2025 年 度		修 士 論 文 概 要			
主査	舟橋 健司	**副査**	本谷 秀堅	**研究室**	舟橋研究室
入学年度	2023 年度	**学籍番号**	35414034	**氏名**	菊地 遼

論文題目 大腸内視鏡動画セット C3VD と内視鏡フード情報を利用したポリープの大きさ推定 (Polyp Size Estimation Using the Colonoscopy 3D VideoDataset and Endoscopic Hood Geometry)

1 はじめに

二次元画像における大きさ推定には三次元情報が不可欠である一方，単眼画像からは透視投影に起因するスケール不定性が生じ，絶対深度を一意に定められない．レーザ内視鏡のように奥行きを明確に取得可能な専用機器も存在するが，導入・運用コストの観点から常時使えるとは限らない．したがって，専用機器を用いずに一般的な内視鏡画像から大きさを推定するには，内視鏡環境に適した参照情報の導入が重要となる．

先行研究では，内視鏡先端に装着されるフードが画像内に一定位置に写り込み，内径・突出長が既知である点に着目した [1]．フード領域の輝度と三次元幾何から反射係数を推定し，FMM（Fast Marching Method）に基づく形状復元と大きさ推定を行った（図 1）．ただし実画像では，撮影条件差やモデル化誤差に起因すると考えられる系統誤差が残り，過小推定傾向を含む誤差が確認された．そこで本研究では，FMM に基づいた枠組みとは別に，機械学習ベースの単眼深度推定枠組みを提案する．さらに，FMM ベースおよび機械学習ベースのいずれにも適用可能なニューラルネットワークによる誤差補正を導入し，ポリープ大きさ推定の精度向上を図る．

2 提案手法

2.1 C3VD を教師とする単眼深度推定

本研究では Colonoscopic 3D Video Dataset（C3VD）[2] を教師データとし，RGB 画像から画素単位の深度を回帰する ResNet-UNet を学習する．C3VD の深度は 16-bit 画素値 $I_{ij} \in [0, 65535]$ として与えられ，深度上限は 100 mm であるため次式のように復号する．

$$Z_{ij} = 100 \cdot \frac{I_{ij}}{65535} \tag{1}$$

損失は画素集合 Ω 上の L_1 損失とする．

$$L_{\mathrm{depth}} = \frac{1}{|\Omega|} \sum_{(i,j)\in\Omega} \left| Z_{ij} - \hat{Z}_{ij} \right| \tag{2}$$

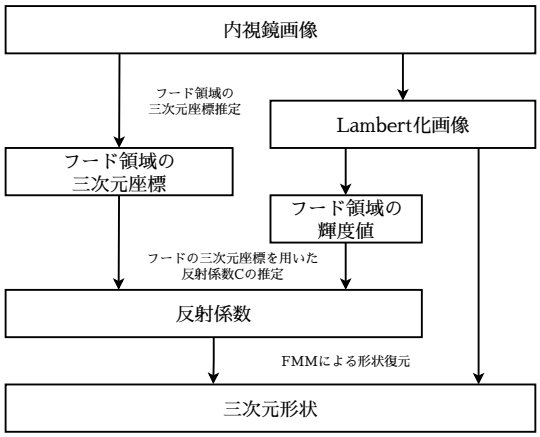


図 1: 従来手法の処理手順

2.2 スケール導入とサイズ算出

学習は C3VD の焦点距離 f_{C3VD} を前提とするため，実画像（焦点距離 f_{real}）へ適用すると深度にスケール差が生じる．ここで f_{real} はフード内径 D_{hood}，突出長 L_{hood}，画像上のフード内径 D_{pixel} から以下のように推定できる．

$$f_{\mathrm{real}} = \frac{D_{\mathrm{pixel}}\, L_{\mathrm{hood}}}{D_{\mathrm{hood}}} \tag{3}$$

焦点距離比により推定深度 $\hat{Z}$ へスケールを導入することで，深度は以下のように得られる．

$$Z_{\mathrm{est}} = \hat{Z}\, \frac{f_{\mathrm{real}}}{f_{\mathrm{C3VD}}} \tag{4}$$

さらに透視投影の相似関係から，実画像上のポリープ径 S_{pixel} と深度 Z_{est} よりポリープの大きさは以下のように算出できる．

$$S_{\mathrm{est}} = \frac{S_{\mathrm{pixel}}\, Z_{\mathrm{est}}}{f_{\mathrm{real}}} \tag{5}$$

処理の全体像を図 2 に示す．

2.3 ニューラルネットワークによる誤差補正

従来手法・提案手法いずれも観測条件差やモデル化誤差により推定誤差が残り得る．そこで推定結果からレーザ内視鏡で計測した正解大きさ S_{gt} を回帰する MLP を導入し，補正後推定値 $\hat{S}$ を得る．

$$\hat{S} = g(\boldsymbol{x}) \tag{6}$$

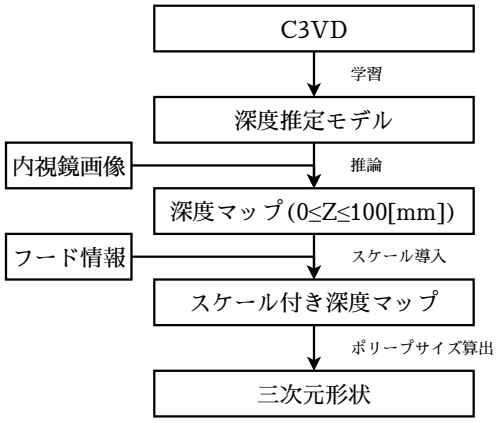
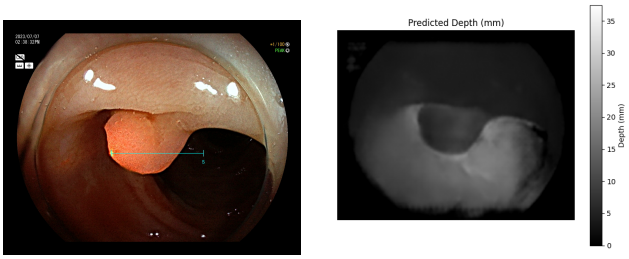


図 2: 提案手法の処理手順



実画像（約 5 mm）　　推定深度マップ

図 3: 実画像に対する深度推定結果

入力特徴量は，従来手法では $\boldsymbol{x}_{\mathrm{conv}} = [S_{\mathrm{est}}, C_{\mathrm{est}}]^\top$，提案手法では $\boldsymbol{x}_{\mathrm{prop}} = [S_{\mathrm{est}}, Z_{\mathrm{est}}, S_{\mathrm{pixel}}]^\top$ とする．MLP は 2 層（16 ユニット →8 ユニット，ReLU）で構成し，損失には平均二乗誤差 (MSE) を用いる．特徴量は学習集合の平均・標準偏差で標準化する．

3　深度推定

深度推定の妥当性は C3VD の真値深度との比較で確認する．その結果，C3VD 上での深度推定誤差（MAE）は約 0.7 mm であった．また実画像ではフード幾何により f_{real} を推定しスケールを導入する．フードの内径 $D_{\mathrm{hood}} = 15\,\mathrm{mm}$，フードの突出長 $L_{\mathrm{hood}} = 4\,\mathrm{mm}$，実画像においてフード内径が $D_{\mathrm{pixel}} = 1080\,\mathrm{px}$ であることから，f_{real} は以下のように得られる．

$$f_{\mathrm{real}} \approx \frac{1080 \times 4}{15} \approx 288\ \mathrm{px} \tag{7}$$

C3VD の焦点距離は $f_{\mathrm{C3VD}} = 769.24\,\mathrm{px}$ であり，比 $\alpha = f_{\mathrm{real}}/f_{\mathrm{C3VD}} \approx 0.3745$ として $Z_{\mathrm{est}} = \alpha \hat{Z}$ を適用する．実画像への適用例を図 3 に示す．フード情報に基づく焦点距離比でスケール導入することで，実環境に整合した深度として扱い，後段のサイズ算出に用いる．

表 1: ポリープ大きさ推定における 4 条件の誤差比較

条件	MAE [mm]	MAPE [%]
条件 1 (従来/NN 補正なし)	3.789	73.49
条件 2 (従来/NN 補正あり)	1.162	23.37
条件 3 (提案/NN 補正なし)	1.561	39.70
条件 4 (提案/NN 補正あり)	0.749	17.36

4　ポリープの大きさ推定

ポリープ大きさ推定はレーザ内視鏡で撮影された実画像 11 例で評価し，指標として MAE と MAPE を用いる．NN 補正は LOOCV により，1 例をテストとして保持し残り 10 例で学習する手順を 11 回繰り返して評価する．

表 1 に 4 条件の誤差比較を示す．提案手法（条件 3）は従来手法（条件 1）より誤差が低減し，さらに NN 補正（条件 4）により MAE 0.749 mm，MAPE 17.36%まで改善した．また，条件 1 と条件 2 から従来手法に対しても NN 補正が有効であり，系統誤差の低減を確認した．

5　むすび

通常内視鏡画像からポリープの大きさを mm 単位で推定するため，C3VD を教師とする単眼深度推定とフード幾何に基づくスケール導入，および NN による誤差補正を組み合わせた枠組みを提案した．また，提案手法に対して NN 補正を行った結果 MAE 0.749 mm を達成し，従来手法および補正なし条件より高い精度で大きさ推定が可能なことを確認した．今後の課題として，臨床データ拡充と，C3VD と実画像の領域差を低減する学習手法の導入が挙げられる．

参考文献

[1] R. Kikuchi, Y. Iwahori, K. Funahashi, M.K. Bhuyan, A. Wang, N. Ogasawara, K. Kasugai, “Polyp Size and Shape Estimation by Using an Endoscopic Hood Information”, Procedia Computer Science, vol.225, pp.2412-2419, 2023.

[2] T.L. Bobrow, M. Golhar, R. Vijayan, V.S. Akshintala, J.R. Garcia, N.J. Durr, “Colonoscopy 3D video dataset with paired depth from 2D-3D registration”, Medical Image Analysis, vol.90, p.102956, 2023.