

小型飛行ロボットのための距離画像カメラを使った 深層学習による物体の位置・向き推定

○中田 知宏, 加藤 清敬 (東京理科大学)

Object position and orientation estimation by deep learning using depth camera for small flying robot

○Tomohiro NAKATA, and Kiyotaka KATO (Tokyo University of Science)

Abstract: This paper proposes an object position and orientation estimation system by deep learning using depth camera for a small flying robot that can be used indoors, in a dark place, or in an environment under low contrast. Depth image data is captured by a depth camera for an object. Then, a 3D model for the object is generated by the depth image. In deep learning, data augmentation is performed using the 3D model for object position and orientation estimation. There, data augmentation processes are performed for the various movement of a small flying robot. The data augmentation processes can greatly reduce the preparation of learning data for the deep learning.

1. 緒論

小型飛行ロボットに、対象物の位置や向き情報を用いて自律的に制御を行う機能を持たせることで、様々な作業の自動化が期待できる。小型飛行ロボットの自律制御システムに関する従来研究としては、主に GPS や RGB カメラを用いたシステムが提案されている¹⁾²⁾。しかし、これらのシステムは屋内や暗所、コントラストのない環境において正常に動作しない可能性がある。

そこで、それら環境でも利用可能な小型飛行ロボットのための対象物の位置や向き情報を用いた自律制御システムを目指す。本論文では、そのシステムに必要な物体の位置・向き推定に関するシステムについて提案する。

このシステムを実現するにあたり、小型軽量の距離画像カメラと深層学習を採用する。距離画像カメラは、屋内や暗所、コントラストのない環境でも利用することができるほか、対象物までの距離を直接正確に測定することができ、制御において対象物までの距離を取得する追加のデバイスは不要となるメリットがある。

深層学習は、特徴量を自動的に抽出し、複雑な距離画像に対しても精度の高い推定を行うことができる。しかし、そのためには大量の学習用データが必要となる。そこで、学習時において 3D モデルの作成や data augmentation として、小型飛行ロボットの動きに対応した拡大・縮小、傾き、移動の処理等を行い、学習用データを大量に生成する。

2. 物体の位置と向きの推定による対象物追従

本研究の提案システムの概要を Fig.1 に示す。

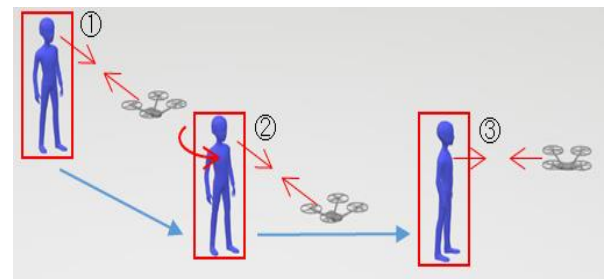


Fig.1: Object position and orientation estimation system

小型飛行ロボットに距離画像カメラを搭載し、これにより距離画像を取得する。位置向き推定部において、この距離画像を用いて深層学習による推論で対象物の位置と向きを推定する。この際、対象物の位置と向きの推定は、予め学習させておいた学習済みモデルにより推定を行う。対象物の位置と向きを推定後、追従制御部において、それらの推定結果と距離画像の距離データを用いて、①→②のような位置の変化、②→③のような向きの変化に応じた追従を行うように小型飛行ロボットを制御する。なお、本研究では、位置向き推定部についての検討を行う。

3. 位置向き推定機能のシステム構成

位置向き推定機能のシステム構成図を Fig.2 に示す。

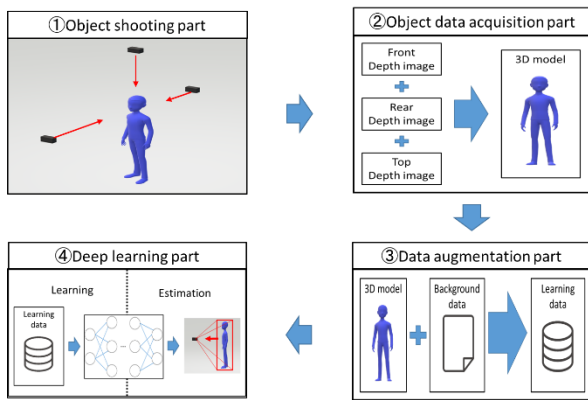


Fig.2: Position and orientation estimation process

位置向き推定部では、まず対象物撮影部において、対象物の前後と上面から 3 枚の距離画像を取得する。次に、対象物データ取得部において、その 3 枚の距離画像を元に対象物の 3D モデルを作成する。その後データ増生成部において、3D モデルによる対象物データと背景データを合成し、深層学習のための学習用データを生成する。この際、data augmentation として、処理を施すことで大量の学習用データを生成する。最後に深層学習において、その学習用データを深層学習プログラムに対応した形式に変換し、対象物の位置と向きについて学習を行う。学習後、未知の入力データに対して対象物の位置と向きについての推論を行い、対象物の位置と向きを出力する。

4. Data Augmentation(データ増生成)

本研究では、物体の位置及び向きの推定に深層学習を利用している。この深層学習において、学習を行うためには大量の学習データが必要となる。そこで、少量の学習データを集め、その学習データにより大量の学習データを増生成する data augmentation を行う。学習データの増生成において、まず距離画像から対象物を抽出し、3D モデルを作成する。その 3D モデルを用いて様々な角度から見た対象物の距離画像を生成し、拡大・縮小、傾き(ロール方向、チルト方向)の処理を施す。そして、背景データ内の様々な位置に処理後の対象物データを移動させて合成し、大量の学習データを生成する。以下に、本研究で用いた data augmentation の手法を述べる。

4.1 対象物抽出

対象物抽出において、まず対象物を映した距離画像を取得する。次に、取得した距離画像において、どの部分に対象物が存在しているか人による確認を行う。確認の様子を Fig.3 に示す。

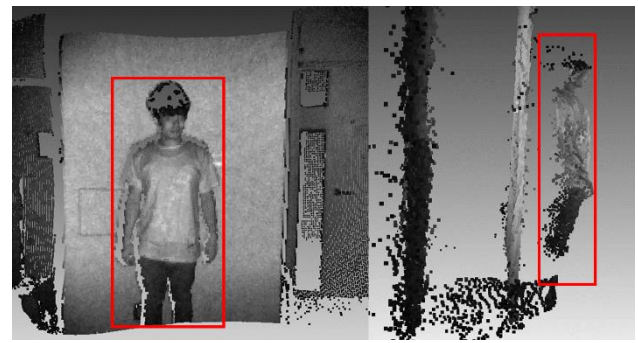


Fig.3: Capture of depth image

Fig.3 より対象物の存在する範囲を確認し、距離画像から対象物のみを抽出する。抽出前の距離画像と抽出後の距離画像を Fig.4 に示す。

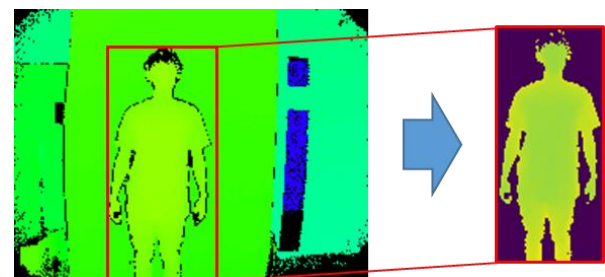


Fig.4: Object extraction

このように対象物の抽出において、対象物と背景の境界は座標に応じて明確に区別できるため、RGB 画像における対象物の抽出よりも容易にそして正確に抽出を行うことができる。

4.2 対象物と背景画像の合成

① 拡大・縮小(ズーム)

実際の距離画像において、距離画像カメラに対して対象物が近くに存在すれば対象物は大きく映り、遠くに存在すれば小さく映る。これを実現するため、抽出した対象物データの拡大・縮小の処理を行う。拡大・縮小の処理のために、まず抽出した対象物データに対して、大きさを変更する処理を行う。この処理は、Python の画像処理ライブラリ Pillow により行う。これにより、距離画像に映る対象物の大きさを変えることができる。しかし、このままでは距離画像における距離データは変更されていないため、対象物までの距離が同一でありながら、対象物の映る大きさが異なるという現実とは異なるデータになってしまう。そこで、拡大や縮小の比率に応じた距離データの変更を行う。物体の大きさは距離に反比例するため、大きさと距離の関係は式(1)となる。

$$s_0 : s_1 = d_1 : d_0 \quad (1)$$

ここで、 s_0 、 s_1 は大きさ、 d_0 、 d_1 は距離を表す。式(1)より、大きさと距離データの変更を行った距離画像を Fig.5 に示す。

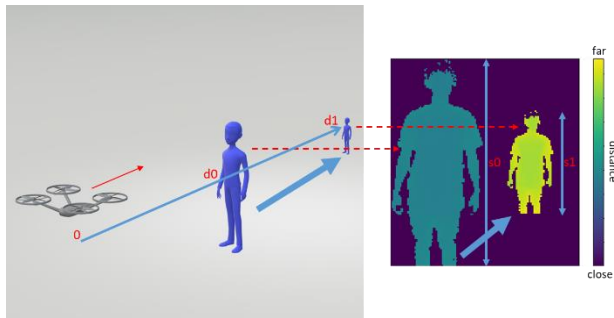


Fig.5: Change in distance

② 傾き

本研究は、小型飛行ロボットへの適用を前提としている。そのため、小型飛行ロボットは移動する際に機体が傾くことを考慮に入れる必要がある。移動により機体が傾くと、機体に固定された距離画像カメラより取得される距離画像において、対象物はロール方向やチルト方向に傾いて映ることになる。そこで、距離画像において対象物が傾いて映っていても対応できるようにするため、抽出した対象物データをもとに傾いた対象物データを生成する。

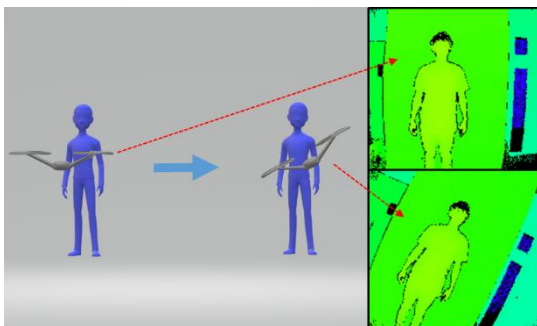


Fig.6: Change in roll

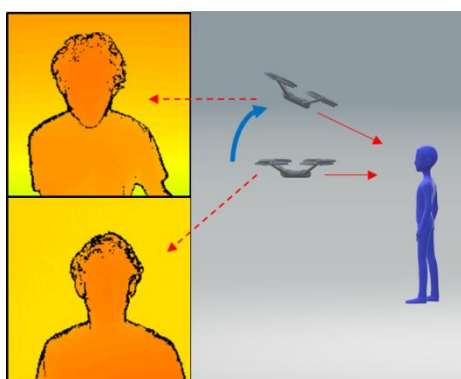


Fig.7: Change in tilt

③ 合成(移動)

抽出した対象物データに、拡大・縮小処理や傾き対応の処理を行い生成された対象物データと、別に用意した対象物の映っていない背景データにより、合成距離画像を生成する。この際、対象物データは背景データ内において様々な位置に

移動させて合成を行う。生成した合成距離画像の一例を Fig.8 に示す。

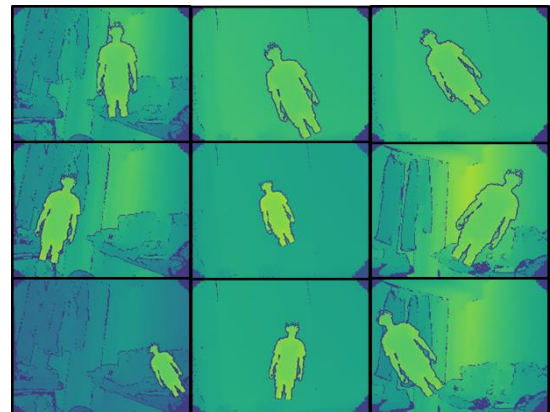


Fig.8: Composite depth image

4.3 学習

データ増生成部において拡大・縮小、傾き、移動の処理を行い生成した学習データにより学習を行い、最終的な学習モデルを構築する。学習データには、訓練データと教師データが含まれており、本研究では訓練データは距離画像、教師データは対象物の位置と向きの情報を使用している。ここで、対象物の位置情報は、距離画像中の対象物における中心座標と大きさ($cx, cy, size$)、向き情報は対象物が向いている角度(deg)である。本研究における学習の流れを Fig.9 に示す。

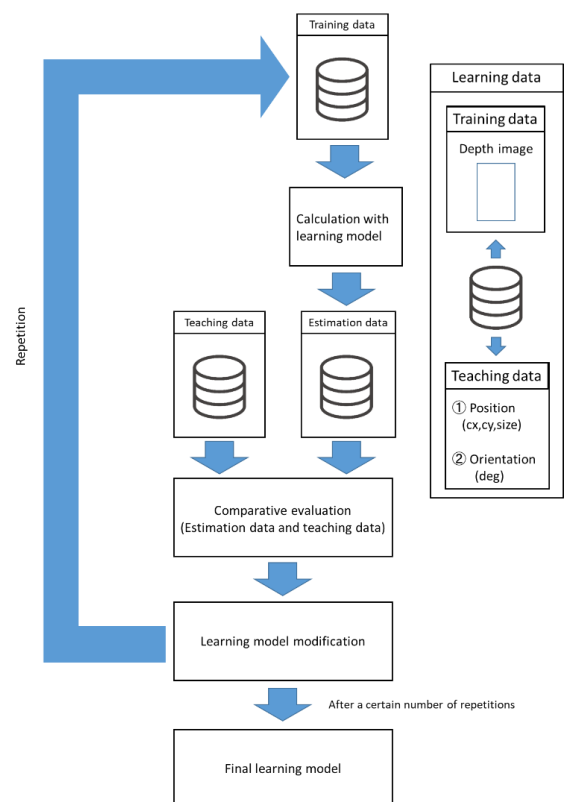


Fig.9: Learning flow

Fig.9 において、まず前処理として学習データを訓練データと教師データに分かる。次に、学習モデルに対して訓練データを入力として与えて計算し、推論データを出力する。その推論データと教師データを比較評価し、誤差が最小になるよう学習モデルを修正する。この流れを一定回数繰り返し、入力データに対して適切に推論を行うことができる最終的な学習モデルを得る。

4.4 推論

Fig.10 に、学習済みモデルを用いた推論の流れを示す。

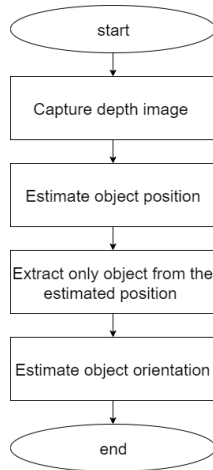


Fig.10: Inference flow

Fig.10 において、まず距離画像カメラより未知の距離画像を取得する。次に、対象物の位置について学習を行った学習モデルを用いて、取得した距離画像に対して推論を行う。その後、推論により取得した対象物の位置情報を用いて距離画像全体から対象物のみを抽出する。この際、対象物は推論で取得した対象物の位置情報の範囲で最も手前に存在する物とする。最後に、抽出した対象物に対して、対象物の向きについて学習を行った学習モデルを用いて推論を行う。

5. 結論

本研究では、小型飛行ロボットのための距離画像カメラを使った深層学習による物体の位置・向き推定システムについて検討を行ってきた。本論文では、距離画像カメラにより取得した距離画像中の対象物を距離に応じて正確に抽出し、その対象物データを用いて data augmentation により学習用データを増生成し学習することで、対象物の位置と向きの推定する方法を提案した。

今後は、3D モデルによる学習用データの作成や対象物の位置や向きの推定精度向上、小型飛行ロボットへのシステム適用を目指す。

参考文献

- [1] AA. Wazzan, F.A. Ali, R.A. Farhan, M.E. Abd “Tour-Guide Robot” , International Conference on Industrial Informatics and Computer Systems (CIICS), 2016
- [2] J.P. Su and K.H. Hsia, “Height Estimation and Image Tracking Control of an Indoor Quad-Rotor Craft via Multi-Vision Systems” , International Journal of Computer, Consumer and Control (IJ3C), vol. 2, 28/36, 2013