

鋼製土木構造物の形状把握のための Neural Radiance Fields と Photogrammetry の比較検討

Comparative Study of Neural Radiance Fields and Photogrammetry for Shape Determination of Steel Structures

○大澤 信哉^{*1}, 齊藤 忠明^{*1}, 原 直人^{*2}
Shinya OSAWA^{*1}, Tadaaki SAITO^{*1} and Naoto HARA^{*2}

^{*1} 株式会社 IHI インフラシステム 新規事業推進部

Assistant Manager, New Business Promotion Division, IHI Infrastructure Systems Co, Ltd.

^{*2} 株式会社 IHI インフラシステム 新規事業推進部 課長

Manager, New Business Promotion Division, IHI Infrastructure Systems Co, Ltd.

キーワード : NeRF; SfM; 社会インフラ; 鋼構造物

Keywords: NeRF; SfM; public infrastructure; steel structure

1. はじめに

公共インフラプロジェクトで新たに建設される鋼製土木構造物は、BIM/CIM (Building/Construction Information Modeling) を活用した効率的なデータ作成が一般的になりつつある。しかし、既存の構造物を更新する場合、データや情報の不足が問題となることがある。たとえ情報が存在しても、経年変化などの要因により、実際の状況と設計情報が一致しないことが多々ある。このため、建設現場では構造物の現状を正確に情報化する技術が求められている。

現在、形状把握の手段として、従来の「人力による巻き尺計測」に加えて、「レーザスキャナによる点群取得」が一般的に採用されている。巻き尺計測は一部では効率的だが、ヒューマンエラーが発生しやすい上、事前準備に時間を要する。それに対して、レーザスキャナはヒューマンエラーが起きにくく、スキャナの設置箇所の周囲全体をデータ化するため、計測漏れのリスクを最小限に抑えられる。このことは事前準備の時間短縮にも繋がる。

しかし、実際の現場ではレーザスキャナが歓迎されない場合もある。鋼製土木構造物の場合、局所的な精度はミリ単位が要求されるが、局所的な精度は巻き尺と比較して十分ではない。レーザスキャナの計測精度は計測対象との距離に依存し、遠い場合はミリ単位の誤差が生じることがある。この課題は、設置箇所の工夫などである程度解決できるが、コストも問題となる。ハードウェアやソフトウェアを含むイニシャルコストが高く、設置・操作・データ操作に特殊なスキルが必要になる。

そのため、より簡単かつ低コストで構造物の現形状を把握できる方法への需要が高まっている。この要求に応えるために、photogrammetry を試験的に導入している。その結果、①容易かつ低コストで実施可能、②質・量共に十分な写真があればレーザスキャナに匹敵する精度が得られる

ことを確認している。これは施工計画などの用途においては十分に採用できることを示唆している。ただし、「十分な写真」を取得する必要があるため、現地での写真撮影が煩雑であり、写真の入手が難しいという課題が存在する。

この課題に対処するため、最近急速に進化している NeRF (Neural Radiance Fields) 技術が有望ではないかと考えたり。しかし、NeRF はまだ発展途上の技術であり、実際の建設プロジェクトへの適用例は限られている。このため、本報告では photogrammetry と NeRF を比較し、その有用性について検討する。

2. 試験条件等

2.1. 試験環境

適用したソフトウェアとハードウェア環境を表1、表2に示す。ソフトウェアはオープンソースを基本とし、NeRF用ソフトとして、InstantNGP²⁾に加え点群出力が可能なNerfstudio³⁾[nerfacto 適用]でも試験を行った。

表1 適用ソフトウェア

	photogrammetry	NeRF	
ソフト	openMVG/MVS	InstantNGP	Nerfstudio
Version	2.0	表記なし	0.3.2

表2 使用ハードウェア

CPU	AMD Ryzen5600X		
Memory	32GB		
Storage	M2.SSD 1TB ×2unit		
GPU	nVidia RTX3060ti (VRAM8GB)		
OS	Windows11		
環境等	Python3.8, CUDA11.7, Anaconda3 etc		

2.2 試験対象

2.2.1 鋼構造物

試験対象は、弊社工場内にある金物とした（図 1）。大きさは、横幅約 6m、奥行約 3.5m、高さ約 2.2m である。材質は、ステンレス製であり、光の反射が強く、光学的手法（例えば赤外線やレーザによる方法）による計測は、やや不向きな材質である。また、試験対象の前面を除く三方には、別の金物が配置されていた（図 2）。



図 1 試験対象の鋼構造物

これらは写真撮影時に障害物となるため、対象物の前面を除いて、写真撮影時に理想的な画角を確保することが難しかった。試験対象が画角に収まりきらないか、あるいは、画角内に障害物が映り込んでいるものもあった。

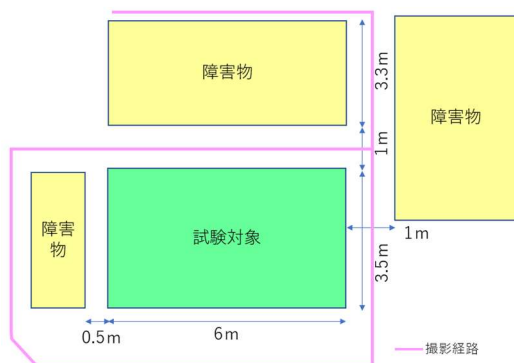


図 2 試験対象周囲の平面配置(概略寸法)

2.2.2 撮影条件

撮影はスマートフォン（iPhone12）のカメラで行った。解像度 4K / HDR オフ / 広角モードで撮影を行った。これらの条件は、事前の予備検討結果を考慮して決定した。

撮影は、試験対象を 3 周（それぞれ画角を水平/上向き/下向き）回って合計 57 枚の撮影を行った（図 2）。

撮影は春の 14 時ごろに行った。晴天だったため、影が比較的くっきりと写っている。画角によっては、対象物が完全に影に入り込んでいるところもある。加えて反射しやすい材質のため、画角によっては同じ部位であっても、見え方（＝反射光）が変わる写真となっていた。

3. 試験方針

3.1 比較項目

実現場に応じ、適切に photogrammetry か NeRF か選定する方法の要因となる比較項目を検討した。

撮影は、施工の種類や進捗によって、計測対象にカメラが近づける場合もあれば、撮影できる時間が限られる場合もある。従って撮影できる写真の枚数が限られる場合を考慮し、比較にあたっての変動パラメータは写真枚数とした。今回撮影した枚数は 57 枚であり、その中から写真を抽出していき最低枚数は 8 枚と設定した（4 方向×2 枚）。これを基準に、パラメータは 8/12/20/32/48/57 枚と設定した。

事前の検討で定性的に把握していた、photogrammetry と NeRF それぞれの特性を考慮し、比較項目は、①処理速度、②再現度、③精度（平坦度）とした。これは、①NeRF の処理速度が明らかに速い、②photogrammetry は写真品質が悪い場合は点群が欠損しがち、ということ把握していたことが背景にある。

3.2 NeRF 処理完了の定義

NeRF は AI の一種であり、入力となる写真に可能な限り近づくアウトプットになるように学習し続ける。処理速度を photogrammetry と比較するには、NeRF を“処理終了”と判断するための指標が必要になる。NeRF の学習度合いを示すパラメータとしてロス率があり、処理終了とみなせるロス率について検討した。ロス率は、画面表示の読み取り及び tensorboard で出力して記録した。

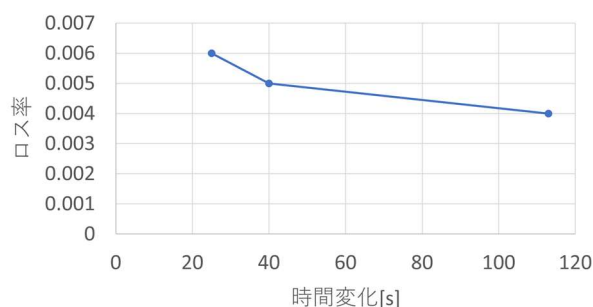


図 3 ロス率の時間変化

図 3 によると、処理開始とともにロス率は急激に下がり、0.006→0.005 に下がるまでに約 15 秒だったところ、ロス率が 0.005→0.004 に下がるまでには約 75 秒と大幅に長くなった。従ってロス率は 0.004 で収束つまり処理終了とみなすことにした。なお、ロス率が 0.005→0.004 に低下しても、NeRF 生成物に大きな差がないことは確認しており、定性的に見ても収束とみなしてよいと考えた。

4. 試験結果

4.1 処理速度

写真枚数を変えたときの、photogrammetry と NeRF それぞれの処理に要した時間を図 4 に示す。なお、photogrammetry の処理時間は MVS (Multi-view stereo) 処理時間とし、NeRF は NeRF 生成処理の時間そのものを計上した。なお、表 3 に示すように一部処理失敗している写真枚数があり、そのデータは欠損している。

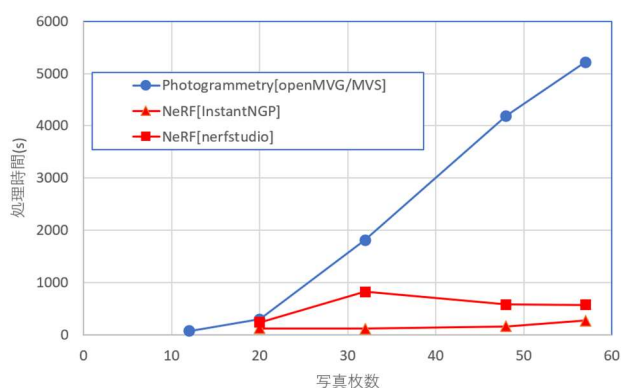


図 4 処理速度の比較

結果、photogrammetry は写真の枚数に比例して処理の時間が増えている。原理上、組み合わせ数に比例して処理時間が増えると予想していたが、今回の結果によると写真枚数に比例して処理時間が増える結果となった。一方、NeRF の場合、写真の枚数が増えても、その処理時間は微増にとどまった。よって、NeRF は写真の枚数に依らず短時間で処理できることが分かった。

4.2 再現度

次に、写真の枚数を変えたときの、現物と比較しての再現度について定性的に比較した。表 3 に結果を整理する。

表 3 再現度の比較

枚数	Photogrammetry		NeRF	
	openMVG/MVS Mesh	InstantNGP NeRF	Mesh	Nerfstudio Mesh
8				
12				
20				
32				
48				
57				

凡例: 左 後 右 欠けている 凸凹大きい 再現性高い

Photogrammetry は、写真枚数と再現度に明確な関係がある。枚数が少ない場合は、ごく一部の再現にとどまり、多くが欠損したメッシュが生成されている。枚数が多い場合

は、全体を再現できているのはもちろん、クオリティの高いメッシュが生成できている。一方、NeRF の場合、写真枚数が少なくてもある程度のクオリティの NeRF/メッシュが生成できている。一方で、写真の枚数が増えても photogrammetry ほどのクオリティ向上は見られない結果となっている。

また、適切な画角での撮影が困難だった側部および背面の再現度は、photogrammetry 及び NeRF とともに高くない。どちらの処理方法であっても、まずカメラの外部パラメータ (すなわち各写真の撮影位置/角度の推定) 算出プロセスを経ることになる。その処理ログを見てみると、側部および背面では採用されていない写真が複数あった。また、今回はその痕跡は見つけられなかったが、撮影位置推定を誤っている可能性もある。これは闇雲に写真の枚数を増やしても、アウトプットの品質向上に寄与できないことがあることを示す。カメラの外部パラメータ算出プロセスを理解し、それに適した撮影が品質向上につながると思われる。

4.3 精度 (平坦度・相似性)

次に精度比較を行った。比較方法は色々考えられるが、面の再現度に着目し、出力点群の平坦度を比較した。なお、instantNGP は点群の出力ができなかったため、NeRF は nerfstudio を使用した。また、写真枚数や処理方法によらず、再現性が高い部位を選択・抽出して比較した。これは、再現性が低い部位の場合、どこからどこまでの点群が、対象とする面であるか判定がつかず、比較そのものが困難になるためである。結果的に、最高の条件で撮影できた (=撮影に可能な限り手間暇をかけることができた) 部位の平坦度を比較した。

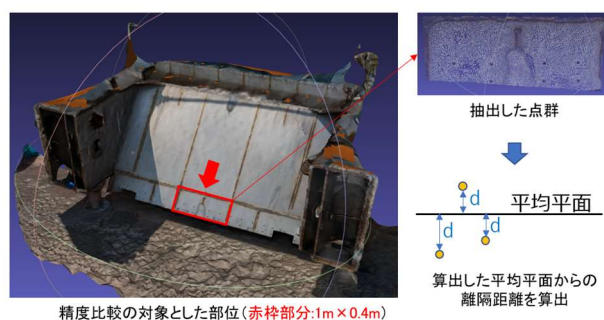


図 5 平坦度比較対象と算出方法

対象部位の点群を抽出し、各点群のスケールが同一となるよう拡大縮小した上で、平均平面を算出し、そこからの離隔距離を算出した。離隔距離の度数分布を整理したものを図 6 に示す。

全体的に見ると、32 枚の NeRF の場合を除き正規分布に近い度数分布となっている。photogrammetry と NeRF とを比較すると、離隔距離の分布範囲は、おおむね

photogrammetryの方が10倍程度良好となっている。すなわち、離隔距離としてはphotogrammetryの方が約10倍優れているといえる。

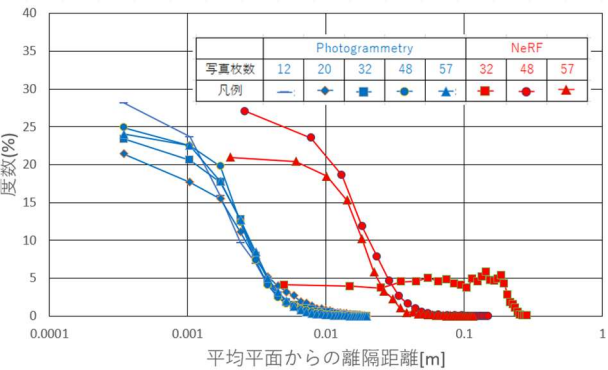
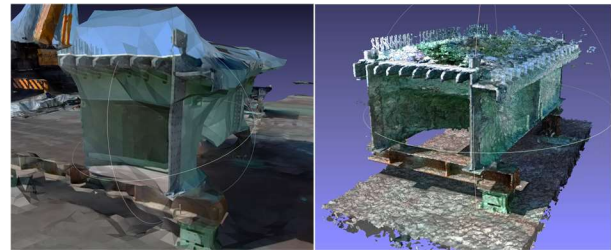


図6 平坦度



Photogrammetry [openMVG/MVS] NeRF [InstantNGP]
図7 360° 動画からのメッシュ生成例

次に、低解像度写真を使った場合の再現度比較を紹介する。図7は、360°カメラによって撮影した動画から、合計185枚の写真を切り出してphotogrammetryとNeRF処理し、出力したメッシュを比較したものである。動画から切り出したことで、写真の解像度は低い(1024x1024)上、若干ボケた写真になっているため、低品質写真での比較に適したケースと考える。このように、低品質写真の場合は、photogrammetryでは欠損の大きいメッシュになることがあるが、NeRFでは細部はともかく全体の再現には優れていることがよくわかる。

5. まとめ

- 1)写真品質が低い場合(撮影枚数が少ない30~40枚以下の場合や解像度がフルHD以下などの、photogrammetry/NeRF処理に適さない写真の場合)は、NeRFの方が良いアウトプットが得られる。
- 2)写真品質が高い場合は、photogrammetryの方が良いアウトプットが得られる。このときの生成点群の精度は、NeRFの10倍程度良好になる場合がある。
- 3)写真の枚数とphotogrammetry処理の時間は比例する。
以上の結果から、
- 4)撮影にリソースを割り当てる余裕に応じて、動画撮影、

静止画撮影、点群計測の順で検討する
5)処理時間に割り当てる余裕に応じて、NeRF、SfMの順で検討する
ことが良いと考える。

以上の結果をもとに、鋼製土木構造物の一つである橋梁実工事において、現地での制約条件を前提として計測方法をどのように選んでいくべきかを検討してまとめた(表4)。橋梁実工事における最初のステップ、つまり現地踏査における計測箇所の選定は、計測者の経験に大きく依存していた。とりわけ保全工事においては、工事ごとに異なる特性があり、設計が進行するにつれて必要な情報が増加する傾向がある。そこで、photogrammetryやNeRFなどの技術を活用することで、工事対象箇所全体の情報を包括的に取得でき、情報漏れを極小化できる。これにより、設計業務が進行する過程で予想外に必要な情報を、現場の制約に左右されずにタイムリーに取得でき、設計業務のフロントローディングに貢献できると考える。

表4 橋梁実工事での調査方法適用案

調査カテゴリ	調査詳細	撮影素材	撮影機材	処理方法
現地踏査	受注直後実施。ごく短時間で現場確認。	動画	360°カメラ	NeRF
簡易計測	中程度のコスト/時間をかけて、中程度の精度で実施する計測。	静止画	ドローン+高解像度カメラ	SfM
		静止画	足場+高解像度カメラ	SfM
詳細計測	コスト・時間を投入して、高精度で実施する計測。	点群	足場+レーザスキャナ	—

6. 展望

鋼構造物建設現場へ適用するに際し、いわゆるBIM/CIM活用として考えると、現時点でのNeRF生成物としてのメッシュは、精度として不十分だと考えている。今後の技術開発に期待したい。一方で、NeRFの優れた特性を活かしたこれまでにない活用方法があるはずである。今後は必ずしもメッシュが必要でなくなることも考えうる。3DモデルありきのBIMCIMではなく、NeRFそのものを効果的に活用することによる、より効率的な設計・建設・維持管理のあり方について検討・提案をしていきたい。

謝辞

本研究は、大阪大学 福田知弘准教授の助言を受けて実施したもので、厚く御礼申し上げます。

【参考文献】

1) Martin-Brualla, Ricardo, et al.(2020), "NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections." arXiv preprint arXiv:2008.02268.
2) Thomas Muller, Alex Evans, Christoph Schied, Alexander Keller(2022), "Instant Neural Graphics Primitives with a Multiresolution Hash Encoding" arXiv:2201.05989
3) Matthew Tancik, et al. (2023), "Nerfstudio: A Modular Framework for Neural Radiance Field Development" arXiv:2302.04264