

3. 予防安全・自動運転のための環境認識

情報・知能工学系 教授 三浦 純

3-1 移動ロボットのための単眼深度推定

単眼深度推定：移動ロボットの制御や自動車の自動運転のためには、周囲の物体（他車や歩行者など）を確実に認識することが必要である。その際、物体の形や位置を知るためには、3次元情報を取得が必要となる。多くの自動運転研究では3次元レーザ距離センサ（LIDAR）を用いているが、一般に高価で重いものとなる。距離計測できるカメラ（Depthカメラ）は遠距離計測には不向きであり、またステレオカメラではキャリブレーション（カメラ間の位置合わせ）が必要となる。そこで、最近の深層学習技術の発展に伴い、一つのカメラで得たカラー画像から直接距離を推定（深度推定）する手法が注目されている。本稿では単眼深度推定を行った結果を紹介する。詳細については参考文献 [山形 2020] を参照されたい。

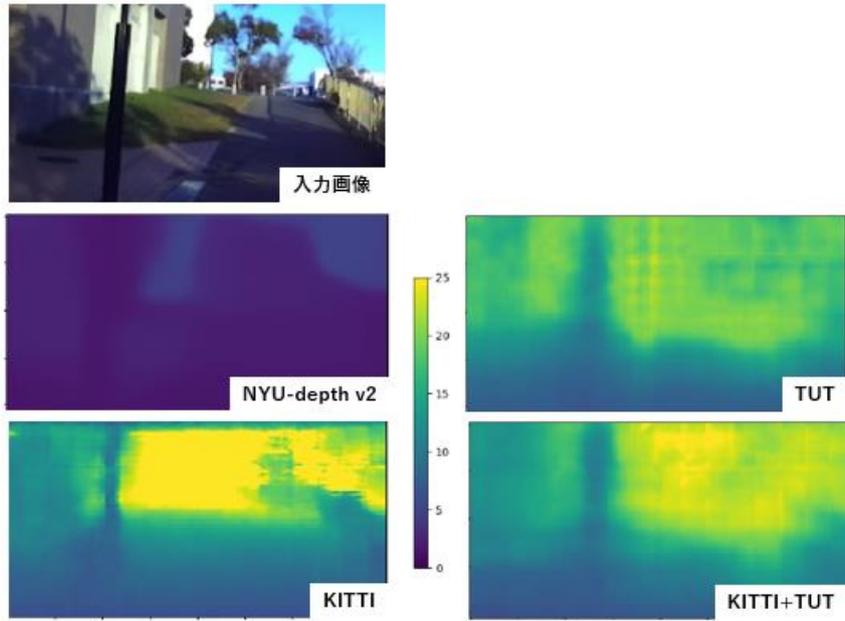
手法とデータセット：深層学習ネットワークはLainaら [Laina 2016] によって提案されたものを再実装して利用した。データセットとして、カラー画像と深度画像が対になったものが必要である。ここでは、公開されているものとして室内環境用のNYU-Depth v2、屋外道路環境用のKITTI、そして独自に構築したもの（TUTデータセット）を利用した。TUTデータセットの構築にあたっては、ステレオカメラを持って大学構内を移動し、カラー画像と距離画像を同時に取得した。前2者のデータセットでは画像対は提供されているが、データ生成に利用したカメラと、実験で用いたカメラとの画角が異なるため、画角を合わせるように画像の切り取りを行った。

実験結果：図3-1-1に異なるデータセットで学習したネットワークを適用したときの比較結果を示す。(a)は左上の入力画像に対する各モデルの深度推定の結果である。また、(b)は正解値が得られているピクセルに対して誤差を計算し、誤差の大きさに対するピクセルの累積分布をプロットしたものである。NYU-Depth v2 データセットで学習したモデルは、屋外環境に対応できていない。KITTI データセットで学習したモデルはおおよそ正しい深度を推定できている。TUTデータセットで学習したモデルは、学習環境と入力画像の環境が類似しているため、さらにより結果が得られている。KITTIとTUTの両方のデータセットを用いて学習したモデルで、もっともよい結果が得られた。これは、KITTIデータセットの多様性とTUTデータセットの環境類似性の相乗効果と考えられる。

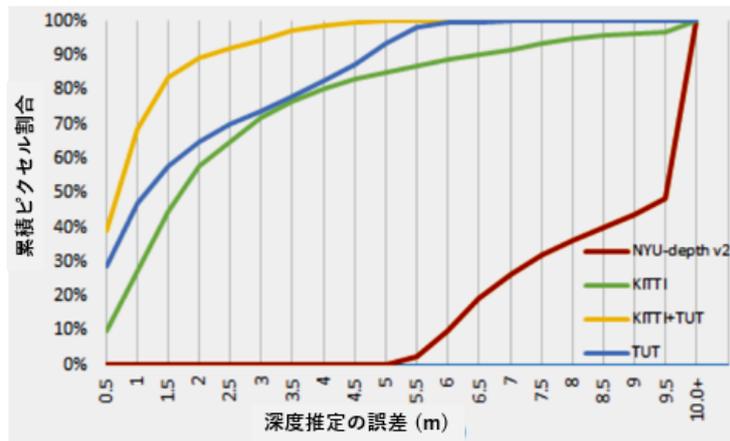
今後の課題：実験結果より、ある程度大きな物体については得られた距離情報を基に回避することが可能と考えられるが、小さな段差や溝などの認識については、精度の評価と必要な改良を行う必要がある。さらに、多様な環境での検証や実ロボットのナビゲーションへの適用が今後の課題である。

[山形 2020] 山形健斗, 三浦 純. ロボットナビゲーションのための深層学習を用いた単眼深度推定, 2020年日本機械学会ロボティクス・メカトロニクス講演会 (Robomech2020).

[Laina 2016] I. Laina et al. Deeper Depth Prediction with Fully Convolutional Residual Networks, arXiv:1606.00373, 2016.



(a) 深度推定結果



(b) 深度推定誤差の分布

図 3-1-1: 深度推定の比較実験の結果