ての原子力発電プラントは活動を停止した。九州電力の川内原子力発電プラントを始め、いくつかの発電プラントは再稼働したが、一次エネルギーの国内供給に占める原子力の割合は依然として低いままである^[1]。これは、事故により規制基準が強化されたが、規制基準を満たすためには設備投資を伴う改修が必要であり、多くのプラントの再稼働の見通しが立たないためである。そして、再稼働の見通しが立たない経済性に劣る規模の小さい炉は廃炉される傾向がある。また、多くの原子力発電プラントの稼働年数が設計当初の運転想定期間である30年から40年に迫っており、電気事業者は高い安全対策費を払い運転期間を延長するか廃炉を進めるかの選択を迫られている^[2]。このように、原子力発電所プラントの廃炉は差し迫った社会問題となっている。

廃炉が決まった原子力発電プラントを解体するためには、放射性汚染物質の慎重な取り扱いを必要とするため、作業員に多大な負担がかかる。このような現状の解体作業を支援するために、拡張現実感（Augmented Reality：AR）を利用する研究が行われている^{[3][4]}。ARは、通常では見ることのできない情報をユーザに知覚させたり、現実世界の位置や方向をユーザに直感的に提示できるなどの特長を有している。このような特長を有するARを専門的な知識と技術が必要な解体作業の支援に用いることにより、効率的かつ安全に作業を行えるようになる可能性がある。また、現場の作業員の高齢化が進み、今後専門技術の継承が難しくなることが予想されるため、新しい作業員でも直感的に理解できるARによる情報提示は有用であると考えられる。

原子力発電プラント内でARを利用するためには、プラント内で作業員の位置と方向をリアルタイムに計測するトラッキングと呼ばれる技術が必要である。このトラッキングには、位置・方向の推定に失敗した場合に、現在のカメラ画像のみを用いて処理を再開させるリローカリゼーションと呼ばれる処理が必要である。しかし、既存の多くのリローカリゼーション手法には、使用可能な領域が事前に撮影した領域に制限されるという問題がある。この問題を解決するため、本研究では、作業環境の三次元形状や色を忠実に再構成した環境再構成モデルによるレンダリング画像を利用したり

ローカリゼーション手法を開発することを目的とする。環境再構成モデルを用いて事前に撮影していない場所でのカメラ画像を模擬的に生成することにより、ARを利用する際のカメラの移動範囲の制限を緩和できると考えられる。これにより、カメラをより自由に動かすことが可能となり、ARに関する専門的な知識を持たないユーザでも、容易にARを用いた作業支援システムを利用できるようになると期待される。

本論文は第1章の序論を含め、5章で構成される。第2章では、本研究の背景と目的について述べ、第3章では、本研究で提案する環境再構成モデルによるレンダリング画像を利用したローカリゼーション手法について述べる。第4章では、本提案手法の性能を評価するために実際の原子力発電プラントを対象として実施した評価実験について述べ、第5章で結論として本研究のまとめと今後の課題を述べる。

第 2 章 研究の背景と目的

2.1 研究の背景

拡張現実感 (Augmented Reality: AR) とはカメラ等で撮影された現実世界の映像に、Computer Graphics(CG) 等の情報を重畳表示することにより、ユーザの知覚を拡張する技術である [5-7]。例えば、図 2.1 に示すように、解体箇所を視覚的に提示したり、放射線のような目視できない情報をユーザに直感的に示すことが可能である [8][9]。CG 等の情報を現実世界の映像に適切に位置を合わせて重畳表示するためには、撮影している環境を基準としたカメラの位置・方向（カメラ姿勢）を計測する必要がある。計測された現在のカメラ姿勢に応じて適切に位置を変化させた表示情報を観察することにより、ユーザは現実世界と表示情報の位置関係を理解できる。この位置関係を直感的に理解できるという AR の特徴を、原子力発電プラントの解体作業支援へ応用することにより、解体作業の効率や安全性を向上できると期待されている [10]。

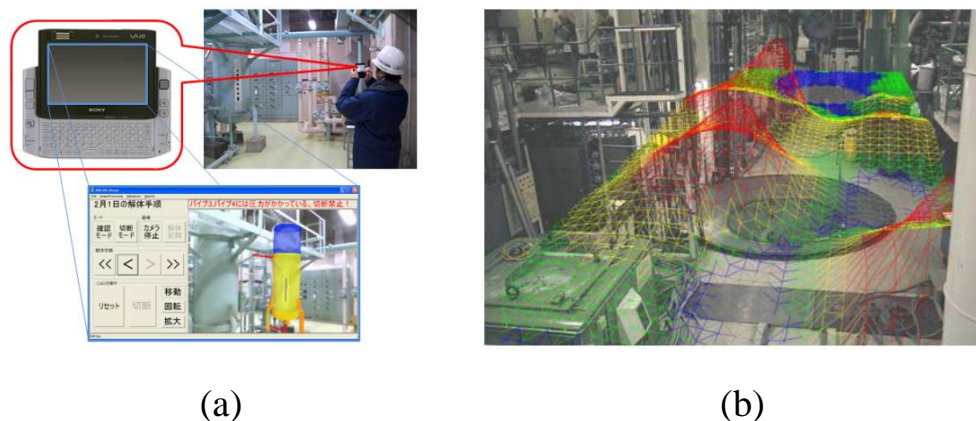


図 2.1: AR を用いたプラント作業支援の例 [8][9]

*

2.1.1 既存のトラッキング手法とその問題点

カメラ姿勢を求める技術はトラッキングと呼ばれる [11]。トラッキングを実現する手法としては、ジャイロセンサ・加速度センサなどの慣性センサを利用する手法、超音波

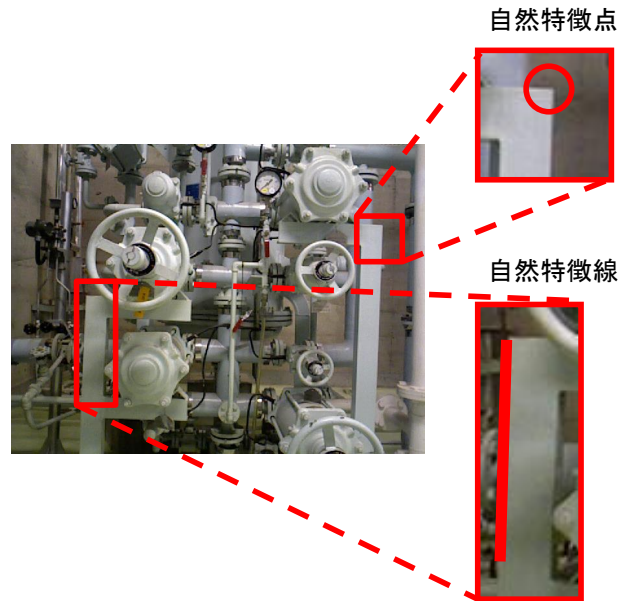
センサ・磁気センサを利用する手法、全地球測位システム (Global Positioning System : GPS) を利用する手法、カメラ等のビジョンセンサを利用する手法等が提案・実現されている^{[12][13]}。しかし、原子力発電プラントのような広域で複雑な形状の物体が多数存在している環境での作業支援を想定した場合、使用できるトラッキング手法は限られる。例えば、ジャイロセンサ・加速度センサは時間が経過するにつれて誤差が蓄積する特徴があり、数時間単位での長時間の作業が行われることもある原子力発電プラントでの保守・解体作業で用いる場合は精度が十分ではない。超音波センサは、広域で複雑な環境では正確に音波の反射波が得られないため、プラント内部の様な環境では正確な計測結果を常に得ることが難しい。磁気センサは機材 1 台あたりの使用可能な範囲が狭いことや機材が高価なことも問題となる。GPS は原子力発電プラント内のような衛星からの電波が届かない場所では利用できない。

一方、ビジョンセンサは原子力発電プラント内部でも使用可能である。そのため、原子力発電プラントでは、ビジョンセンサを使用したビジョンベースの手法を利用することが望ましい。ビジョンベースの手法では、カメラで撮影した画像から図 2.2 に示すような人工マーカや、椅子や操作パネルなどの環境に予め存在する物体、図 2.3 に示すような環境に存在する輝度の変化が大きい点や線等の自然特徴を認識し、これらを利用してカメラ姿勢を推定する。



図 2.2: 人工マーカの例

人工マーカを用いた手法では環境を基準とした三次元座標が既知である人工マーカをカメラで撮影し、画像処理と幾何計算により環境を基準としたカメラ姿勢を求める^[14]。具体的には、環境を基準とした人工マーカの三次元位置と撮影画像上の二次元位置から、Perspective-n-Point(PnP) 問題を解くことによりカメラ姿勢を推定する^[15]。人工マーカを用いた手法は照明の変化などの外乱に強く、安定して利用できるという利



Copyright (C) 2017 Japan Atomic Energy Agency

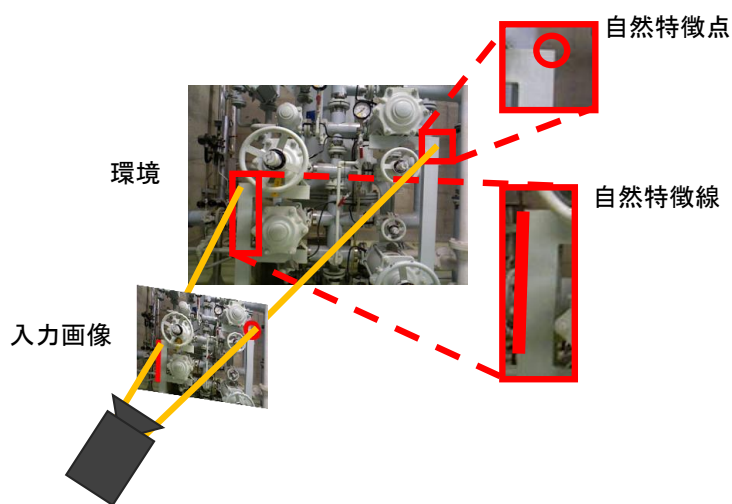
図 2.3: 自然特徴の例

点があるが、広域な原子力発電プラント内部で使用する場合、マーカの設置とその三次元位置の計測に手間がかかるという問題がある。しかし、ARを用いた作業支援システムを実用化するためには、準備に要する手間は少ない必要がある。また、設置した人工マーカの認識を安定させるためには人工マーカをある程度大きくする必要があり、人工マーカ自体が作業の妨げになる可能性もある。そのため、事前準備の手間や環境整備の労力を考慮すると、環境に存在する物体や自然特徴を利用したトラッキングを用いることが望ましい。

環境に存在する物体をトラッキングに利用する手法では、トラッキングを利用する環境全体や環境内に存在する物体のモデルを保持しておき、撮影画像から得た物体の特徴とモデルの特徴の対応付けから物体の姿勢を認識し、その物体を用いてカメラ姿勢を推定する。様々な手法が存在するが、物体の輪郭線とモデルの輪郭線を対応付ける手法がよく用いられる^{[16][17]}。これらの手法では、環境に存在する物体の3D Computer-aided design(CAD)モデルを予め作成もしくは取得しておき、3DCADモデルから抽出した輪郭を撮影画像上に投影した結果と、撮影画像上で認識された物体の輪郭の差を最小化するカメラ姿勢を推定する。多くの場合、3DCADモデルの輪郭を撮影画像上に投影する際に、直前のカメラ姿勢を利用することにより対応付けの時間短縮と精度の向上を図っている。環境に存在する物体を認識する手法は環境や物体のモデルを利用する

ため、モデルベースの手法とも呼ばれる。モデルベースの手法は、精度が高く安定性が高いという特長があり、ロボットピッキングなど工業方面での応用が期待されている^[18]。しかし、この手法では予め環境や物体のモデルを取得する必要があるが、環境や物体の精巧なモデルの取得が困難な場合も多い。また、モデルを手作業で作る事も考えられるが、手間がかかるために実用的ではなかった。

自然特徴を利用したトラッキング手法では、環境に存在する点や線などの自然特徴を認識し、これらを利用してカメラ姿勢を推定する。この手法では、環境中の自然特徴の特徴量や三次元位置などの情報を取得・保存し、図 2.4 に示すように、環境を撮影した画像から認識される自然特徴と保存されている自然特徴を対応付け、この対応付けられた自然特徴の三次元位置と画像上の二次元位置を利用することによりカメラ姿勢を推定する。一般的にはまず、何らかの手法で最初のカメラ姿勢を決定する。その後、現在のカメラ姿勢は直前の画像を撮影した時のカメラ姿勢に近いと仮定し、図 2.5 に示すように、予め三次元位置の情報とともに保存された自然特徴を直前のカメラ姿勢を用いて現在のカメラの画像上に投影した際の 2 次元座標 p を求める。この p の周辺で外見が最も似た自然特徴 p' を探索し、対応する自然特徴とする。そして、対応付けられた p と p' の間の距離を誤差（再投影誤差）として、最小化するカメラ姿勢を求める。



Copyright (C) 2017 Japan Atomic Energy Agency

図 2.4: 自然特徴を用いたカメラ姿勢の推定

自然特徴を利用したトラッキングを実行するためには、自然特徴の三次元位置を知る必要がある。自然特徴の三次元位置を求める手法には、レーザ距離計測機などを用

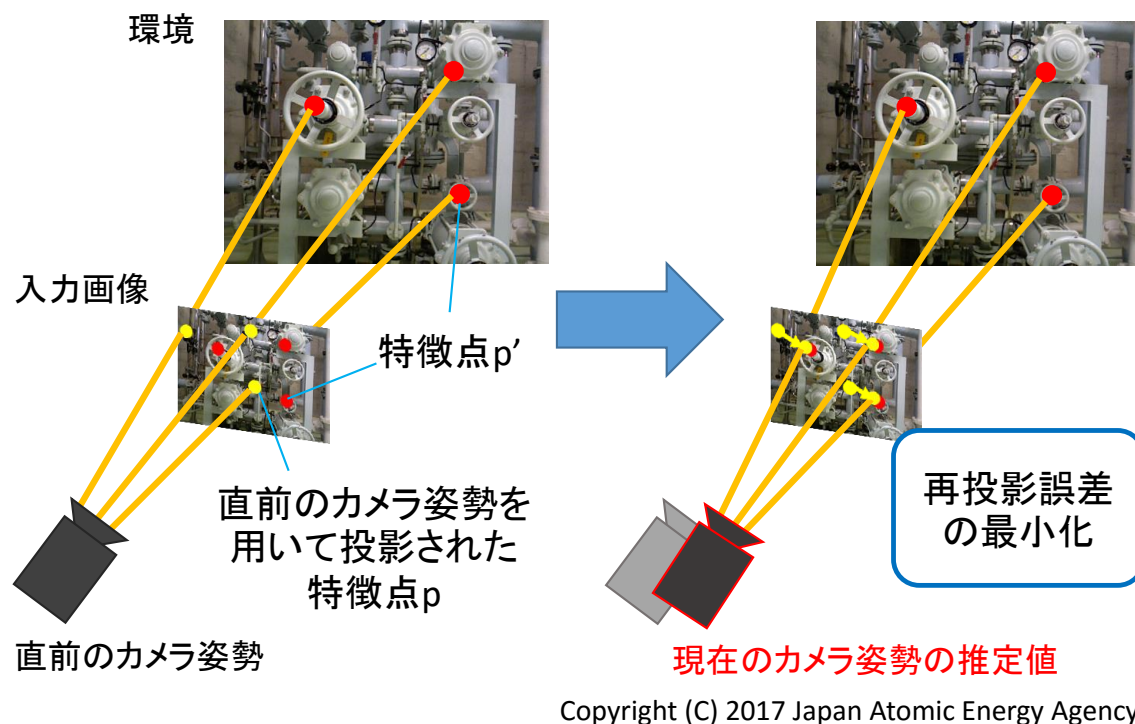


図 2.5: 再投影誤差の最小化による現在のカメラ姿勢の推定

いて計測する手法や、撮影した複数のカメラ画像から推定する手法などがあるが、カメラのみで実行できる後者の手法がよく用いられる。この手法の一種として Structure from Motion(SfM)^[19] が知られている。多くの場合、SfM では異なる視点から撮影した複数の画像から 2 フレーム毎に特徴点マッチングを行い、基礎行列を求める。基礎行列によりカメラの内部、外部パラメータを求めることができ、画像上の特徴点と三次元空間上の特徴点の位置関係を求めることができる。2 フレーム毎に計算した後、全体のフレームに三次元空間上の特徴点を投影し、投影された特徴点と画像上の特徴点との再投影誤差を最小化することにより特徴点の三次元座標と全フレームのカメラ姿勢の計算の精度を向上させる。これにより、環境の三次元マップを取得できる。SfM には、町全体の構造を計算する大規模な研究^[20] や高い計算コストを削減するための研究^[21] など多くの研究が存在する。これまで、SfM は計算量が多いために基本的にはオフラインで実行されてきたが、カメラ姿勢の推定と環境のマップの生成を同時に行う SLAM(Simultaneous Localization And Mapping)^[22-24] と呼ばれる手法が、AR のトラッキングやロボットの自己位置推定問題に利用できるために近年注目を浴びるようになってきた。SLAM では撮影画像から抽出した自然特徴の三次元位置を三角測量に似た原理を用いて計算し、この環境中の自然特徴の三次元配置を環境のマップとする。そしてこ

のマップに対するカメラ姿勢を推定する。このマップ生成をカメラ姿勢推定処理のバックグラウンドプロセスとして並列に実行することにより、カメラ姿勢推定の処理を常に高速に行えるようにした代表的手法が Parallel Tracking and Mapping(PTAM)^[25] である。PTAMでは画像の一部の情報(特徴点)のみを用いており、環境の疎なマップを生成・利用していたが、画像のより多くのピクセルの情報を用いることでより正確なカメラ姿勢推定とより密な環境のマップの構築を実現した DTAM^[26] が提案された。DTAMでは環境のマップとして密な三次元モデルが得られる。近年は Microsoft の Kinect カメラのような赤外線センサにより環境の距離情報も取得できる RGB-D カメラが登場し、KinectFusion^[27] などの手法によりセンサ情報から環境を再構成した再構成モデルをリアルタイムで取得できるようになった。この分野の近年の発展は目覚ましく、メモリと HDD のスワッピングにより理論的に無限の領域を再構成できる InfiniTAM^{[28][29]} などの手法によって、より広域な環境でより精巧なモデルを構築できるようになってきた。このような自然特徴を利用したトラッキングは事前準備などのユーザの負担が少ないという利点を持つが、図 2.6 に示すように、トラッキングに使用する自然特徴が他の物体の陰に隠れてしまうオクルージョンと呼ばれる状況が頻発する環境では動作が不安定になる。また、照明条件が変化したり、環境を撮影するカメラの角度が大きく変化した場合、特徴点の対応付けが失敗する場合が多くなることも問題となる。

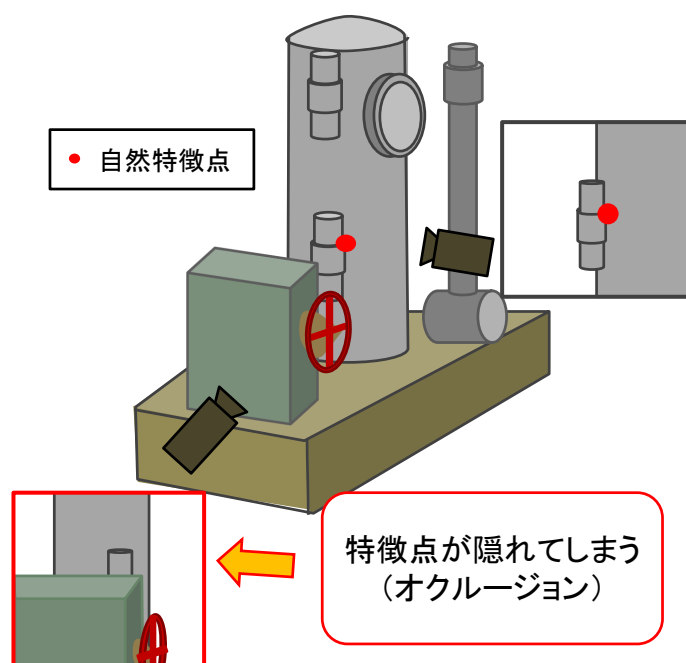


図 2.6: オクルージョンの例

ここまで述べてきたようなビジョンベースのトラッキング手法では、多くの場合、直前のカメラ姿勢と現在のカメラ姿勢の変化は小さく、直前に取得したカメラ画像と現在取得されているカメラ画像の差が小さいという仮定を利用している。しかし、カメラを急に異なる方向に向けた場合など、この仮定が成り立たなくなる場合がある。この時、再投影誤差の最小化問題が正しい解（カメラ姿勢の真値）に収束せず、間違っ了解に収束する場合がある。この場合、以降のトラッキングにおいてこの間違っ了解を用いて再投影誤差を計算することになり、カメラ姿勢の推定に繰り返し失敗する状態に陥ることになる。この状態をロストと呼ぶ。ロスト時に、トラッキングとは別に、現在の画像のみの情報を利用してカメラ姿勢を推定する手法をリローカリゼーションと呼ぶ。本研究では直前のカメラ姿勢を利用して現在のカメラ姿勢を推定する手法をトラッキング、直前のカメラ姿勢を利用せずに現在のカメラ姿勢を推定する手法をリローカリゼーションと呼ぶ。トラッキングとリローカリゼーションの関係を図 2.7 に示す。実装上は最小化後の再投影誤差の値が予め定めた閾値よりも大きくなった場合には、間違っ了解に収束し、トラッキングが失敗してロストしたと判断し、リローカリゼーションを実行する。

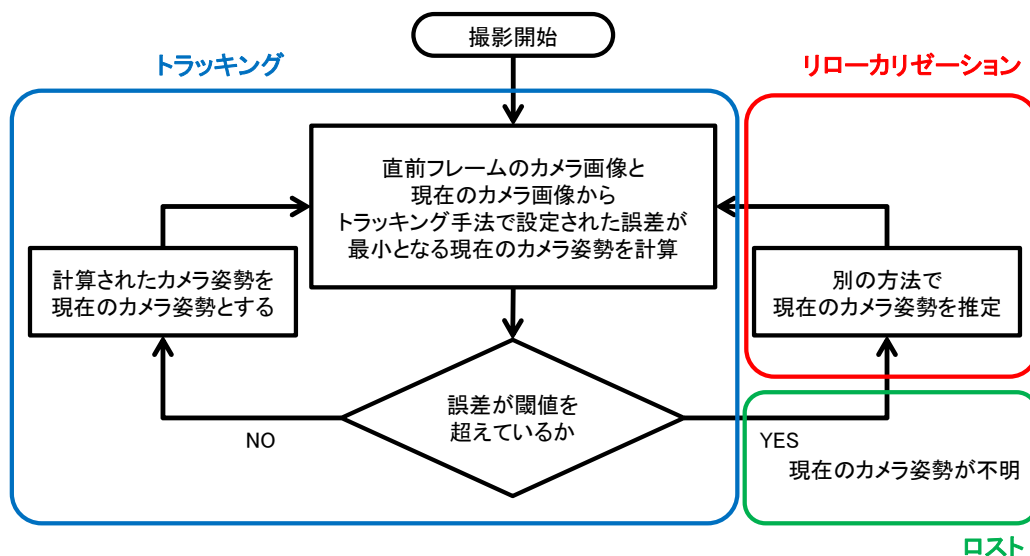


図 2.7: トラッキングとリローカリゼーションの関係

2.1.2 既存のリローカリゼーション手法とその問題点

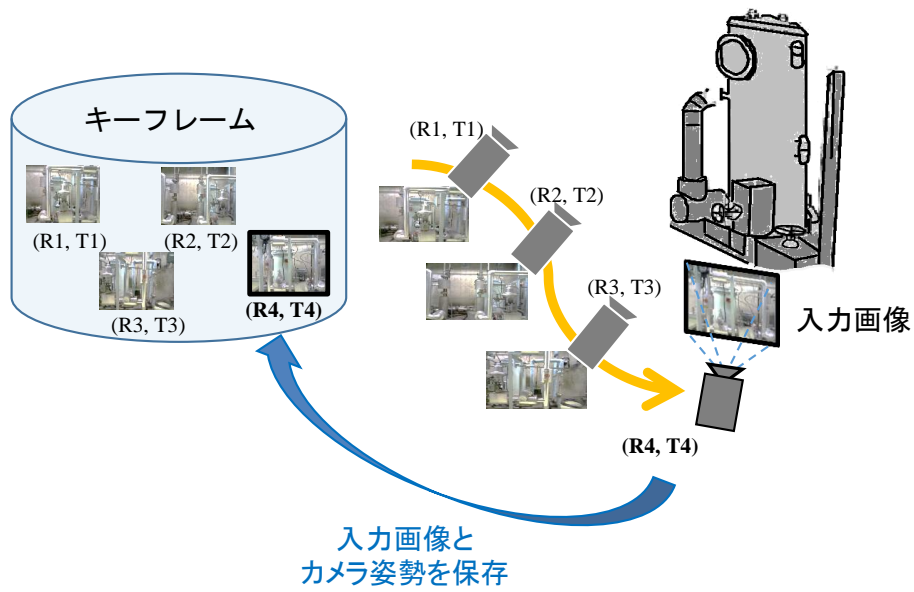
実際に AR を用いて解体作業を支援する場合、現場作業員にカメラを速く動かさないうようにしてもらうことは難しい。作業中にカメラを置いたり別の場所を撮影する際

に、カメラを急に大きく動かしてしまうことは十分起こり得ることであり、そのたびにロストしてトラッキングできないならば AR を用いた実用的な作業支援は実現できない。実際の解体作業を想定すると作業員が自由にカメラを動かすことができる必要がある。そのため、原子力発電プラント内で自由にカメラを動かせるリローカリゼーション手法の実現が重要となる。

2.1.1 項で述べたように、原子力発電プラント内ではビジョンセンサを使用することが望ましい。そのため、原子力発電プラント内での AR による作業支援の実現のためのリローカリゼーション手法もビジョンセンサを利用したビジョンベースの手法を使用することが望ましい。ビジョンベースのリローカリゼーション手法の主な手法としては、キーフレームベースの手法とキーポイントベースの手法がある。以下で各リローカリゼーション手法の概要と問題点を述べる。

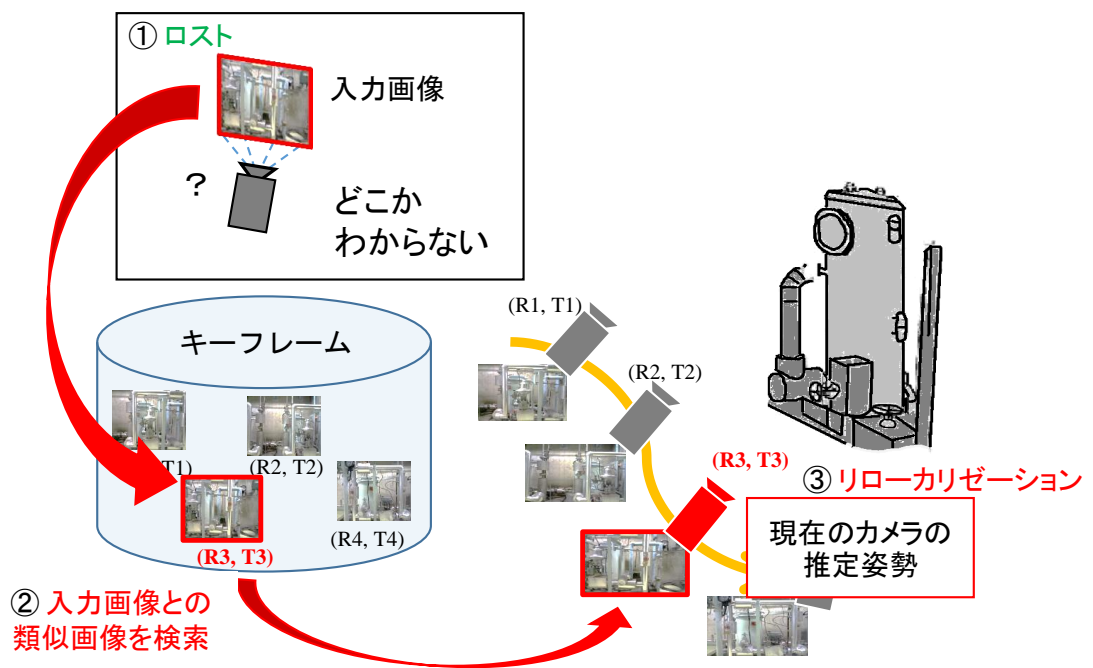
2.1.2.1 キーフレームベースのリローカリゼーション手法

キーフレームベースの手法では、図 2.8 に示すように、事前もしくはトラッキングが成功している際に、撮影した画像から個々の画像の区別に利用できるその画像を代表する特徴を抽出し、その画像を撮影した時のカメラ姿勢と組にしてキーフレームとして保存しておく。そして図 2.9 に示すように、ロストした際にまず、入力画像から抽出した特徴と保存しておいた画像の特徴を比較し、入力画像と最も類似した特徴を持つキーフレームを抽出する。2つのカメラ姿勢が近い場合、撮影される画像も類似する。従って、現在の入力画像と類似したキーフレームが存在すれば、そのキーフレームのカメラ姿勢は現在のカメラ姿勢に近い可能性が高い。そこで、入力画像と類似したキーフレームと組になっているカメラ姿勢を入力画像のカメラ姿勢の近似値とする。そして、この近似値を誤差の最小化処理を行う際の初期値として利用することによりトラッキングを再開する。類似画像を検索する比較的単純な手法としては、入力画像とキーフレームの縮小画像の全ピクセルの輝度の差の二乗の和 (Sum of Squared Difference : SSD) を計算し、最も SSD が小さいキーフレームを探す手法^[30] や、縮小画像の正規化相互相関により類似度を計算する手法^[31] などがある。これらは処理が単純で実装が容易であるが、処理時間が掛かる。また、Fast Appearance-Based Mapping(FAB-Map)^{[32][33]} はベイズの定理を利用して画像間の類似度を確率的に計算する手法である。この手法は自然言語処理の分野の概念である Bag Of Words を利用して画像を特徴点の集まりとして表現し、入力画像とキーフレームが同じ特徴点を含んでいるという情報から類似度を確率的に計算する。FAB-Map の利点として、同じ特徴点が映っていれば同じ環境



Copyright (C) 2017 Japan Atomic Energy Agency

図 2.8: キーフレームの保存



Copyright (C) 2017 Japan Atomic Energy Agency

図 2.9: 類似画像検索によるカメラ姿勢の推定

と判断されるため、多少カメラの姿勢が変化しても同じ環境を撮影している画像であると認識できることが挙げられるが、本手法は屋外などの広域で使用することを想定しており、撮影位置の差が小さい場合、撮影される画像内に写る特徴点の構成がほぼ同じとなるため、画像の違いを認識することが難しい。一方、画像の中からランダムに選んだピクセルを、ランダムな閾値で二値化して符号化することにより、類似画像検索時に必要な情報処理量を削減し、検索の高速化を図った Randomized Fern^{[34][35]} と呼ばれる手法も存在する。本手法では、図 2.10 に示すように、画像からランダムに複数のピクセルを選び、RGBD 各チャンネルの値が各チャンネル毎に事前にランダムに設定された閾値より大きければ 1、小さければ 0 として符号化する。ここで、ピクセル内の 4 つのチャンネルの値と 4 つの閾値の集まりを fern と表現する。fern 毎に 4 ビットのコードを生成し、同じ座標で異なるコードとなるピクセルの割合である Block-wise Hamming Distance(BlockHD) を画像間の相違度として計算する。リローカリゼーションを行う際には、現在の入力画像と全キーフレームとの BlockHD を計算し、BlockHD が小さい上位 5 つのキーフレームを入力画像と類似している画像の候補として選出する。Randomized Fern は画像全体を比較するのではなく、コード化した結果のみを比較するため処理が高速である。また、符号化の際に閾値と比較して大きいか小さいかのみで 0 か 1 かを決定しているため、同じ個所を撮影した際に環境の輝度が微妙に変化しても、コードは大きくは変化しない。そのため、カメラ姿勢が多少変化してもコードの変化が小さく、同じ位置で撮影していると認識できる。

しかし、ここまで述べたキーフレームベースの手法は原理上、過去にキーフレームを撮影した場所の近傍にしか復帰できない。すなわち、図 2.11 に示すように、現在のカメラが過去にキーフレームを取得するために撮影した領域から離れた位置にある時は、入力画像と最も類似した画像を抽出したとしても、その抽出された画像を取得した時のカメラ姿勢と現在のカメラ姿勢は大きく異なる。そのため、そのカメラ姿勢を使用してトラッキングを再開して情報を重畳表示したとしても、適切な情報提示ができない可能性が高い。

この問題を緩和する手法として、Gee らが提案した環境の三次元再構成モデルを用いる手法が存在する^[36]。この手法では環境の三次元再構成モデルを予め作成し、様々なカメラ姿勢からこのモデルを撮影した画像を生成する。これをキーフレームとして保存し、入力画像との SSD が小さい順に上位複数枚のキーフレームを選び、これらのキーフレームを生成した時のカメラ姿勢を取得する。そして、これらのカメラ姿勢のキーフレームと入力画像との SSD による重み付け平均を計算することにより推定カメ

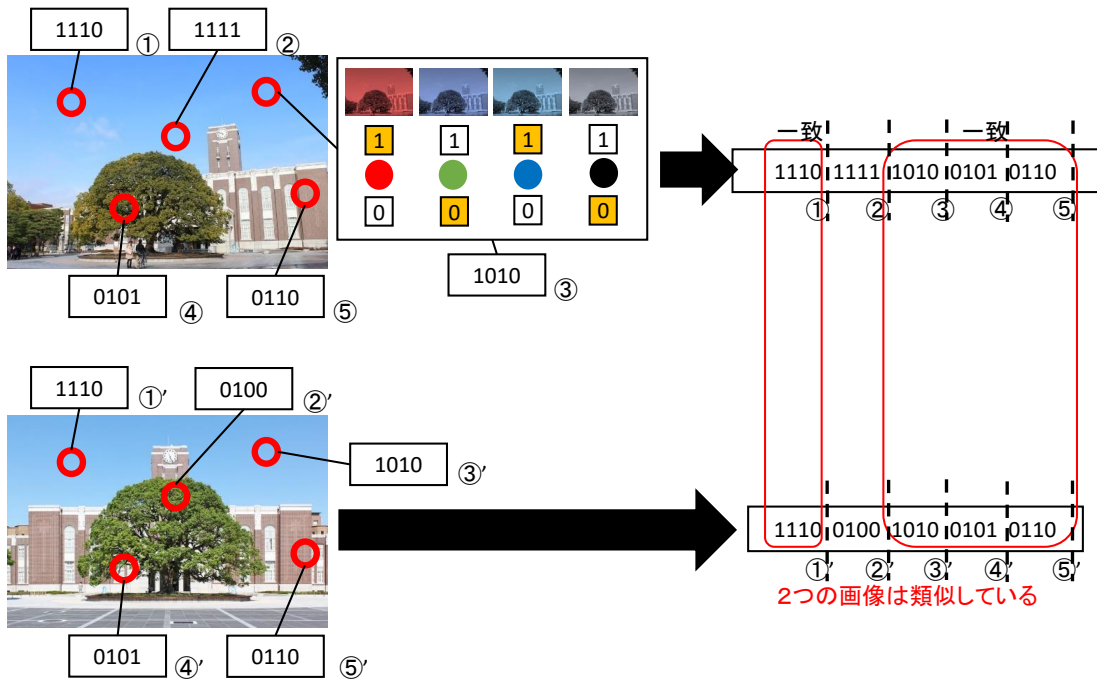
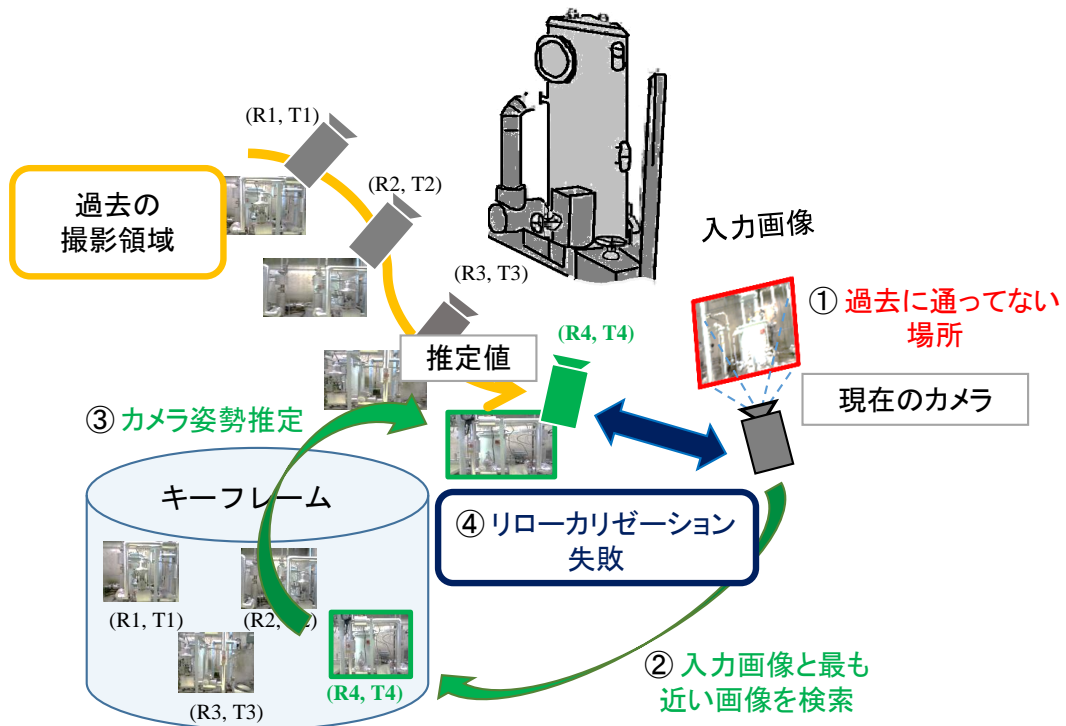


図 2.10: Randomized Fern における相違度の計算



Copyright (C) 2017 Japan Atomic Energy Agency

図 2.11: キーフレームベース手法を用いた場合の復帰可能範囲の限界

ラ姿勢を計算する。この Gee らの手法では、実際のカメラで環境を撮影した姿勢とは異なる姿勢でも三次元再構成モデルを用いてカメラ画像を生成し、生成した画像とその時のカメラ姿勢をキーフレームとして保存しておくことにより、実際のカメラで環境を撮影した姿勢と異なる姿勢であってもリローカリゼーションを実行できる可能性がある。Gee らの手法は小さな室内環境では有用性が示されているが、複雑なプラント環境では、オクルージョンなどにより撮影が難しい領域が多いため、図 2.12 に示すように環境の三次元再構成モデルが取得できずに欠損となる領域が増えるため、この手法はそのままでは適用できない。すなわち、SSD のような単純な手法では、欠損のあるモデルを用いて生成した画像と実際の入力画像との違いが大きくなり、同じ姿勢から撮影した画像であっても異なる画像であると認識する可能性が高くなり、入力画像と類似したキーフレームの探索が難しくなると考えられる。また、画像生成の際にカメラを配置した場所にしか復帰できないため、カメラの設置の仕方でもリローカリゼーション可能な範囲やカメラ姿勢推定の精度が変化する。そのため、対象とする環境ごとに適切なカメラの設置の仕方を検討する必要があるが、AR を用いた支援システムを利用する現場でそのような検討を適切に行うことは難しいという点も課題となる。



Copyright (C) 2017 Japan Atomic Energy Agency

図 2.12: 環境再構成モデルの欠損領域

2.1.2.2 キーポイントベースのリローカリゼーション手法

キーポイントベースの手法では、事前に撮影した画像から自然特徴点などの局所的な特徴（キーポイント）を抽出し、その特徴量と三次元位置をキーポイントデータベースとして保存しておく。リローカリゼーション実行時には、入力画像から抽出されたキー

ポイントと類似するキーポイントをキーポイントデータベースから探して対応付けを行い、対応付けられたキーポイントの画像上の二次元位置と三次元位置から PnP 問題を解くことにより入力画像のカメラ姿勢を推定する。事前に SfM などを用いて特徴点の三次元位置と各画像を取得した際のカメラ姿勢を計算しキーポイントデータベースに保存しておく手法がよく用いられる。このタイプの手法ではキーポイントと画像上の二次元特徴点を高速かつ高い成功率で対応付けることが重要であり、木構造データベースを利用する手法や^[37]、予めキーポイントをグループ分けしておき画像上の特徴点の特徴量と近いグループから対応するキーポイントを探索する手法^[38]などがある。これらの手法ではカメラ画像上のキーポイントとデータベース内のキーポイントを対応づけした結果から RANSAC^[39]によりランダムに PnP 問題を解くのに必要な最小数の対応付けを繰り返し選択し、それらを用いて求めたカメラ姿勢の精度を再投影誤差の大きさなどを用いて推定し、最も精度が高いと期待されるカメラ姿勢を推定する。キーポイントベースの手法はカメラ画像の一部のみを利用した手法であるため、原子力発電プラント内部のような複雑な環境では環境の一部が互いに隠蔽されるオクルージョンが結果に悪影響を及ぼしやすいという問題がある。

2.1.2.3 その他のリローカリゼーション手法

その他のリローカリゼーション手法として環境内に存在する物体の三次元モデルを利用したカメラ姿勢推定手法も存在する。SLAM++^[40]は椅子や机などの環境内に存在する物体の三次元モデルを予め用意しておき、RGB-D カメラで取得した画像を変換して得た三次元点群との位置合わせを行い、環境中の物体の姿勢を推定する。その後、環境内の物体の配置を表すマップ（グローバルマップ）を作成する。リローカリゼーション時には新たな入力画像から作成したマップ（ローカルマップ）と元のグローバルマップとのマッチングを行い現在のカメラ姿勢を推定する。この手法では環境に存在する物体のモデルを予め用意する必要がある。しかし、複雑なプラント環境では形状の種類が多いため、数多くのモデルを用意する必要があり労力が大きいという問題がある。

また、環境内の物体の姿勢を推定する手法としてテンプレートマッチングを行う手法^[41]もある。この手法では環境に存在する物体の三次元モデルを予め作成もしくは取得しておき、様々な角度からその三次元モデルを見た時の画像（テンプレート画像）を作成し、輝度の微分とデプスの微分情報により物体の三次元形状や外観を表現する LINEMOD^[42]を用いた入力画像とのテンプレートマッチングを行う。入力画像とマッ

チングしたテンプレート画像を作成した時のカメラの角度から、カメラを基準とした物体の姿勢を推定する。これらのようなモデルを利用した手法は古くからあるが、環境に存在する物体のモデルを用意するコストが大きいという問題があり、これまでは実用には至らなかった。しかし、カメラ姿勢を安定かつ正確に推定できる手法が多く、モデルが写っていればカメラを自由に動かすことができるという特長はプラント解体作業支援に向いている。

また、近年は機械学習を利用したカメラ姿勢推定手法も活発に研究され、多くの成果が挙げられている。カメラ姿勢が分かっている RGB-D 画像を使用して、RGB-D 画像を訓練データ、そのカメラ姿勢を教師データとして Regression Forest を学習させ、入力画像のカメラの姿勢を推定する手法^[43]がある。またニューラルネットワークを用いた手法^[44]も存在する。RGB 画像を訓練データ、そのカメラ姿勢を教師データとして Convolutional Neural Network(CNN) を学習させ、CNN により入力 RGB 画像のカメラ姿勢を推定する。しかし、どちらの手法でも事前に撮影した領域から遠く離れた場所でカメラの姿勢を推定することは困難である。また、学習のために大量のデータを用意する必要があるという問題もある。

2.2 研究の目的と意義

2.1.2 項で述べたように、ビジョンセンサを利用したりローカリゼーション手法には様々な手法が存在する。しかし、原子力発電プラントの解体作業支援を実現するためのローカリゼーション手法は、AR に関する専門知識を持たない人がユーザとして想定されるため、撮影時にユーザが自由にカメラを動かした場合でも利用できる必要がある。また、放射線などにより長時間現場に滞在することは難しく、現場での事前準備は少ないことが望ましい。そして、解体する物体や配管が環境内のどこにあるかをユーザに正確に指示できる程度の精度が必要となる。しかし、現時点では原子力発電プラント内でこれらの要求を完全に満たす手法は存在しない。先行研究^[45]によりキーフレームベース手法の Randomized Fern は実際の原子力発電プラント内においても正確かつ高速なカメラ姿勢の推定が可能であることを示されているが、この手法には事前に撮影していない領域では利用できないという問題がある。しかし、2.1.2.1 条で述べたように、環境の三次元再構成モデルを用いることによりこの問題を解決できる可能性がある。ただし、現時点で廃炉が計画されている原子力発電プラントは、建設されてから 30 年以上経過したものが多く、環境の 3DCAD モデルが存在しておらず、ま

た、手作業で改めて 3DCAD モデルを作成することは、膨大な労力が必要となるため困難である。しかし、2.1.1 項で述べたように、近年、環境の高精度かつ密な再構成モデルを取得できる手法が確立されてきた。そこで、本研究ではこのような環境再構成モデルを用いて、従来の手法が持つカメラの移動範囲が制限されるという問題を緩和するリローカリゼーション手法を開発することを目的とする。

本研究の提案手法では、原子力発電プラント内で環境再構成モデルを取得し、これを用いてレンダリングにより事前に撮影していない領域からも画像を生成し、キーフレームベースの手法と同様に類似画像検索処理によりカメラ姿勢の推定を行う。これにより、事前に撮影していない領域においてもリローカリゼーションが可能となると期待される。

本研究の新規性として以下の 2 点が挙げられる。

- 環境再構成モデルによるレンダリング画像を用いたリローカリゼーション手法の開発及び、実際の原子力発電プラント内での性能の評価
- 実際の原子力発電プラントで生成したレンダリング画像に対応できる類似画像検索手法の開発

現時点では、環境再構成モデルによるレンダリング画像を用いたリローカリゼーション手法を実際の原子力発電プラントで実行して評価した研究はほぼ存在しない。また、実際の原子力発電プラントでは、図 2.12 に示したような環境再構成モデルが取得できずに欠損となる領域や背景が存在するため、入力画像を撮影した際と同じ場所からレンダリングにより生成した画像が入力画像と異なると判断される場合がある。そこで、本研究ではこの問題に対処できる類似画像検索手法を開発する。

本提案手法により、リローカリゼーションで復帰可能な領域が限られるという制限を緩和できれば、AR を用いた支援システムを利用する際に、カメラを自由に動かすことが可能になる。これによりリローカリゼーションなどに関する専門知識を持たない人でも容易に利用可能なシステムが実現可能になる。このことは実際のプラント解体作業に支援システムを導入するにあたり重要であり、将来的には AR による解体作業支援システムの実用化に寄与できると期待される。

第 3 章 環境再構成モデルを用いたリローカリゼーション手法の提案

本章では、まず、本研究で提案する環境再構成モデルを用いたリローカリゼーション手法の概要を述べ、次に、提案手法の詳細を述べる。最後に本提案手法の予想される利点と欠点について述べる。

3.1 提案手法の概要

本節では、提案手法である環境再構成モデルを用いたリローカリゼーション手法の概要を述べる。まず、図 3.1 に、提案手法で扱うデータと全体の処理の流れを示す。2.2

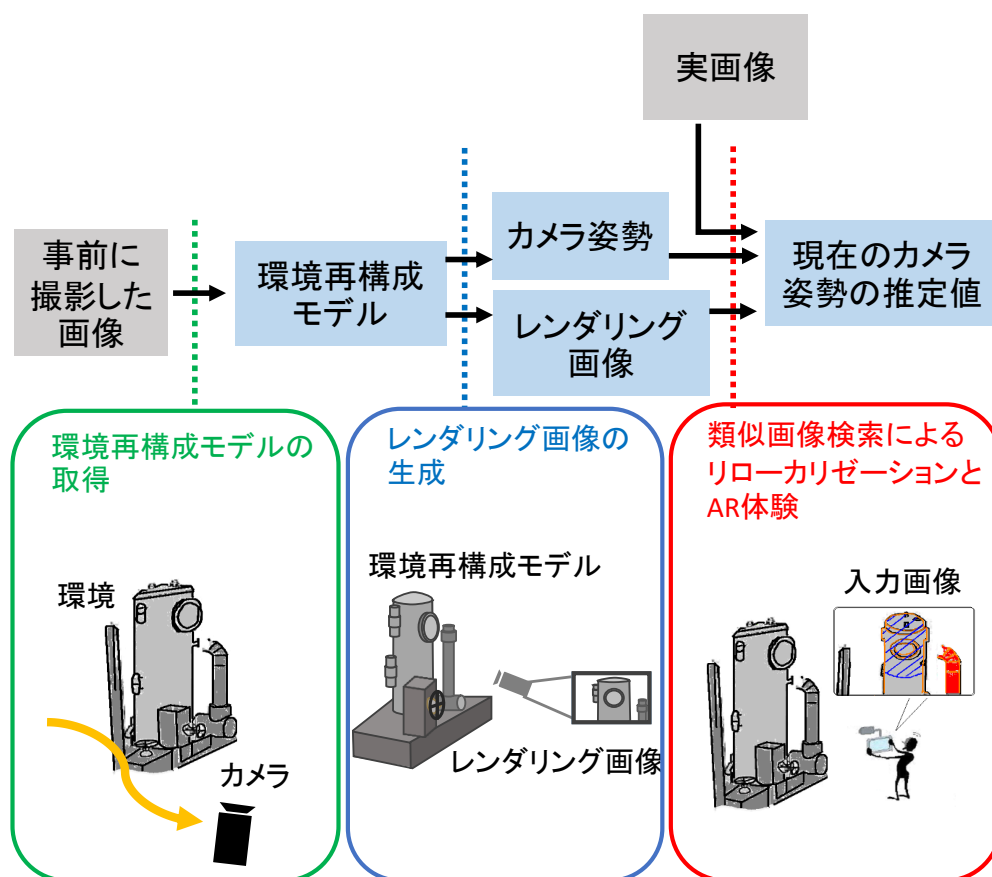
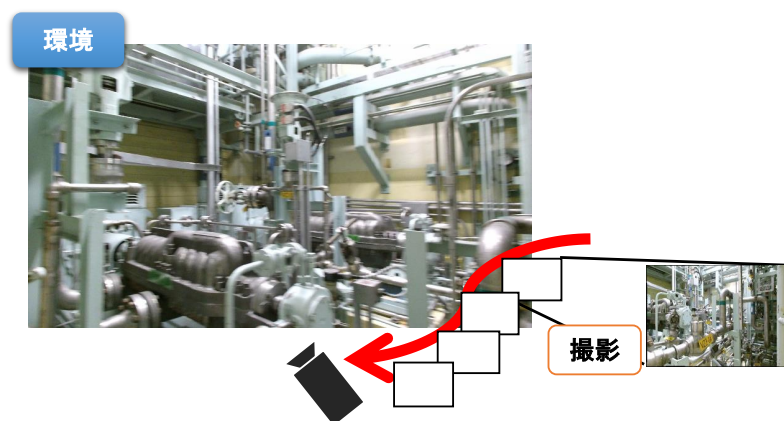


図 3.1: 提案手法で扱うデータと処理の流れ

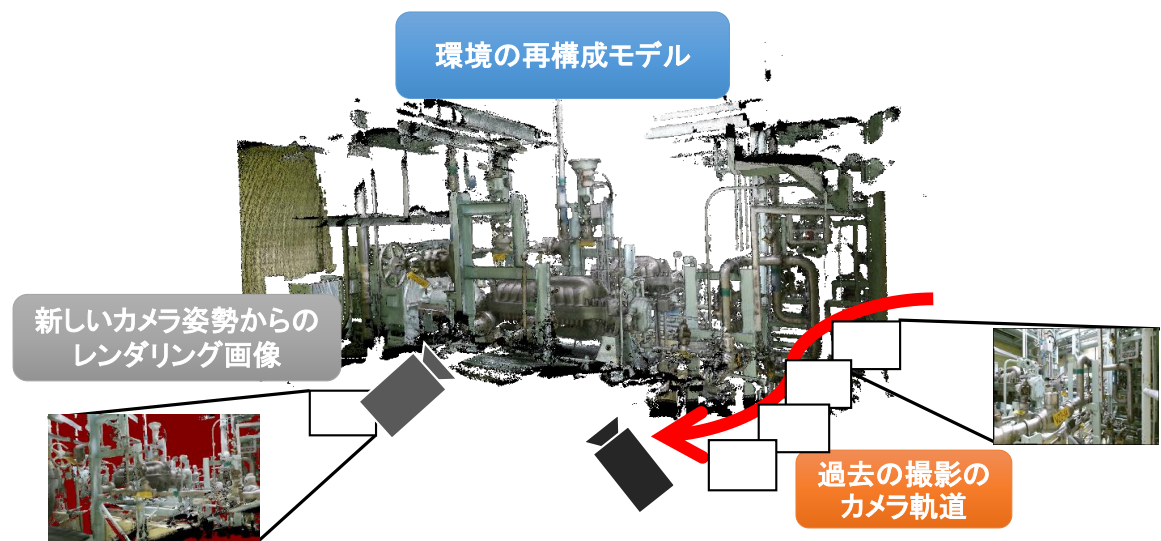
節で述べたように、本研究の提案手法はキーフレームベースの手法である。ただし、本提案手法では、図3.2に示すように予め環境を撮影しておき、図3.3に示すように事前に撮影した画像から作成した環境再構成モデルを様々なカメラ姿勢からレンダリングして得られる画像と、その時のカメラ姿勢を組にしてキーフレームデータベース(キーフレームDB)に保存する。そして、実環境をカメラで撮影して得た画像(以下、実画像)とキーフレームDB内の環境再構成モデルを用いて生成したレンダリング画像(以下、レンダリング画像)を比較し、実画像と類似するレンダリング画像を生成するカメラ姿勢を求める。この推定したカメラ姿勢を利用してトラッキングを再開させ、ARによるプラント解体作業支援などのAR体験を実現する。本提案手法は、事前に環境を撮影した時のカメラ姿勢とは異なるカメラ姿勢からもレンダリング画像を生成することにより、事前に実環境では撮影していない領域でもリローカリゼーションを成功させることができる。なお、本提案手法では、環境再構成モデルを生成するためにプラント環境を事前に撮影する際と、リローカリゼーションを利用するためにプラントを撮影する際の両方で、RGB画像とカメラから環境までの距離情報を持つデプス画像を同時に取得できるRGB-Dカメラを使用することを想定する。



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.2: トラッキング対象とする環境の撮影

しかし、環境再構成モデルを用いて生成したレンダリング画像を使用することにより、いくつかの問題が発生することが予想される。まず、配管などが多数配置されたプラント環境では、物が他の物の陰になって隠れるオクルージョンが多発するため、カメラに写らない領域が多くなることが予想される。そのため、プラント環境内には存在するにも関わらず、環境再構成モデルには含まれない欠損領域の割合が多くなり、実



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.3: 環境再構成モデルを用いたレンダリング画像の生成

画像に映っている物が環境再構成モデルを用いて生成されたレンダリング画像には現れない場合が頻発することが予想される。また、現在の技術では、環境再構成モデルの色と実際のプラント環境の色を正確に合わせることが難しいため、環境再構成モデルの色と実画像上の色が異なるなどの問題も考えられる。これらの問題により、実画像と類似するレンダリング画像を正しく検索することが難しくなる。そのため、本提案手法を実現するためには、以上に述べたような、実画像とレンダリング画像の間の差異に起因する問題を解決する必要がある。

3.2 キーフレームデータベースの構築

本節では事前準備としてキーフレーム DB を構築する処理について述べる。まず図 3.4 にキーフレーム DB を構築する処理の全体の流れを示す。キーフレーム DB を構築する処理は、以下の 4 つのステップで構成される。

1. まず RGB-D カメラを用いて環境を撮影し、得られた画像を用いて環境再構成モデルを作成する。
2. 次に、リローカリゼーション処理を実行する必要がある領域、例えば、解体作業支援で AR を用いる際にリローカリゼーションが必要となる領域（作業領域）を

設定し、その領域内で位置 (x, y, z) と方向 (yaw, pitch, roll) の 6 パラメータを変化させてカメラ姿勢をサンプリングする。

- 2でサンプリングされたカメラ姿勢から環境再構成モデルを撮影した時のレンダリング画像を生成する。
- サンプリングされた全てのカメラ姿勢とそのカメラ姿勢でのレンダリング画像を組にしてキーフレーム DB に格納する。この時、レンダリング画像そのものではなく、後の類似画像検索処理で利用する特徴のみをレンダリング画像から抽出して格納する。

以下で各処理の詳細を述べる。

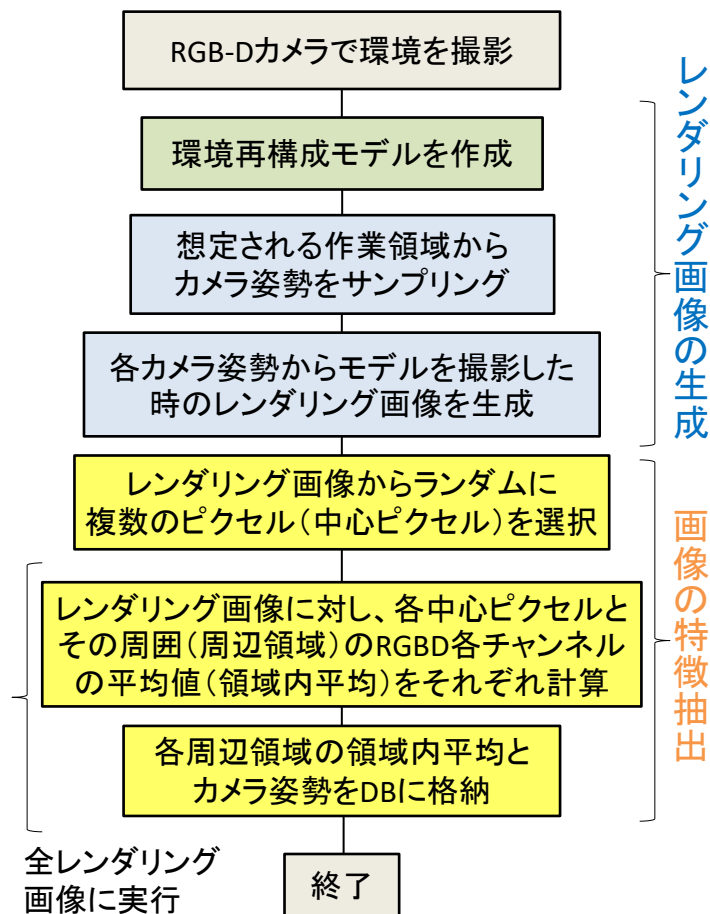
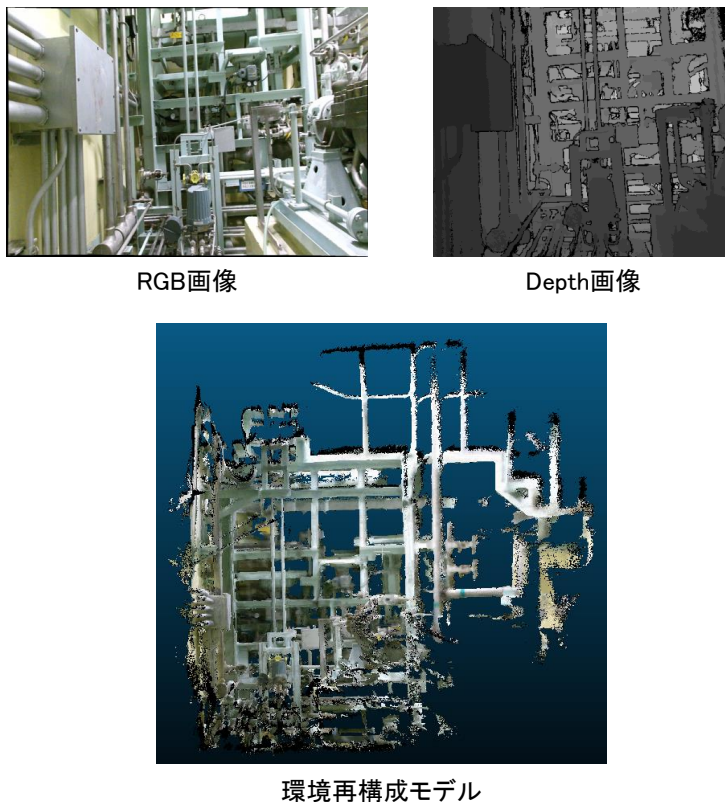


図 3.4: キーフレームデータベースの構築処理の流れ

3.2.1 環境再構成モデルの作成

2.1.1 項で述べた環境再構成技術を利用して、環境の撮影画像から環境再構成モデルを作成する。本提案手法では、キーフレーム DB から実画像と類似したレンダリング画像を正確に検索するために、密で色情報を持ち、環境の三次元構造が高精度で再現された環境再構成モデルを利用することが望ましい。そこで、本研究では、より精度の高い環境再構成モデルを得るために、RGB-D カメラを用いた手法であり、広域で密な環境再構成モデルを生成できる InfiniTAM を利用する。InfiniTAM では、撮影した RGB 画像とデプス画像から図 3.5 に示すような環境再構成モデルを生成すると同時に、各実画像を取得した時のカメラ姿勢を推定する。InfiniTAM では、環境を撮影する前にキャリブレーションによりカメラのレンズの歪み補正を行う必要があるが、本研究では歪み補正済みのカメラを使用することを前提とする。本研究のキャリブレーションの詳細は 4.3.1 項で述べる。



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.5: RGB 画像とデプス画像及び三次元環境再構成モデル

3.2.2 レンダリング画像の生成と特徴量の計算

次に環境再構成モデルを用いてレンダリング画像を生成する。レンダリング画像を生成する際には、プラント解体作業時などで実際のカメラが移動する領域（作業領域）内にカメラを繰り返し設置し、そのカメラ姿勢から環境再構成モデルを撮影した画像を生成する。このカメラを設置した場所でのみリローカリゼーション時にカメラ姿勢を推定してトラッキングを再開できるため、作業領域全体にカメラを設置しなければならない。そこで、図3.6に示すように、環境再構成モデルを読み込んだ後に作業領域を設定し、その領域全体からカメラ姿勢のサンプリングを行い、それぞれのカメラ姿勢からのレンダリング画像を生成する。このサンプリングの仕方により、リローカリゼーションを成功させることができる範囲や類似画像検索の処理時間などが変わる可能性がある。サンプリングの方法については3.2.3項で詳細に検討する。

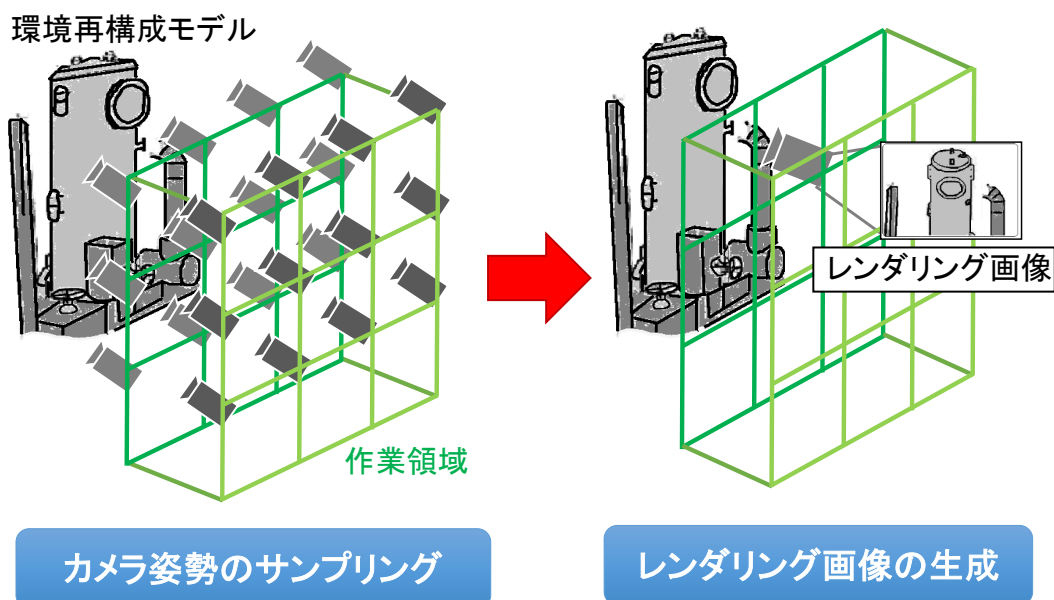


図 3.6: カメラ姿勢のサンプリングとレンダリング画像の生成

次に、レンダリングの際のカメラのパラメータの設定について述べる。カメラによる撮影をコンピュータ上で再現するためには、設置するカメラの視野角など様々なパラメータを実際に撮影で利用するカメラのパラメータに合わせる必要がある。本提案手法で使用を想定している RGB-D カメラは、図 3.7 に示すように、通常の RGB カメラに加え、ピクセル毎に環境までの距離を測定するデプスセンサを内部に持つ。一般の RGB-D カメラは、RGB カメラとデプスセンサが別々に装着されており、それらの焦点位置が異なっている。そこで、予め使用する RGB-D カメラのキャリブレーション

を行い、RGBカメラとデプスセンサの内部パラメータと外部パラメータを取得しておく。ここでの内部パラメータとはカメラの主点、焦点距離、歪み係数を指し、外部パラメータとはRGBカメラ座標系 F_c とデプスセンサ座標系 F_d の間の並進移動ベクトル t と回転行列 R を指す。そして、図 3.8 に示すような透視投影モデルを用いて、これらの内部パラメータによってレンダリングの際のRGBカメラとデプスセンサの解像度と視野角を実際のRGB-Dカメラと一致させる。そして、環境再構築モデルを用いてレンダリングする際には、両カメラの焦点位置のズレを考慮した上で、RGB画像用とデプス画像用に2回レンダリングし、図 3.9 に示すようなそれぞれのレンダリング画像を取得する。

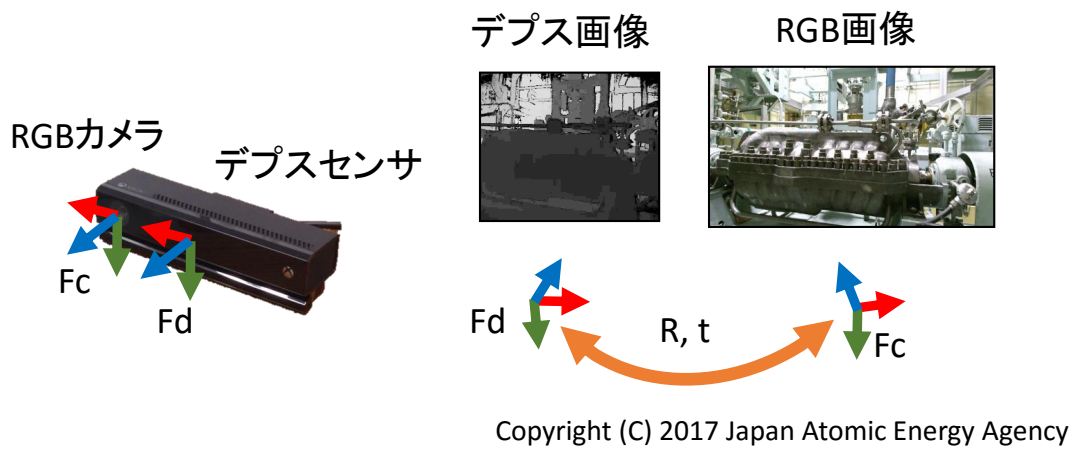


図 3.7: RGB-D カメラの RGB カメラとデプスセンサ

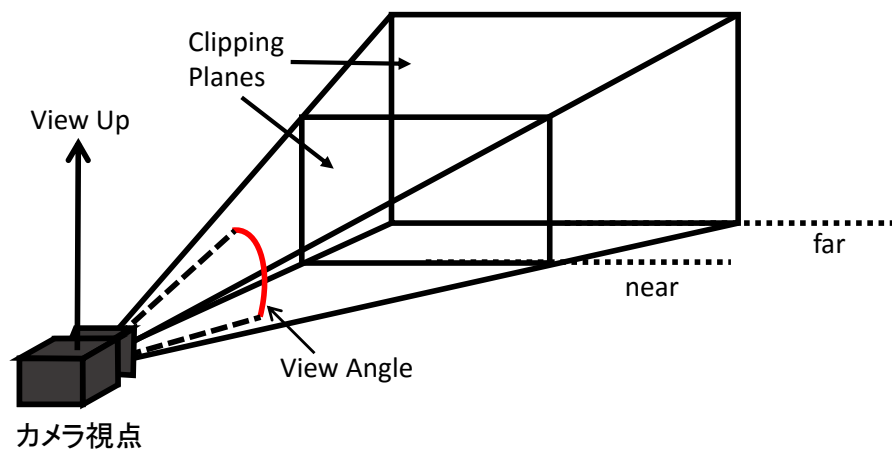
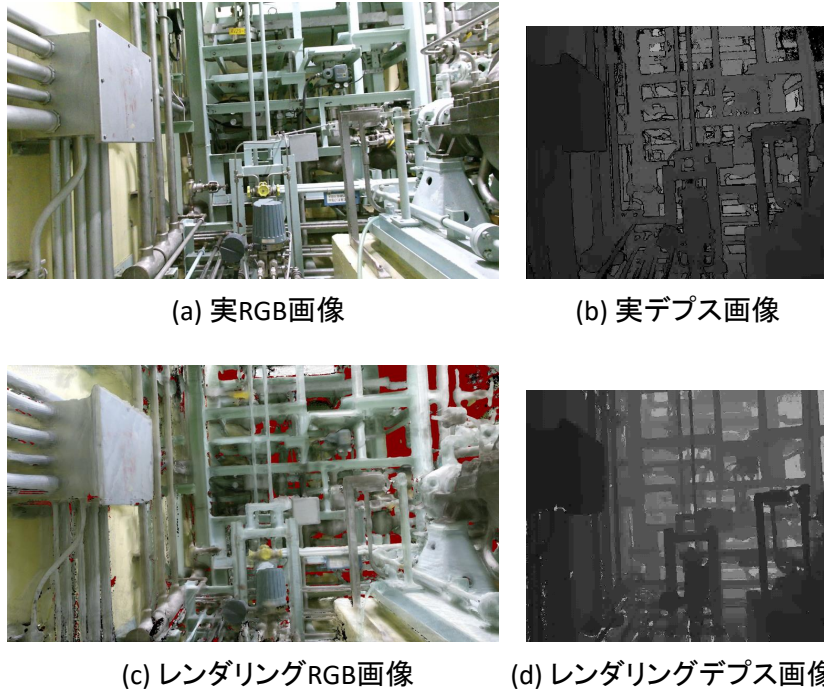


図 3.8: カメラの透視投影モデル

なお、レンダリングの際には、環境光を無効とし、実RGB画像から得られた色を直接用いることにより、レンダリング画像で実画像に近い色が得られるようにする。レ



Copyright (C) 2017 Japan Atomic Energy Agency

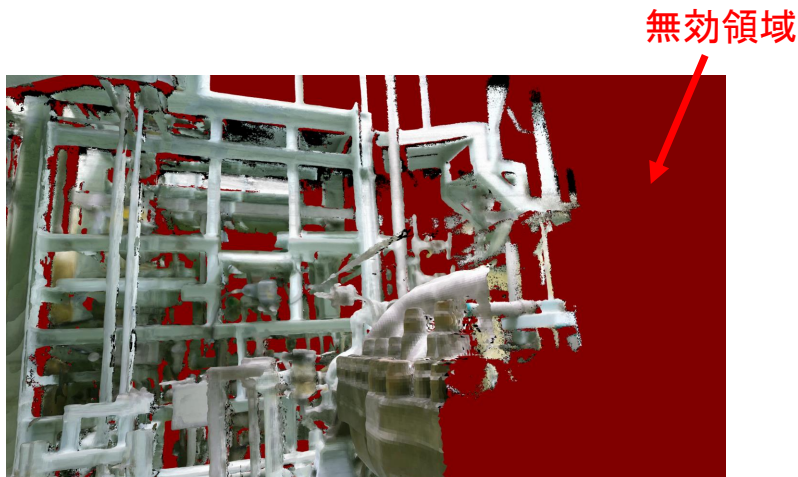
図 3.9: 実画像とレンダリング画像

レンダリングデプス画像はデプス画像用カメラのデプスバッファから取得する。デプスバッファはカメラの画像の各ピクセルの奥行き情報を格納している。デプスバッファの値は一般に図 3.8 の clipping plane の near で 0、far で 1 になるように変換されている。レンダリング画像の生成に必要なものはカメラから環境再構成モデルまでの距離であり、式 3.1 で求めることができる。

$$Z = \frac{\text{near} * \text{far}}{(D * (\text{far} - \text{near}) - \text{far})} \quad (3.1)$$

ここで、Z は各ピクセルにおける環境再構成モデルまでの距離 (モデルが写っていないピクセルでは背景を表す処理系に依存する値)、D はデプスバッファの値、far、near はそれぞれ clipping plane の far と near のカメラからの距離となる。以上の処理によりレンダリングデプス画像を生成する (実画像と異なり、レンダリング画像には図 3.10 の赤い領域のような環境再構成モデルが写らない領域が存在する。この背景が写る領域を以降では無効領域と呼ぶ)。

次に、図 3.4 の後半に示したように、生成されたレンダリング画像に対して画像の特徴を抽出する処理を実行し、カメラ姿勢と抽出した特徴を組みにしてキーフレーム DB に保存する。この処理は以下の 3 つのステップで構成される。



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.10: レンダリング画像の無効領域

1. 最初のレンダリング画像からランダムに複数のピクセル（中心ピクセル）を選ぶ。
2. 中心ピクセルとその周辺のピクセル（周辺領域）の RGBD それぞれの平均値を画像の特徴として計算する。この結果、中心ピクセル毎に RGB3 色とデプスで計 4 個の平均値が得られる。以降、これらの平均値を領域内平均と呼ぶ。
3. 全てのレンダリング画像に対して、1 で決定したものと同一座標の中心ピクセルの周辺領域から領域内平均を計算し、カメラ姿勢と組にしてキーフレーム DB に格納する。

レンダリング画像で周辺領域の領域内平均を計算する際の具体的な処理の流れを図 3.11 に示す。環境再構成モデルを用いて生成したレンダリング画像からキーフレーム DB を作成する場合は、図 3.10 に示したレンダリング画像の無効領域が後の類似画像検索に悪影響を及ぼす可能性がある。そこで、本提案手法ではランダムに選んだピクセルのみを利用するのではなく、その周囲の情報も利用する。これにより、選ばれたピクセルが無効領域であってもその周囲に有効な領域があればそれを利用できる。また同じ環境を撮影している場合でも、環境再構成モデルの色が実画像の色と異なる可能性や環境再構成モデルの歪みや位置ずれにより実画像と位置がずれる可能性があるが、一定の領域の平均値を用いることによりこれらの違いによる類似画像検索への影響を緩和できると考えられる。また、本提案手法では周辺領域に対して信頼度という指標を導入し、後の類似画像検索の処理では、信頼度が予め設定された閾値以上

の周辺領域の領域平均のみを用いる。周辺領域内の無効領域の処理や信頼度の詳細は3.3.2項で述べる。

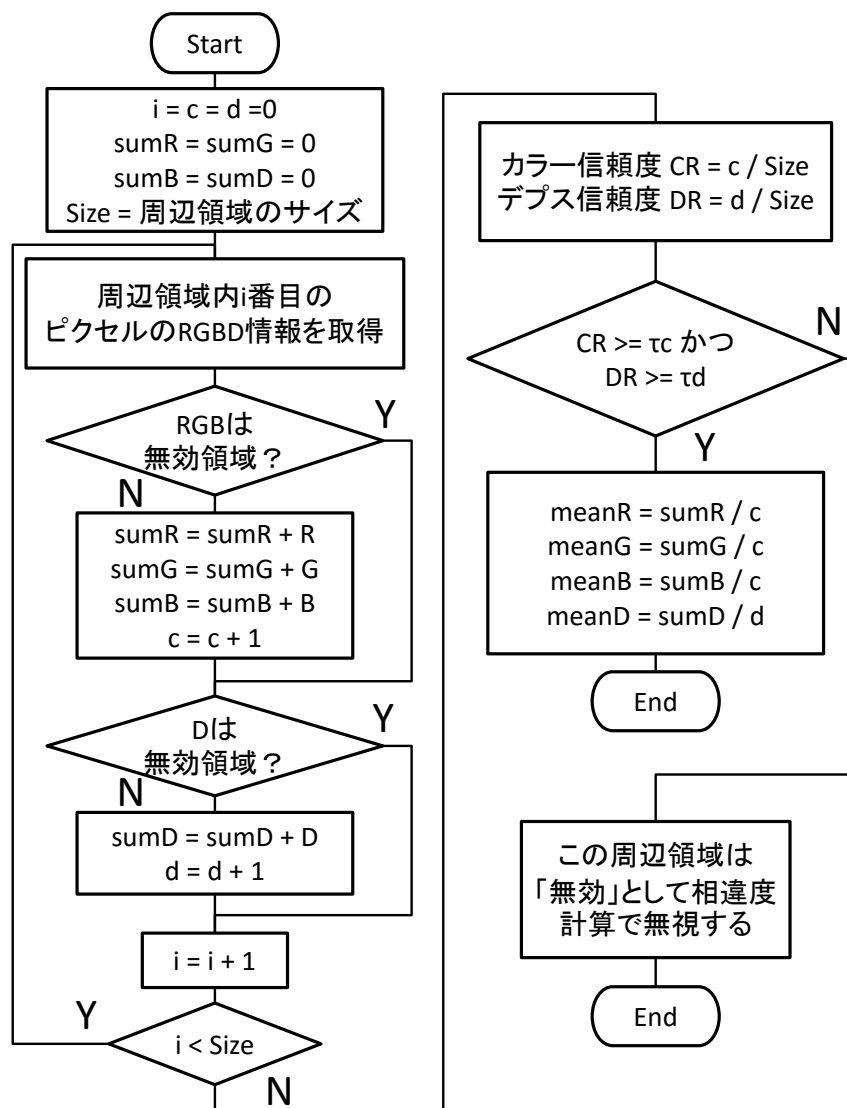


図 3.11: レンダリング画像における周辺領域の領域内平均の計算処理の流れ

3.2.3 レンダリング間隔の設定

3.2.2項で述べたように、カメラ姿勢のサンプリングの仕方により、本提案手法のローカリゼーションが可能な範囲や処理時間が変わる可能性がある。そこで本研究では、等間隔にカメラ姿勢のパラメータを変化させてサンプリングする等間隔サンプリングと、あるカメラ姿勢のレンダリングデプス画像のデプス値のばらつきを利用して次のカメラ姿勢のパラメータを決定する深度地標準偏差サンプリングの2種類のサン

プリング方式を検証する。これらのサンプリング方式を用いて、予め定めた作業領域内でカメラの位置 (x, y, z) と方向 (yaw, pitch, roll) の6パラメータを設定する。本項ではこれら2種類のサンプリング方式によりカメラ姿勢のパラメータを設定する方法について述べる。

等間隔サンプリングでは、作業領域内でカメラ姿勢の各パラメータを一定のステップで変化させ、その時のカメラ姿勢をサンプリングする。等間隔サンプリングではステップを十分小さく設定すれば、実際のカメラが作業領域内のどこにある場合でも、リローカリゼーション時にトラッキングを再開させるのに必要な精度のカメラ姿勢の推定が可能となると期待される。しかし、ステップが小さすぎる場合、サンプリングするカメラ姿勢の数が膨大となり、レンダリングに時間がかかる。一方、ステップが大きすぎる場合、リローカリゼーションに失敗する可能性が高くなる。

適切なステップの大きさは環境やリローカリゼーションの後に利用するトラッキングの誤差最小化手法に依存することが予想されるため、環境や手法に合わせてステップ間隔を変更する方が望ましい。そこで、リローカリゼーション実行後のトラッキングの誤差最小化手法が最適解に収束しやすい環境ではステップを大きく、局所解に陥りやすい環境ではステップを小さくして密にサンプリングすることにより、効率の良いサンプリングが可能になると期待される。そこで本研究では、レンダリングデプス画像のデプス値のばらつきを表すと考えられる標準偏差を利用したサンプリング方式として、深度値標準偏差サンプリングを提案する。本サンプリング方式では、あるカメラ姿勢でのレンダリング画像の生成時に、レンダリングデプス画像の無効領域を除いて標準偏差を計算し、この値を元に次のサンプリングのステップの大きさを決定する。誤差最小化手法はトラッキング手法により異なるが、例えば、RGB-Dカメラを利用したトラッキングの代表的手法である KinectFusion や RGB-D SLAM^{[46][47]} ではカメラ姿勢の推定に Iterative Closest Point(ICP) アルゴリズム^[48] を利用している。ICP アルゴリズムでは、2つの三次元点群から最も近い点を対応点として選び、対応点間の距離が最小となる変換行列を求める。デプス値の標準偏差が大きいということはデプス値のばらつきが大きい、つまり複雑な形状をした環境である。このような平坦ではない三次元的に特徴のある環境では、ICP を用いるトラッキングは成功しやすい。そのため、サンプリングのステップを大きくしても最適解に収束しやすい。逆にデプス値の標準偏差が小さい環境、すなわち起伏の少なく三次元的な特徴の少ない環境では ICP を利用したトラッキングは失敗しやすい。そのため、ステップを小さくすることにより最適解に収束しやすくする。そこで、本研究では、式3.2を用いてサンプリング

のステップ S を決定する手法を提案する。

$$S = a \times D \times C \quad (3.2)$$

ここで、 a は係数、 D はレンダリングデプス画像の無効領域を除くデプス値の標準偏差、 C は基準となる一定のステップを表す。式 3.2 により、複雑な環境ではステップを大きくして疎にカメラ姿勢のサンプリングを行い、平坦な環境ではステップを小さくして密にサンプリングを行う。ただし、式 3.2 を用いてステップを決定する方法は、例えば平面を異なる角度から撮影する場合など、同じ環境であっても異なるステップでサンプリングする場合があります。最適な方法であるとは言えず、本来はレンダリングデプス画像から三次元点群を求めた結果に平面をフィッティングさせ、その平面と点群の各点との偏差を用いるなどの方法によりデプス値のばらつきを表現する必要がある。しかし、広域な作業領域でサンプリングを行う場合、平面フィッティングを繰り返し実行すると処理に時間が掛かるため、本研究では、処理を高速化するために、単純にデプス画像の標準偏差を用いる。また、係数 a は実験的に決定する。この詳細は 4.3.4 項で述べる。

3.3 類似画像検索とカメラ姿勢の推定

3.3.1 類似画像検索の概要

類似画像を検索する処理では、実画像から抽出された特徴とキーフレーム DB に格納された各キーフレームの特徴を比較し、実画像と類似した特徴を持つレンダリング画像をキーフレーム DB から抽出する。

リローカリゼーション実行時の類似画像検索処理の流れを図 3.12 に示す。この処理は、以下の 3 つのステップで構成される。

1. キーフレーム DB を作成する場合と同様に、RGB-D カメラで取得した実画像から各周辺領域の領域内平均を計算する。この時、キーフレーム DB を作成する際に設定した中心ピクセルと同じ座標の周辺領域を用いる。
2. キーフレーム DB の全てのキーフレームに対して、図 3.13 に示すように、事前準備で計算されたレンダリング画像の領域内平均と実画像の領域内平均を比較することにより、実画像とキーフレームの相違度を計算する。

3. キーフレーム DB 内の全てのキーフレームの中で相違度が小さい上位5つのキーフレームのカメラ姿勢を提案手法による現在のカメラ姿勢の推定値の候補とする。

RGB 画像の RGB 各チャンネルは 0 から 255 の値を取る場合が多く、デプス画像の D チャンネルはカメラや処理系に依存するが、距離を mm で表した値となり、0 から数千の値となる場合が多い。そこで、本提案手法はデプス画像の最大値が 255 になるように 256 段階に線形変換して領域内平均を計算する。なお、中心ピクセルの数や周辺領域の大きさなどのパラメータについては実験的に決定する。この詳細は 4.3.6 項で述べる。以下で各処理の詳細を述べる。

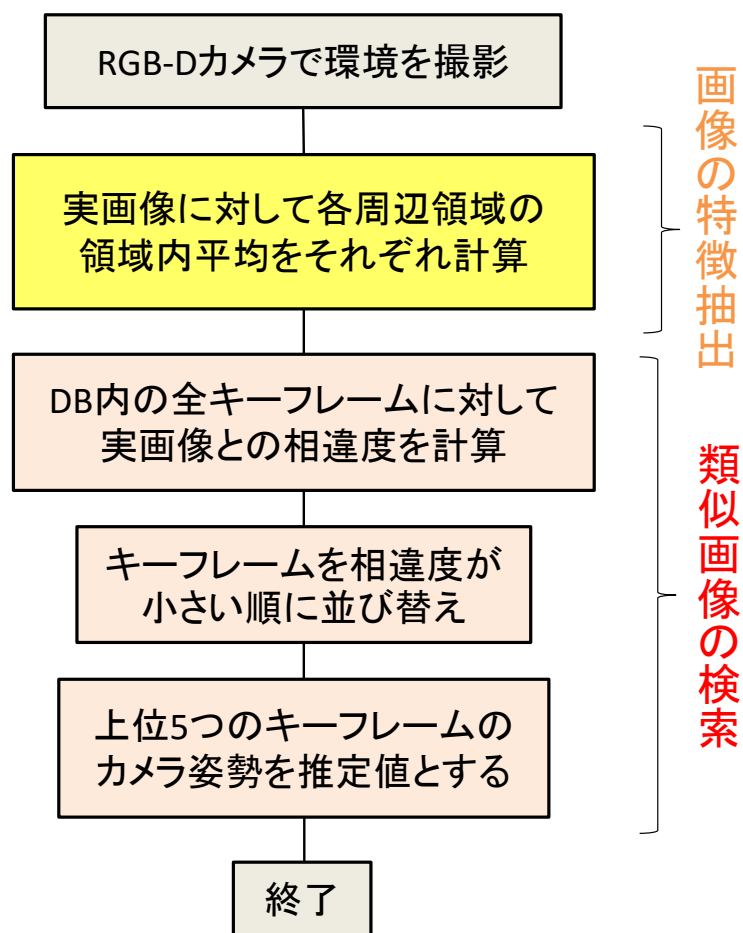
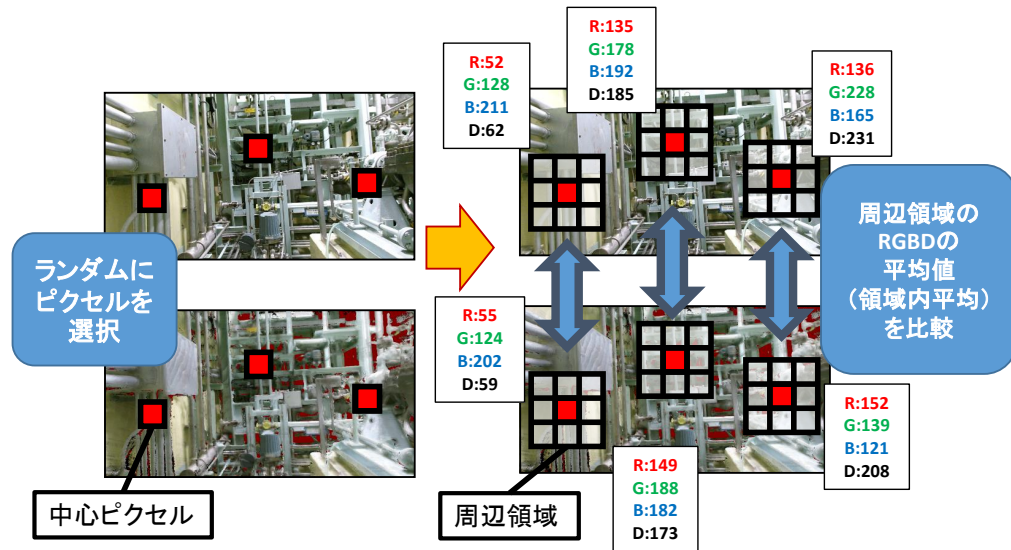


図 3.12: 類似画像検索処理の流れ

3.3.2 画像特徴の抽出と無効領域の処理

3.2.2 項で述べたように、レンダリング画像には背景が写る無効領域が存在する場合があります、これにより単純にレンダリング画像と入力実画像の RGBD 情報を比較した場



Copyright (C) 2017 Japan Atomic Energy Agency

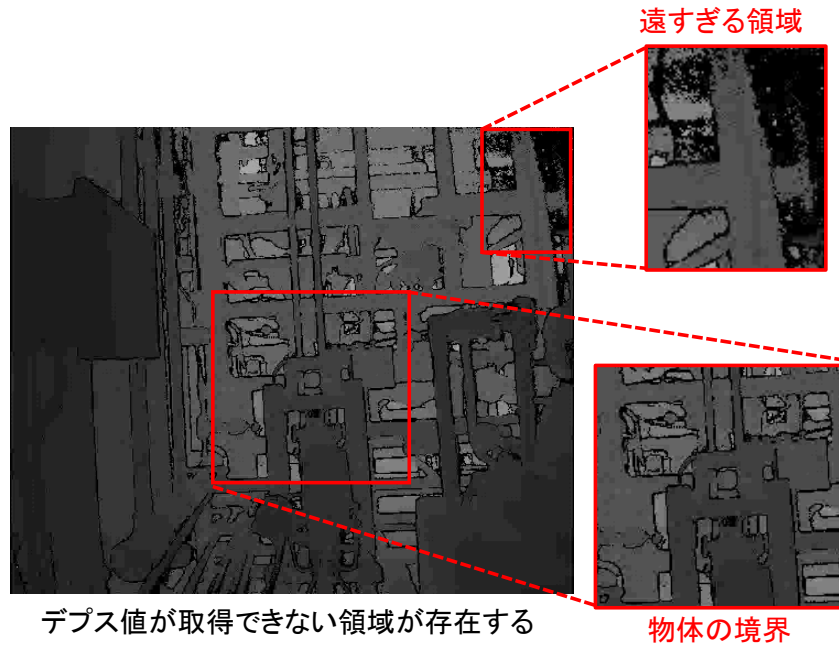
図 3.13: 周辺領域を用いた画像間の相違度の計算

合は相違度が大きくなる可能性がある。一方、図 3.14 に示すように実デプス画像にもデプスセンサの距離情報の取得範囲の限界によりデプス値が取得できない無効領域が存在するため、レンダリングデプス画像との相違度が大きくなる可能性がある。そのため、入力実画像とレンダリング画像の特徴抽出や相違度を計算する処理ではこのような無効領域を考慮し、排除する必要がある。

そこで、本提案手法では周辺領域の領域内平均を求める際に、信頼度を求め、信頼度が予め決定した閾値以上の周辺領域のみを画像間の相違度の計算に用いる。まず、実画像で周辺領域の領域内平均を計算する際の処理の流れを図 3.15 に示す。ここで、信頼度 R を式 3.3 で定義する。

$$R = \frac{V}{P} \quad (3.3)$$

V は周辺領域内の有効ピクセル（無効領域ではないピクセル）数、 P は周辺領域内の全ピクセル数とする。信頼度が閾値未満の周辺領域は「無効」として、以降の画像間の相違度の計算では排除される。周辺領域の信頼度が閾値以上となる場合は有効ピクセルのみを用いて領域内平均を計算し、以降の画像間の相違度の計算で利用する。図 3.15 に示す実画像の処理の流れからわかるように、実 RGB 画像においては無効領域は存在しないため、信頼度の計算は行わない。すなわち、本提案手法における信頼度は、レンダリング画像のカラー信頼度、デプス信頼度及び実画像のデプス信頼度の 3 種類となる。本提案手法では周辺領域を用いることにより、図 3.16 に示すように多少の無



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.14: 実デプス画像の無効領域

効領域が含まれる場合でも周囲の有効ピクセルの情報を用いて領域内平均を計算でき、これにより類似画像検索の精度を維持できると考えられる。そして、信頼度を用いて有効ピクセルが少ない周辺領域を排除することにより、相違度計算に信頼できないノイズ情報が含まれないようにして類似画像検索の精度の低下を防ぐ。

3.3.3 相違度の計算

本項ではここまで述べた領域内平均を用いた画像間の相違度の計算方法について述べる。実画像とキーフレームのレンダリング画像との相違度の計算の流れを図 3.17 に示す。同じ座標の実画像の周辺領域とレンダリング画像の周辺領域の両方が、図 3.11、図 3.15 で示した処理で「無効」と判定されなかった場合のみ、この周辺領域の情報は画像間の比較において信頼できる情報であると判断し、領域内平均の二乗誤差を計算する。「無効」でない周辺領域の領域内平均の二乗誤差の和を計算し、「無効」でない周辺領域の数で割り、これを相違度として用いる。相違度 E は

$$E = \frac{\sum_{i=1}^{Size} s_i ((R_{reni} - R_{reali})^2 + (G_{reni} - G_{reali})^2 + (B_{reni} - B_{reali})^2 + (D_{reni} - D_{reali})^2)}{\sum_{i=1}^{Size} s_i} \quad (3.4)$$

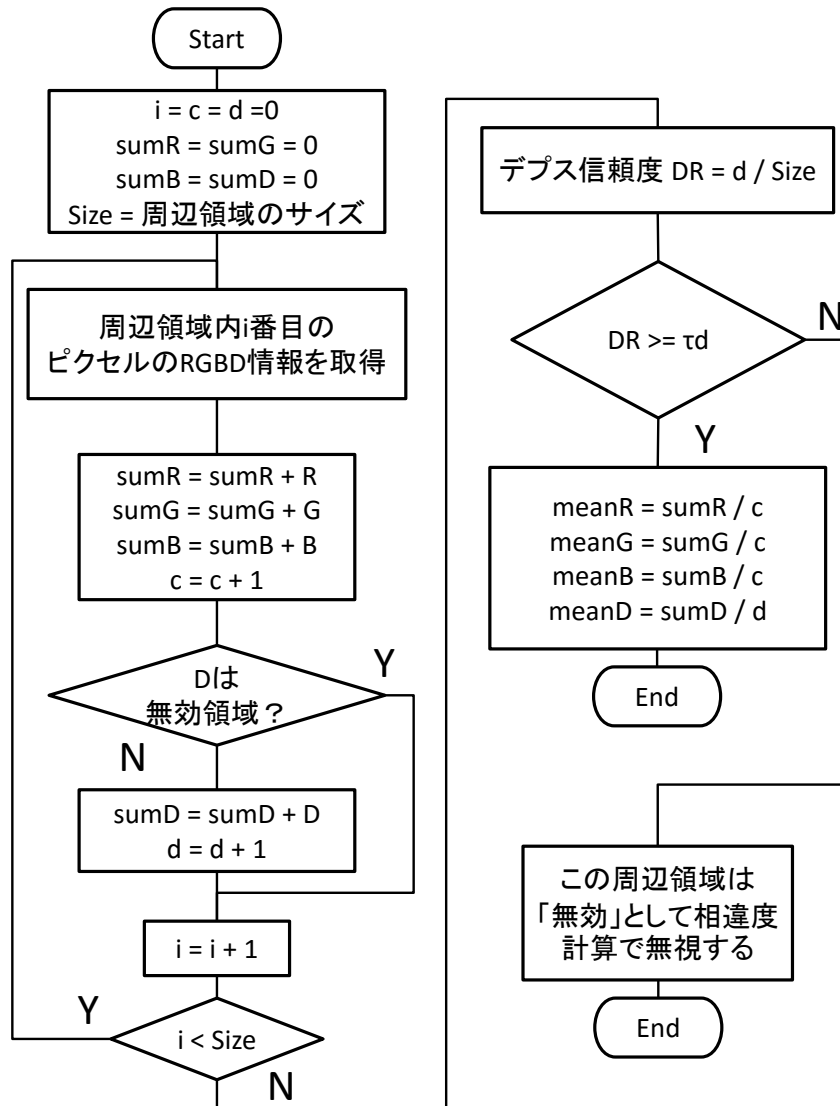
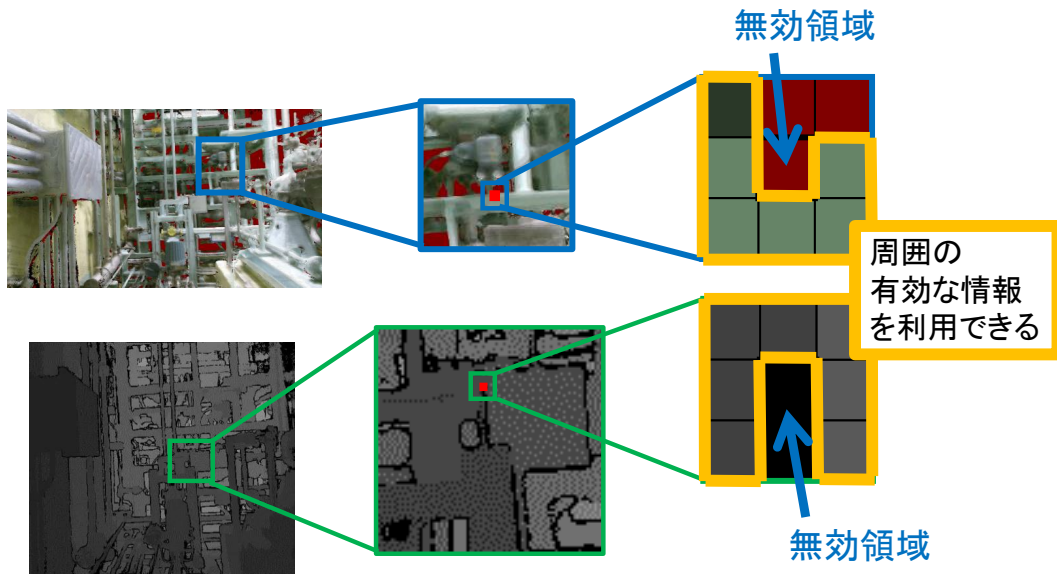


図 3.15: 実画像における周辺領域の領域内平均の計算処理の流れ



Copyright (C) 2017 Japan Atomic Energy Agency

図 3.16: 周辺領域による有効ピクセルの利用

$$s_i = \begin{cases} 0 & (\text{Invalid}) \\ 1 & (\text{otherwise}) \end{cases} \quad (3.5)$$

で表される。 R_{reni} はレンダリング画像の i 番目の周辺領域の R 値の平均値、 R_{reali} は実画像の i 番目の周辺領域の R 値の平均値を表す。 G_{reni} 、 B_{reni} 、 D_{reni} はそれぞれレンダリング画像の i 番目の周辺領域の G 値、 B 値、 D 値の平均値を表す。 G_{reali} 、 B_{reali} 、 D_{reali} はそれぞれ実画像の i 番目の周辺領域の G 値、 B 値、 D 値の平均値を表す。Size は周辺領域の総数である。 s_i は実画像とレンダリング画像の i 番目の周辺領域のどちらか一方でも「無効」の場合は 0、それ以外の場合は 1 となる。

3.3.4 カメラ姿勢の推定

次に、本提案手法では実画像と類似したレンダリング画像を生成した際のカメラ姿勢を推定カメラ姿勢とする。具体的には、キーフレーム DB 内の全キーフレームに対し実画像との相違度を計算して、相違度が小さい上位 5 つのキーフレームのカメラ姿勢を推定カメラ姿勢の候補とする。この相違度はランダムに選ばれる中心ピクセルの位置により変わるため、実行の度に相違度の順位が多少変動する可能性がある。この影響を緩和するため、上位複数枚を推定値とし、これらの推定値のどれかでトラッキングを再開させることができればリローカリゼーション成功とする。推定値の候補と

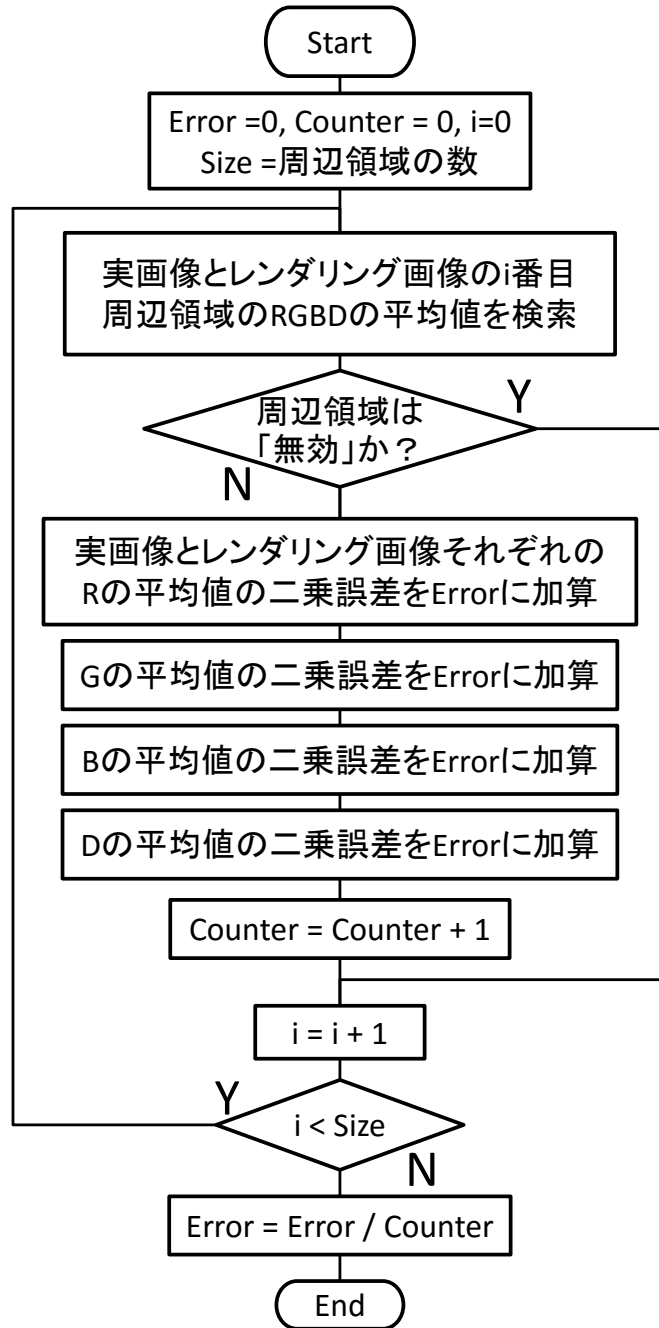


図 3.17: 画像間の相違度の計算の流れ

して選ぶキーフレーム数が多いとリローカリゼーションに成功する可能性が増えるが、処理時間が増えるため、フレームレートや提案手法の処理時間を考慮して、精度と速度の両方を維持できる数として5枚程度が候補の数として望ましいと考えられる。

3.4 提案手法の予想される利点と欠点

本提案手法は、環境再構成モデルを用いて事前に撮影していない領域でもレンダリング画像を生成することにより、2.2節で述べたように、Randomized Fernのような多くのキーフレームベースの手法が持つカメラ移動範囲の制限を緩和できると期待される。また、本提案手法は画像内の比較的広い範囲から領域内平均を求めるため、多少の環境再構成モデルの歪み・位置ずれ、実画像との色の微妙な違いがある場合でも領域内平均は大きくは変化しないことが予想され、その悪影響を緩和できることが期待できる。さらに、信頼度を用いた無効領域の処理により、本提案手法はレンダリング画像と実画像の無効領域により画像間の正確な比較が困難になるという問題にも対応できると考えられる。また、深度値標準偏差サンプリングにより、リローカリゼーションが成功しやすい複雑な環境では疎にキーフレームを取得して高速化を図り、失敗しやすい環境では密にキーフレームを取得して精度を向上させる効率の良いキーフレームDBの構築が期待できる。

一方で、想定する作業領域が広がるとキーフレームの数も増えるため、キーフレームDBの構築や類似画像検索の処理にかかる時間が増加する。また、環境再構成モデルがレンダリング画像内に写ってない領域ではリローカリゼーションを成功させることができない。他にも、レンダリング画像内の大半の領域でモデルが欠損している場合は、実画像と比較して正しい類似画像を検索することは困難であると考えられる。本提案手法は原理的には原子力発電プラント環境以外でも利用可能だが、画像内のRGBやデプスの値が殆ど変化しない環境、例えば一様な壁などでは、提案手法の類似画像検索処理では画像間の違いを認識できず、正確なカメラ姿勢は推定できない。また、現時点では、環境再構成モデルは、自動的に更新されないため、解体や保守作業が進むに従って環境が変化した場合などは、環境再構成モデルを再度作成し、キーフレームDBを更新しなければならないという問題が残る。

第 4 章 提案手法の評価

本章では、まず、本研究のリローカリゼーション手法の評価の方法について述べる。次に、評価に用いるデータセットの作成方法やリローカリゼーション手法のパラメータの設定方法について述べる。次に、実際の原子力発電プラント内でリローカリゼーションを実行する場合を対象に、本研究の提案手法と既存手法（Randomized Fern）の性能を評価した実験について述べる。

4.1 評価の概要

本研究では、実際の原子力発電プラント内で RGB-D カメラにより取得した実画像を対象に提案手法と既存手法によるリローカリゼーションを実行し、その成否と処理に要する時間で各リローカリゼーション手法の性能を評価する。

まず、リローカリゼーションの成否の判断方法について述べる。本研究では、リローカリゼーション手法により推定されたカメラ姿勢と実際のカメラ姿勢（真値）との差を計算し、その差が一定の閾値以下ならばリローカリゼーションに成功したと判断する。実際に AR による作業支援等を実行する場合は、リローカリゼーションによるカメラ姿勢の推定値をトラッキングの初期値として誤差最小化手法を実行し、得られた結果をカメラ姿勢の推定値として用いることが多い。また、誤差最小化手法としては、ICP を用いることが多い。そこで、本研究での評価では、これらを模擬し、リローカリゼーションで得たカメラ姿勢を初期値として ICP を実行した結果（補正後推定値）を対象に成否を判断する。ICP を用いた補正の詳細は 4.2.3 項で述べる。

以上に述べた方法によりリローカリゼーションの成否を判断するためには、カメラ姿勢の真値を知る必要がある。そこで本研究では、Infinitam を用いて、カメラを低速で移動させながら取得した RGB-D 画像を処理し、その際に得られるカメラ姿勢を用いて、実画像とその画像を取得した時のカメラ姿勢の組を多数作成し、リローカリゼーションの評価に用いる。以下、この実画像とカメラ姿勢の組の集合をデータセットと呼ぶ。リローカリゼーション手法の評価のためのデータセットは、リローカリゼーション手法で使用するキーフレーム DB を作成する際に使用するデータベース作成用データセット（以下、DB 用データセット）と、リローカリゼーション時に入力画像として

使用する入力用データセット（以下、入力用データセット）の2種類のデータセットが必要となる。DB用データセットは、事前準備で対象となる環境をカメラを用いて撮影した結果得られる画像群に相当し、入力用データセットは、ARシステムを使用時にカメラから得られるオンラインの画像群に相当する。本研究では、入力用データセットのカメラ姿勢の真値 $[R_g|t_g]$ と、リローカリゼーション手法で推定されたカメラ姿勢 $[R_e|t_e]$ の補正後推定値 $[R_i|t_i]$ の差を求め、予め定めた閾値以内であるかを調べることにより、リローカリゼーションの成否を判断する。

本研究では実際の原子力発電プラント内で作成したデータセットを用いて、提案手法と、既存手法として先行研究^[45]により実際の原子力発電プラント内で正確かつ高速なカメラ姿勢の推定が可能であることが示された Randomized Fern の性能を評価し、比較する。

4.2 評価の方法

本節ではリローカリゼーション手法の評価方法について述べる。

4.2.1 評価に用いる指標

本研究ではリローカリゼーション手法を評価する指標として復帰率と処理時間を採用した。復帰率は、評価対象とした全画像に対する復帰に成功した画像の割合であり、リローカリゼーション手法によりどの程度正しくカメラ姿勢を推定できるかを評価するために用いる。処理時間は、1フレームあたりのリローカリゼーションに要する時間であり、実際の作業支援を想定すると重要な指標である。

本研究では、復帰率 R を式 4.1 で定義する。

$$R = \frac{S}{A} \quad (4.1)$$

S は入力用データセットの中でリローカリゼーションが成功したフレームの数、 A は入力用データセットの全フレーム数を表す。

次に、リローカリゼーションの成否の判定に用いる2つのカメラ姿勢の差を計算する方法を説明する。カメラ姿勢は変換行列 $[R|t]$ で表される。 R は回転行列、 t は並進移動ベクトルを表す。2つのカメラ姿勢の差は移動誤差と回転誤差で表される^[49]。まず、リローカリゼーション手法により推定されたカメラ姿勢 $[R_e|t_e]$ の補正後推定値 $[R_i|t_i]$

と真値とするカメラ姿勢 $[R_g|t_g]$ の移動誤差は式 4.2 で表される。

$$d = \|t_g - t_i\| \quad (4.2)$$

式 4.2 は 2 つのカメラ位置の間のユークリッド距離を求める式である。また、補正後推定値の回転行列 R_i と真値の回転行列 R_g の差分行列 R_d は

$$R_d = R_g R_i^T \quad (4.3)$$

と計算される。この差分行列 R_d を用いて、回転誤差は式 4.4 で導出される。

$$\theta_{R_d} = \arccos\left(\frac{\text{Trace}(R_d) - 1}{2}\right) \quad (4.4)$$

ここで、 θ_{R_d} は回転行列 R_d の回転角度を表す。また、 $\text{Trace}(R_d)$ は行列 R_d の主対角成分の総和である。

本研究では、各リローカリゼーション手法により推定された 5 つのカメラ姿勢に対し順番に ICP による補正を行い、補正後推定値とカメラ姿勢の真値に対して式 4.2 および式 4.4 で求められる差を計算し、移動誤差と回転誤差の両方が閾値以下となった時点でリローカリゼーションに成功したと判定する。

次に、処理時間について説明する。本研究の評価では、入力用データセットから実画像を読み込み、実画像と全キーフレームとの相違度を計算し、最も類似した上位 5 フレームを抽出するまでに要する時間を入力 1 フレームあたりの処理時間と定義する。そして、入力用データセットの全フレームを処理するのに要した時間をフレーム数で割った値を評価に用いる処理時間とする。

4.2.2 評価用データセットの作成

本項では、評価に用いるデータセットの作成方法について述べる。データセットのカメラ姿勢の真値は、RGB-D カメラで撮影した RGB 画像とデプス画像を用いて InfiniTAM で環境再構成モデルを作成する際に推定値として取得できる撮影画像のカメラ姿勢を模擬的に利用する。4.1 節で述べたように、評価には DB 用データセットと入力用データセットの 2 種類のデータセットが必要である。本研究では、DB 用データセットを用いてリローカリゼーション手法が推定したカメラ姿勢と入力用データセットのカメラ姿勢（入力の実画像のカメラ姿勢の真値に相当）の差を計算してリローカリゼーションの成否を判定する。ここで、DB 用データセットのカメラ姿勢と入力用データセットのカメラ姿勢の差を計算するためには、これら 2 種類のデータセットのカメラ姿勢の

座標系が共通である必要がある。InfiniTAM ではカメラ姿勢は最初のフレームを座標系の原点として計算されるため、図 4.1 に示すように、撮影の開始時のカメラ姿勢が異なる場合、カメラが物理的に同じ姿勢でも、2つのデータセットのカメラ姿勢 $[R|t]$ の回転行列と並進移動ベクトルは異なる値となる。そこで、撮影する環境内に固定点を設定し、この固定点から撮影を開始することにより、各データセットの始点のカメラの姿勢が同じになるようにした。

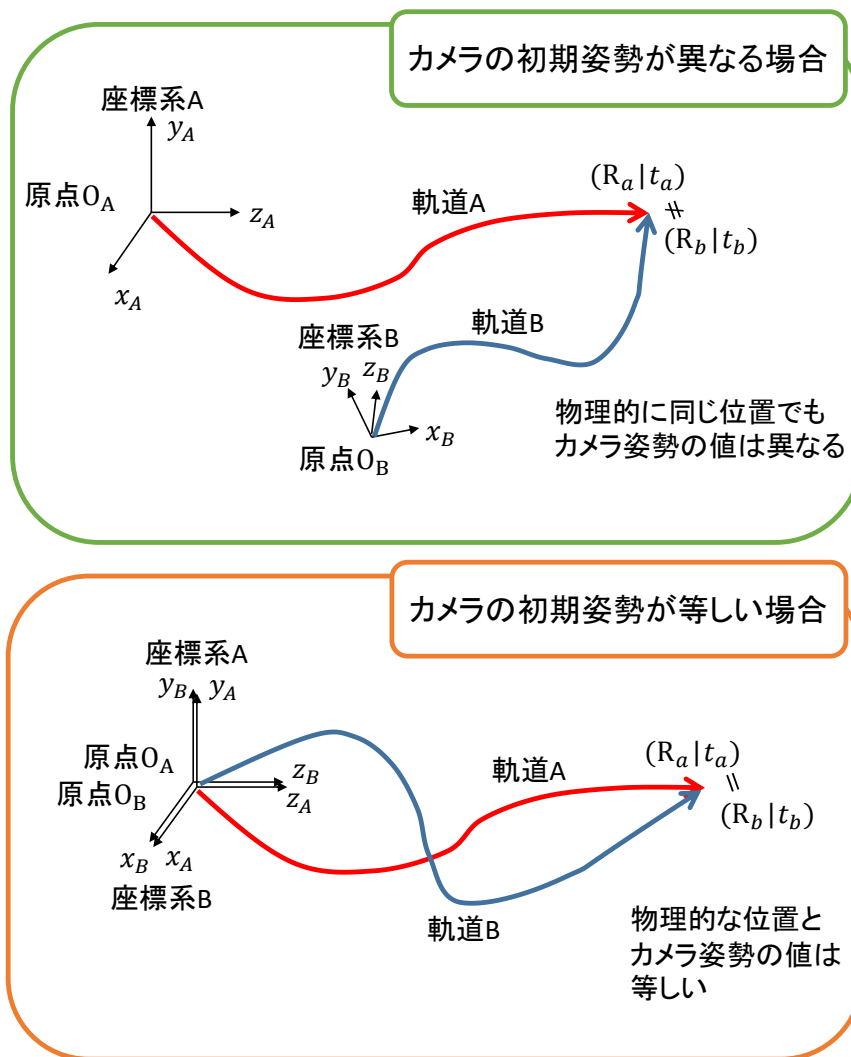


図 4.1: 初期姿勢の違いによる各データセットの座標系の違い

データセット用の画像を撮影する際には、図 4.2 に示すように、環境内に三脚を固定し、これを固定点として撮影の始点のカメラ姿勢が同じになるようにした。そして、カメラを動かして周囲の環境を撮影し任意の位置で撮影を終了した。この時の撮影時の様子を図 4.3 に示す。その後、撮影したカメラの軌道毎に InfiniTAM を用いてカメラ姿

勢を得た。1つのカメラ軌道で撮影された、連続した RGB 画像とデプス画像とカメラ姿勢を1つのデータセットとする。そして、同じ領域で同じ撮影の始点から作成したデータセットから入力用データセットと DB 用データセットを選んだ。



Copyright (C) 2017 Japan Atomic Energy Agency

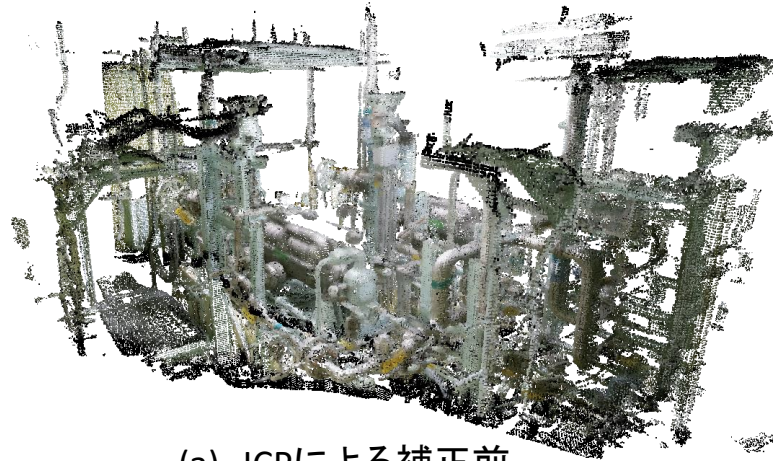
図 4.2: 環境内での三脚の設置の様子



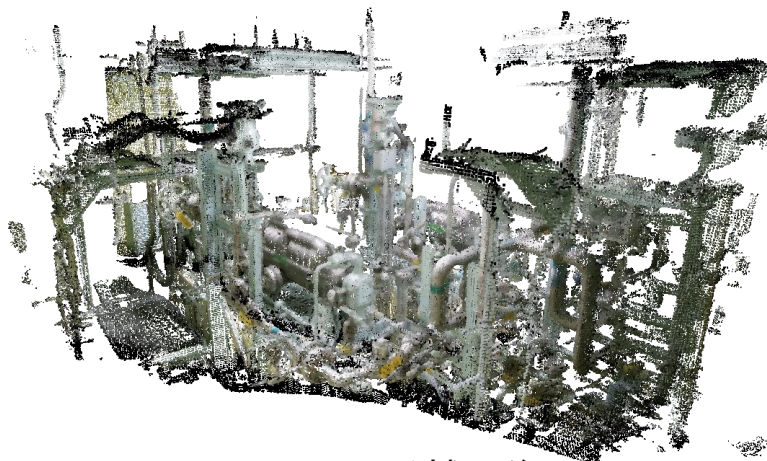
Copyright (C) 2017 Japan Atomic Energy Agency

図 4.3: データセット用画像の撮影時の様子

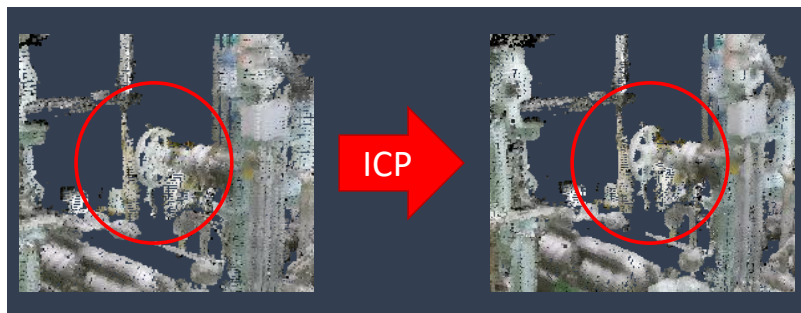
しかし、三脚のみで始点のカメラ姿勢を厳密に等しくすることは実際には困難である。そこで、図 4.4 に示すように、入力用データセットと DB 用データセットの撮影画像から InfiniTAM により得られた環境再構成モデル同士で ICP を行うことにより、二つの環境再構成モデルの位置合わせを行い、二つの環境再構成モデルの座標系を統一する。InfiniTAM で得られる環境再構成モデルとカメラ姿勢は同じ座標系であるため、二つの環境再構成モデルの座標系を統一することにより各データセットのカメラ姿勢の座標系も統一できる。



(a) ICPによる補正前



(b) ICPによる補正後



Copyright (C) 2017 Japan Atomic Energy Agency

図 4.4: ICP を用いた環境再構成モデルの座標系の統一

4.2.3 Iterative Closest Point(ICP) アルゴリズムによる推定値の補正

本項では、リローカリゼーション手法によるカメラ姿勢の推定値を ICP を用いて補正する処理について述べる。4.1 節で述べたように、ICP を用いたカメラ姿勢の推定値の補正は、リローカリゼーションの結果を初期値としたトラッキングの誤差最小化計算を模している。ICP は二つの点群の位置合わせを行うアルゴリズムであり、図 4.5 に示すように、入力用データセットの実画像に対してリローカリゼーション手法がカメラ姿勢の推定値 $[R_e|t_e]$ を返した時、入力用データセットの実デプス画像とカメラ姿勢の推定値 $[R_e|t_e]$ から得られる点群 E と、実デプス画像とその実デプス画像のカメラ姿勢の真値 $[R_g|t_g]$ から得られる点群 G を用いて ICP を行う。そして、ICP により得られた二つの点群間の変換行列 M を用いて、カメラ姿勢の推定値 $[R_e|t_e]$ を変換する。このようにして得られたカメラ姿勢が ICP により補正されたカメラ姿勢 $[R_i|t_i]$ 、すなわち、補正後推定値となる。

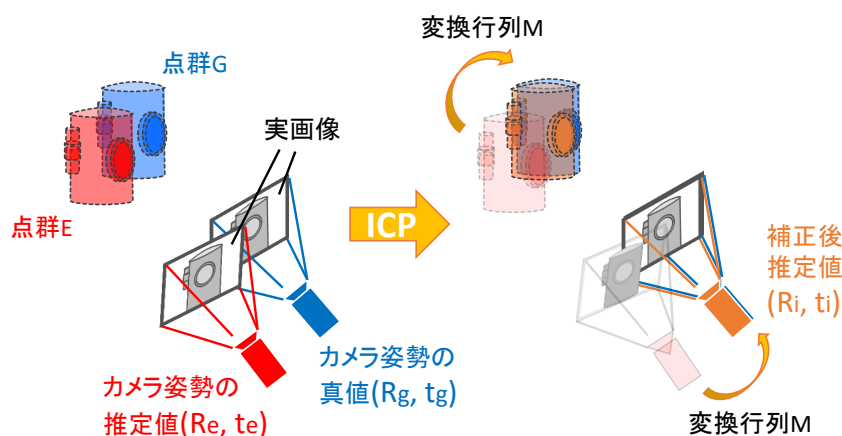


図 4.5: ICP を用いたカメラ姿勢の推定値の補正

4.2.2 項で述べたように、入力用データセットのカメラ姿勢の座標系は DB 用データセットのカメラ姿勢や環境再構成モデルの座標系と等しいため、入力用データセットのカメラ姿勢の真値 $[R_g|t_g]$ から得られる点群 G は環境再構成モデルに相当する。つまり、ICP による位置合わせは、カメラ姿勢の推定値 $[R_e|t_e]$ から得られる点群 E を環境再構成モデルに対して位置合わせすることを意味する。そしてこれはトラッキング手法の持つ環境のマップに対してのカメラ姿勢の推定を模している。

4.3 提案手法を評価するための環境

本節では本研究で使用したソフトウェアとハードウェア、作成したデータセットや各リローカリゼーション手法の具体的な実装について述べる。

4.3.1 評価に用いたソフトウェアおよびハードウェア

本研究で評価に使用したPCの仕様を表4.1に、利用したライブラリを表4.2に示す。本研究で利用したICPはPoint Cloud Library^[50]を用いて実装した^[51]。ICPのパラメータについては、実際のリローカリゼーション時のカメラの移動速度を考慮して、点群間の対応点の探索距離の閾値を0.3m、処理時間を考慮してイテレーション数を10とした。また、撮影に使用したRGB-Dカメラは図4.6に示すMicrosoft Xbox One Kinectを使用した。Kinectカメラの仕様を表4.3に、寸法・重量を表4.4に示す^[52]。表4.3に示すように、カメラからの距離を表す深度の取得範囲の最大値は4.5mであるが、深度取得範囲の境界付近では赤外線センサによる深度取得の精度が落ちることを考慮して、本研究では4mまでのデプス値を使用し、それ以上離れたデプス値は信頼性に欠けるため使用しないこととした。また、本研究で使用したKinectカメラのキャリブレーションにはiai_kinect^[53]を用いた。

表 4.1: 評価に使用したPCの仕様と開発環境

プロセッサ	Intel(R) Core(TM) i7-2600K
CPU	3.40 GHz
メモリ	16.0 GB
OS	Windows 10 Pro(64bit)
GPU	GTX-580



図 4.6: Microsoft Xbox One Kinect の外観

表 4.2: 開発に使用したソフトウェアおよびライブラリ

機能	利用ライブラリ
開発環境	Visual Studio 2013
開発言語	C++
レンダリング画像の生成	VTK 7.0
ICP	PCL 1.7.2
画像の歪補正	OpenCV 3.1.0

表 4.3: Xbox One Kinect のセンサの仕様^[52]

	解像度	FPS	深度の取得範囲	FOV
RGB カメラ	1920 × 1080	30		70 × 60
デプスセンサ	512 × 424	30	0.5~4.5m	

表 4.4: Xbox One Kinect の寸法・重量^[52]

寸法 (長さ × 幅 × 高さ)	重量
24.9 cm × 6.6 cm × 6.7 cm	約 1.4 kg

4.3.2 リローカリゼーションの成否判定の閾値の設定

次に、リローカリゼーションの成否判定に用いる閾値について述べる。本研究では、補正後推定値とカメラ姿勢の真値の間の移動誤差と回転誤差が閾値以下であればリローカリゼーションが成功したと判断する。

本研究ではこれらの誤差は AR による表示情報と現実世界の位置ずれに相当する。そのため、許容される誤差は AR によるアプリケーションに依存する。本研究で想定する AR アプリケーションは原子力発電プラントの解体作業支援であり、具体的には解体するタンクや配管の指示などの情報支援である。そこで、実際の原子力発電プラント内の配管の太さやタンクの大きさを考慮して、タンクや配管の解体箇所の指示の際の表示情報の位置ずれの許容誤差を想定して表 4.5 に示す値に閾値を設定する。

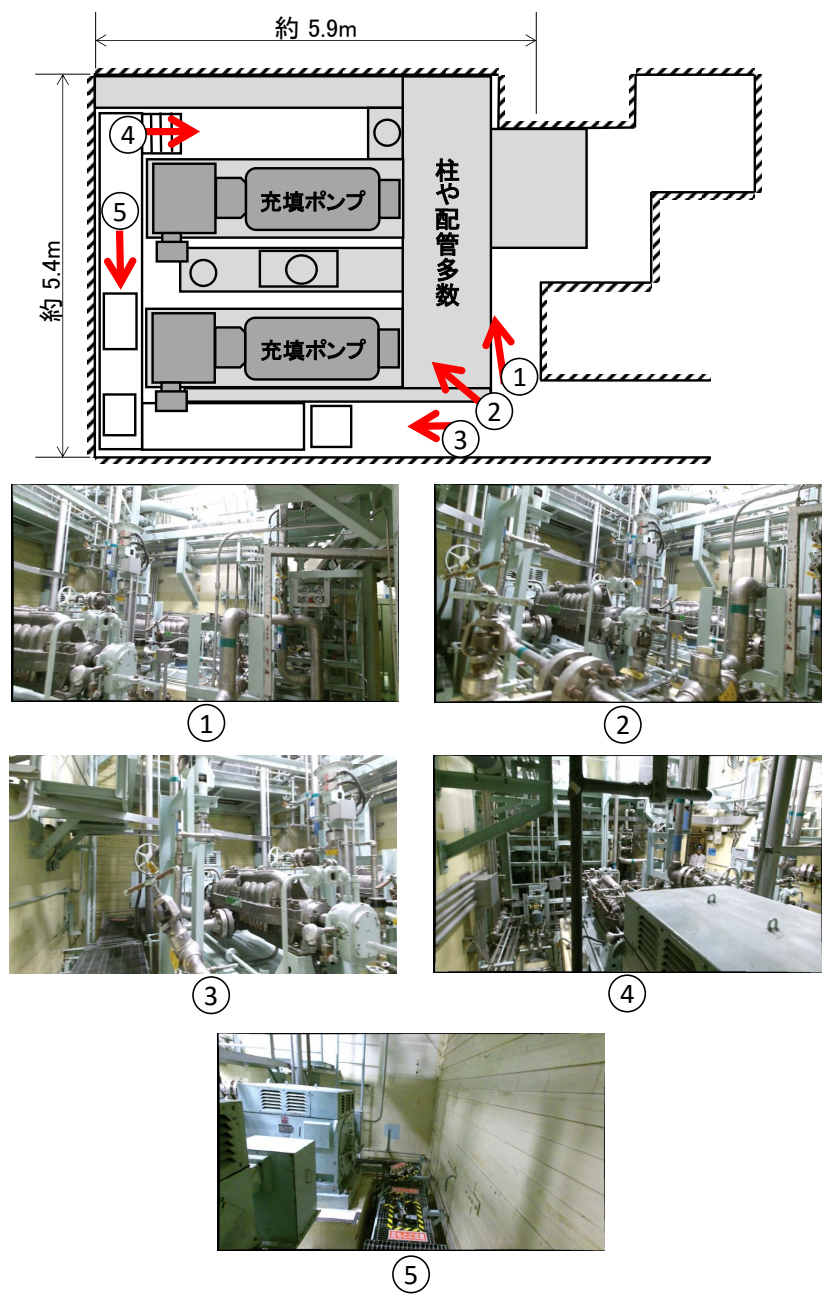
表 4.5: リローカリゼーションの成否判定の閾値

	移動誤差 (m)	回転誤差 (度)
成否判定の閾値	0.2	20

4.3.3 作成した評価用データセット

次に本研究で実際に作成したデータセットについて述べる。本研究では原子炉廃止措置研究開発センターの充填ポンプ室内で撮影を行った。撮影の際は InfiniTAM で正確なカメラ姿勢の真値を得るために Kinect カメラをゆっくりと動かして撮影を行った。図 4.7 に示すように、充填ポンプ室の内部は充填ポンプや多数の配管が複雑に配置され、人が移動出来るスペースも限られている。本研究では充填ポンプ室内の三か所に三脚を設置してデータセット用の撮影を行った。この三か所の撮影領域を図 4.8 から図 4.10 に示す。また、それぞれの撮影領域での DB 用データセットと入力用データセットの撮影時のカメラ軌道を図 4.11 から図 4.13 に示す。このように、これらの撮影領域でそれぞれ DB 用データセットと入力用データセットを取得した。取得した各 DB 用データセットの撮影領域とフレーム数を表 4.6 に、各入力用データセットの撮影領域とフレーム数を表 4.7 に示す。なお、各データセットの 1 フレームは RGB 画像とデプス画像の組を表す。

本研究の評価では、提案手法が事前に撮影していない領域でもリローカリゼーションが可能なことを示すために、入力用データセットのカメラ軌道の一部を DB 用データセットのカメラ軌道が通過していない領域にも設定して撮影した。図 4.8 に示す Area1 は、撮影された領域は最も狭いが配管や柱などの複雑な形状から壁などの平坦な環境まで様々な環境を含む。図 4.9 に示す Area2 では充填ポンプを中心に撮影しており、カメラを大きく左に向けるなどのカメラの回転を多く含むデータセットとした。図 4.10 に示す Area3 では、図 4.13 に示すように、入力用データセットのカメラ軌道の大部分が DB 用データセットのカメラ軌道とは異なる領域を通過しており、Randomized Fern のような事前に通過した場所のみに復帰できる手法では復帰率が低くなると予想される。作成した DB 用データセットを用いて InfiniTAM を利用して作成した環境再構成モデルを図 4.14 から図 4.16 に示す。以降ではこれらの環境再構成モデルをそれぞれ、モデル 1、モデル 2、モデル 3 と呼ぶ。なお、InfiniTAM を利用して作成される環境再構成モデルのスケールは実世界のスケールとは異なる。そこで、大きさが分かっている物体を撮影して InfiniTAM を利用して環境再構成モデルを作成し、該当する物体のスケールを実世界のスケールに合わせるための比率を求め、その比率を用いて他のモデルのスケールを調整した。そのため、InfiniTAM により得られた環境再構成モデルの大きさを 8 倍にしている。



Copyright (C) 2017 Japan Atomic Energy Agency

図 4.7: 充填ポンプ室の様子

表 4.6: DB 用データセットの撮影領域とフレーム数

	撮影領域	フレーム数
DB 用データセット 1	Area1	1,255
DB 用データセット 2	Area2	1,087
DB 用データセット 3	Area3	1,077

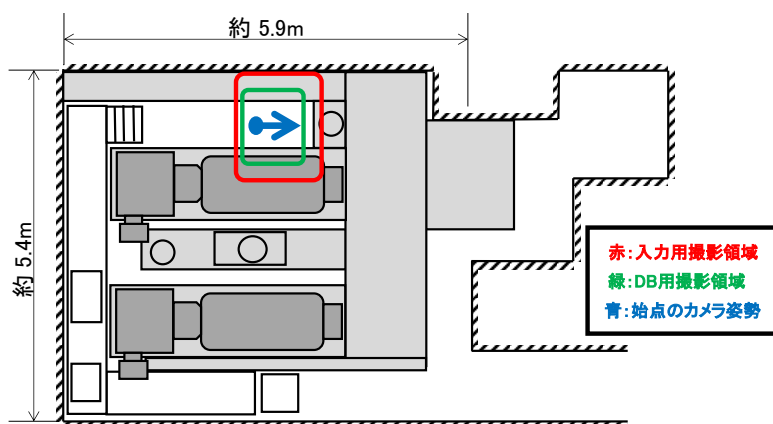


図 4.8: Area1 の撮影領域

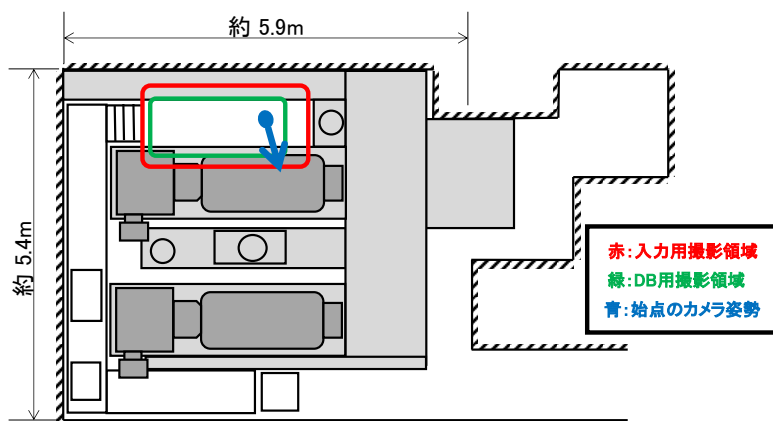


図 4.9: Area2 の撮影領域

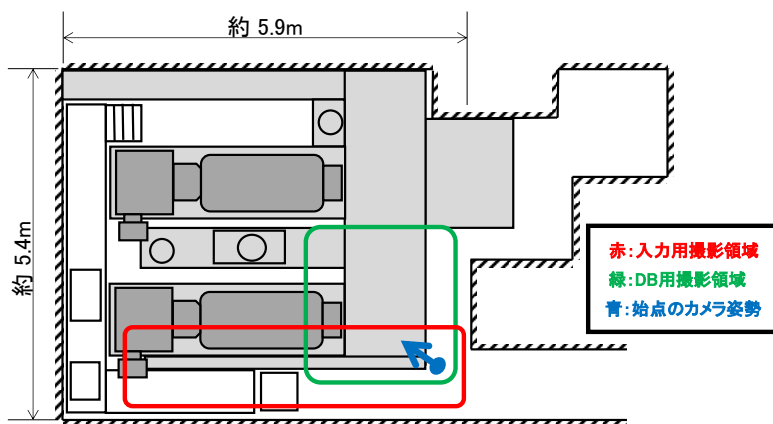
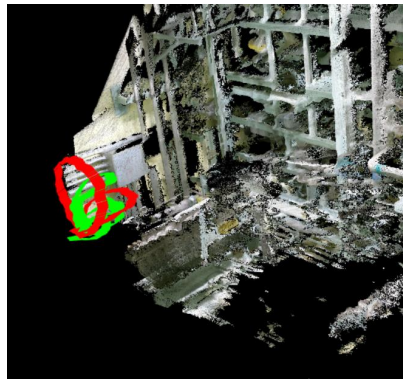


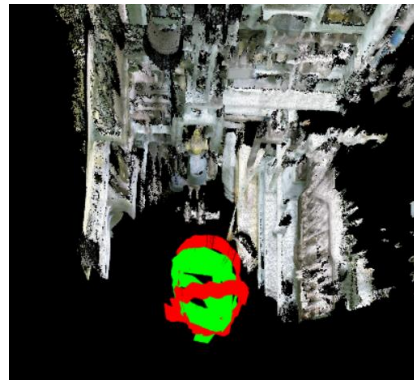
図 4.10: Area3 の撮影領域



(a) front



(b) side



(c) top

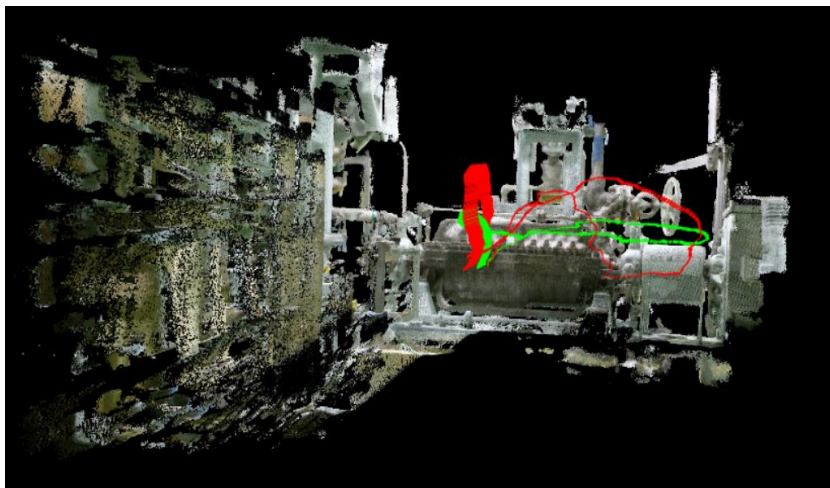
緑: DB用データセットのカメラ軌道
赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.11: Area1 の撮影時のカメラ軌道

表 4.7: 入力用データセットの撮影領域とフレーム数

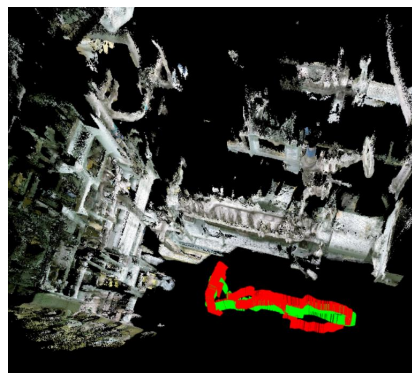
	撮影領域	フレーム数
入力用データセット 1	Area1	1,122
入力用データセット 2	Area2	1,382
入力用データセット 3	Area3	792



(a) front



(b) side

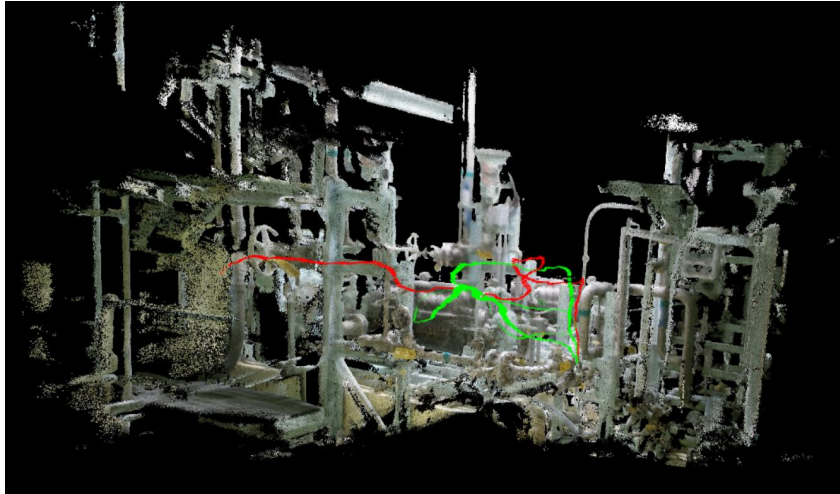


(c) top

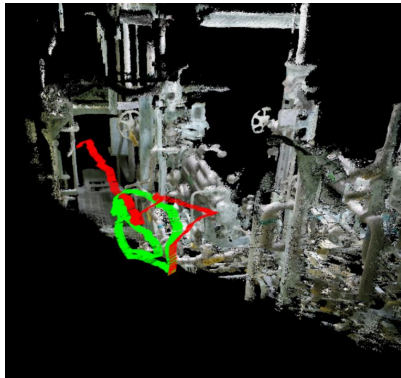
緑: DB用データセットのカメラ軌道
赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

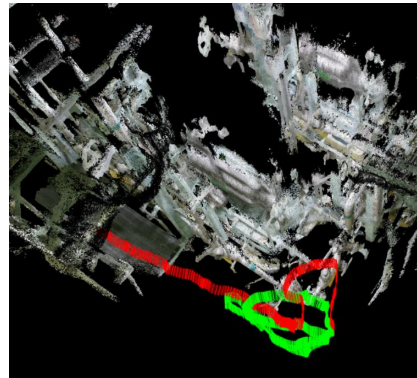
図 4.12: Area2 の撮影時のカメラ軌道



(a) front



(b) side

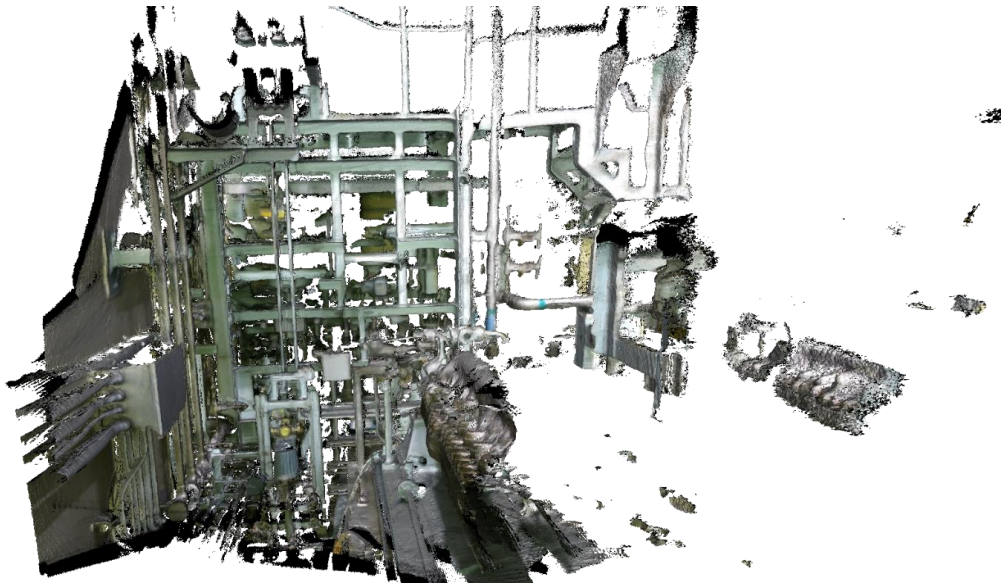


(c) top

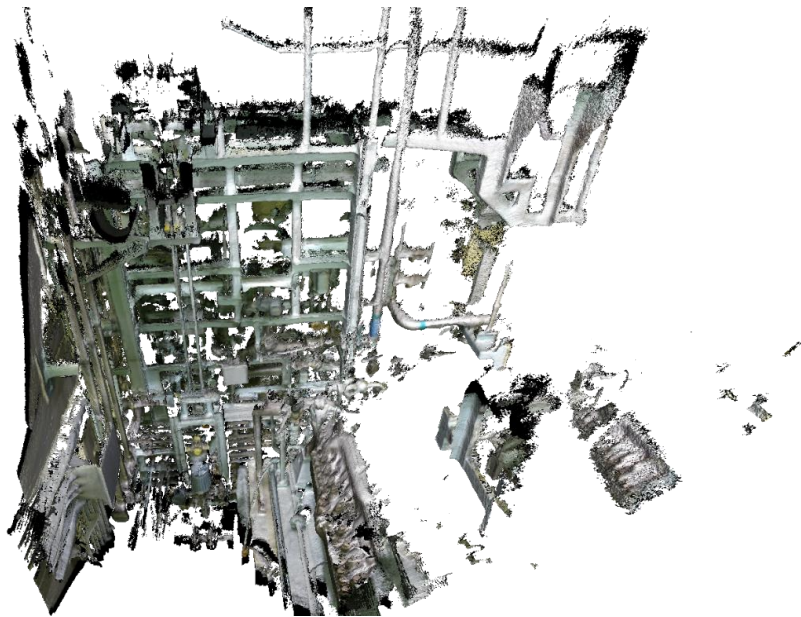
緑: DB用データセットのカメラ軌道
赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.13: Area3 の撮影時のカメラ軌道



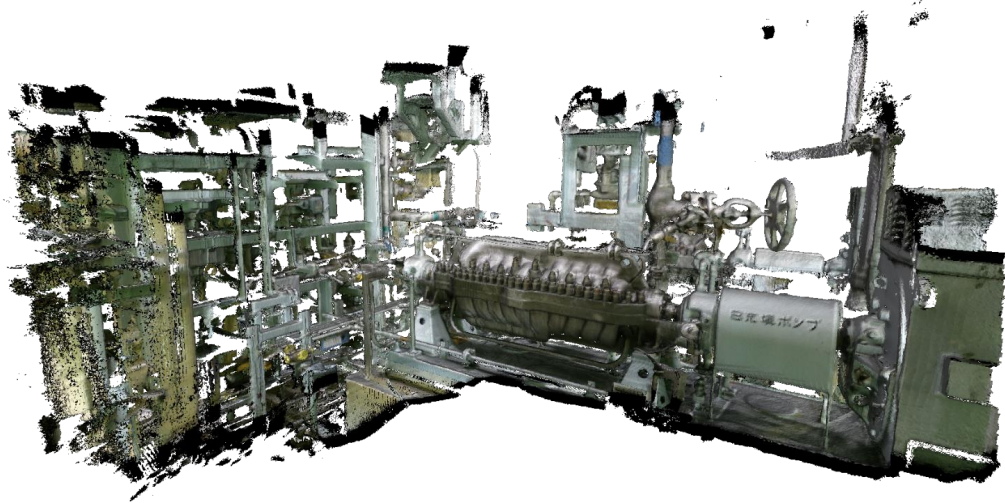
(a) front



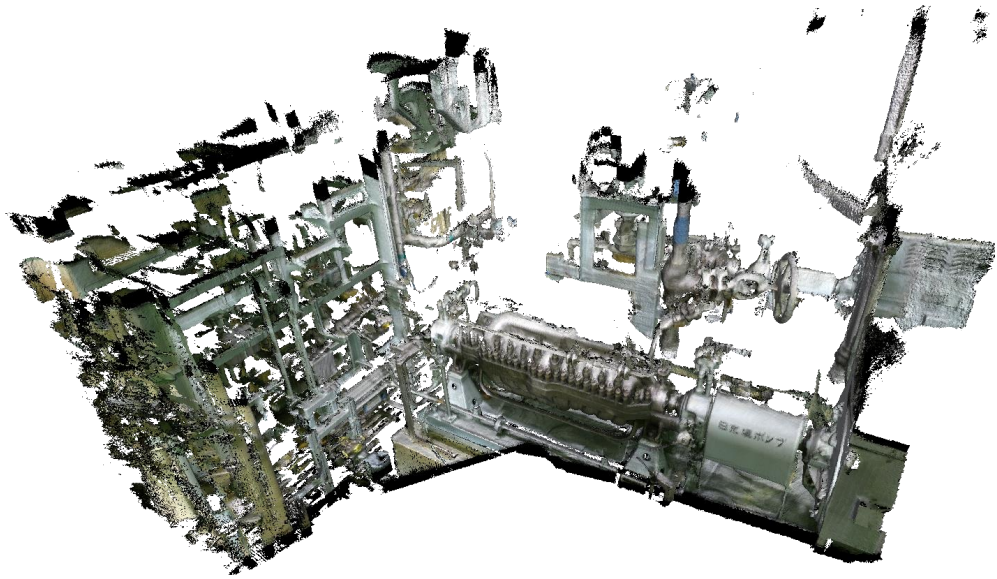
(b) top

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.14: モデル 1



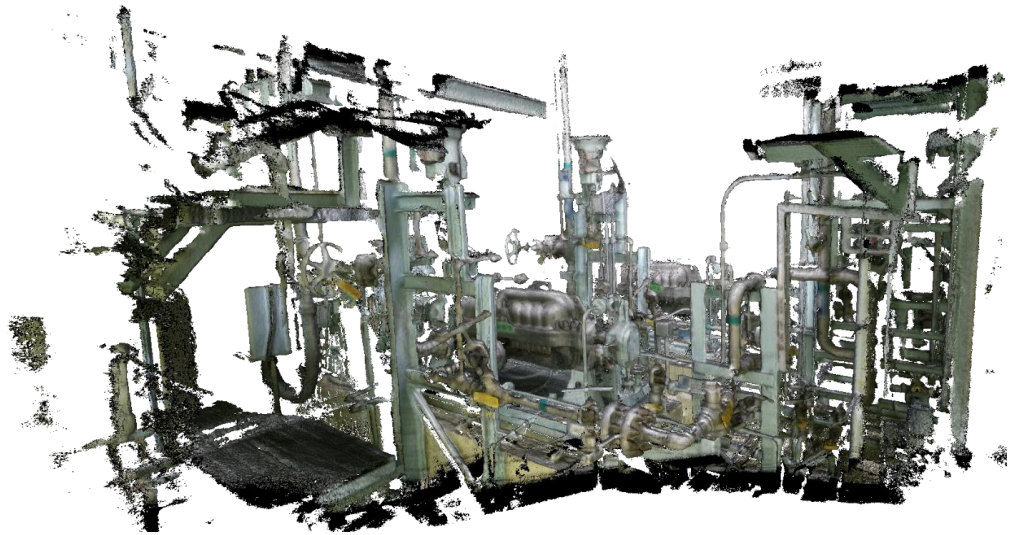
(a) front



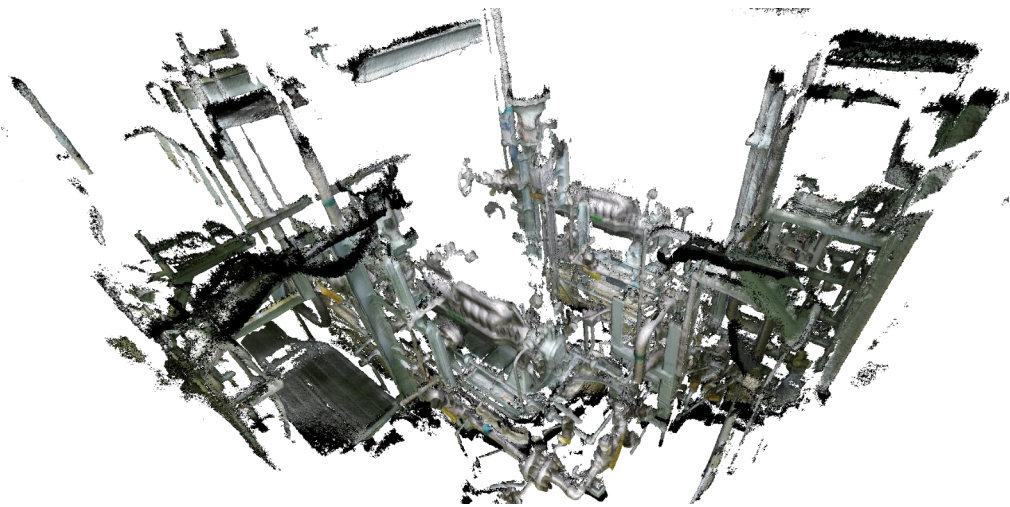
(b) top

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.15: モデル 2



(a) front



(b) top

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.16: モデル 3

4.3.4 レンダリング画像の生成

本項では、作成した環境再構成モデルを用いたレンダリング画像の生成について述べる。まず、カメラ姿勢のサンプリングに関するパラメータについて述べる。本研究の評価では、図 4.17 に示すカメラ姿勢の 6 パラメータ x 、 y 、 z 、 yaw 、 $pitch$ 、 $roll$ の範囲を、図 4.18 に示すように入力用データセットのカメラ軌道全体をカバーできるように設定してカメラ姿勢をサンプリングする。

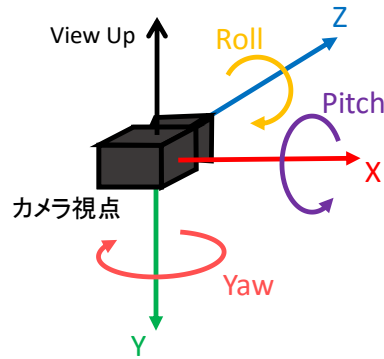


図 4.17: カメラ姿勢の 6 パラメータ

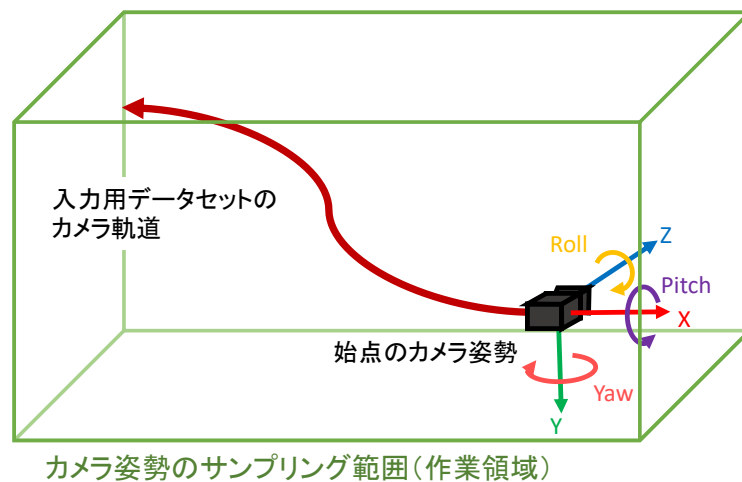


図 4.18: カメラ姿勢のサンプリングの範囲

実際の評価の際に設定したパラメータの範囲を表 4.8 に示す。本研究では図 4.17 の矢印の向きがカメラの各パラメータの正の向きを表し、撮影の始点のカメラ姿勢が各撮影領域の座標系の原点となる。なお、Area1 は軸に平行に領域の境界を設定し、Area2、Area3 でのサンプリングでは入力用データセットのカメラ軌道をサンプリングの範囲内に含めるために、Area2 では $z=-0.2x+b$ 、Area3 では $z=-0.4x+b$ として z の値を設定

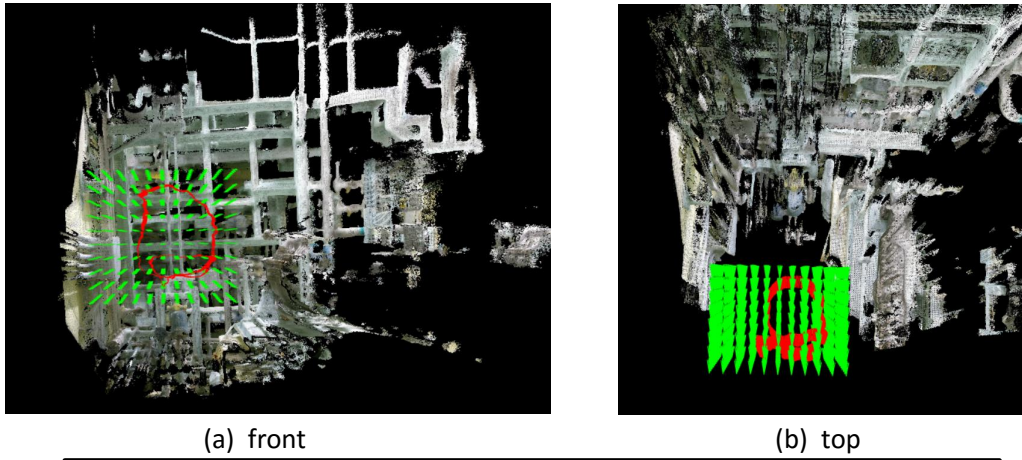
した。本研究では、3.2.3 項で述べたように、等間隔サンプリングと深度地標準偏差サンプリングの2種類のサンプリング方式を実装して検討するが、以下で各サンプリング手法のパラメータの設定について述べる。

表 4.8: 各撮影領域のサンプリング時のカメラ姿勢のパラメータの範囲

		x(m)	y(m)	z(m)	b(m)	roll(度)	pitch(度)	yaw(度)
Area1	max	0.5	0.1	0.6	N/A	0	0	10
	min	-0.5	-0.8	-0.1	N/A	0	0	-10
Area2	max	1.2	0	0.32	0.3	0	0	10
	min	-0.1	-0.6	-0.44	-0.2	0	-10	-90
Area3	max	0.1	-0.1	1.68	0.6	0	0	5
	min	-2.7	-0.6	-0.34	-0.3	0	0	-10

等間隔サンプリングの場合、4.3.2 項で述べたりローカリゼーションの成否判定の閾値を考慮して、サンプリングを行う範囲内のどこにカメラがある場合でもりローカリゼーションが成功となるカメラ姿勢がキーフレーム DB 内に含まれるようにするために、ステップの大きさは x、y、z の場合は 0.1m、roll、pitch、yaw の場合は 5 度とした。ただし、Area2 の場合のみ、pitch、yaw のステップを 10 度とした。各撮影領域で等間隔サンプリングによりサンプリングしたカメラ姿勢と入力用データセットのカメラ軌道（InfiniTAM を用いて取得した真値に相当）を図 4.19 から図 4.21 に示す。入力用データセットのカメラ姿勢を内包するようにカメラ姿勢がサンプリングされていることが分かる。

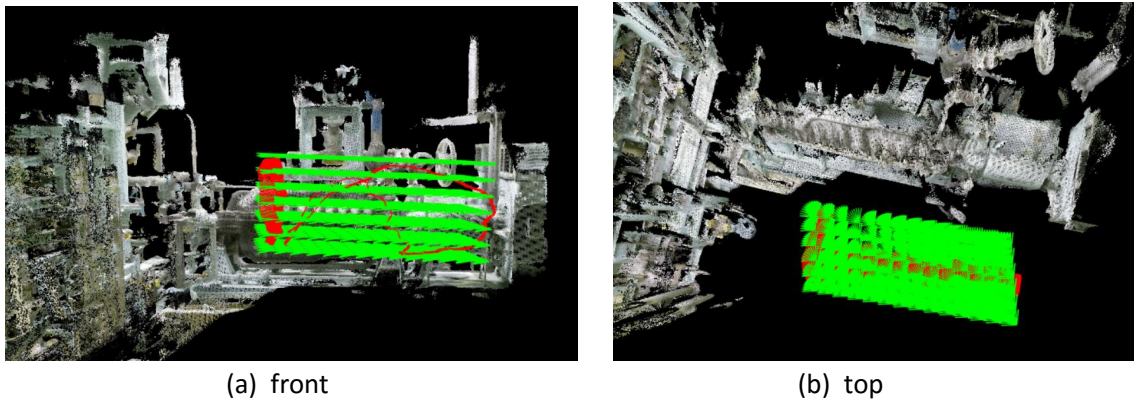
次に、深度値標準偏差サンプリングの際は 3.2.3 項で述べた式 3.2 によりステップの大きさを決定する。事前検討の結果から、本研究で使用するデータセットでは式 3.2 のレンダリングデプス画像のデプス値の標準偏差 D が 200 から 1000 程度の値をとること分かったため、全体でのキーフレーム総数を等間隔サンプリングの場合と同等の範囲に含めるために、式中の基準となるステップ C は等間隔サンプリング時のステップの大きさと等しくした上で、係数 a を $1/200$ 、 $1/400$ 、 $1/600$ 、 $1/800$ の 4 パターンとして、深度値標準偏差サンプリングを実行した。一例として、Area3 で深度値標準偏差サンプリングにより得られたカメラ姿勢と入力用データセットのカメラ軌道を図 4.22 に示す。図 4.22 より、深度値標準偏差サンプリングの場合、 a の値が減少するにつれてステップ間隔が小さくなり、サンプリングされるカメラ姿勢の数が増加することが分かる。Area3 の場合は、左に進むにつれてカメラ姿勢の密度が上がっていく。これは、



緑: サンプルングされたカメラ姿勢、赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

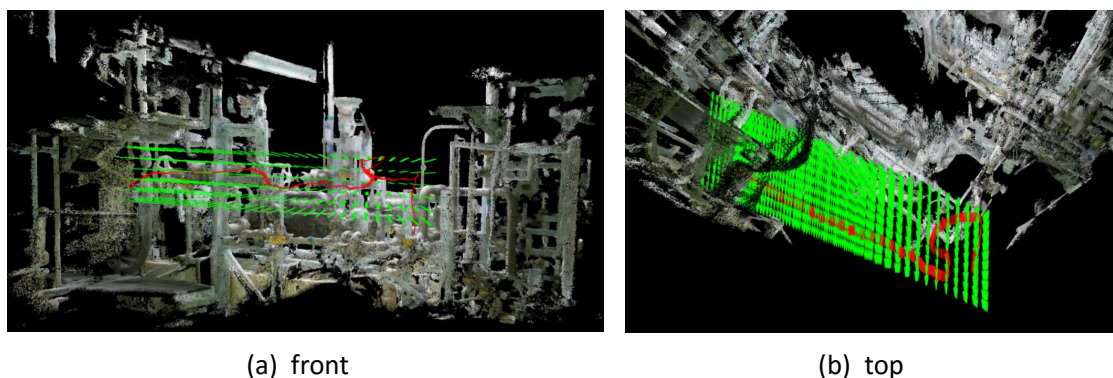
図 4.19: Area1 で等間隔サンプルングされたカメラ姿勢と入力用データセットのカメラ軌道



緑: サンプルングされたカメラ姿勢、赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.20: Area2 で等間隔サンプルングされたカメラ姿勢と入力用データセットのカメラ軌道



緑: サンプルされたカメラ姿勢、赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

図 4.21: Area3 で等間隔サンプルされたカメラ姿勢と入力用データセットのカメラ軌道

Area3 の左の奥にはタンクなどの比較的平面の多い物体が配置されていたことと、奥の方はモデルの欠損領域となっておりカメラから遠い部分が欠損していたためデプス値のばらつきが小さいことが原因と考えられる。

各サンプリング手法でカメラ姿勢をサンプリングし、レンダリングにより生成した画像のフレーム数を表 4.9 に示す。また、等間隔サンプリング時に全フレームのレンダリングにかかる時間を表 4.10 に示す。1 フレーム当たり約 0.6 秒程度の時間がかかった。

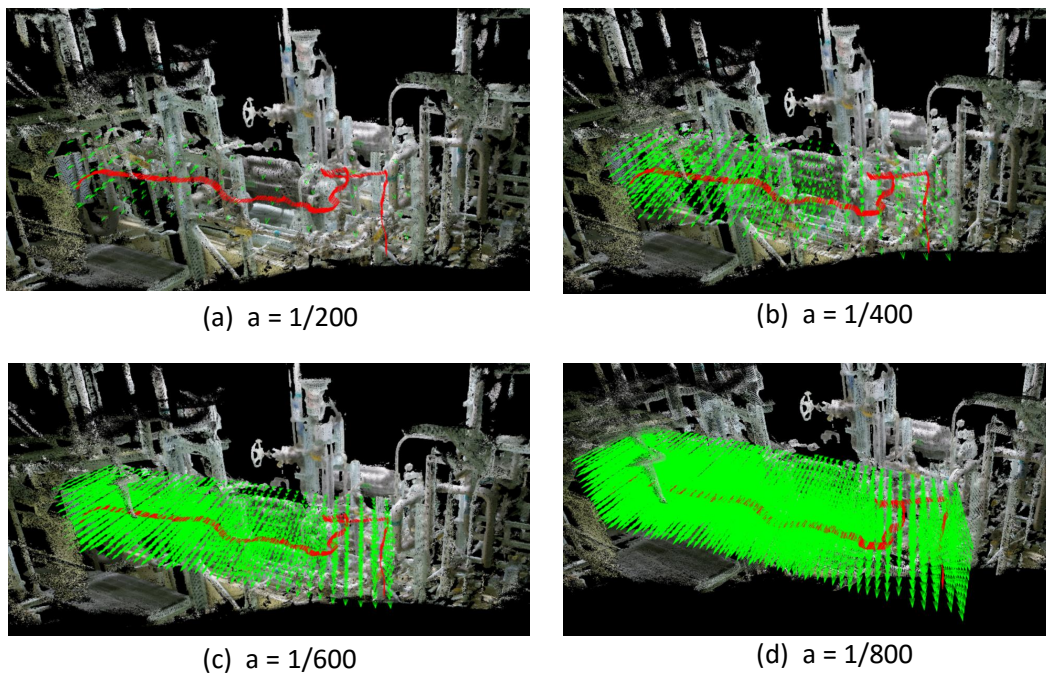
表 4.9: 各サンプリング手法の場合のレンダリング画像数

	等間隔	深度 : 1/200	深度 : 1/400	深度 : 1/600	深度 : 1/800
モデル 1	4,400	58	785	3,713	10,558
モデル 2	12,936	187	4,562	42,833	N/A
モデル 3	6,960	292	3,950	18,297	57,191

表 4.10: 等間隔サンプリング時に全フレームのレンダリングにかかる時間

使用したモデル	フレーム数 (等間隔)	全フレームの生成にかかる時間 (m)
モデル 1	4,400	45
モデル 2	12,936	117
モデル 3	6,960	75

本研究ではレンダリング画像の生成の際に The Visualization Toolkit(VTK)^[54] を用



緑: サンプルされたカメラ姿勢、赤: 入力用データセットのカメラ軌道

Copyright (C) 2017 Japan Atomic Energy Agency

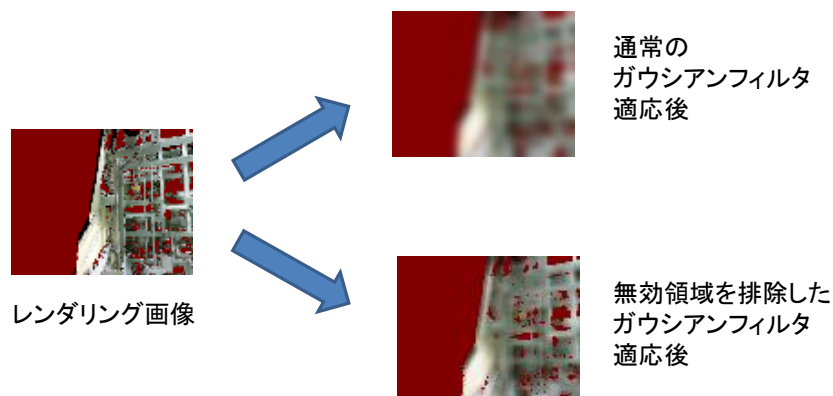
図 4.22: Area3 において深度値標準偏差サンプリングにより得られたカメラ姿勢と入力用データセットのカメラ軌道

いており、その設定では、レンダリングデプス画像を生成する際の各ピクセルの環境再構成モデルまでの距離 Z を表す式 3.1 において、 far が 1000m で $near$ が 0.01m である。4.3.3 項と本項で作成した全てのデータセットの RGB 画像とデプス画像およびレンダリング画像は容量の削減と高速化のために解像度を 100×80 とした。

4.3.5 本研究で使用する Randomized Fern の実装

本項では既存手法として評価する Randomized Fern の本研究での実装について述べる。本研究で Randomized Fern を使用する際には、画像からランダムに選ばれるピクセル数を表す $fern$ 数を、一般的に使用されている 500 とした。一方、本来 Randomized Fern ではキーフレームを保存する際に、BlockHD により既に保存しているキーフレームと類似した画像を省く処理を行うことにより類似画像検索の高速化を図っている。しかし、本研究で使用する DB 用データセットのフレーム数があまり大きくなく、また、Randomized Fern が達成できる最大の復帰率と本研究の提案手法の復帰率を比較するために、全ての画像を省かずにキーフレームに含めることとした。

また、オリジナルの Randomized Fern は画像に無効領域が存在する場合を想定していないため、提案手法と同じ条件で比較するために無効領域を排除する処理を追加した。具体的には、本研究では画像を比較する際、選ばれたピクセルが比較対象の画像のどちらか、もしくは両方で無効領域の場合、そのピクセルは BlockHD の計算の際に排除する。また、Randomized Fern では画像の符号化の前に画像に対し平滑化するためのガウシアンフィルタを適用するが、この時にレンダリング画像の無効領域も含めてガウシアンフィルタを適用すると結果が実画像と大きく異なったものになる。そこで、ガウシアンフィルタのカーネル内に無効領域のピクセルが存在する場合、そのピクセルを排除してフィルタ処理を行うように変更した。例として、図 4.23 に無効領域を持つレンダリング画像に対し、OpenCV^[55] で実装されている一般的なガウシアンフィルタを適用した場合と本研究で用いる無効領域を考慮したガウシアンフィルタを適用した場合の結果を示す。図 4.23 が示すように、本研究で用いるガウシアンフィルタの場合は、無効領域は平滑化されず有効領域に影響を与えない。なお、本研究ではガウシアンフィルタの σ は 2.5 とした。



Copyright (C) 2017 Japan Atomic Energy Agency

図 4.23: 一般的なガウシアンフィルタと本研究で用いるガウシアンフィルタ

4.3.6 提案手法のパラメータの決定

本項では本研究の提案手法で用いるパラメータについて述べる。3.3.2 項で述べたように、本提案手法の類似画像検索処理では信頼度が閾値以上の周辺領域のみを画像間の相違度の計算に用いる。そのため、信頼度の閾値を事前に決定する必要がある。本提案手法では、信頼度は、レンダリング画像のカラー信頼度 CR_{ren} 、レンダリング画像のデプス信頼度 DR_{ren} 及び実画像のデプス信頼度 DR_{real} の3つを指す。この閾値が小さい時は周辺領域内に無効領域が多く含まれている場合でも残りの有効ピクセルを利用する。一方、閾値が大きい時は周辺領域に少しでも無効領域が存在する場合、相違度計算に用いられない。それぞれの閾値 τC_{ren} 、 τD_{ren} 、 τD_{real} の適切な設定は環境依存で不明であるため、0.0001、0.3、0.5、0.7、1.0 の5段階で変化させ、提案手法の復帰率が最も高くなる閾値を決定する。ここで、閾値が0.0001の場合は、本研究の評価の際の周辺領域の大きさでは、周辺領域の全てピクセルが無効領域の場合のみ排除される。一方、閾値が1.0の場合、周辺領域に無効領域が1ピクセルでも存在すれば、その周辺領域は相違度の計算に用いない。本研究では3つの信頼度の閾値の内の2つを固定して残りの1つを変化させたときの復帰率の傾向から、最も復帰率が高くなる閾値を決定する。この信頼度の閾値の決定の際には、最も狭いが様々なプラント内の環境を含む Area1 の入力用データセット1と、モデル1を用いて等間隔サンプリングにより生成されたレンダリング画像を使用した。また、中心ピクセル数は500、周辺領域のサイズは、1では中心ピクセルのみとなり、大きすぎると処理時間が問題となるため5とする。周辺領域のサイズは周辺領域の辺の大きさを表し、この場合は周辺領域に含まれるピクセル数は25となる。

3種類の閾値の全てを5段階で変化させる場合、その組み合わせは全部で125通りとなる。そこで、125通りの中で復帰率の高い順の上位5つの閾値の組み合わせを調べ、それらの閾値の組み合わせの時の閾値 $\tau_{C_{ren}}$ 、 $\tau_{D_{ren}}$ を固定して、残りの $\tau_{D_{real}}$ を変化させた時の復帰率を見る。この時の提案手法の復帰率を図4.24に示す。この時、復帰率上位5つの組み合わせの内の2つは $\tau_{C_{ren}}$ と $\tau_{D_{ren}}$ が同じ値であったため、図中には4つの組み合わせしか示していない。図4.24から $\tau_{D_{real}}$ は復帰率に大きく影響することが分かる。 $\tau_{D_{real}}$ を高くすると復帰率が低下することから、実デプス画像では周辺領域内に無効領域が含まれる場合でも、周辺領域内の残りの有効ピクセルを用いた方が類似画像検索において効果的であると考えられる。原子力発電プラント内のデータセットの場合、実デプス画像の無効領域の多くは図3.14に示したように物体と物体の境界であった。物体の境界付近のデプス値は大きく変化する場合が多く、領域内平均の値が境界付近の物体のデプス値と異なる可能性が高いため、画像の区別に役立つ特徴的な情報であると期待できる。4つのグラフのいずれも $\tau_{D_{real}}$ が0.5の時、復帰率が最大であった。

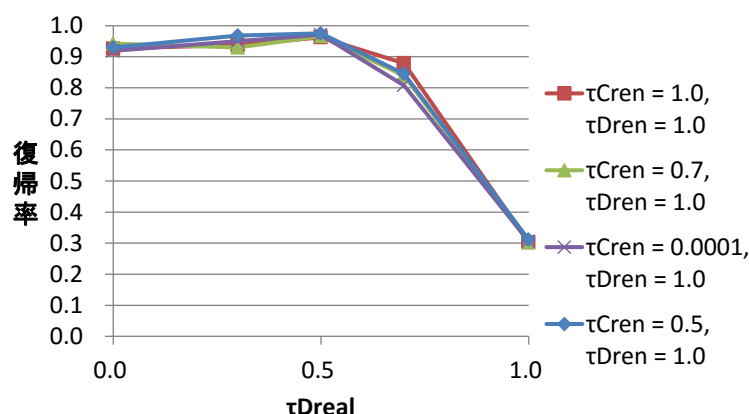


図 4.24: $\tau_{D_{real}}$ を変化させた場合の提案手法の復帰率の変化

次に同様に復帰率上位5つの組み合わせの時の $\tau_{C_{ren}}$ 、 $\tau_{D_{real}}$ を固定して、 $\tau_{D_{ren}}$ を変化させた時の復帰率を見る。結果を図4.25に示す。図4.25より、 $\tau_{D_{ren}}$ を高くすると復帰率も高くなる傾向にあることが分かる。この結果から、レンダリングデプス画像の場合は周辺領域内に無効領域が少しでも存在する場合は、その周辺領域を用いない方が復帰率が高くなることが分かる。レンダリング画像の場合、周辺領域内に無効領域と有効な領域が共に存在するのは環境再構成モデルと背景の境界である。多くの場合、背景となるのは撮影を行っていない領域か遠すぎるなどでRGB-Dカメラで計測できなかった領域、もしくはオクルージョンなどによるモデルの欠損領域である。そ

して、図 2.12 に示したように、モデルの欠損領域付近は形状が正しく取得されていない場合が多く、そのデプス値はノイズとして類似画像検索の妨げになると考えられる。いずれのグラフも τD_{ren} が 1.0 の時が復帰率が最も高くなった。

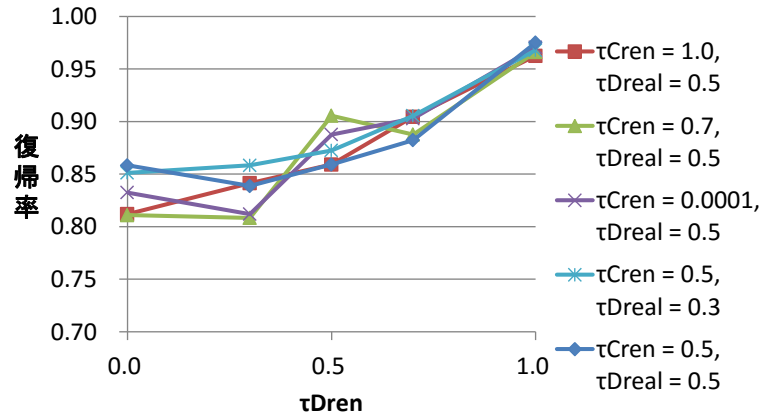


図 4.25: τD_{ren} を変化させた場合の提案手法の復帰率の変化

最後に復帰率上位 5 つの組み合わせの時の τD_{real} 、 τD_{ren} を固定して τC_{ren} を変化させた時の復帰率を見る。結果を図 4.26 に示す。復帰率上位 5 つの組み合わせの内、4 つは τD_{real} と τD_{ren} が同じ値であったため、図 4.26 には 2 つのグラフしか示していない。図 4.26 より、全体的に復帰率が高く、 τC_{ren} を変化させても復帰率はほとんど変化しないことが分かる。つまり、環境再構成モデルと背景の境界領域の RGB 情報を使う場合と使わない場合ではあまり差がない。この結果より、デプス情報に比べると RGB 画像の境界領域の情報の復帰率に対する影響は小さいと考えられる。2 つのグラフで復帰率が最大になるのは τC_{ren} が 0.5 の時であった。

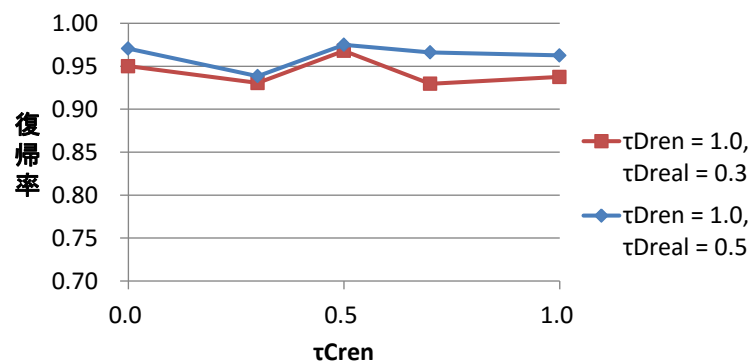


図 4.26: τC_{ren} を変化させた場合の提案手法の復帰率の変化

以上の結果から、復帰率が最も高くなる信頼度の閾値 τC_{ren} 、 τD_{ren} 、 τD_{real} として、本研究では、表 4.11 に示す値を用いることとする。最後に τD_{real} を 0.5 に固定した時の

$\tau_{C_{ren}}$ と $\tau_{D_{ren}}$ を変化させた時の復帰率を図 4.27 に、 $\tau_{D_{ren}}$ を 1.0 に固定した時の $\tau_{C_{ren}}$ と $\tau_{D_{real}}$ を変化させた時の復帰率を図 4.28 に、 $\tau_{C_{ren}}$ を 0.5 に固定した時の $\tau_{D_{real}}$ と $\tau_{D_{ren}}$ を変化させた時の復帰率を図 4.29 に示す。これらの図からも実画像のデプス信頼度 DR_{real} の閾値 $\tau_{D_{real}}$ の復帰率に対する影響は大きく、レンダリング画像のカラー信頼度 CR_{ren} の閾値 $\tau_{C_{ren}}$ の影響は小さいことが分かる。

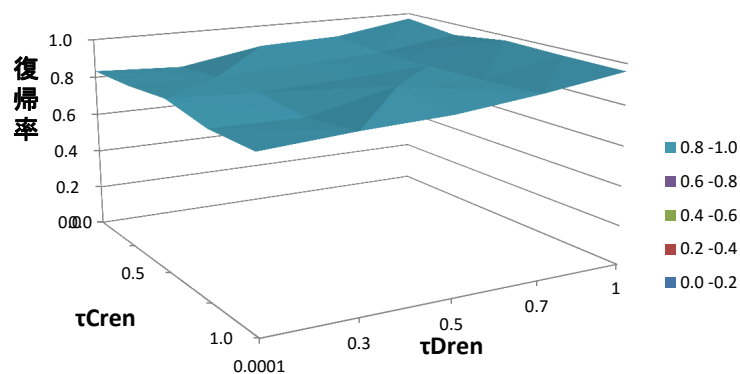


図 4.27: $\tau_{C_{ren}}$ と $\tau_{D_{ren}}$ を変化させた場合の提案手法の復帰率の変化

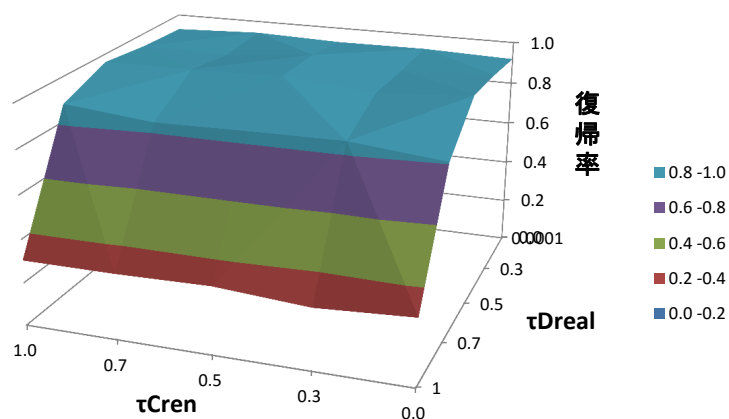


図 4.28: $\tau_{C_{ren}}$ と $\tau_{D_{real}}$ を変化させた場合の提案手法の復帰率の変化

表 4.11: 本研究の信頼度の閾値の設定

$\tau_{C_{ren}}$	$\tau_{D_{ren}}$	$\tau_{D_{real}}$
0.5	1.0	0.5

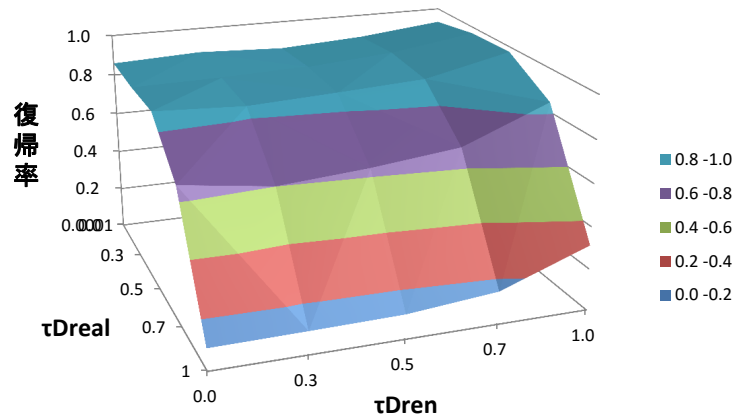


図 4.29: τD_{real} と τD_{ren} を変化させた場合の提案手法の復帰率の変化

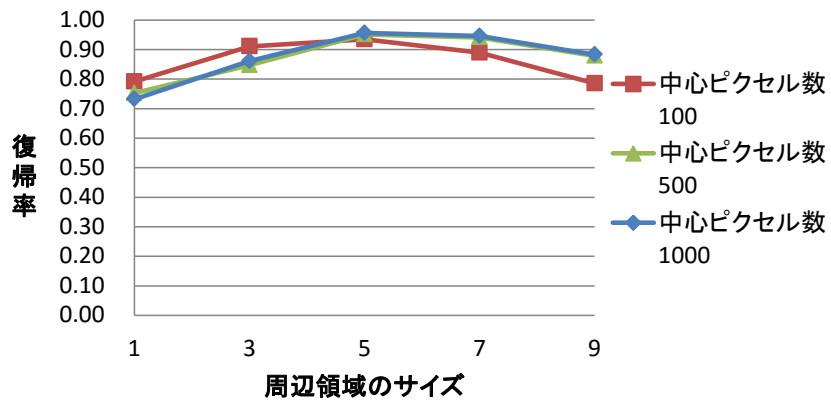
4.4 評価結果と考察

本節では各リローカリゼーション手法の評価結果と考察を述べる。まず、提案手法の中心ピクセル数と周辺領域のサイズなどのパラメータを変化させたときの評価結果と考察を述べる。次に、Randomized Fern の評価結果を示し、提案手法と比較する。最後に、提案手法の2種類のサンプリング方式による結果を比較し考察を述べる。以下では、提案手法や Randomized Fern で処理の対象とするピクセルを選ぶ際にランダムに選ぶ影響により、結果がばらつくことがあることを考慮して、各手法を3回実行し、その平均を結果として示す。

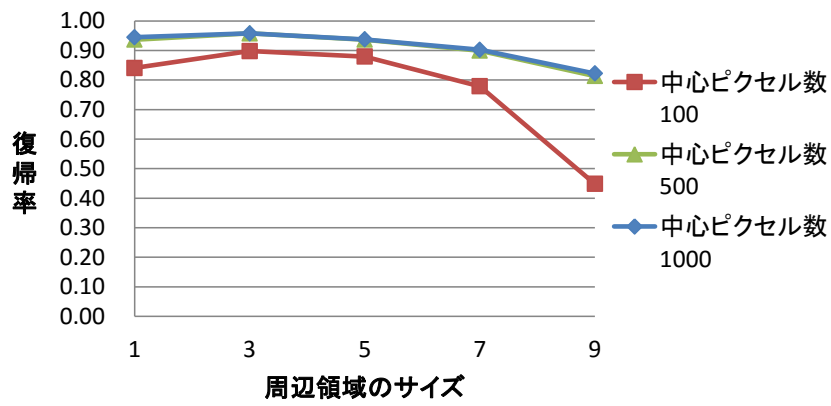
4.4.1 中心ピクセル数と周辺領域のサイズの影響の評価結果と考察

各撮影領域のデータセットを用いた場合の中心ピクセル数と周辺領域のサイズを変化させた時の提案手法の復帰率を図 4.30、処理時間を図 4.31 に示す。リローカリゼーションの成否判定の閾値は表 4.5、入力用データセットのフレーム数は表 4.7、環境再構成モデルを用いて生成したレンダリング画像のフレーム数は表 4.9、信頼度の閾値は表 4.11 に示した値を使用した。サンプリング方式は等間隔サンプリングとした。

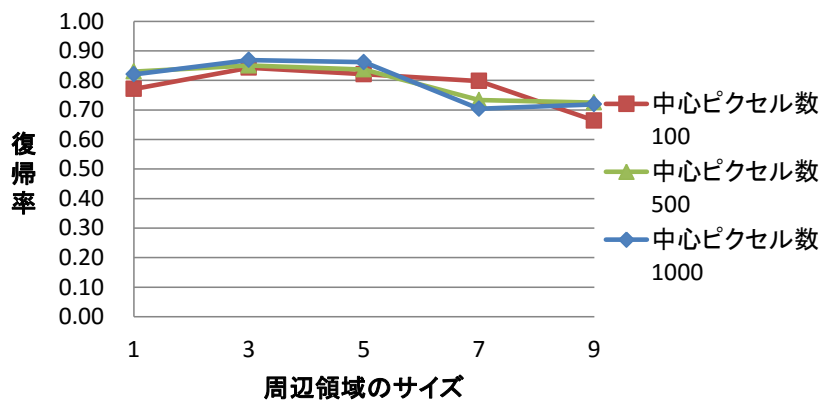
図 4.30 より、提案手法の復帰率は、中心ピクセル数が少ないと低下するが、一定以上ならばあまり影響を受けないことが分かる。これは、中心ピクセル数が多くなると画像間の相違度を計算する際に情報量が増えて精度が向上するが、中心ピクセル数が一定を超えると画像中のほぼ全てのピクセルが中心ピクセルと周辺領域により含まれることになるためと考えられる。周辺領域のサイズに関しては、撮影領域毎に異なる傾向を示しているが、周辺領域のサイズが1、つまり周辺領域を使わずに中心ピクセル



(a) area1

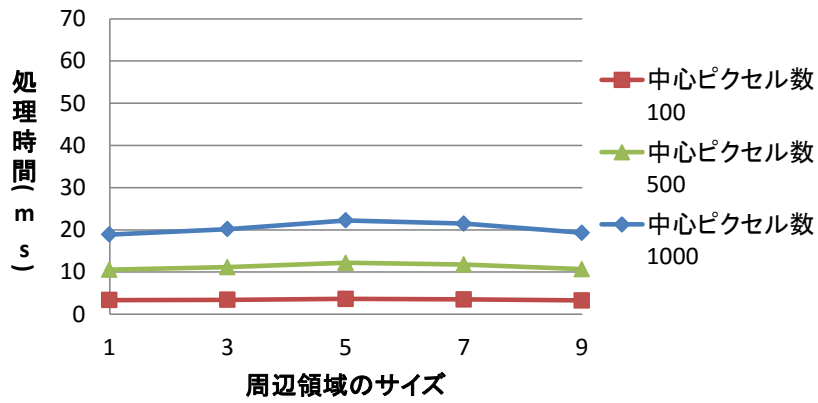


(b) area2

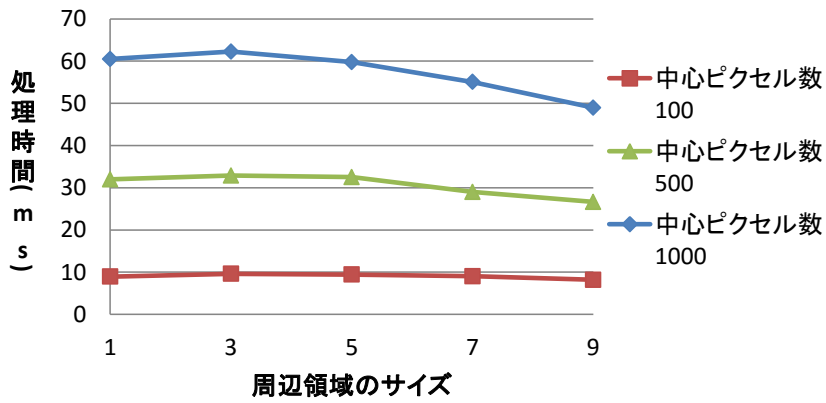


(c) area3

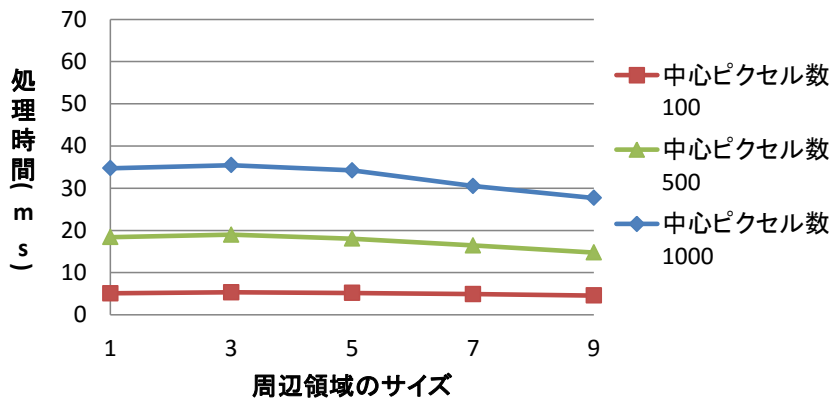
図 4.30: 等間隔サンプリングを用いた場合の提案手法の復帰率



(a) area1



(b) area2



(c) area3

図 4.31: 等間隔サンプリングを用いた場合の処理時間

のみを利用する場合に比べて、周辺領域を使う場合、特に周辺領域のサイズが3や5の場合の方が復帰率が高い傾向がある。特に Area1、Area3 が顕著で、無効領域を多く含む環境ほどこの傾向が強いことが分かった。この結果より、周辺領域と信頼度を用いることにより、最大の復帰率得られるため、環境再構成モデルと背景の境界の情報を上手く処理して画像の区別に有効に利用できると考えられる。ただし、周辺領域のサイズが大きくなりすぎる場合、周辺領域により多くの無効領域が含まれる可能性があり信頼度が低下する可能性がある。また、周辺領域毎の領域内平均に差が出にくくなり、類似画像検索の精度が低下すると考えられる。これらが周辺領域のサイズが9の時に復帰率が低下する理由であると考えられる。

次に、図 4.31 より、提案手法の処理時間は中心ピクセル数が増えると、それに比例して処理時間も増加することが分かる。これは、中心ピクセル数が増えると画像の特徴抽出や相違度計算の処理量が増えるためと考えられる。しかし、周辺領域のサイズが増えた時、一定以上になると処理時間が減少する。本来、周辺領域のサイズが増える場合、周辺領域の領域内平均を求める処理にかかる時間が増えるはずである。これは、周辺領域が大きくなると信頼度が閾値を越えることが難しくなるためであり、実際に周辺領域のサイズが1や3の時と比べると7や9の時は多くの周辺領域が信頼度の閾値を越えることができずに、相違度計算時に「無効」として排除されていた。この排除による処理量の減少が処理時間の減少の原因と考えられる。また、撮影領域毎に処理時間を比較すると、処理時間はレンダリング画像のフレーム数に比例していることもわかる。

以上の結果より、本研究の原子力発電プラント内で作成したデータセットを用いた場合、周辺領域のサイズはどの撮影領域でも安定して高い復帰率を維持できる3か5、中心ピクセル数は周辺領域のサイズが3か5の場合に安定して復帰率が高い500か1000が望ましい。中心ピクセル数が1000の場合の方が復帰率は高いが、500の場合も復帰率はあまり低下せずに処理時間を削減できることが分かる。以降では、提案手法については、その時の撮影領域やサンプリング方式に対して最も復帰率が高くなる組み合わせの場合を中心に議論する。

4.4.2 既存手法との比較

次に、本項では Randomized Fern の評価結果を示し、提案手法の評価結果と比較する。Randomized Fern と比較することにより、環境再構成モデルを用いることによりカメラの移動範囲の制限を緩和できるという提案手法の利点を確認する。本研究では、

Randomized Fern のキーフレーム DB を DB 用データセットの実画像で作成し、入力用データセットを入力とした時の復帰率と処理時間を評価対象とする。3つの撮影領域のデータセットを用いた場合の Randomized Fern の復帰率と処理時間の結果を表 4.12 に示す。また、使用した DB 用データセットのフレーム数を表 4.6、中心ピクセル数を Randomized Fern の fern 数と同じ 500 とした場合の提案手法の最大の復帰率と処理時間を表 4.13 に示した。

表 4.12: Randomized Fern の復帰率と処理時間

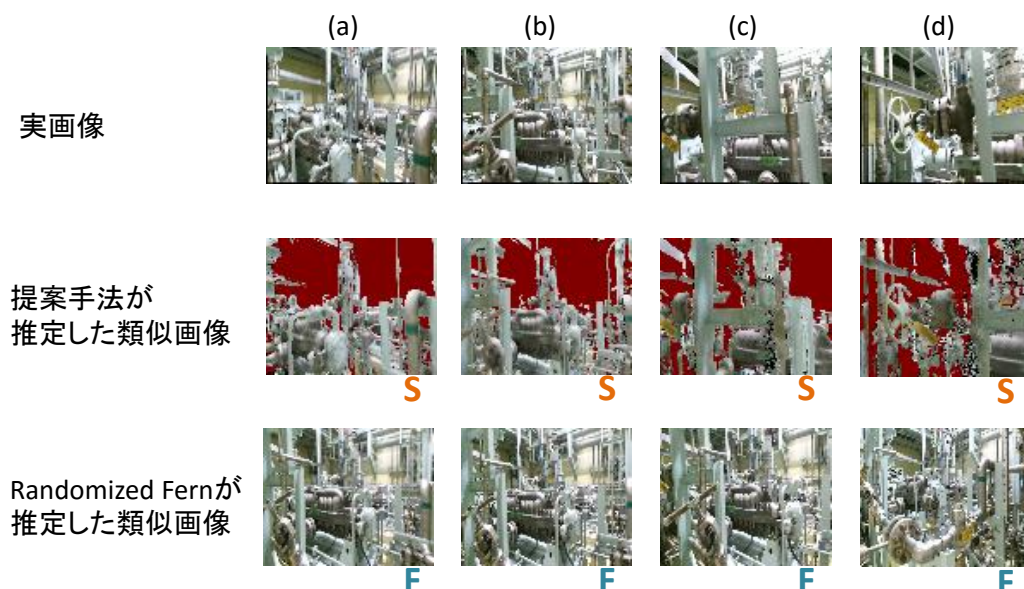
	復帰率	処理時間 (ms)
Area1	0.743	13.5
Area2	0.965	12.7
Area3	0.376	13.0

表 4.13: 中心ピクセル数が 500 の場合の提案手法の最大の復帰率と処理時間

	周辺領域のサイズ	復帰率	処理時間 (ms)
area1	5	0.952	12.2
area2	3	0.958	32.9
area3	3	0.851	19.0

表 4.12 より、特に Area3 の復帰率が非常に低いことが分かる。これは 4.3.3 項で述べたように、Area3 の入力用データセットのカメラ軌道の大部分が DB 用データセットのカメラ軌道から大きく離れているためである。実際に図 4.32 に示す提案手法と Randomized Fern が類似画像として抽出した画像の例から分かるように、類似画像を抽出してもキーフレーム DB には実画像を撮影している領域の画像は含まれないため、リローカリゼーションに失敗する。なお、図中の S はリローカリゼーションに成功した結果であることを表し、F は失敗したことを表す。この結果と提案手法の結果から、Randomized Fern のような手法の持つ、キーフレームを取得する際に撮影していない領域に復帰できないという問題を、提案手法は緩和できることが分かる。また、Area1 においても提案手法の方が高い復帰率を示した。Area1 には全体的に似た色調や形状を持つ環境があり、ここを撮影しているフレームでは Randomized Fern は画像間の違いを識別できずにリローカリゼーションに失敗していた。また、Area1 も入力用データセットのカメラ軌道の一部が DB 用データセットのカメラ軌道から大きく離れており、

これも Randomized Fern の失敗の原因と考えられる。一方、Area2 では Randomized Fern が僅かに提案手法を復帰率で上回っていた。図 4.33 の (a) では入力用データセットのカメラ軌道が DB 用データセットのカメラ軌道よりも広い範囲を通過しているため、Randomized Fern はリローカリゼーションに失敗している。一方、(b) ではサンプリング時の pitch 角が足りていないことが提案手法が失敗している主な原因と考えられる。今回の Area2 のデータセットは (b) の環境のフレームの方が多かったため、Randomized Fern が復帰率で提案手法を上回ったと考えられる。Area2 も入力用データセットのカメラ軌道の一部は DB 用データセットのカメラ軌道から離れているが、Area1 や Area3 と異なり比較的近かったため、ICP による補正によってリローカリゼーションを成功させることができたと考えられる。処理時間に関しては等間隔サンプリングにより得られるレンダリング画像のフレーム数より DB 用データベースのフレーム数が少ないため、Randomized Fern の方が提案手法より高速であった。

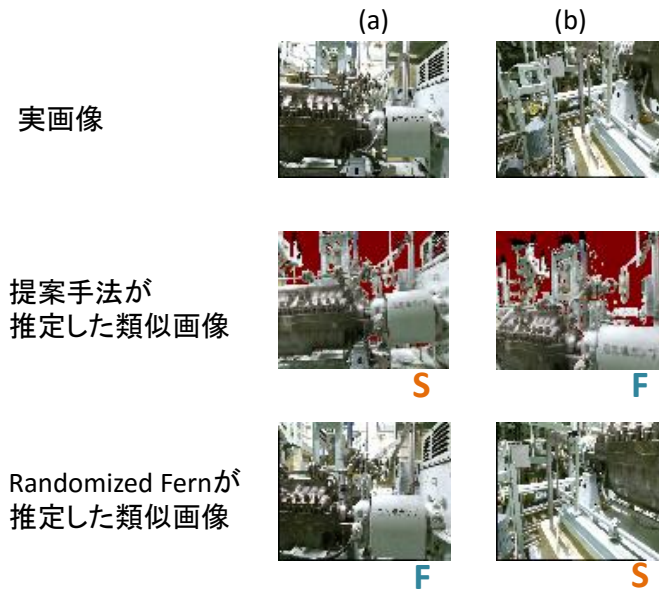


Copyright (C) 2017 Japan Atomic Energy Agency

図 4.32: 提案手法と Randomized Fern が類似画像として抽出した画像の例 (Area3 を処理対象とした場合)

4.4.3 類似画像検索に Randomized Fern を用いた場合の評価

ここまでの評価では、図 3.1 に示す提案手法の全体の流れの中のレンダリング画像に対する類似画像検索処理で、周辺領域や信頼度を用いた本研究独自の手法を用いたが、ここで Randomized Fern を用いることも考えられる。つまり、環境再構成モデルを用



Copyright (C) 2017 Japan Atomic Energy Agency

図 4.33: 提案手法と Randomized Fern が類似画像として抽出した画像の例 (Area2 を処理対象とした場合)

いて生成されたレンダリング画像から Randomized Fern により実画像との類似画像を抽出する方法である。そこで、この場合の Randomized Fern と提案手法を比較することにより、本提案手法の類似画像検索処理のレンダリング画像に対する有効性を評価する。

レンダリング画像を用いてキーフレーム DB を作成し、入力用データセットを入力とした類似画像検索に Randomized Fern を用いた場合の復帰率と処理時間と、中心ピクセル数を Randomized Fern の fern 数と同じ 500 とした場合の提案手法の最大の復帰率と処理時間を表 4.14 に示す。

表 4.14: 類似画像検索に Randomized Fern を用いた場合と提案手法の復帰率と処理時間

	Randomized Fern		提案手法	
	復帰率	処理時間 (ms)	復帰率	処理時間 (ms)
Area1	0.764	25.6	0.952	12.2
Area2	0.914	56.1	0.958	32.9
Area3	0.617	34.2	0.851	19.0

表 4.14 より、類似画像検索に Randomized Fern を用いた場合と比較して、全ての撮影領域で提案手法はより高い復帰率を示した。図 4.34 から図 4.36 に各撮影領域を処理

の対象とした場合の、類似画像検索に提案手法と Randomized Fern を用いた場合の結果の例を示す。図 4.34 の (a)、(b) や図 4.35 の (a)、図 4.36 に示すように、提案手法が正しい画像を検索でき、Randomized Fern が失敗した場合は、レンダリング画像に多くの無効領域が含まれる傾向があることが分かる。Randomized Fern の場合、無効領域の割合が高くなると画像間の比較に利用できないピクセルが増え、類似画像検索の精度が低下する。一方、提案手法は周辺領域や信頼度を用いることにより、多くの無効領域が含まれるレンダリング画像の場合でも、正確に類似画像を検索しリローカリゼーションを成功できることが示された。また、図 4.34 の (d) は実画像と似ている結果であるが、どちらの手法もリローカリゼーションは失敗していた。これは、画像中央の鉄骨が占める領域が全体的に緑の色調でデプス値もほぼ同じであり、このような領域が大部分を占める画像同士を正確に区別することがどちらの手法にとっても困難であることが原因と考えられる。一方、図 4.34 の (c) が成功しているのは、画像中の右下の充填ポンプの画像を占める割合が比較的大きく、この部分が画像の区別に有効であったためと考えられる。また、図 4.35 の (b) で失敗している原因はサンプリング時に pitch 角のサンプリングの範囲が十分ではなく、ICP による補正をしても正解が無かったことが原因と考えられる。また、中心ピクセル数が 500 の時、提案手法は Randomized Fern よりも高い復帰率と早い処理時間を実現でき、提案手法は類似画像検索の際にレンダリング画像による無効領域などの問題に対処できていると言える。提案手法の方が高速であるのは提案手法では、Randomized Fern でのガウシアンフィルタ処理のような前処理を行っていないためであると考えられる。

4.4.4 サンプリング方式の比較

本項では 3.2.3 項で述べた 2 種類のサンプリング方式による提案手法の評価結果を比較する。

深度値標準偏差サンプリングを実行した場合の各撮影領域の提案手法の復帰率を図 4.37、処理時間を図 4.38 に示す。この時、中心ピクセル数は 4.4.1 項で最大の復帰率を示した 1000 とした。また、中心ピクセル数が 1000 の時の等間隔サンプリングを用いた提案手法と比較した場合の復帰率を図 4.39 に、処理時間を図 4.40 に示す。

深度値標準偏差サンプリングの場合、図 4.22 に示したように、ステップの係数 a が $1/200$ の時はステップ間隔が大きすぎるために、サンプリングされるカメラ姿勢が少なく復帰率が低下したと考えられる。しかし、図 4.39 に示すように、 a が $1/400$ 以下の場合は等間隔サンプリング時と比較しても復帰率に大きな差が無いことが分かる。そ

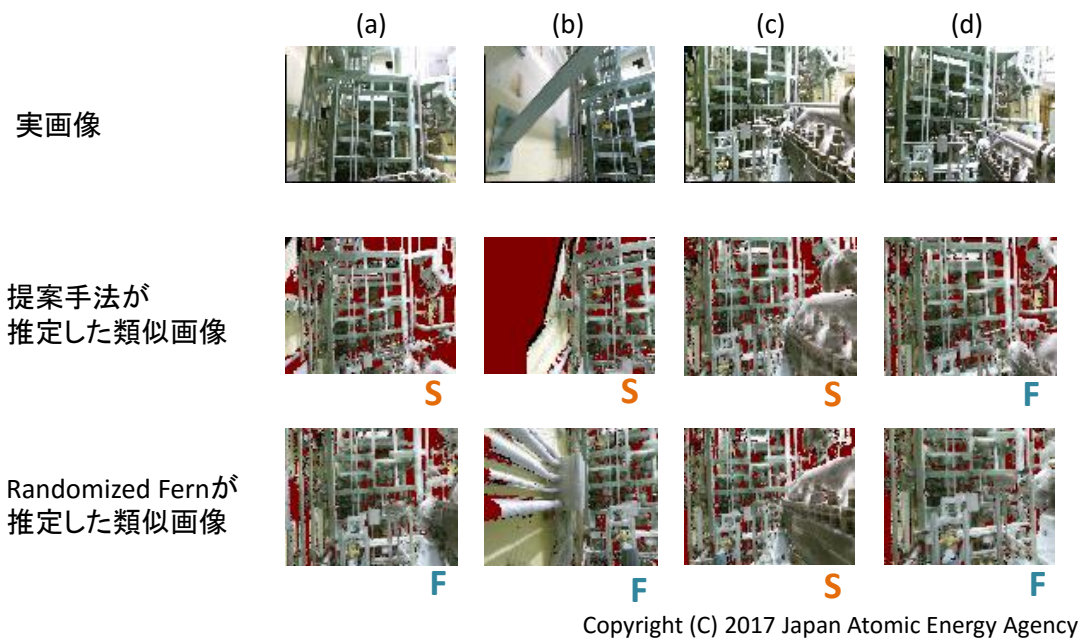


図 4.34: 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似画像として抽出された画像の例 (Area1 を処理対象とした場合)

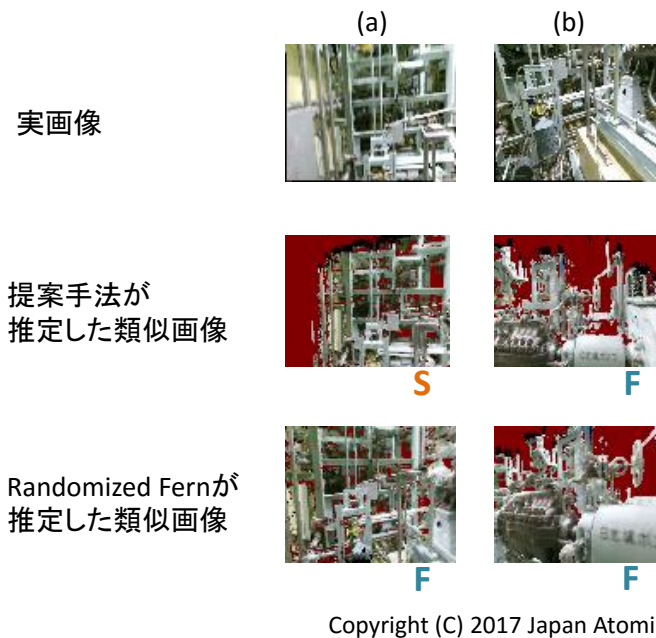
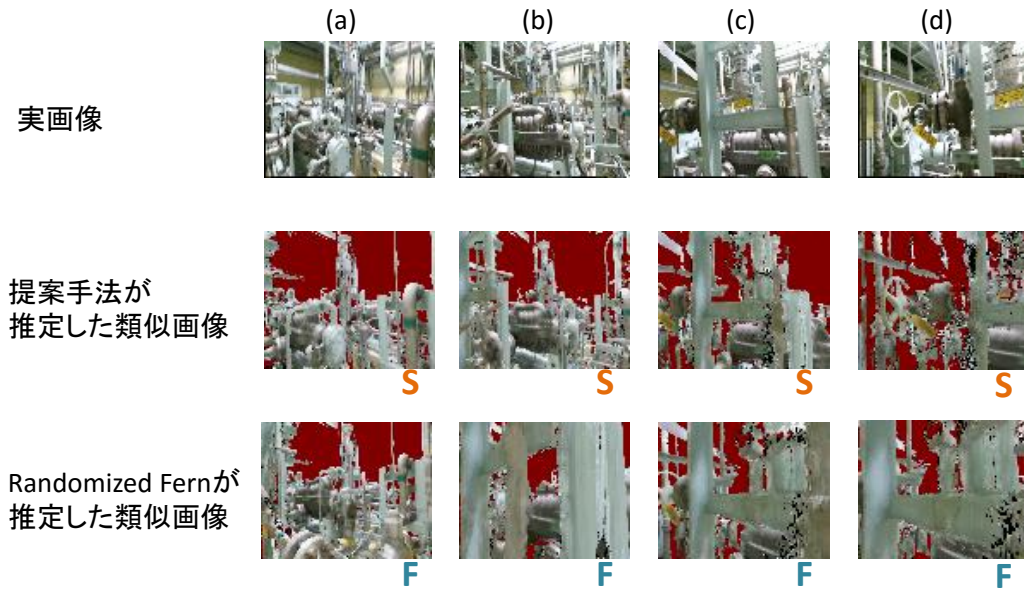


図 4.35: 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似画像として抽出された画像の例 (Area2 を処理対象とした場合)



Copyright (C) 2017 Japan Atomic Energy Agency

図 4.36: 提案手法と類似画像検索に Randomized Fern を用いた場合の結果で類似画像として抽出された画像の例 (Area3 を処理対象とした場合)

して、a が 1/400 の場合は等間隔サンプリング時と比較して、表 4.9 に示したようにレンダリング画像のフレーム数が減少し、それにより図 4.40 に示すように大幅に処理時間が減少しており、復帰率を維持したまま処理時間を削減できた。また、Area3 の場合は、図 4.22 の左側で密にサンプリングされており、これにより等間隔サンプリング時には失敗していた場所でもリローカリゼーションに成功していたため、等間隔サンプリングの場合より復帰率が向上していた。この結果により、深度値標準偏差サンプリングを利用することにより、少ないフレーム数でも高い復帰率が維持できる効率的なサンプリングができる可能性が示された。3.2.3 項で述べたように、本研究の深度値標準偏差サンプリングは単純化された手法であるが、環境の複雑さによるデプス値のばらつきを利用することの有効性が示された。深度値標準偏差サンプリングにより、等間隔サンプリングと比較して、同じキーフレーム DB の量でもより多様なパラメータの範囲でのカメラ姿勢のサンプリングが可能となり、カメラの移動可能領域の拡大やカメラ姿勢推定の精度の向上などが期待できる。

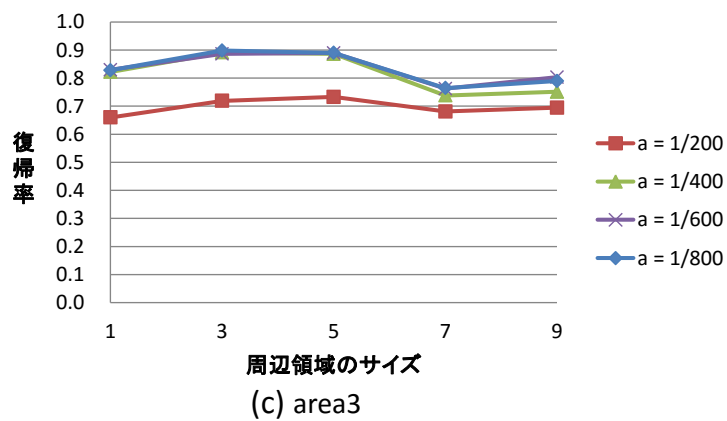
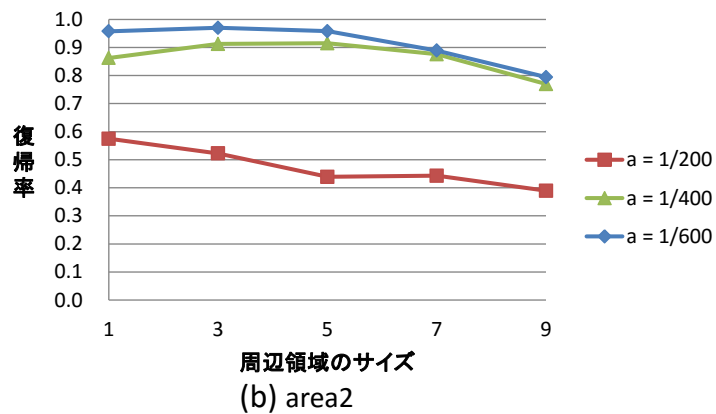
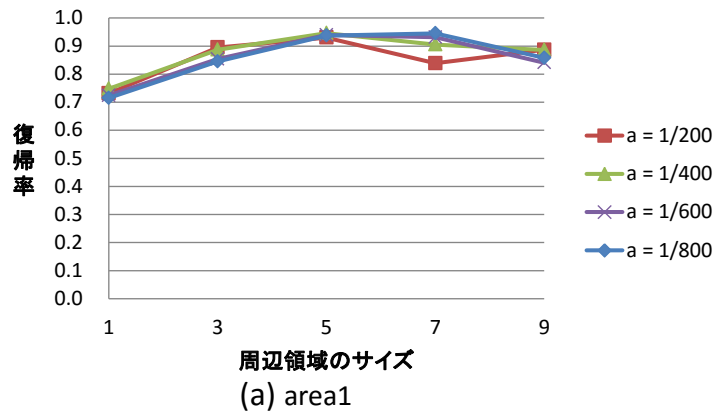


図 4.37: 深度値標準偏差サンプリングを用いた場合の提案手法の復帰率

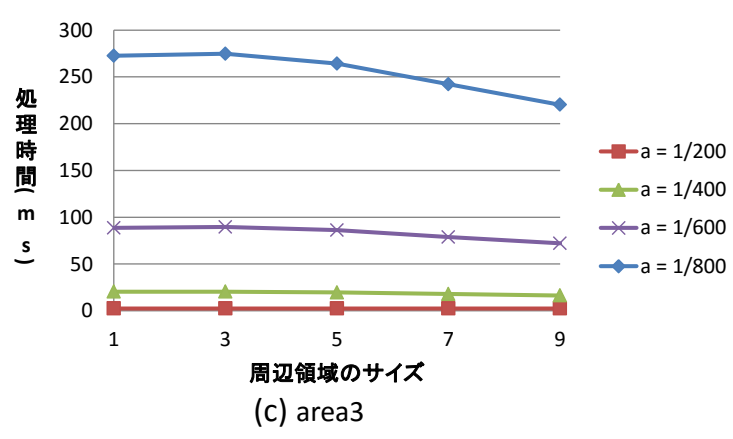
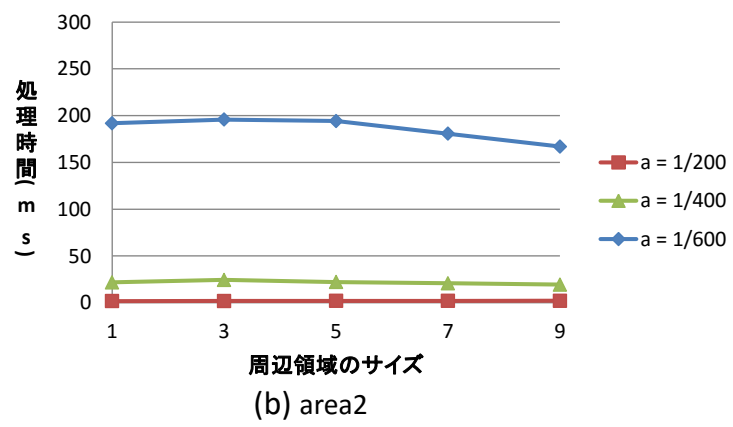
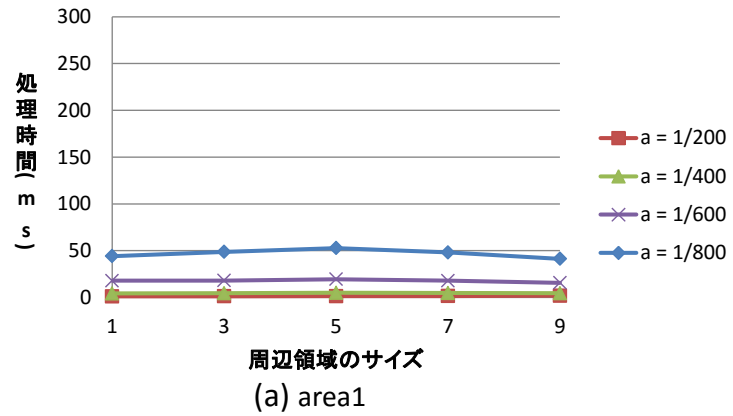


図 4.38: 深度値標準偏差サンプリングを用いた場合の処理時間

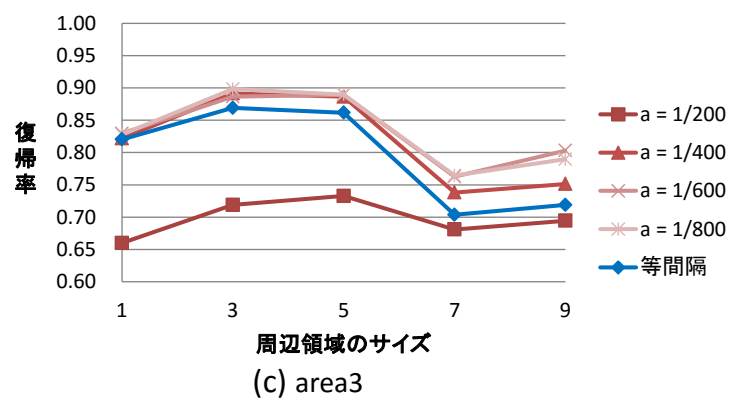
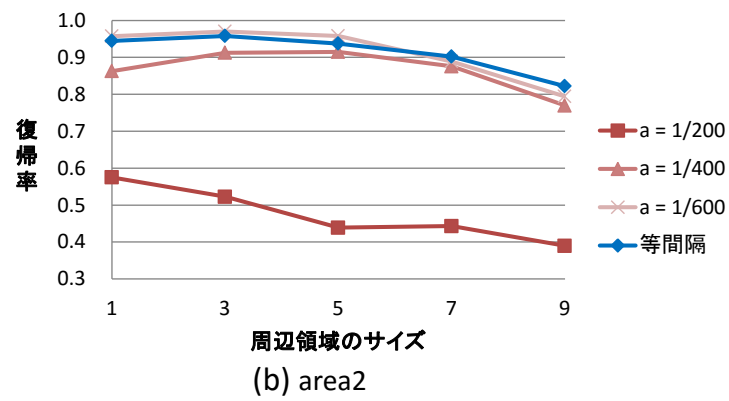
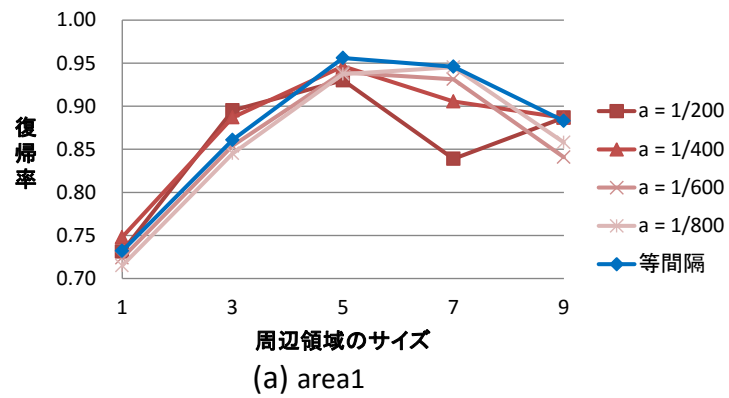


図 4.39: 等間隔サンプリングと深度値標準偏差サンプリングの復帰率の比較

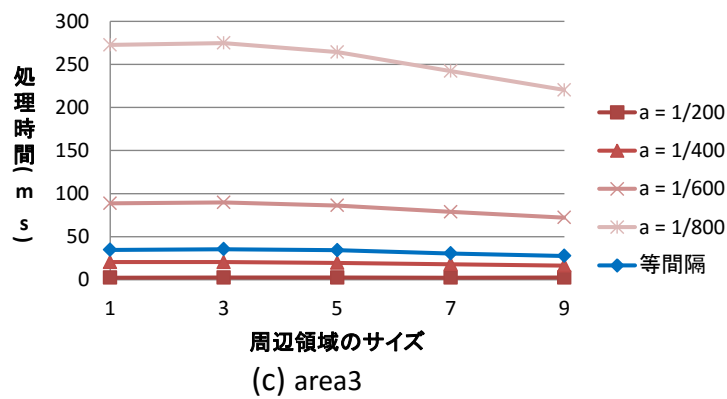
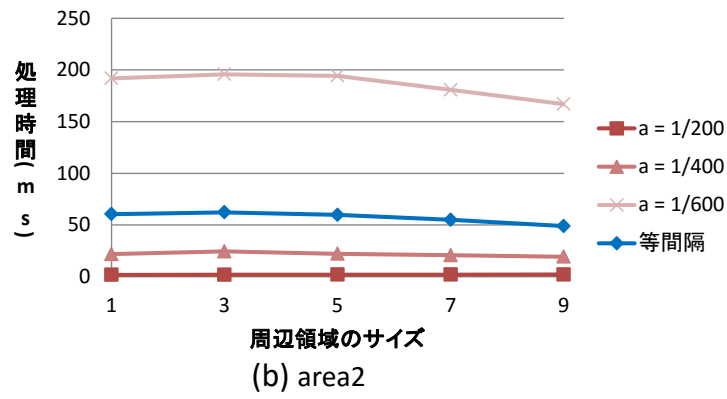
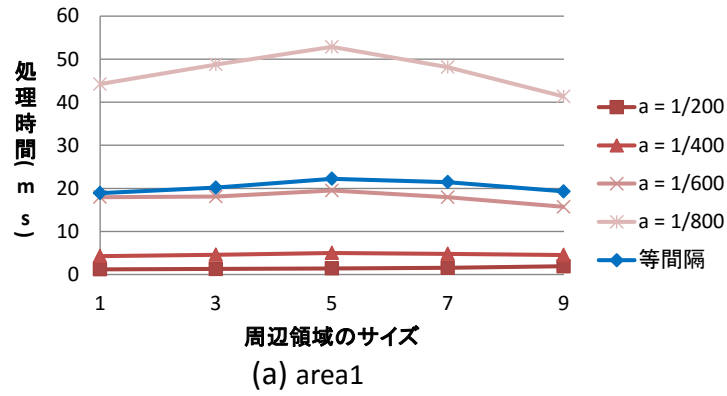


図 4.40: 等間隔サンプリングと深度値標準偏差サンプリングの処理時間の比較

第 5 章 結論

原子力発電プラント内で AR を用いた解体作業支援を実現するためには、従来のリローカリゼーション手法が持つ、事前に撮影した領域でしか利用できないという問題を解決する必要がある。そこで、本研究では撮影画像から環境の色と三次元構造を再構成した環境再構成モデルによるレンダリング画像を利用し、カメラの移動範囲が制限されるという問題を緩和するリローカリゼーション手法を開発することを目的とした。

提案手法では、環境再構成モデルを用いて、事前に撮影していない領域で撮影した画像をレンダリングにより模擬的に生成して利用する。しかし、レンダリング画像を利用する場合、環境再構成モデルの欠損などによる実画像とレンダリング画像の間の差異に起因する問題を解決する必要がある。そこで、本提案手法では類似画像検索処理に周辺領域や信頼度を導入することにより、環境再構成モデルの歪み・位置ずれ、色の微妙な違いがある場合の悪影響や、環境再構成モデルが写らない無効領域の悪影響を緩和することを試みた。

また、本提案手法ではレンダリングの際のカメラ姿勢のサンプリングの仕方により、リローカリゼーションが可能な範囲や処理時間が変化する。そこで本研究では、等間隔にカメラ姿勢のパラメータを変化させてサンプリングする等間隔サンプリングと、あるカメラ姿勢のレンダリングデプス画像のデプス値のばらつきを利用して次のカメラ姿勢のパラメータを決定する深度地標準偏差サンプリングの 2 種類のサンプリング方式を実装して比較した。

本研究では、原子炉廃止措置研究開発センターの充填ポンプ室内で Kinect カメラにより取得した RGB 画像とデプス画像を対象に提案手法と既存手法によるリローカリゼーションを実行し、その成否と処理に要する時間で各リローカリゼーション手法の性能を評価した。

原子力発電プラント内で提案手法と Randomized Fern を評価した結果、環境再構成モデルを用いて生成したレンダリング画像を用いることにより、提案手法は事前に撮影していない領域でもリローカリゼーションを成功させることができた。この結果より、Randomized Fern のような既存手法が持つカメラの移動範囲の制限の問題を緩和できたことが示された。周辺領域や信頼度を導入することにより、提案手法はレンダリング画像に多くの無効領域が含まれる場合でも、正確に実画像と類似したレンダリ

ング画像を検索しリローカリゼーションを成功させることができることも確認できた。また、等間隔サンプリングの場合と比較して、深度値標準偏差サンプリングの場合は高い復帰率を維持したまま処理時間を削減できた。これにより、環境の複雑さによるデプス値のばらつきを利用することの有効性が示された。

今後の課題としては、想定する作業領域が広くなるとレンダリング画像の数も増えるため、キーフレーム DB の構築や類似画像検索の処理にかかる時間が増加する問題がある。現在はキーフレーム DB 全体を検索しているが、木構造データベースやハッシュテーブルなどを用いてデータベースの構造を改良することにより類似画像検索の処理時間を削減できると期待できる。また、深度値標準偏差サンプリングの係数を自動的に決定する方法や、サンプリングの範囲を自動的に決める方法を考案することも課題となる。現時点では、環境再構成モデルは自動的に更新されないため、実際の AR による解体作業支援では、解体や保修作業が進むに従って環境が変化した場合に、環境再構成モデルを再度作成し、キーフレーム DB を更新しなければならない。これを解決するためには、環境再構成モデルやキーフレーム DB の更新手法や、多少の環境変化に対してもロバスト性の高い類似画像検索処理を導入するなどの工夫が必要である。

謝 辞

本研究を進めるにあたり、研究会等での数々の助言に加え、惜しみない激励とご指導ご鞭撻を頂いた下田宏教授に深く感謝を申し上げます。時折学生部屋で励ましの言葉や差し入れを頂いたりしたことも大変励みになりました。

研究の進め方、プログラミング、論文執筆、うどんの打ち方等の様々な面でご指導とお力添えを頂いた石井裕剛准教授に心より感謝致します。ARの研究は自身の進路にも影響するほどに面白い研究であったと思います。こう思えるのも先生のご指導のおかげです。

先生方には研究のみならず、就職活動など様々な面でもご指導頂き、また学会発表などのチャンスも頂き、自分でも大きな成長を遂げられたと自負しております。重ねてお礼申し上げます。

同じARチームとして、プラント内でのデータセット作成、アプリケーションの作成、論文執筆等で、様々な助力を頂いた工学部電気電子工学科4回生の久留島隆史君、原園友規君に深く感謝致します。同じ分野を研究する仲間がいることは非常に心強かったです。

また、論文執筆の際に全面的にサポートをして頂いたエネルギー科学研究科修士2回生のBingrong Huang君、1回生の緒方省吾君に深くお礼申し上げます。忙しい中様々な面でサポートしてもらい非常に助かりました。

研究のみならず、様々なことで支えて頂き、共に学び、ともに遊び、有意義な時間を過ごせるようにして下さったエネルギー情報学研究室の修士二回の浦山大輝君、下中尚忠君、遠藤竜太君、杉田耕介君、修士一回の上田樹美さん、大橋由暉君、辻雄太君、岩崎達郎君、工学部電気電子工学科4回生の日下部曜君、皆様に心より感謝を申し上げます。おかげで研究室での生活は研究も含めて非常に楽しく感じました。

細かな事務手続きや研究生活全般で、様々な支援を頂いた普照郁美秘書と山田美保秘書に深く感謝を申し上げます。いろいろと気にかけて頂きありがとうございました。

お忙しい中、研究に協力して下さった原子炉廃止措置研究開発センターの皆様に、改めてお礼申し上げます。

最後にここまで支えて下さった家族や友人にこの場を借りて、お礼申し上げます。

参考文献

- [1] 経済産業省: 資源エネルギー庁ホームページ, http://www.enecho.meti.go.jp/about/whitepaper/2016pdf/whitepaper2016pdf_2_1.pdf, Accessed January 29, 2017.
- [2] 株式会社三菱総合研究所: 原子力安全研究本部ホームページ, http://www.mri.co.jp/opinion/column/atomic_energy/aenergy_20160809.html, Accessed February 1, 2017.
- [3] Weida Yan, Shuhei Aoyama, Hirotake Ishii, Hiroshi Shimoda, Tran T. Sang, Solhauglars Inge, Toppe aleksander Lygren, Johnsen Terje, Masanori Izumi: Development and Evaluation of a Temporary Placement and Conveyance Operation Simulation System Using Augmented Reality, *Nuclear Engineering and Technology*, 44(5), pp.507-522 (2010).
- [4] Hirotake Ishii: Plant maintenance and dismantling work support based on three-dimensional scanning technology, *International Journal of Nuclear Safety and Simulation*, 4(2), pp.97-104 (2013).
- [5] Ihsan Rabbi, Sehat Ullah: A Survey on Augmented Reality Challenges and Tracking, *Acta Graphica*, 1(2), pp.29-46(2013).
- [6] Bruce H. Thomas: A Survey of Visual Mixed, and Augmented Reality Gaming, *ACM Computers in Entertainment*, 10 pp.16-52(2012) .
- [7] Jorge Fuentes-Pacheco, Juan Manuel Rendon-Mancha: Visual simultaneous localization and mapping: a survey, *Artificial Intelligence Review* 43(1), pp. 55-81 (2015).
- [8] 中井 俊憲: 拡張現実感技術を利用した原子力プラントの解体支援手法の提案と評価, エネルギー科学研究科修士論文, 京都大学 (2006).

- [9] Rune Vabo, Leon Piotrowski, Grete Rindahl: 3D representation of radioisotopic dose rates within nuclear plants for improved radioprotection and plant safety, *International Journal of Nuclear Safety and Simulation*, 1(2), pp.127-133 (2010).
- [10] Hirotake Ishii : Augmented Reality: Fundamentals and Nuclear Related Applications, *International Journal of Nuclear Safety and Simulation*,1(4), pp. 316-327(2010).
- [11] Eric Marchand, Hideaki Uchiyama, Fabien Spindler: Pose Estimation for Augmented Reality: A Hands-On Survey. *IEEE Transactions on Visualization and Computer Graphics*, Institute of Electrical and Electronics Engineers, 22 (12), pp.2633 - 2651(2016).
- [12] 唐津豊、中澤仁、高汐一紀、徳田英幸: 加速度センサを利用したマーカレス AR によるセンサ情報可視化システム, *電子情報通信学会信学技報*,110, pp.79-84 (2010).
- [13] Joseph Newman, David Ingram, Andy Hopper: Augmented Reality in a Wide Area Sentient Environment, *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR) 2011*, pp.77-86 (2011).
- [14] Hirokazu Kato, Mark Billinghurst: Marker Tracking and HMD Calibration for a Video based Augmented Reality, *2nd IEEE and ACM International Workshop*, pp.85-94 (1999).
- [15] Vincent Lepetit, Francesc Moreno Noguera, Pascal Fua: EPnP: An Accurate $O(n)$ Solution to the PnP Problem *International Journal of Computer Vision*, pp.155-166 (2009).
- [16] Harald Wuest, Didier Stricker: Tracking of industrial objects by using cad models. *Journal of Virtual Reality and Broadcasting*, 4(1), pp.155-164(2007).
- [17] Wenjun Zhu, Peng Wang, Fudong Li, Jianhua Su, Hong Qiao: Real-time 3D Model-based Tracking of Work-piece with Monocular Camera, *Proceedings of the IEEE/SICE International Symposium on System Integration (SII)*, pp.777-782 (2015).

- [18] Amit Agrawal, Yu Sun, John Barnwell, Ramesh Raskar: Vision Guided Robot System for Picking Objects By Casting Shadows, *The International Journal of Robotics Research* , 29(2-3), pp.155-173(2010).
- [19] Frank Dellaert, Steven M. Seitz, Charles E. Thorpe, Sebastian Thrun: Structure from Motion without Correspondence, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, pp.557-564(2000).
- [20] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, Richard Szeliski: Building Rome in a Day, *Communications of the ACM*, 54(10), pp.105-112(2011).
- [21] Changchang Wu: Towards Linear-time Incremental Structure From Motion, *3DTV-Conference, 2013 International Conference on*, pp.127-134 (2013).
- [22] Andrew J. Davison: Real-Time Simultaneous Localisation and Mapping with a Single Camera, *Proceedings Ninth IEEE International Conference on Computer Vision (ICCV)*, pp.1403-1410 (2003).
- [23] Andrew J. Davison, Ian D. Reid, Nicholas D. Molton, Olivier Stasse, MonoSLAM: Real-Time Single Camera SLAM, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6) pp.1052-1067 (2007).
- [24] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, John J. Leonard: Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age, *IEEE Transactions on Robotics*, 32(6), pp.1309-1332 (2016).
- [25] Georg Klein, David Murray: Parallel Tracking and Mapping for Small AR Workspaces, *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR) 2007*, pp.225-234 (2007).
- [26] Richard A. Newcombe, Steven J. Lovegrove, Andrew J. Davison: DTAM: Dense Tracking and Mapping in Real-Time, *UK Computer Vision (ICCV), IEEE International Conference on 2011*, pp.2320 - 2327 (2011).

- [27] Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon: KinectFusion: Real-Time Dense Surface Mapping and Tracking, IEEE International Symposium on Mixed and Augmented Reality 2011, Science and Technology Proceedings, pp.26 -29 (2011).
- [28] Olaf Kahler, Victor Adrian Prisacariu, Carl Yuheng Ren, Xin Sun, Philip Torr, David Murray: Very High Frame Rate Volumetric Integration of Depth Images on Mobile Device IEEE Transactions on Visualization and Computer Graphics, 21(11), pp.1241-1250 (2015).
- [29] GitHub, Inc: InfiniTAM, <https://github.com/victorprad/InfiniTAM/tree/c8ba1872746a4f490568613e331069a9166a855f> Accessed December 10, 2016.
- [30] Georg Klein, David Murray: Improving the Agility of Keyframe-Based SLAM, European Conference on Computer Vision (ECCV), 2008, pp.802-815 (2008).
- [31] Christian Pirchheim, Gerhard Reitmayr: Homography-Based Planar Mapping and Tracking for Mobile Phones, 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp.27-36 (2011).
- [32] Mark Cummins, Paul Newman: FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance, The International Journal of Robotics Research 2008, 27(6), pp.647-665 (2008).
- [33] Arren, Michael, W. Maddern, Stephen: openfabmap, <http://code.google.com/p/openfabmap/http://code.google.com/p/openfabmap/>, Accessed February 3, 2015.
- [34] Ben Glocker, Jamie Shotton, Antonio Criminisi, Shahram Izadi: Real-Time RGB-D Camera Relocalization, Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on, 1(4), pp.173-179 (2013).
- [35] Ben Glocker, Jamie Shotton, Antonio Criminisi, Shahram Izadi: Real-Time RGB-D Camera Relocalization via Randomized Ferns for Keyframe Encoding, Visualization and Computer Graphics, IEEE Transactions on, 21(5), pp. 571-583 (2015).

- [36] Andrew P. Gee, Walterio Mayol-Cuevas: 6D Relocalisation for RGBD Cameras Using Synthetic View Regression, Proceedings of the British Machine Vision Conference (BMVC), (2012).
- [37] Arnold Irschara, Christopher Zach, Jan-Michael Frahm, Horst Bischof: From Structure-from-Motion Point Clouds to Fast Location Recognition, IEEE Conference on Computer Vision and Pattern Recognition, pp.2599-2606 (2009).
- [38] Torsten Sattler, Bastian Leibe, Leif Kobbelt: Fast ImageBased Localization using Direct 2D-to-3D Matching, IEEE International Conference on Computer Vision, pp.667-674 (2011).
- [39] Martin A. Fischler, Robert C. Bolles: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, Communications of the ACM, 24(6), pp.381-395 (1981).
- [40] Renato F. Salas-Moreno, Richard A. Newcombe, Hauke Strasdat, Paul H. J. Kelly, Andrew J. Davison: SLAM++: Simultaneous Localisation and Mapping at the Level of Objects, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1352-1359 (2013).
- [41] Stefan Hinterstoisser, Vincent Lepetit, Slobodan Ilic, Stefan Holzer, Gary Bradski, Kurt Konolige, Nassir Navab: Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes, Asian Conference on Computer Vision (ACCV), pp.548-562 (2012).
- [42] Stefan Hinterstoisser, Stefan Holzer, Cedric Cagniart, Slobodan Ilic, Kurt Konolige, Nassir Navab, Vincent Lepetit: Multimodal Templates for Real-Time Detection of Texture-Less Objects in Heavily Cluttered Scenes, IEEE International Conference on Computer Vision (ICCV), pp.858-865 (2011).
- [43] Jamie Shotton, Ben Glocker, Christopher Zach, Shahram Izadi, Antonio Criminisi, Andrew Fitzgibbon: Scene coordinate regression forests for camera relocalization in RGB-D images, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2930-2937 (2013).

- [44] Alex Kendall, Matthew Grimes, Roberto Cipolla: PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization, IEEE International Conference on Computer Vision (ICCV), pp.2938-2946 (2015).
- [45] Taro Kimura, Hiroki Tokumaru, Hirotake Ishii, Hiroshi Shimoda, Yuya Kouda: Evaluation and Improvement of a Relocalization Method for Camera Tracking in Nuclear Power Plants, Proceedings of International Symposium on Socially and Technically Symbiotic Systems (STSS), pp.407-416 (2015).
- [46] Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren, Dieter Fox: RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments, In The International Journal of Robotics Research 31(5), pp.647-663 (2012).
- [47] Liang Zhang, Peiyi Shen, Jieqiong Ding, Juan Song, Jingwen Liu, Kang Yi: An Improved RGB-D SLAM Algorithm based on Kinect Sensor, IEEE International Conference on Advanced Intelligent Mechatronics (AIM), pp.555-562 (2015).
- [48] P.J. Besl, Neil D. McKay: A Method for Registration of 3-D Shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2), pp.239-256 (1992).
- [49] Antoine Petit, Guillaume Caron, Hideaki Uchiyama, Eric Marchand: Evaluation of Model based Tracking with TrakMark Dataset, 2nd Int. Workshop on AR/MR Registration, Tracking and Benchmarking, (2011).
- [50] Radu Bogdan Rusu, Steve Cousins, Willow Garage: 3D is here: Point Cloud Library (PCL), IEEE International Conference on Robotics and Automation (ICRA)
- [51] Dirk Holz, Alexandru-Eugen Ichim, Federico Tombari, Radu B. Rusu, and Sven Behnke: Registration with the Point Cloud Library PCL A Modular Framework for Aligning 3D Point Clouds, IEEE Robotics & Automation Magazine 22(4), pp.110-124 (2015).
- [52] Microsoft Inc: Kinect ハードウェア, <https://developer.microsoft.com/ja-jp/windows/kinect/hardware>, Accessed February 1, 2017.

- [53] Thiemo Wiedemeyer: IAI Kinect2, https://github.com/code-iai/iai_kinect2 Accessed June 12, 2016.
- [54] Will Schroeder, Ken Martin, Bill Lorensen, The Visualization Toolkit (4th ed.), Kitware Inc, (2006).
- [55] G. Bradski and A. Kaehler. Learning OpenCV: Computer vision with the OpenCV library. O ' Reilly Media Inc, (2008).