

法線ベクトル推定AIの開発

Development of Artificial Intelligence for Normal Vector Estimation

福留祐太

Yuta Fukudome

電子・有機素材研究所 電子システムグループ

製造業における深刻な人手不足への対応として、外観検査の自動化が強く求められている。本研究では、照明条件の異なる複数画像から物体表面の法線ベクトルを推定するAIモデルを開発した。既存手法であるPS-FCNを基盤としつつ、損失関数にMAE（平均角度誤差）を採用することで、推定精度と学習安定性の向上を図った。その結果、合成データによる学習と公開ベンチマークデータセットを用いた評価を通じて、開発したAIモデルが従来法を上回る精度と高速な推論性能を有することを確認した。

1. はじめに

鳥取県では、人口減少に伴う人手不足が深刻化しており、県内製造業において、生産性の向上を目的とした工程の自動化が強く求められている。中でも、外観検査は依然として多くの現場で熟練作業者の目視に依存しているのが現状であり、作業者の習熟度や主観的判断のばらつきが製品品質の安定性を損なう要因となっている。

外観検査の自動化において、対象物の凹凸や曲面の微細な変化を捉えるには、表面の法線ベクトル情報が非常に有効である。従来は、異なる照明条件下で撮影された画像を複数（通常3枚以上）用いて法線ベクトルを推定するフォトメトリックステレオ法^{1,2)}が活用されてきた（図1）。この手法は、物体表面がランバート反射に従うことを前提とした理論モデルに基づいており、比較的高い精度で法線ベクトルを算出可能である。しかし、金属など鏡面反射の強い材質や複雑な照明環境では精度が低下しやすく、誤差が増大する傾向がある。

このような精度の低下を抑えるためには、光源の数を増やすことで情報の冗長性を確保し、反射特性のばらつきを平均化する方法が用いられる。しかし、実際の生産現場では、照明装置の設置スペースやコ

ストの制約から、多数の光源を使用することは現実的ではない。

そこで本研究では、従来法の複雑な演算処理や多光源環境を不要とすることで、処理時間の短縮とシステム構成の簡素化を実現することを目的とし、照明条件の異なる複数画像から物体表面の法線ベクトルを高精度かつ高速に推定可能なAIモデルを開発した。

$$I_1 = \rho n \cdot l_1$$

$$I_2 = \rho n \cdot l_2$$

$$I_3 = \rho n \cdot l_3$$

I_i : 点pにおける光源*i*の明るさ
 ρ : 反射率
 n : 法線ベクトル
 l : 光源ベクトル

法線ベクトル*n*と反射率 ρ が定まる

図1 フォトメトリックステレオ法（光源数 N=3）

2. 実験方法

2.1 AI 学習用データセット

本研究では、既存のフォトメトリックステレオ推定ネットワークであるPS-FCN³⁾で採用された高精度な法線ベクトルを含む合成データセットを用いて、AIモデルの学習を行った。具体的には、3次元形状

に Blobby Shape Dataset⁴⁾と Sculpture Shape Dataset⁵⁾を使用し、表面の反射特性には MERL BRDF database⁶⁾を適用した。これらを組み合わせ、様々な光源方向下でレンダリングを行うことで、現実的な非ランバート表面を模擬した画像群と真値法線データを生成した(図2)。これにより、多様な照明条件や材質に対応したAI学習用データが構築できた。

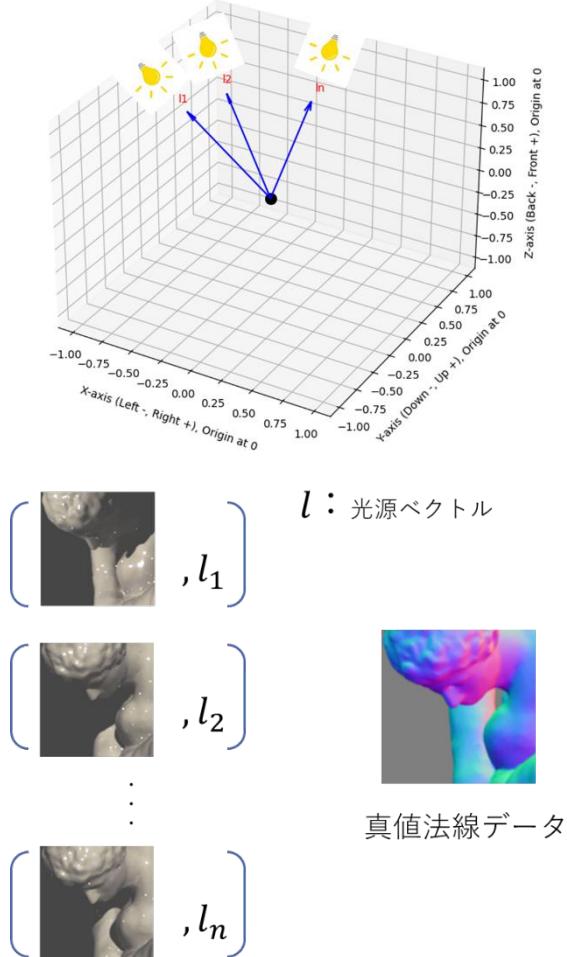


図2 AI学習用データセット

2.2 AI評価用データセット

学習済みモデルの評価には、公開ベンチマークとして広く用いられている DiLiGenT Benchmark Dataset⁷⁾を使用した(図3)。このデータセットは10種類の実在の物体に対して、複数の照明条件下で撮影された画像と、それに対応する正確な法線を提供しており、モデルの汎化性能を検証するのに適している。なお、評価は式(1)で示される MAE(平均角度誤差)を指標として行った。

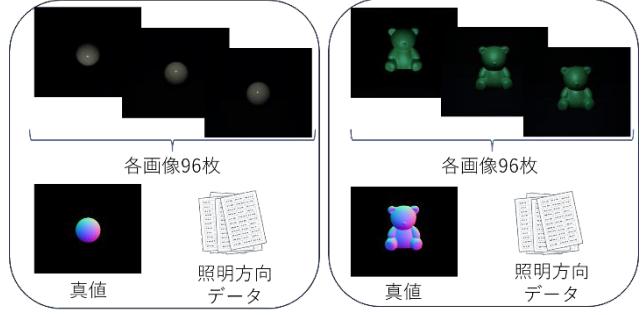


図3 AI評価用データセット

$$MAE = \frac{1}{n} \sum_{p=1}^n \cos^{-1} (\vec{n}_{tp} \cdot \vec{n}_{ep}) \quad \cdots (1)$$

n : 画素数

\vec{n}_{tp} : 画素 p における単位法線ベクトルの真値

\vec{n}_{ep} : 画素 p における単位法線ベクトルの推定値

2.3 AIモデルの設計と学習

本研究では、PS-FCNをベースに、損失関数を改良することで法線ベクトル推定の精度向上を図った。PS-FCNは、共有重みの特徴抽出器、複数画像からの特徴統合を行う結合層、各画素に対応する法線ベクトル(法線マップ)を出力するデコーダから構成される。

当初は、表現力の強化を目的に、特徴抽出器の構造に変更を加えたが、学習が不安定となり、評価指標である MAE が既存構造と比較して改善されなかった。そのため本研究では、PS-FCN のネットワーク構造を維持しつつ、損失関数を見直す方針とした(図4)。

損失関数には、従来の L2 ノルム、コサイン類似度ではなく、MAE を採用した。一般的な L2 ノルムによる損失関数は、ベクトルの大きさの違いに敏感であり、単位ベクトルである法線推定では方向の違いを正確に評価しにくい。一方、コサイン類似度は方向の近さに着目するものの、最大値が 1(完全一致の場合)に固定されるため、学習勾配が小さくなりやすく、学習が停滞する傾向がある。これに対して MAE は、損失が角度の大きさそのもので表され、誤差の大きい方向に対して適切に学習を促すことができると考えた。

学習には PyTorch を用い、最適化手法として Adam を採用した。学習時には、入力画像の枚数や照明順序のランダム化などのデータ拡張を行い、ネットワークの汎化性能を高めた。

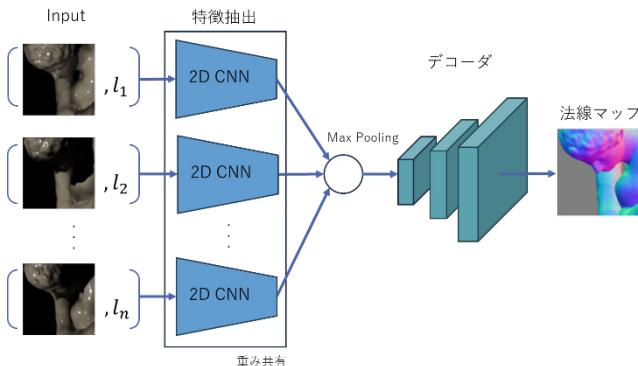


図 4 開発した AI のネットワークアーキテクチャ

3. 結果と考察

本研究で開発した AI モデルを用いて、DiLiGenT Benchmark Dataset に対して法線ベクトル推定を行い、その結果として得られた 10 種類の法線マップを図 5 に示す。画像からは、物体の形状に沿った法線方向が正確に再現されており、提案手法が有効に機能していることが視覚的に確認できる。



図 5 推定した法線マップ

3.1 処理速度の評価

処理速度について DiLiGenT Benchmark Dataset を対象に評価を行った。結果を表 1 に示す。従来のフォトメトリックステレオ法では、法線ベクトルの算出に 60 s 以上を要していたのに対し、本研究で開発した AI モデルでは、10 種類の画像に対して、モデル読み込みを含めても平均 3 s 以内、読み込みを除いた純粋な推論処理は 1 s 未満で完了することが確認された。これは従来手法に比べて 10 倍以上の高速化に相当し、実用化に向けた大きな進展である。また、評価に使用した PC の構成を表 3 に示す。

表 3 評価に用いた PC の構成

項目	性能
OS	Windows11 Pro 64bit
GPU	NVIDIA GeForce RTX 4090 / 24 GB
CPU	AMD Ryzen 9 9950X 16-Core Processor 4.30 GHz
メモリ	128 GB

3.2 精度の評価

開発した AI モデルの精度を、DiLiGenT Benchmark Dataset を用いて評価した結果を表 2 に示す。全 10 カテゴリのうち、「harvest」と「pot2」を除く 8 カテゴリにおいて、先行研究である PS-FCN の結果を上回る MAE (平均角度誤差) を達成した。

特に「ball」カテゴリでは、PS-FCN の 2.82 に対して本モデルは 2.39 と大きく誤差を下回り、高い精度が確認された。「bear」「cat」「cow」などのカテゴリでも一貫して良好な推定結果が得られ、従来手法に対する優位性が示された。

表 1 処理速度の評価結果

ラベル	ball	bear	buddha	cat	cow	goblet	harvest	pot1	pot2	Avg. (初回除く)
時間(s)	24.514	0.483	0.468	0.496	0.501	0.515	0.47	0.476	0.452	2.879

表 2 精度の評価結果

ラベル	ball	bear	buddha	cat	cow	goblet	harvest	pot1	pot2	reading	Avg.
誤差 (提案手法)	2.39	5.60	7.17	5.53	6.61	8.26	16.22	7.44	7.68	13.19	8.01
誤差 (PS-FCN)	2.82	7.55	7.91	6.16	7.33	8.6	15.85	7.13	7.25	13.33	8.39

一方で、「harvest」および「pot2」のカテゴリでは、PS-FCN と比べて MAE が若干大きくなる結果となり、形状や材質に依存した課題が残ることが分かった。全体の平均 MAE は、PS-FCN の 8.39 に対し本モデルは 8.01 であり、総合的に見て精度の向上が確認された。

3.3 今後の課題

今後の課題として、実物体を用いた精度検証の実施が挙げられる。特に、フォトメトリックステレオ法によって取得された高精度な法線ベクトルを真値として用い、本モデルによる推定結果との比較を通じて、定量的な精度評価を進めていく予定である。また、より複雑かつ多様な形状を対象とした評価を行うことで、モデルの形状依存性や一般化性能の限界を明らかにし、信頼性の高い推定手法の確立を目指す。

4. おわりに

本研究では、従来のフォトメトリックステレオ法に代わり、照明下で撮影された画像から直接法線ベクトルを推定する AI モデルを開発した。PS-FCN をベースに、損失関数として MAE（平均角度誤差）を新たに導入することで、推定精度の向上と学習の安定性を図った。

DiLiGenT Benchmark Dataset による評価の結果、全 10 カテゴリ中 8 カテゴリにおいて、先行モデル (PS-FCN) を上回る精度を達成した。また、処理時間についても従来法と比べて大幅な高速化が確認され、実用的な応答速度を有していることが明らかとなった。

今後は、フォトメトリックステレオ法による真値を用いた実物体での精度検証や、形状・照明条件の多様化を含むさらなる検証を進め、製造現場への応用を見据えた実用化に向けて取り組んでいく。

文 献

- 1) Woodham, R. J., Photometric method for determining surface orientation from multiple images, Optical Engineering, Vol.19, No.1 1, 1980, pp.139 139-144.
- 2) Silver, W. M., Determining shape and reactance using multiple images, Master thesis, Massachusetts Institute of Technology, 1980.
- 3) Guanying Chen, Kai Han, Kwan-Yee K. Wong, PS-FCN: A Flexible Learning Framework for Photometric Stereo, ECCV 2018.
- 4) Johnson, M.K., Adelson, E.H.: Shape estimation in natural illumination, CVPR 2011.
- 5) Wiles, O., Zisserman, A., Silnet: Single-and multi-view reconstruction by learning from silhouettes, BMVC, 2017.
- 6) Matusik, W., Pfister, H., Brand, M., McMillan, L.: A data-driven reflectance model, SIGGRAPH, 2003.
- 7) Shi, B., Wu, Z., Mo, Z., Duan, D., Yeung, S.K., Tan, P.: A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo, IEEE TPAMI, 2018.