

オンタイム生成結果提示による三次元再構成撮影支援手法

森越 彩楓¹ 五十嵐 健夫¹ 樋口 啓太¹

概要：三次元再構成において、入力画像は結果の品質に大きな影響を持つため撮影の支援を行う研究が多数行われてきた。従来の研究では、撮影対象の物体の位置や大きさに基づいて、リアルタイムに撮影位置へのガイドが行われてきた。この手法ではユーザはガイドに従えば誰でも撮影できるという利点はあるが、照明環境や対象物体により細部がことなることから、撮影画像により高品質な再構成結果を得ることが保証されない。そこで本研究では、対象物体の撮影中に再構成結果を提示することにより、撮影者は撮影が足りていない領域を理解し、より効果的な撮影を行うことができるという仮説に基づきオンタイム提示システムを開発した。本システムは、撮影した画像に対して自己位置推定を適用し、Instant NGPにより再構成をすることによりオンタイム提示を実現した。撮影者は任意のタイミングでこれまでの写真を基にした再構成結果を確認することができ、クオリティを上げるためにどこを撮影すればよいかを知ることができる。

1. はじめに

複数の2次元画像から対象物の3次元形状を復元する、フォトグラメトリやNeural Radiance Fields (NeRF)のような三次元再構成手法において、復元結果の品質は入力画像に依存する。そのため撮影者は近接画像間で特徴を一致させることや、対象物を全方位から撮影することなど、撮影中には確認が難しいテクニックを要求される。

このような背景から、三次元再構成のための撮影を支援するシステムが提案されている。例えばAndersenらは、AR HMDに内蔵されたデプスセンサで物体の大まかな形状を取得し、形状から計算される撮影ガイドをARで表示するシステムを提案した[1]。Langguthらは対象物全体をカバーした撮影が行えるよう、次に撮影すべき視点を撮影画像から得られる疎な再構成結果をもとに計算し、ユーザを誘導するシステムを提案した[2]。これらのシステムでは物体の大まかな形状をリアルタイムに計算する一方で、高品質な再構成結果は確認できない。これは、多くの三次元再構成手法では、高品質な再構成結果を得るのに、情報の複雑さや解像度によっては数時間以上かかるためである。一方で高品質な再構成結果を撮影中に確認できることで、再構成に失敗している領域や、文字や柄のような細かい表現の品質を確認することができる。

本研究では数秒程度の処理で三次元再構成結果を得られるInstant Neural Graphics Primitives (NGP)[3]を用いて以下のような手順で撮影を行う手法を提案する。

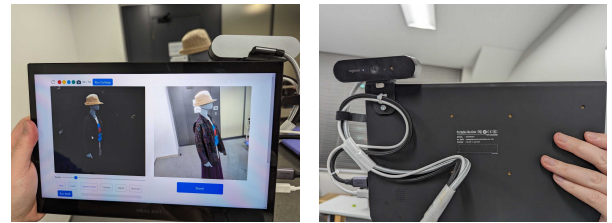


図 1: 提案システムの外観

- (1) ユーザが新しい画像(群)を撮影する。
- (2) 撮影画像に自己位置推定を適用する。
- (3) 自己位置推定が成功した画像を用いてInstant NGPで三次元再構成を行う。
- (4) 再構成結果を提示し、結果を基にユーザが新たな画像(群)を撮影する。

本稿では、この手順を実装したシステムについて紹介する。

2. 撮影支援システム

我々が実装した撮影支援システムを図1に示す。システムは撮影機器としてウェブカメラを使用し、またユーザはPCに繋がれたタッチディスプレイ上でUIの操作を行う。まず初めにユーザは対象物全体を20~30枚程度撮影し、それらを用いて再構成を行い、ベースとなるモデルを得る。続いて、再構成結果を確認しながら撮影すべき領域を特定し、撮影する。その後、更新された再構成結果をもとに次に撮影すべき領域を探す。ユーザはこの反復を繰り返すことで、より高品質な再構成結果を取得できる。本節では、三次元再構成のための処理と、ユーザインタフェースにつ

¹ Preferred Networks, 東京

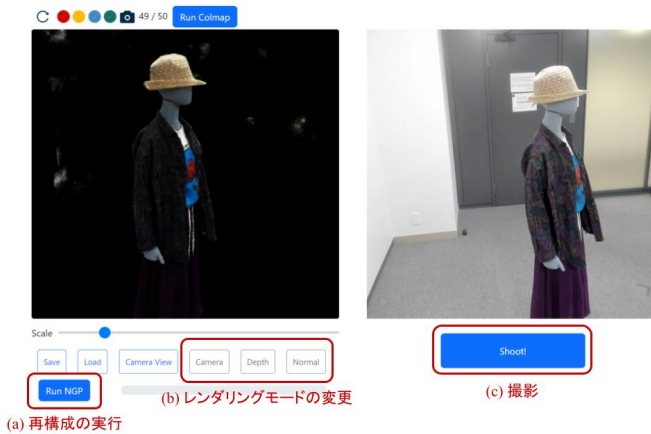


図 2: ユーザインタフェース

いて説明する.

2.1 オンタイム三次元再構成

Instant NGP への入力は、画像とカメラの位置・姿勢のセットである。そのため、オンタイムに三次元再構成を行うには高速な自己位置推定が必要である。本研究では Structure from Motion (SfM) 用のソフトウェアである COLMAP[4] を用いて、漸増的にカメラの位置・姿勢の推定を行うことでこれを実現した。Instant NGP と COLMAP を使用することで、撮影済みの画像が 100 枚程度であれば、画像が 1 枚追加されてから 1 分程度で再構成結果が更新される。

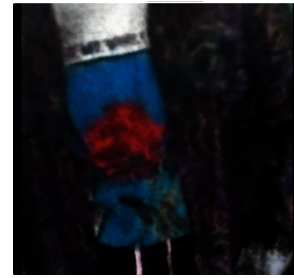
2.2 ユーザインタフェース

本システムは再構成結果の確認画面と、ウェブカメラのビュー画面の 2 画面を 1 つのページに表示する (図 2)。再構成結果の確認画面では、再構成結果を画面のタッチで自由に動かすことが可能である。さらに再構成の実行を制御する機能 (図 2(a)) や再構成結果のレンダリングモードを変更する機能 (図 2(b)) が実装されている。ウェブカメラのビュー画面には、PC に繋がれたウェブカメラの映像がリアルタイムに表示され、ユーザは対象の撮影を行うことができる (図 2(c))。

本システムはブラウザ上で動作し、撮影から再構成結果の確認までを一貫して行うことができる。まずユーザはウェブカメラのビュー画面を用いて対象物を撮影する。続いて、再構成結果の確認画面をタッチ操作で視点を変えながら再構成結果を眺めて、撮影が足りない領域を理解する。ユーザは、満足のいく再構成結果を得るまでこの反復を繰り返す。

3. まとめと今後の課題

本研究では、撮影中に密な再構成結果を表示することで、



(a) 初期画像のみでの再構成結果.



(b) システムありの結果. (c) システムなしの結果.

図 3: 再構成結果の比較

ユーザが撮影・再構成結果の確認を反復的に行うことを支援するシステムを実現した。ユーザは本システムを用いることで、画像が足りていない領域を理解し、より有効な撮影を行うことができるようになる。

図 3 は、本システムを使用して撮影を行った場合と、システムを用いずに撮影を行った場合での再構成結果である。初期画像として 30 枚の画像が登録された状態で撮影を開始し、それぞれの手法で対象物を追加で 20 枚ずつ撮影した。システムを用いることで、服の文字や柄などの細かい部分がより鮮明に復元できていることが確認できる。今後は、再構成結果の品質やユーザ体験に関するユーザ実験や、システムの改善を行なっていく予定である。

謝辞 株式会社 Preferred Networks の橋本泰輔さん、加藤大晴さん、田中悠輔さん、野口敦裕さん、高木士さんの支援に感謝いたします。

参考文献

- [1] Daniel Andersen, Peter Villano, and Voicu Popescu. Ar hmd guidance for controlled hand-held 3d acquisition. *IEEE transactions on visualization and computer graphics*, Vol. 25, No. 11, pp. 3073–3082, 2019.
- [2] Fabian Langguth and Michael Goesele. Guided capturing of multi-view stereo datasets. In *Eurographics (Short Papers)*, pp. 93–96, 2013.
- [3] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.*, Vol. 41, No. 4, pp. 102:1–102:15, July 2022.
- [4] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.