

ドローンを用いた自由視点映像の生成に関する研究

日大生産工(院)

○市木 亮介

日大生産工

矢澤 翔大

日大生産工

新妻 清純

日大生産工

黒岩 孝

1. はじめに

地理情報システム (GIS:Geographic Information System) や拡張現実 (AR:Augmented Reality)等の運用では、実在する様々な構造物の寸法や配置等、膨大な3次元情報の取得が必要となる。その解決策として、例えば深度情報を非接触で測定できるLiDAR (Light Detection and Ranging)等の使用が考えられるが、一般に高価であり、測定にPC等が必要となるため簡易な運用は難しい。一方、近年、単眼カメラを用いて撮影した複数のデジタル画像から物体の3次元構造を推定する手法が活発に研究されており、LiDARに比べ手軽な運用が期待されるが、対象とする測定物の周囲をくまなく撮影する必要があり、作業は煩雑である。ドローンであれば、測定対象の中心にカメラを向けながら周囲を巡回するNose in Circleという飛行法で測定対象の周囲を容易に撮影できると思われるが、手動で正確に操縦するのは難しい。そこで著者らは、ドローンを正八角形の軌道で自律飛行することにより測定対象の周囲を撮影した動画像に対し、SfM(Structure from Motion)²⁾を適用することで、3次元形状を復元出来るか検討を行った³⁾。その結果、測定対象の形状は比較的良く再現するものの、より再現性を改善するには多くのデータを必要とし、処理に時間がかかるという問題点がわかった。そこで本研究では、機械学習を取り入れたアルゴリズムを導入することで、より高精度に物体の3次元構造を推定できるか検討するとともに、自律飛行の軌道をより簡略化した場合に3次元形状の推定が可能かについても検討を行う。

2. 実験方法

動画像の撮影には、簡易なプログラミングで自動飛行できる小型のドローン(Ryze Tech.社 Tello)を使用する。ドローンの視野角(FOV)は82.6[deg]で、35mmフルサイズのカメラに換算して24mmレンズ程度の範囲を撮影でき、720p・30[fps]の動画を取得できる。図1は、撮

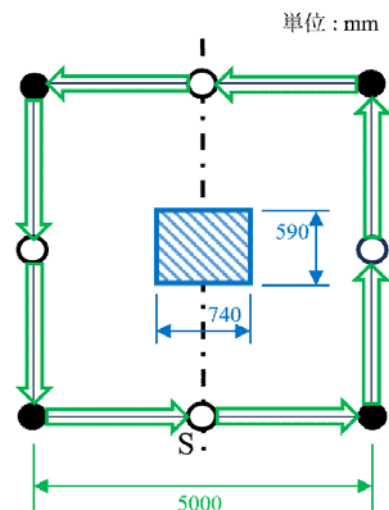


図1 撮影方法の概略図



図2 測定対象の外観

影方法の概略図を示す。ここでは、測定対象を囲む様、ドローンを以下の手順で正形状に自律飛行させる。ドローンは、同図中の位置Sから離陸し、高度約1.7[m]まで上昇したのちホバリングを行う。提案する飛行方法は2種類で、ここではそれぞれ方法1,2と呼ぶ。先ず方法1については、図1の白丸の位置でホバリングしたまま、ヨー軸方向に機首を±45度ずつ回転させる首降り動作を行い、黒丸の位置ではホバリングしたままヨー軸方向に機首を-90度回転(左回転)させる。また、矢印の区間では右方向に横滑り動作を行う。次に、方法2について

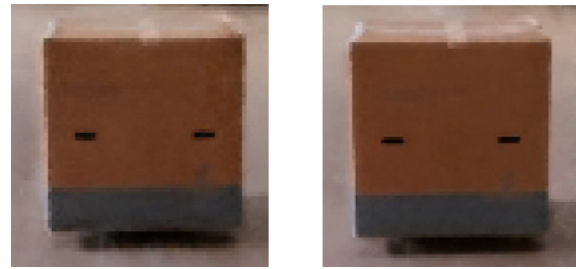
は、同図中の白丸の位置である程度の時間ホバリングを行った後、黒丸の位置まで右方向に横滑り動作を行う。黒丸の位置では、ホバリングしたまま機首を撮影対象の中心に向け、ヨー軸方向に機首を ± 45 度ずつ回転させる首降り動作を行った後、機首を対象物の方向に向けて白丸の位置まで右方向に横滑り動作を繰り返す。また、いずれの方法もドローンが離陸してから再びSに戻るまでを1周の飛行とする。なお図1においてハッチング部分は測定対象であり、図2にその外観を示す。ここでは、重量物の運搬に使用される強化段ボールの箱を使用した。大きさは高さ約830[mm]、幅約740[mm]、奥行き約590[mm]である。表面に装飾は殆どなく、凹凸も少ない形状をしている。ここで得られた動画の解析には、機械学習の技術を応用することで、複数の画像から高精度に物体の3次元形状を推定できるNeRF(Neural Radiance Fields)⁴⁾を使用する。オリジナルのNeRFでは、学習時間がGPU1枚で1日から2日程度かかるため、ここではInstant NGP(Neural Graphics Primitives)⁵⁾と呼ばれる、NeRFにおける学習時間を大幅に短縮したアルゴリズムを用いる。

3. 実験結果

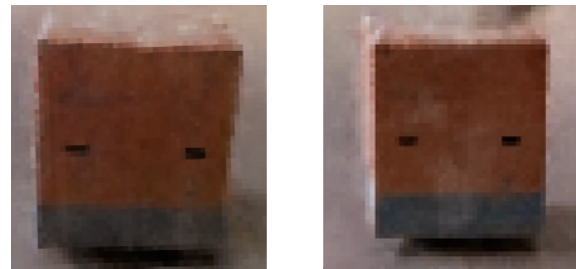
図3は、比較のために手動でNose in Circleを飛行して得られた動画から復元した3次元形状を比較した結果である。同図(a)及び(b)は、1周の飛行で得られた動画から、それぞれ50又は100枚の静止画を取り出して作成した自由視点映像について、視点を測定対象の正面が見える位置にした場合を示す。同図より、いずれの場合もほぼ測定対象の形状を再現しており、静止画の枚数が増えるとやや輪郭が明瞭になっていることもわかる。一方図4は、方法1の飛行時に、1周の自律飛行で得られた動画から作成した場合の結果である。50枚の静止画を用いた場合は、手動で飛行させた場合とほぼ同様に形状が再現されているが、100枚の場合は若干形状に歪みが生じている。また、図5は、方法2の飛行時に、1周の自律飛行で得られた動画から作成した場合の結果である。50枚の場合も100枚の場合も同様の結果となり、図4と比べて形状の歪みはほぼ無い。

4. まとめ

ドローンを正形状に自律飛行させて測定対象の周囲を撮影し、NeRFにより3次元形状を復元出来るか検討を行った結果、測定物体の形状を良く再現できることが分かった。また、手動で飛行させた場合と比較して、形状の歪み



(a) 50枚 (b) 100枚
図3 3次元形状の比較
(手動で飛行させた場合)



(a) 50枚 (b) 100枚
図4 3次元形状の比較
(方法1で自律飛行させた場合)



(a) 50枚 (b) 100枚
図5 3次元形状の比較
(方法2で自律飛行させた場合)

が生じていたが、飛行方法によっては改善される可能性があることが分かった。

参考文献

- 1) 河村幸二:精密工学会誌,Vol.79, No.5, pp.388-391 (2013)
- 2) S.Ullman: Proc.R.Soc.Lond, 203(1153), pp.405-426 (1979)
- 3) 市木, 矢澤, 新妻, 黒岩:2023IEICE 総合大会,D-12-30
- 4) Ben Mildenhall et al.: NeRF:Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV2020 (2020)
- 5) Müller, Thomas, et al. "Instant neural graphics primitives with a multiresolution hash encoding." ACM Transactions on Graphics (ToG) 41.4 (2022): 1-15.