

## 動く物体の 3 次元形状を高精度に計測する 「ニューラルインバースレンダリング手法」を開発 ー デジタルツインや映像制作などへの応用に期待 ー

### 【ポイント】

- わずか 3 枚の投影パターンで、動く物体の 3 次元形状を高精度・高解像度で再構成する手法を開発
- ニューラルインバースレンダリングで対象の動きと 3 次元形状を同時に最適化
- ダイナミックに変化するシーンの 3 次元形状が必要とされる、多様な産業・研究分野への応用に期待

### 【概要】

東京科学大学 (Science Tokyo) 工学院 情報通信系の浦川雄気大学院生 (研究当時)、渡辺義浩准教授の研究チームは、運動する物体を高精度に 3 次元計測できる新しいニューラルインバースレンダリング (用語 1) 技術を提案しました。

パターン投影を用いた 3 次元計測では、**構造化光法** (用語 2) と呼ばれる方式が広く使われており、その代表的手法である**位相シフト法** (用語 3) では高精度な計測が可能です。しかしこの手法では複数のパターンを順に投影・撮像する必要があるため、対象が動くとカメラとプロジェクタの画素の対応関係が崩れ、正確な形状を得られないという課題があります。このため、従来の位相シフト法は静止対象にのみ用いられていました。

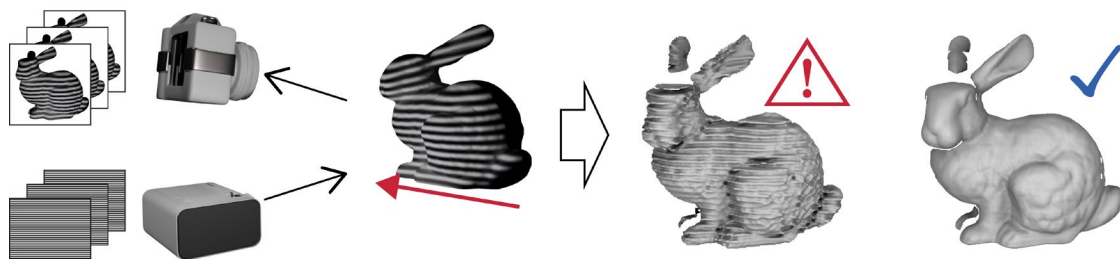
そこで本研究では、物体の 3 次元形状や、その動きを正確にモデル化する**変位場** (用語 4) などのシーン情報をニューラルネットワークで表現したうえで、インバースレンダリングの枠組みで最適化する手法を提案しました。この手法では、物体の動きと形状を同時に推定することで、投影・撮像画像における対象の運動によるずれを補正しながら 3 次元形状を再構成します。さらにこの最適化を、1 台のプロジェクタと 2 台のカメラによる複数視点構成の下で実現する枠組みを構築しました。これにより、従来の位相シフト法でも用いられていた標準的な正弦波パターンをわずか 3 枚投影するだけで、動く物体の 3 次元形状を高精度かつ高解像度に再構成することに成功しました。

本成果は、10 月 23 日付 (ハワイ標準時) の International Conference on Computer Vision (ICCV) 2025 で発表されました。

## ●背景

3次元計測技術は、製造現場での品質検査や、文化財保存のためのデジタルアーカイブ、現実空間を仮想空間上に再現するデジタルツイン、映像制作分野における身体や表情、衣服の動きまでを記録するパフォーマンスキャプチャなど、幅広い分野で活用されています。

このような3次元計測技術の一つに、プロジェクタとカメラで構成されたシステムを用い、パターンを投影することで3次元形状を取得する「構造化光法」があります。なかでも、少ない投影枚数で静止物体の高精度・高解像度な形状取得が可能な手法として、正弦波パターンを用いる「位相シフト法」が広く利用されています。しかし位相シフト法では複数の正弦波パターンを順に投影・撮像するため、計測中は対象が静止していることが前提となっています。そのため対象が動くと、各パターンを撮像するたびに物体の位置が変わり、プロジェクタとカメラの画素の対応関係がずれてしまうため、従来の



形状復元手法では正確な形状を再現できないという問題がありました（図1）。

図1 従来の複数パターン投影による3次元計測（左）では、運動物体に対して形状誤差が発生していた（中央）。この問題を解決するため、高精度かつ高解像度な形状再構成を実現する手法を提案した（右）。

## ●研究成果

本研究では、ニューラルネットワークの高い表現力を活用したインバースレンダリングの枠組みを活用するとともに、1台のプロジェクタと2台のカメラによる3視点の情報を組み合わせることで、運動物体の複雑な3次元形状を高精度に再構成する手法を提案しました。本手法では、シンプルな正弦波パターンをわずか数枚投影するだけで、高精度な形状再構成を実現しています。

提案したニューラルインバースレンダリングは、物体の3次元形状、動きを表す変位場、物体表面の反射率、残差成分（投影パターン以外の光や投影パターンが2次反射することで生じる）という四つのシーン情報を同時に最適化するように設計されています（図2左）。特に、変位場を用いて、3次元空間上の位置や法線を複数のフレーム間で整合させることで、動きによって発生する形状の誤差を効果的に補正することが可能になっています。

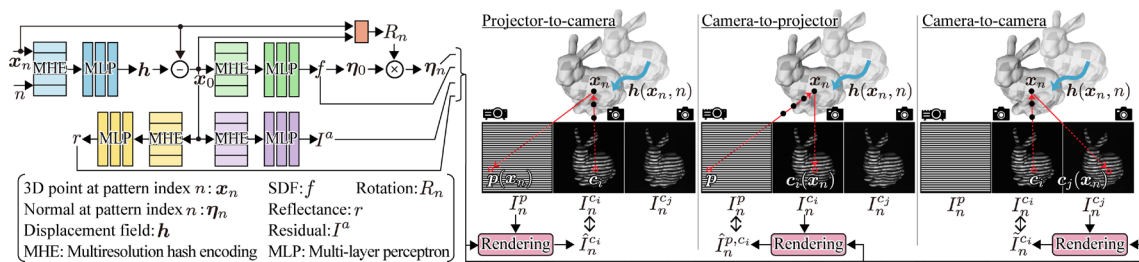


図2 動く物体の3次元計測を実現する本研究の提案手法の概要。(左)変位場に基づく動きのモデルを組み込んだネットワーク全体の構成。(右)ネットワークの出力結果を用いて、1台のプロジェクタと複数のカメラからなる構成で画像を生成する様子。生成された画像と実際に撮影・投影された画像との誤差を最小化することで、シーンの各パラメータを最適化する。

さらに最適化の過程では、実際に撮像・投影された画像と、推定された四つのシーン情報から生成される画像の誤差が最小化されるように、ニューラルネットワークを学習させます。このとき、1台のプロジェクタと2台のカメラによる3視点構成を活かして、1) プロジェクタの投影画像からカメラの撮像画像を生成、2) カメラの撮像画像からプロジェクタの投影画像を生成、3) 1台のカメラの撮像画像から別視点のカメラの画像を生成、という3通りの画像生成を行ったうえで、生成画像をそれぞれ対応する実際の画像と比較します(図2右)。これによって、3次元形状や変位場などの再構成精度を向上させるとともに、投影パターンの枚数が少なくても安定して再構成ができる手法となっています。

実験では、投影パターンの枚数や、パターンの空間周波数、物体の速度などを変化させながら、提案手法の評価を行いました。その結果、物体が動いている場合でも、わずか3枚の標準的な正弦波パターンで、平均誤差0.23 mm程度の高精度な形状再構成を達成し、静止対象に対する計測と同程度の精度を維持できることを確認しました。また従来手法と比較しても、動的シーンでの再構成誤差を低減できることを実証しました(図3)。さらに、剛体や非剛体を問わず、さまざまな形状や動きに対して本手法が形状を正確に再構成できることを実証しました(図4)。

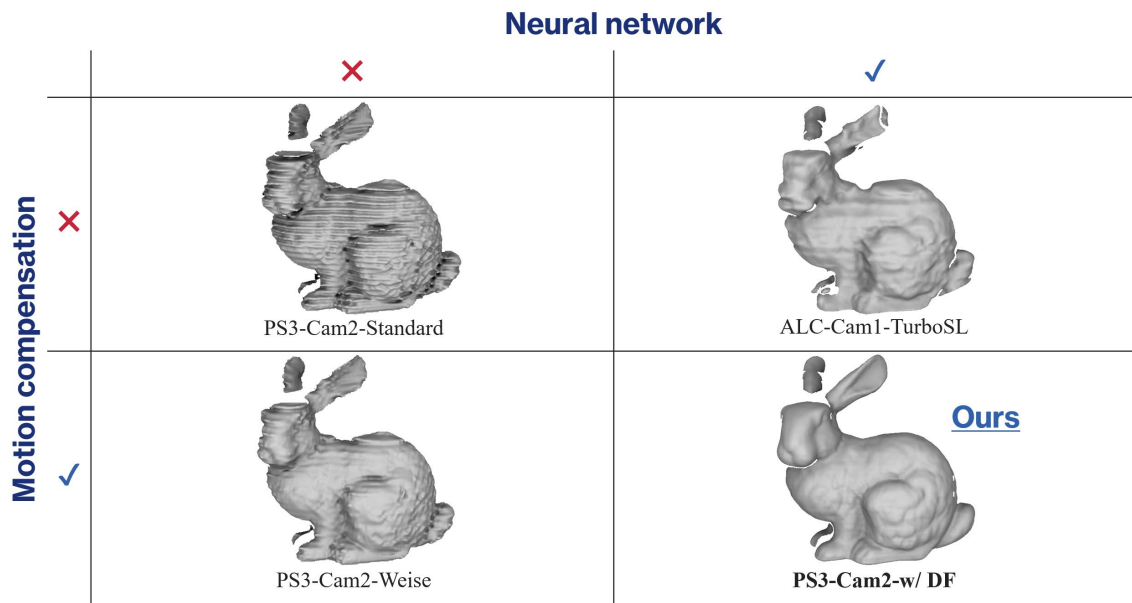


図3 提案手法と従来手法との比較結果。右下が提案手法による再構成結果、それ以外は従来手法による形状再構成を示す。手法は、モーション補償およびニューラルネットワークの利用有無に基づいて分類した。従来手法では、周期的な縞状のアーティファクトや形状の歪みが見られるが、提案手法では、わずかな枚数のシンプルな位相シフトパターンのみを用いた場合でも、高い精度と安定性を実現している。

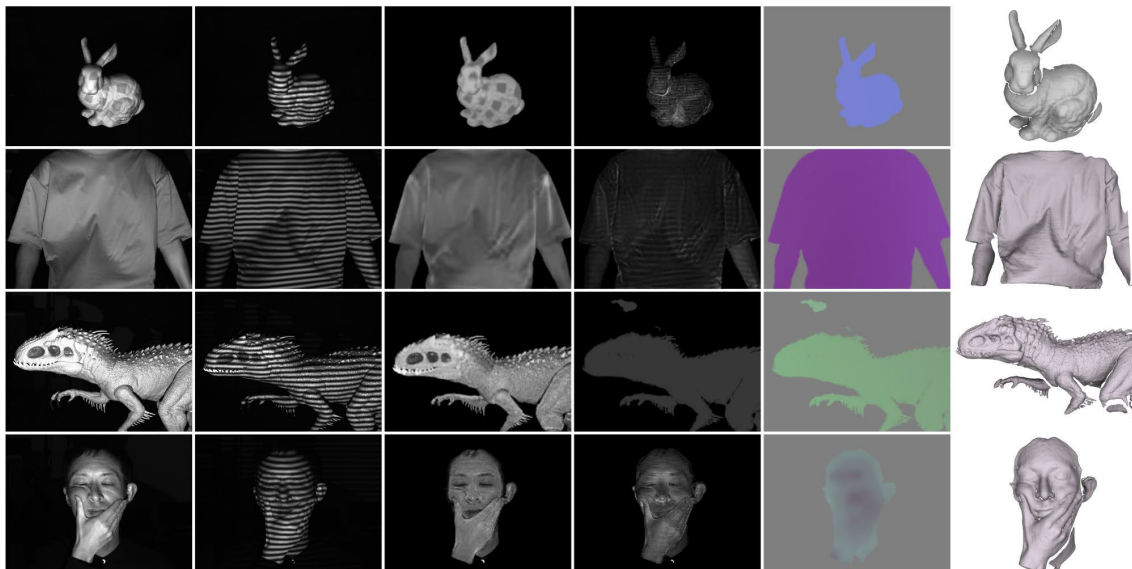


図4 提案手法の評価実験。左列から順に、通常照明下での対象の写真(入力として使用しない)、1台目のカメラの1枚目の画像、ネットワークによって再構成された反射率、残差成分、変位場(x, y, z方向の運動をRGBの各色で表現)、3次元形状。(各対象が運動する結果を動画にしたものを下記のYouTube動画で公開)



動画 再構成した運動物体の3次元形状の動画像

[Neural Inverse Rendering for High-Accuracy 3D Measurement of Moving Objects](#) | Watanabe Laboratory, YouTube

### ●社会的インパクト

3次元計測は古くから研究・開発が進んできた技術で、すでに幅広い応用で利用されています。本研究で提案した技術は、複数枚の投影パターンを用いる構造化光法では避けられなかった動体計測時の誤差を克服し、動く物体の3次元形状を高精度かつ高解像度に再構成できる点で画期的です。ダイナミックに変化するシーンの3次元形状を必要とする、製造業や文化財保存、デジタルツイン、映像制作、エンターテインメント、スポーツ分析、ロボットビジョン、拡張現実などの応用において新たな展開を生み出すと期待されます。

### ●今後の展開

今後は、再構成される3次元形状の精度向上や処理の高速化など、性能面の強化を目指します。また本研究では、同時に最適化される情報のうち、3次元形状を主な成果として注目してきましたが、副次的に得られる変位場からの動き情報も、応用の場面で活用できる重要な情報として期待できます。そのため、変位場のさらなる精度向上にも取り組む予定です。さらに、製造、分析、エンターテインメントなどの分野において、本技術が切り拓く新たな応用を具現化していく予定です。

## ●付記

本研究は科学技術振興機構（JST）未来社会創造事業（JPMJMI24H1）の支援を受けて行われました。

## 【用語説明】

- (1) **インバースレンダリング**：撮像された画像から、その撮像シーンを構成する幾何情報（形状）や反射特性などを逆推定する技術。特に、ニューラルネットワークを用いて推定する手法をニューラルインバースレンダリングと呼ぶ。
- (2) **構造化光法**：プロジェクタからパターン光を物体に投影し、その反射パターンをカメラで撮像することで、物体の形状を三角測量により求める3次元計測技術。1枚のパターンを利用するタイプと、複数枚のパターンを利用するタイプがある。
- (3) **位相シフト法**：複数枚の正弦波パターンを順に投影し、得られた画像の位相情報から形状を算出する手法。複数枚のパターンを利用する構造化光法の代表的な方式。
- (4) **変位場**：画像上の各画素が観測した位置に対応する3次元空間での動きを表すベクトル場情報。物体表面の時間的変化を表現するために用いられる。

## 【学会情報】

学会名：*International Conference on Computer Vision 2025*

タイトル：Neural Inverse Rendering for High-Accuracy 3D Measurement of Moving Objects with Fewer Phase-Shifting Patterns

発表者：Yuki Urakawa, Yoshihiro Watanabe

## 【研究者プロフィール】

渡辺 義浩（ワタナベ ヨシヒロ） Yoshihiro WATANABE

東京科学大学 工学院 情報通信系 准教授

研究分野：拡張現実、コンピュータビジョン、デジタルアーカイブ、インタラクション

## 【お問い合わせ先】

（研究に関すること）

東京科学大学 工学院 情報通信系 准教授

渡辺 義浩

Email：watanabe@ict.eng.isct.ac.jp

TEL：045-924-5474

(報道取材申し込み先)

東京科学大学 総務企画部 広報課

取材申し込みページ： <https://www.isct.ac.jp/ja/001/media>

Email： [media@adm.isct.ac.jp](mailto:media@adm.isct.ac.jp)

TEL：03-5734-2975 FAX：03-5734-3661