

2021年度 修士論文

論文題目

単一内視鏡画像からのポリープ形状復元と
血管情報を利用したその大きさ推定

(Polyp Shape Recovery from Single Endoscope Image and
Size Estimation Using Blood Vessel Information)

指導教員

舟橋 健司 准教授

名古屋工業大学大学院 工学研究科 工学専攻 情報工学系プログラム

2020年度入学 32414017番

江本 峻

目次

第 1 章	はじめに	1
第 2 章	背景知識	6
2.1	内視鏡環境下におけるモデル	6
2.1.1	点光源	6
2.1.2	透視投影	7
2.1.3	Lambert 反射	7
2.2	光学的制約式と幾何学的制約式	9
2.2.1	光学的制約式	9
2.2.2	幾何学的制約式	11
第 3 章	提案手法	13
3.1	血管領域抽出	14
3.2	血管領域の傾き補正	15
3.3	血管領域の幅取得	19
3.4	ランバート画像生成	20
3.5	三次元形状復元	22
3.6	反射係数の推定	25
第 4 章	実験	27
4.1	シミュレーションによる実験	27
4.1.1	ポリープシミュレーションの作成	27
4.1.2	血管領域シミュレーションの作成	29
4.1.3	血管領域の幅取得	30
4.1.4	反射係数の推定及び形状復元	34
4.2	実画像による実験	36
4.2.1	内視鏡画像の前処理	36
4.2.2	血管領域の幅取得	39
4.2.3	反射係数の推定及び形状復元	39
第 5 章	むすび	44
	謝辞	46
	参考文献	47
	発表論文リスト	52

第1章 はじめに

大腸における腫瘍性ポリープは癌化する恐れがあるため、内視鏡を用いて癌となる悪性ポリープを早期発見し、摘出することが重要である。腫瘍性ポリープの良性、悪性の識別材料として、ポリープの大きさが10mm以上の場合、がんを含む可能性が高いことが示されている。そのため、診断医師は内視鏡画像からポリープの大きさを推定し、ポリープの良性、悪性を識別する必要がある。しかし、内視鏡環境が特殊である上に、内視鏡では二次元画像しか得られないため、経験を積んだ診断医師であっても良性、悪性の識別は困難であると言われている。そこで、ポリープの大きさを推定する手法として大きさを1-9mmまたは10mm以上という二分類問題に置き換えてこの問題を解く論文[1]がある。この論文ではポリープの正確な大きさや形状を求めている。しかし、実際には大きさだけでなく形状といった定量的な情報も取得できる方が診断の上では好ましい。そこで、内視鏡画像からポリープの三次元形状復元を行うことで大きさや形状を推定する。その結果、診断医師の負担を軽減することが可能となり、結果として診断時間の短縮や良性、悪性の識別精度の向上が期待できる。

三次元形状復元を行う手法の一つとして、複数のカメラで撮影された画像に含まれる対象の距離情報をカメラの特性および位置関係から推定するステレオ法と呼ばれる手法[2]がある。また、カメラが複数台ない場合でも単眼カメラを移動させ撮影した複数の視点による画像を使用して、環境のマッピングを行うのと同時に周囲に対するカメラの位置と向きをStructure-from-Motion（以下、SFM）に基づいて計算するVisual Simultaneous Localization and Mapping（以下、Visual SLAM）と呼ばれる手法[3]が提案されている。近年さらにそこから発展させたロバストで機能豊富なVisual SLAM技術として、ORB-SLAM3[4]、OpenVSLAM[5]、RTABMap[6]なども提案されている。これらの手法は複数の画像間で共通して見える点の画像中

の位置からカメラ位置と点の位置を推定している。画像中の点のマッチングは、特徴ベースマッチングと呼ばれるものがあり、画像からのエッジやコーナーなどの特徴点を検出し、その周囲の局所領域に対して局所特徴量を特徴ベクトルとして定義し、局所特徴量の距離に基づいて画像間のマッチングを行う。特徴点の検出には一般的に FAST algorithm[7] が用いられており、Visual SLAM では勾配ヒストグラムに基づく特徴量すなわち SIFT[8] や SURF[9] がよく利用されている。これらは最もわかりやすい特徴量であると言われているが計算量が多いという問題がある。異なる画像間で抽出した特徴量はツリーベースのアルゴリズム [10] によって対応づけられる。また、2つの特徴ベクトルを利用することで SIFT や SURF よりも計算コストを下げ同様に特徴量を求めることができる顕著な手法として、BRIEF[11], ORB[12], BRISK[13], FREAK[14] がある。2つの特徴ベクトルは Locality Sensitive Hashing アルゴリズム [15] を用いて、非常に効率的にマッチングさせることができる。しかし、これら Visual SLAM は複数の画像間において対応関係を得ることのできない部分では形状復元が行えないという問題がある。実際に、内視鏡環境においてポリープを含む腸内は水分を多く含むため表面が滑らかであり、高密度かつ高精度な対応付けをすることが困難である。また、一般的な Visual SLAM ではこのような特徴点を利用した手法であるが、輝度勾配の大きいより密な点を利用することで特徴点検出をすることなくカメラ位置を推定する LSD-SLAM[16] と呼ばれる手法もある。しかし、この手法でも対象物体が完全な剛体であるという前提を必要とし、非剛体物体までは拡張されていない。内視鏡環境では肉壁の非剛体運動が含まれやすいため、この前提は成立しないことが多いと考えられる。

三次元形状復元を行う他の手法として、画像の陰影情報から物体の形状を復元する Shape-from-Shading[17] (以下, SFS) と呼ばれる手法がある。この SFS では、物体表面の反射係数及び照明条件を既知とした場合において、画像の濃度と形状との関係を表した光学式を三次元形状に関する偏微分方程式を用いて表し、これらから得られた方程式を解くことによって形状復元を行う。また、この光学的な手法では単一の画像から復元を行うことが可能であるため、画像間対応点に大きく依存し疎な復元結果になりがちな SfM に比べて密な復元結果を得ることができる。そのため、三次元形状復元を行う場合において、復元対象となる物体に関してより高い復

元精度が必要な場合に SFS が活用されることが多い。その一例として論文 [18] の手法が挙げられる。この手法では、復元対象の物体が視線と平行な照明条件下における Lambert 反射である場合において、この問題は光を光線として扱うときの基礎方程式として知られている Eikonal 方程式として見る事が可能であり、この Eikonal 方程式を解くことで、物体の三次元形状を復元できることを示したものである。しかし、Eikonal 方程式を解く場合の問題点として、反復計算による計算コストの高さが挙げられている。そこで、上記で示した Eikonal 方程式の高速な解法として、Fast Marching Method [19] (以下、FMM) と呼ばれる手法が提案された。FMM は平行光線、平行投影、Lambert 反射、単調凸物体を前提条件とし、Lambert 画像を視点から一番近い画素から順に濃淡情報を用いて視点から遠い方向に向かって Eikonal 方程式を解いて形状復元を行う手法である。さらに、Kimmel らは物体の陰影情報から Eikonal 方程式を求め、それを FMM によって解くことで高速に物体の三次元復元を行う手法 [20] を提案している。この手法は平行投影を仮定していたが、論文 [21] や論文 [22] によって高さの更新式を透視投影の条件で立てることで透視投影への拡張がなされた。なお、論文 [21] は物体の形状を求める処理全体を反復しながら最終的な形状を求めていく手法であるため、処理時間が遅いという問題が発生してしまう。これに対して論文 [22] は初期点から逐次的な処理によって画像全体の復元を行っており反復処理を行わないため、高速な処理が可能となっている。しかし、この手法では物体の復元を進めるほど誤差が蓄積していくため、初期点から遠い箇所ほど復元の精度が落ちてしまい、復元の精度が初期点に大きく依存してしまっているという欠点がある。また、これらの手法では画像内における相対的な大きさしか求められていない。

これら三次元形状復元の手法を内視鏡画像に応用した様々な研究が行われている。内視鏡画像は点光源、透視投影という環境で撮影される画像であるため、岩堀らは平行光線を点光源に拡張した手法 [23] を提案した。さらに、論文 [24] では光学的制約式だけでなく幾何学的制約式も用いて目的関数を作成し、最適化を行うことで形状復元する手法を提案し復元精度を向上させた。そして、論文 [25] で 1 光源モデルから 2 光源モデルへと拡張がなされた。この手法により復元精度は向上したが、初期点をもとに形状復元を行なっているため初期点のパラメータを正しく設定しないと

正しい形状を推定できず、さらに最適化を用いているため処理に時間がかかるという問題があった。そこで論文 [26] では、光学的制約式と幾何学的制約式がともに同じ座標の高さを求めているため算出される結果が等しいという条件から Eikonal 方程式に基づいた式を導出することで従来手法よりも高精度で高速に復元できる手法を提案している。しかし、SFS の欠点として、単一の画像を用いた場合での形状復元では、画像内の対象物体の大きさがわかる特別な指標が存在しない場合、復元された物体の絶対的な大きさを求めることが原理的に困難な点が挙げられる。そのため、これらの手法では物体表面の反射係数を既知としているが、実際に復元された物体の絶対的な大きさを求めるにはこの反射係数を推定する必要がある。反射係数を推定する研究として、2枚の内視鏡画像から大きさが既知の医療用縫合糸を参照して、画像間でのカメラ奥行き移動量を用いることで反射係数を推定する手法 [27] がある。この手法により2枚以上の画像がある場合には反射係数を推定することができたが、実際に内視鏡画像において縫合糸が存在する画像は少なく実際に活用できる場が限られてしまう。そこで、論文 [28] は大きさが既知の参照物体を縫合糸ではなく血管にすることでカメラ奥行き移動量を推定している。血管であれば多くの内視鏡画像に含まれており、血管の大きさについても成人ではさほど違いがないことから参照物体としている。これらの手法では、カメラ奥行き移動量を正確に計算するために画像間で参照物体の高精度な対応付けを行う必要があるが、ポリープを含む腸内は表面が滑らかであり、エッジなどがほとんど存在しないため、高精度な対応付けが出来ていないという問題がある。加えて、肉壁の非剛体運動が含まれやすい内視鏡環境において正確なカメラ移動量を計算することは困難である。非剛体運動を含む環境において血管領域のグラフ構造によるマッチング手法 [29] も提案されているが、ノイズを多く含む内視鏡環境ではグラフ抽出の精度が低く、課題が残されている。また、1枚の内視鏡画像から物体の反射係数を推定する手法 [30] も提案されている。この手法では、カメラに対して水平な血管領域を検出し、血管の大きさを既知とすることで、その血管領域におけるカメラからの奥行きと反射係数を求めている。そこから血管領域に隣接する肉壁の奥行きと反射係数を最適化によって推定することで形状復元と絶対的な大きさ推定を行なっている。しかし、この手法では血管領域が内視鏡画像内にカメラに対して水平に位置する箇所が存在しなけ

ればならない。

本論文では、血管領域の幅をミリメートル単位で既知として、得られた画像から血管領域の幅をピクセル値で取得することで、ミリメートル単位とピクセル単位の変換基準を得ることができる。それを用いて三次元形状復元される画像全体の幅を推定し、その値と三次元形状復元した結果との誤差が小さくなるように反射係数を最適化により求める。また、内視鏡環境は透視投影であるため、血管領域がカメラに対して傾いている場合にその写り方が変化し、ピクセル単位の血管領域の幅が変わってしまう。そこで、血管領域がカメラに対して傾いている状況であっても、任意の反射係数を与えたFMMにより相対的な大きさの形状復元を行い血管領域の傾きを求めることで、血管領域の傾き補正を行う。これにより、内視鏡画像内にカメラに対して水平な血管領域が存在しない状況であっても、1枚の内視鏡画像から物体表面の反射係数が推定できる。そして得られた反射係数を用いたポリープ形状復元及び絶対的な大きさ推定の手法を提案する。

第2章 背景知識

本章では背景知識として、内視鏡環境モデルにおける点光源、透視投影、Lambert 反射について説明する。また、その環境において成り立つ式から導出される光学的制約式、幾何学的制約式について説明する。

2.1 内視鏡環境下におけるモデル

内視鏡環境下で撮影されたモデルは以下の条件を仮定する。

- 撮影環境は点光源・透視投影である
- 光源およびカメラレンズの中心が原点 $(0, 0, 0)$ に存在する
- Lambert 画像を使用する

ここでは、照明モデルの点光源、投影モデルの透視投影、反射モデルの Lambert 反射について説明する。

2.1.1 点光源

点光源とは、位置だけが定まり大きさを持たない光源のことである。実際には点光源というものは存在しないが、光源の大きさが光源から物体までの距離に比べて小さいときに理想的に光源が点であるとみなすことができる。点光源モデルにおける光源方向の単位ベクトル \mathbf{s} は式 (2.1) として表される。ここで、 (X_s, Y_s, Z_s) は光源の座標、 (X, Y, Z) は観測点の座標である。

$$\mathbf{s} = \frac{[X_s - X, Y_s - Y, Z_s - Z]}{\sqrt{(X_s - X)^2 + (Y_s - Y)^2 + (Z_s - Z)^2}} \quad (2.1)$$

また、表面法線方向の単位ベクトル \mathbf{n} は式 (2.2) として表される。ここで、 (p, q) は X, Y 方向の各傾きパラメータである。

$$\mathbf{n} = \frac{[p, q, -1]}{\sqrt{p^2 + q^2 + 1}} \quad (2.2)$$

$$p = \frac{\partial Z}{\partial X}$$

$$q = \frac{\partial Z}{\partial Y}$$

このとき、輝度値 E は光源からの距離の逆二乗則に従った式 (2.3) から求めることができる。ここで、 C は物体の表面反射に基づく反射係数、 r は光源から観測点までの距離である。

$$E = C \frac{(\mathbf{s} \cdot \mathbf{n})}{r^2} \quad (2.3)$$

$$r = \sqrt{(X_s - X)^2 + (Y_s - Y)^2 + (Z_s - Z)^2} \quad (2.4)$$

2.1.2 透視投影

透視投影とは、三次元物体を二次元平面で表現するための投影法の一つである。透視投影では、視点から物体までの視線が一つの視点に集められるという特徴があり、対象物を目で見た像と近い表現が得られる。また、同じ大きさの物でも視点からの奥行きが小さい場合は物体が大きく、奥行きが大きい場合は物体が小さく投影される。透視投影のモデルを図 2.1 に示す。透視投影における画像面での x, y 座標は、視点からの奥行き Z によって変化し、式 (2.5) として表される。ここで f は焦点距離、 X, Y, Z は物体の三次元座標である。

$$\begin{aligned} x &= \frac{X}{Z} f \\ y &= \frac{Y}{Z} f \end{aligned} \quad (2.5)$$

2.1.3 Lambert 反射

物体をカメラ等から観測した場合、視線方向ベクトル \mathbf{v} 、光源方向ベクトル \mathbf{s} 、法線ベクトル \mathbf{n} の関係は図 2.2 のような反射モデルで表される。ここで、 i は入射角、 e は反射角、 g は位相角である。このような反射モデルにおいて、画像の各画素にお

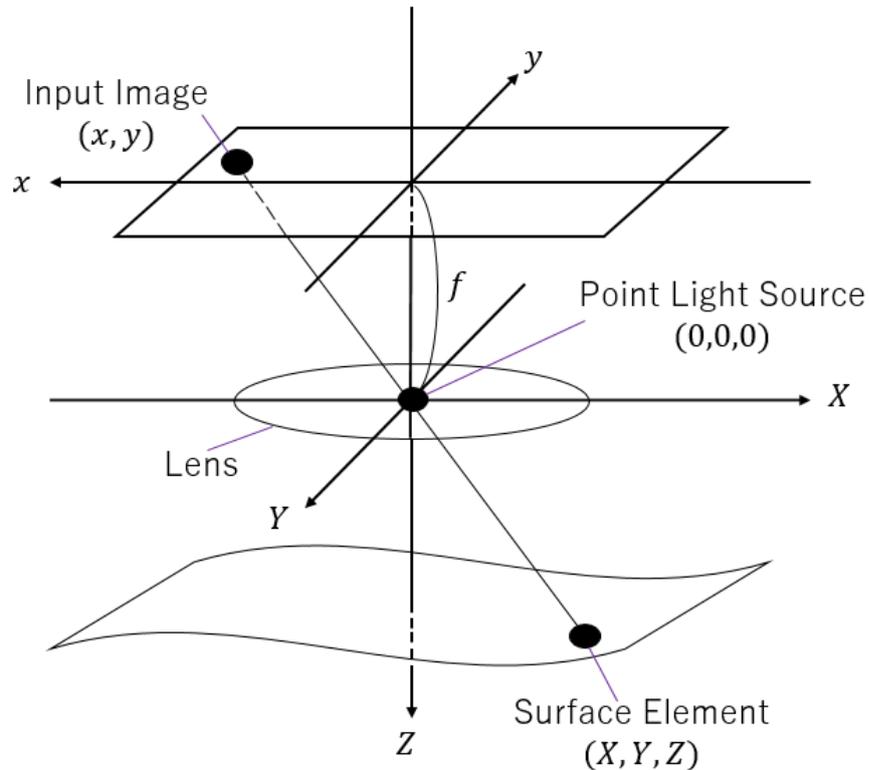


図 2.1: 透視投影モデル

ける輝度値 E は反射モデル関数 ϕ と光源の光の強度を示す光源照度 l および反射係数 C を用いて, 式 (2.6) によって表される.

$$E = l \cdot C \cdot \phi(\mathbf{s}, \mathbf{n}, \mathbf{v}) \quad (2.6)$$

また, 反射モデルの中でも物体面全てにおいて拡散反射成分のみで形成されているものを Lambert 反射モデルという. 拡散反射成分は物体面全ての輝度値が視点の角度に関係なく同じ値となり, 入射する光束に対する反射した光束の割合を表す反射率が全て一様になっている. Lambert 反射モデルにおける輝度値 E は式 (2.7) によって表される.

$$E = C(\mathbf{s} \cdot \mathbf{n}) \quad (2.7)$$

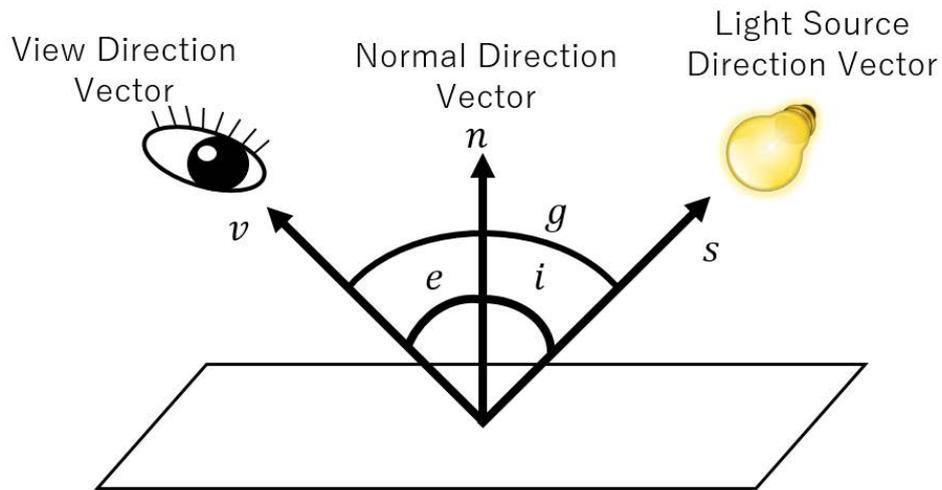


図 2.2: 反射モデル

2.2 光学的制約式と幾何学的制約式

論文 [26] では、光学的制約式と幾何学的制約式がともに同じ座標の高さを求めているため算出される結果が等しいという条件から Eikonal 方程式に基づいた式を導出することで従来手法よりも高精度で高速に復元できる手法を提案している。本研究における物体形状復元の手法でも論文 [26] の形状復元手法を継承する。そのため、ここでは光学的制約式と幾何学的制約式について説明する。

2.2.1 光学的制約式

光学的制約とは、輝度値 E を表す式 (2.3) から視点からの奥行きを求める式を導出する方法である。光学的制約モデルを図 2.3 に示す。内視鏡環境下で撮影されたモデルでは、光源およびカメラレンズの中心が原点 $(0, 0, 0)$ に存在すると仮定しているため、光源方向ベクトル s を表す式 (2.1) から式 (2.8) を得ることができる。

$$\mathbf{s} = \frac{[-X, -Y, -Z]}{\sqrt{X^2 + Y^2 + Z^2}} \quad (2.8)$$

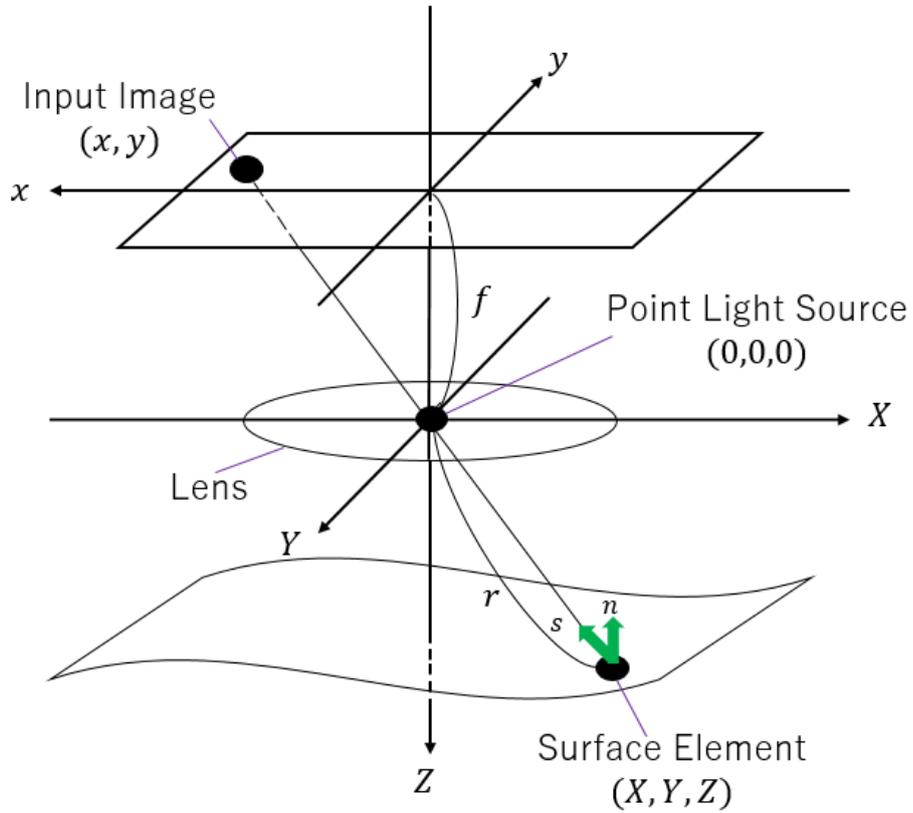


図 2.3: 光学的制約モデル

また同様に、表面法線ベクトル \mathbf{n} を表す式 (2.2) から式 (2.9)、光源から観測点までの距離 r を表す式 (2.4) から式 (2.10) を得ることができる。

$$\mathbf{n} = \frac{[p, q, -1]}{\sqrt{p^2 + q^2 + 1}} \quad (2.9)$$

$$r = \sqrt{X^2 + Y^2 + Z^2} \quad (2.10)$$

輝度値 E を表す式 (2.3) に対して、式 (2.8)、式 (2.9)、式 (2.10) を代入し、変形すると式 (2.11) となる。

$$E = C \frac{-pX - qY + Z}{(X^2 + Y^2 + Z^2)^{\frac{3}{2}} (p^2 + q^2 + 1)^{\frac{1}{2}}} \quad (2.11)$$

ここで、透視投影における画像面での x, y 座標を求める式 (2.5) を X, Y について解いた式 (2.12) を代入することで、式 (2.13) を得る。

$$\begin{aligned} X &= \frac{x}{f}Z \\ Y &= \frac{y}{f}Z \end{aligned} \quad (2.12)$$

$$E = C \frac{f^2(-px - qy + f)}{(x^2 + y^2 + f^2)^{\frac{3}{2}} Z^2 (p^2 + q^2 + 1)^{\frac{1}{2}}} \quad (2.13)$$

式 (2.13) を Z について解くと、式 (2.14) となる。

$$Z_1 = \sqrt{\frac{CV(-px - qy + f)}{E(p^2 + q^2 + 1)^{\frac{1}{2}}}} \quad (2.14)$$

ただし、式 (2.14) を簡略化するために V として式 (2.15) を定義する。

$$V = \frac{f^2}{(x^2 + y^2 + f^2)^{\frac{3}{2}}} \quad (2.15)$$

なお、 Z の値の平方根を取る場合に正と負の二つの解が存在するが、負の値はカメラに映ることがないため、正の値が採用されることは自明である。

2.2.2 幾何学的制約式

幾何学的制約とは、ある注目点から微小量移動した点での変化量から視点の奥行きを導出する方法である。 ΔZ だけ移動したときの幾何学的制約におけるモデルを図 2.4 に示す。幾何学的制約における Z を式 (2.16) のように定義する。ここで、 t は奥行きが未知の画素であり、 k は奥行きが既知の t と隣り合っている画素を示す。

$$Z_t = Z_k + p\Delta X + q\Delta Y \quad (2.16)$$

$$\Delta X = X_t - X_k$$

$$\Delta Y = Y_t - Y_k$$

ここで、透視投影における三次元座標 X, Y と画像面座標 x, y の関係式 (2.12) を式 (2.16) に代入することで、式 (2.17) を得る。

$$Z_t = Z_k + p\left(\frac{x_t}{f}Z_t - \frac{x_k}{f}Z_k\right) + q\left(\frac{y_t}{f}Z_t - \frac{y_k}{f}Z_k\right) \quad (2.17)$$

Z_t, Z_k についてそれぞれまとめると式 (2.18) となる.

$$Z_t \left(-\frac{x_t}{f} p - \frac{y_t}{f} q + 1 \right) = Z_k \left(-\frac{x_k}{f} p - \frac{y_k}{f} q + 1 \right) \quad (2.18)$$

さらに, Z_t について解くと式 (2.19) となる.

$$Z_t = \frac{Z_k(-px_k - qy_k + f)}{-px_t - qy_t + f} \quad (2.19)$$

この式 (2.19) を用いることで, 視点からの奥行きを計算することができる.

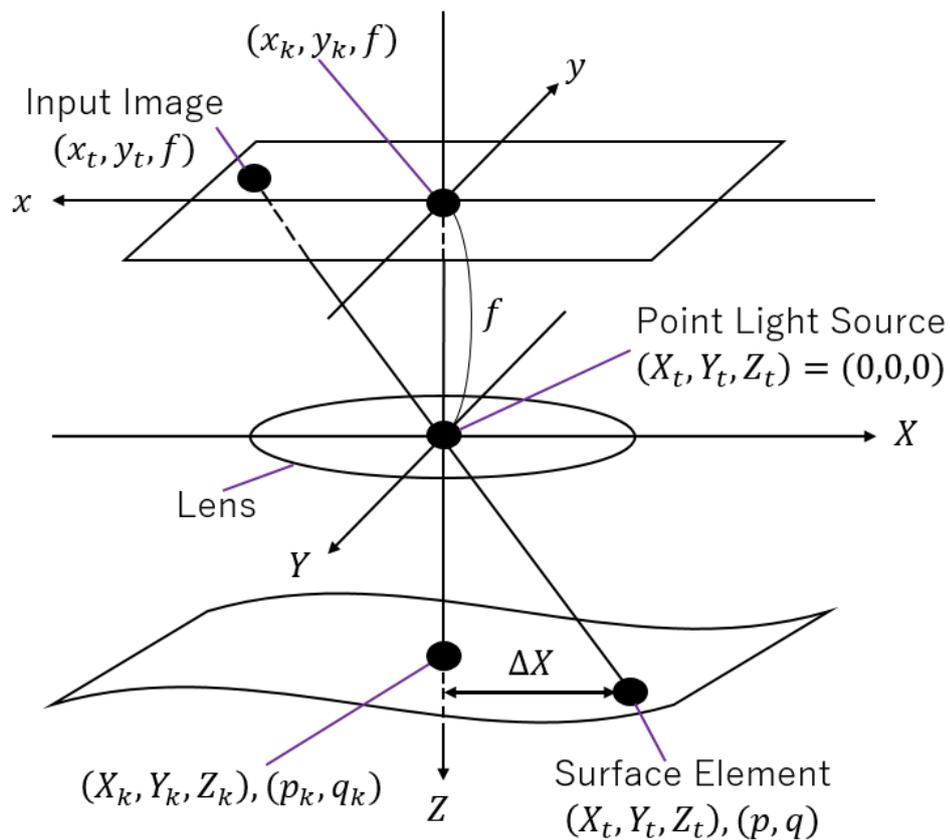


図 2.4: 幾何学的制約モデル

第3章 提案手法

血管領域の幅取得にはU-Netを用いた血管検出を行う。そこから可能な限りポリープに近い血管領域を選択し、その箇所の血管領域の幅をピクセル単位で取得する。なお、血管が水平になっていない場合でもその箇所の傾きを補正することでその箇所の血管領域の幅を正確に取得する手法も提案する。血管領域の幅をミリメートル単位で既知とすることで、ピクセル単位とミリメートル単位の変換基準を得ることができる。そして、復元対象のピクセルサイズとこの変換基準から復元対象の予測値を定義する。最後に、定義した復元対象の予測値と一致するように最適化によって反射係数 C を推定する。これまでの1枚の内視鏡画像から絶対的な大きさと形状を得る先行研究では血管が水平になっているところを検出して形状復元に利用する手法[30]が提案されているが、本論文では血管が水平になっていない場合でも絶対的な大きさと形状を得られるように拡張している。提案手法全体の流れを図3.1に示す。

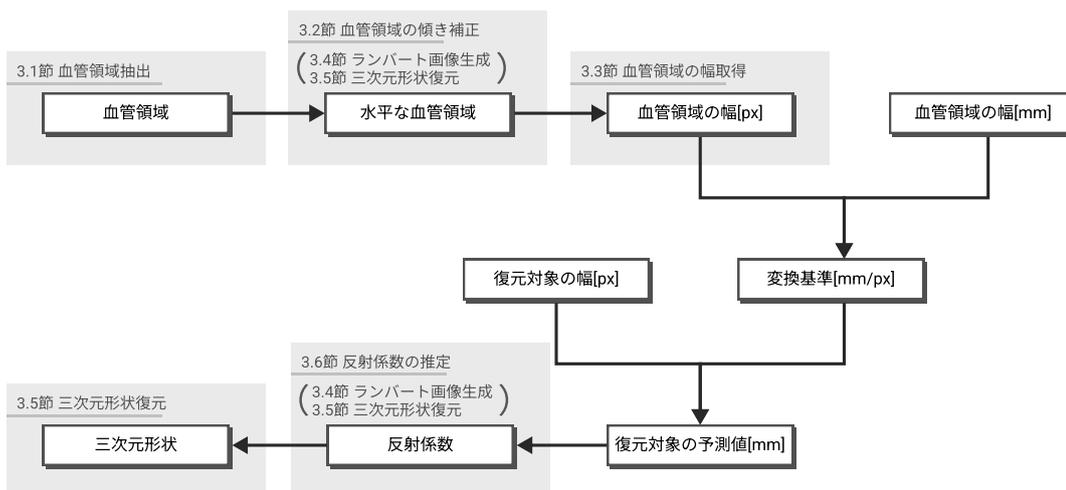


図 3.1: 提案手法の流れ

3.1 血管領域抽出

血管領域を検出する手法に、Frangi Filter[31]を用いた文献[32]がある。しかし、本手法で血管領域を参照物体としてポリープの大きさを推定するため、対象の領域をピクセルの精度で検出する必要がある。そこで、医用画像処理分野において、実用されているディープニューラルネットワーク (DNN:Deep Neural Network) を用いる。DNN の代表的なアーキテクチャの一つに畳み込みニューラルネットワーク (CNN:Convolution Neural Network) がある。通常の CNN によって行われる画像のクラス分類では、畳み込み層が対象の局所的な特徴を抽出する役割を担い、プーリング層が対象の全体的な位置情報をぼかす役割を担う。そのため、より深い層ほど抽出される特徴はより局所的になり、またその特徴の全体的な位置情報はより曖昧になる。CNN による検出処理の多くは画像中の対象を発見し、対象の大きさや位置を同定することを目的としている。一方の領域分割は、画像中の個々の対象の領域をピクセルの精度で同定することを目的とする。つまり、プーリング層でぼかされた局所的特徴の位置情報を元画像上でピクセル単位で復元する必要がある。そこで、対象の局所的特徴と全体的位置情報の両方を結合して学習させるために開発された U-Net[33] を用いて血管領域分割を行う。U-Net の中でも、文献[34] で用いられる Retina U-Net によって血管領域分割を行う。本研究の目的はポリープの形状復元であり、血管抽出はあくまで手段であるため、変更は必要最小限とし、今回扱う内視鏡画像データ (OLYMPUS 製) では、畳み込み層を 5 層から 3 層への変更のみが妥当であった。

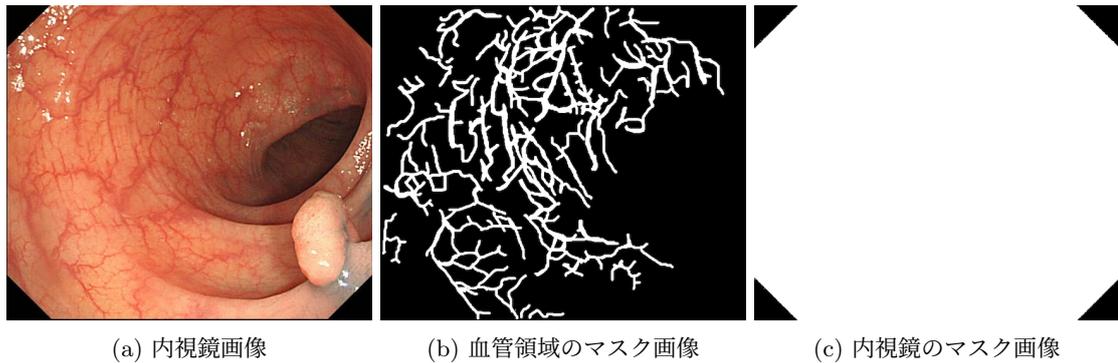
以下に手順を示す。

Step 1. 図 3.2a にある学習用の画像から図 3.2b にあるマスク画像を作成する。

Step 2. 学習用の画像と Step1 のマスク画像から 48×48 ピクセルの 9500 個のランダムパッチを抽出する。

Step 3. Step2 のパッチから血管領域を特徴とさせて学習を行う。

Step 4. Step3 の学習データを元にテストデータから学習させた特徴を検出する。



(a) 内視鏡画像

(b) 血管領域のマスク画像

(c) 内視鏡のマスク画像

図 3.2: 学習用画像

図 3.2c にある Field Of View(FOV) 境界と血管領域を区別するために、一部または完全に FOV の外側にあるパッチも抽出する。また、元の画像の同じ部分を含む複数のパッチが抽出される場合もあるが、それ以上のデータ拡張は行わない。

3.2 血管領域の傾き補正

内視鏡環境下における撮影環境は透視投影であるため、同じ大きさの物でも視点からの奥行きが小さい場合は物体が大きく、奥行きが大きい場合は物体が小さく投影される。つまり、図 3.3a のように血管領域がカメラに対して水平に位置している場合、3.3 節で述べる手法で血管領域の幅を推定することができる。しかし、図 3.3b のように血管領域がカメラに対して傾きがある場合、その領域の幅を正しく測定することができない。そこで、ここでは血管領域が水平に位置していない場合でも正しい血管領域の幅を測定する手法を提案する。

血管領域が水平に位置していない場合において、図 3.4 に示すように血管領域上のある点の表面法線ベクトル \vec{OP} は Z 軸に対して天頂角 α 、方位角 β をもっているとする。ここで、ベクトル \vec{OP} を Z 軸に合わせるためにアフィン変換による回転を行う。これを血管領域上の他の任意の表面法線ベクトルでも同様に変換することで、血管領域全体をカメラに対して水平に位置させることができる。なお、天頂角 α 、方位角 β は血管領域箇所の傾き p 、 q から求める。物体の表面法線ベクトル \mathbf{n} を天頂

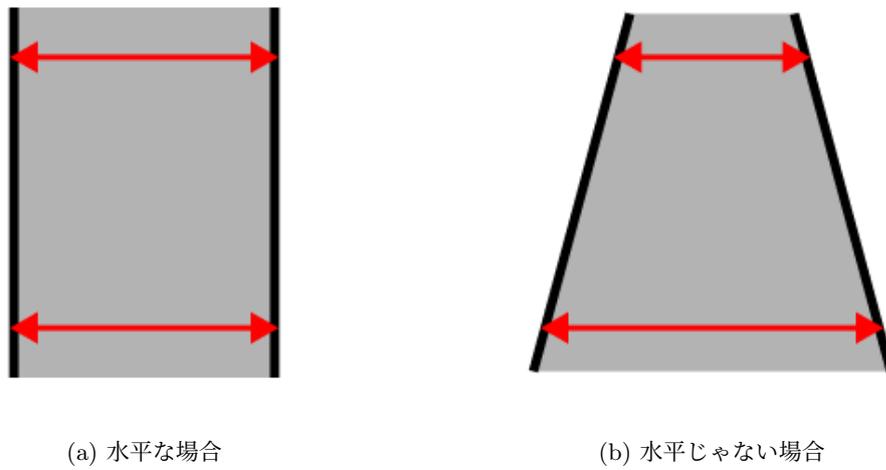


図 3.3: 血管領域の写り方

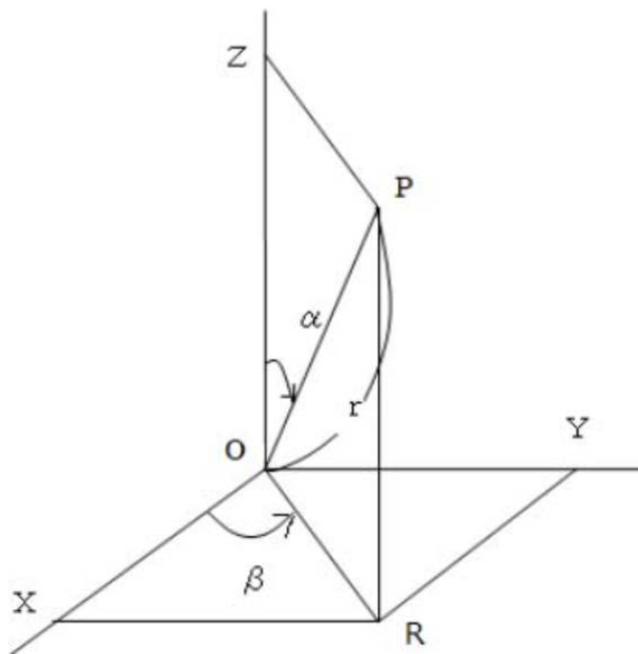


図 3.4: アフィン変換の角度 α , β

角 α , 方位角 β を用いて表すと, 式 (3.1) となる.

$$\mathbf{n} = \begin{bmatrix} \sin \alpha \cos \beta \\ \sin \alpha \sin \beta \\ \cos \alpha \end{bmatrix} \quad (3.1)$$

また、傾き p , q を用いて表すと、式 (3.2) となる.

$$\mathbf{n} = \begin{bmatrix} \frac{-p}{\sqrt{p^2+q^2+1}} \\ \frac{-q}{\sqrt{p^2+q^2+1}} \\ \frac{1}{\sqrt{p^2+q^2+1}} \end{bmatrix} \quad (3.2)$$

式 (3.1), 式 (3.2) から α , β について解くと、式 (3.3) となる.

$$\begin{aligned} \alpha &= \cos^{-1} \left(\frac{1}{\sqrt{p^2+q^2+1}} \right) \quad (0 < \alpha < \pi) \\ \beta &= \cos^{-1} \left(\frac{p}{\sqrt{p^2+q^2}} \right) \quad (0 < \beta < \pi) \end{aligned} \quad (3.3)$$

3.4 節による手法で血管領域箇所におけるランバート画像を生成し、3.5 節による手法で形状復元することで、血管領域箇所各点の X 軸, Y 軸方向の傾き p , q が計算される. そこから式 (3.3) を用いて天頂角 α , 方位角 β を計算する.

このとき、3.4 節による手法で反射率を一様化しているが、実画像では血管領域箇所の反射率が一様化されず濃い色で残ってしまうことがある. そうなると SFS では色が暗い箇所を高さが低い箇所と認識するため、血管領域だけ凹んだ形状が得られてしまい、血管領域の傾きを正しく推定できない. そこで、形状復元する血管領域は局所的な狭い範囲であるため、曲面ではなく平面に近似できると考える. つまり、形状復元して得られた形状を平面に近似することで血管領域の形状を再定義する. 平面であれば透視投影における視点ベクトルとの交点が一意に求まるため、画像面座標 x, y から三次元座標を一意に求めることができる. 平面の方程式は p_{00} , p_{10} , p_{01} を係数として、式 (3.4) を用いる.

$$Z = p_{00} + p_{10}X + p_{01}Y \quad (3.4)$$

また、アフィン変換による回転の中心は原点となるため、本来血管領域が位置していた奥行き Z がずれてしまう恐れがある. 透視投影モデルにおいては視点からの奥行きが小さい場合は物体が大きく、奥行きが大きい場合は物体が小さく投影されるという特徴があるため、奥行き Z がずれてしまうと参照物体の写り方が変化してしまい大きさが変わってしまう. 本論文では、ポリープが位置する底面を高さのペー

スとするため、参照物体も同じ高さに存在するものを使用することが望ましい。そこで、アフィン変換による回転の中心をポリープに最も近い点を基準点とし、その基準点を中心に回転を行う。アフィン変換の手順を以下に示す。

Step 1. ベクトル \vec{OP} を基準点となるベクトル $-\mathbf{T}$ だけ並進させる。これによりベクトル \mathbf{T} が原点に移動するため、ベクトル \mathbf{T} 中心の回転操作が可能となる。

Step 2. ベクトル \vec{OP} を Z 軸に対して角度 $-\beta$ だけ回転させる。これはベクトル \vec{OP} が XZ 平面に写像されることを意味する。

Step 3. ベクトル \vec{OP} を Y 軸に対して角度 $-\alpha$ だけ回転させる。これはベクトル \vec{OP} が Z 軸に写像されることを意味する。

Step 4. ベクトル \vec{OP} を基準点となるベクトル \mathbf{T} だけ並進させる。これはベクトル \vec{OP} を元の位置に戻すことを意味する。

また、血管領域上の他の任意の表面法線ベクトルも同様に変換することで、ベースとなる高さに合わせて血管領域を水平に位置させることができる。

ここで使用したアフィン変換を式 (3.5), 式 (3.6), 式 (3.7), 式 (3.8) に示す。

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -T_x \\ 0 & 1 & 0 & -T_y \\ 0 & 0 & 1 & -T_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.5)$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos(-\beta) & -\sin(-\beta) & 0 \\ \sin(-\beta) & \cos(-\beta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.6)$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & 0 & \sin(-\alpha) \\ 0 & 1 & 0 \\ -\sin(-\alpha) & 0 & \cos(-\alpha) \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3.7)$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3.8)$$

ただし、基準点となるベクトル \mathbf{T} は式 (3.9) とする。

$$\mathbf{T} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (3.9)$$

血管領域の各点において、式 (3.5) の並進を適用したのち、式 (3.6)、式 (3.7) の2つの回転を適用する。その後、式 (3.8) の並進を適用することで、ベースとなる高さに合わせた状態でカメラに対して水平に位置させた血管領域形状を得る。そこから式 (2.5) を用いて、透視投影における画像面での x, y 座標を計算し、3.3 節の線形回帰によって画像面上の血管領域の幅 W_{mm} を計算する。そして、式 (3.10) を用いて画像面上の血管領域の幅をピクセル単位 W_{px} に変換する。

$$W_{\text{px}} = \frac{W_{\text{mm}} R_{\text{px}}}{S_{\text{mm}}} \quad (3.10)$$

ただし、 R_{px} は入力画像の解像度、 S_{mm} は透視投影における画像面の大きさを表す。

3.3 血管領域の幅取得

三次元形状復元される画像全体の幅をミリメートル単位で知るために、ピクセル単位をミリメートル単位に変換する必要がある。本論文では血管領域の幅をミリメートル単位で既知としているため、血管領域の幅をピクセル値で取得できれば変換基準を得ることができる。ここでは血管領域の幅をピクセル単位で取得する手順について説明する。まず、U-Net を用いて血管領域のセグメンテーションを行い、そこから血管領域の輪郭を抽出する。そして、血管の輪郭を直線の方程式に近似することで、その直線間の距離から血管の幅を取得する。

血管領域の輪郭を直線方程式に近似するために、式 (3.11)、(3.12) を定義する。ここで、2つある輪郭のピクセル座標をそれぞれ (x_1, y_1) , (x_2, y_2) , y 切片をそれぞれ β_0 , β_1 , 傾きを β_2 とする。直線間の距離を計算するために傾きは同じものとする。

$$y_1 = \beta_0 + \beta_2 x_1 \quad (3.11)$$

$$y_2 = \beta_1 + \beta_2 x_2 \quad (3.12)$$

また、上記線形モデルと実測値との誤差の二乗和はそれぞれ式 (3.13), (3.14) になる.

$$\epsilon_1 = \sum_{i=1}^n \{y_1^i - (\beta_0 + \beta_2 x_1^i)\}^2 \quad (3.13)$$

$$\epsilon_2 = \sum_{i=1}^m \{y_2^i - (\beta_1 + \beta_2 x_2^i)\}^2 \quad (3.14)$$

これらの誤差が最小になるように式 (3.15) で最適化を行う.

$$\beta_0^{opt}, \beta_1^{opt}, \beta_2^{opt} = \arg \min_{\beta_0, \beta_1, \beta_2} (\epsilon_1 + \epsilon_2) \quad (3.15)$$

最適化によって得られた $\beta_0, \beta_1, \beta_2$ を用いて 2 直線間の距離をピクセル単位で測定する.

3.4 ランバート画像生成

SFS による三次元形状復元において、鏡面反射成分が含まれている場合や色の異なる箇所があり反射率が一様ではない場合、復元精度が大きく低下してしまう. そのため、SFS による三次元形状復元では Lambert 反射を仮定したアプローチが多く提案されている. 提案手法でも形状復元に Lambert 反射を仮定し、内視鏡画像に対して Lambert 画像生成を行う.

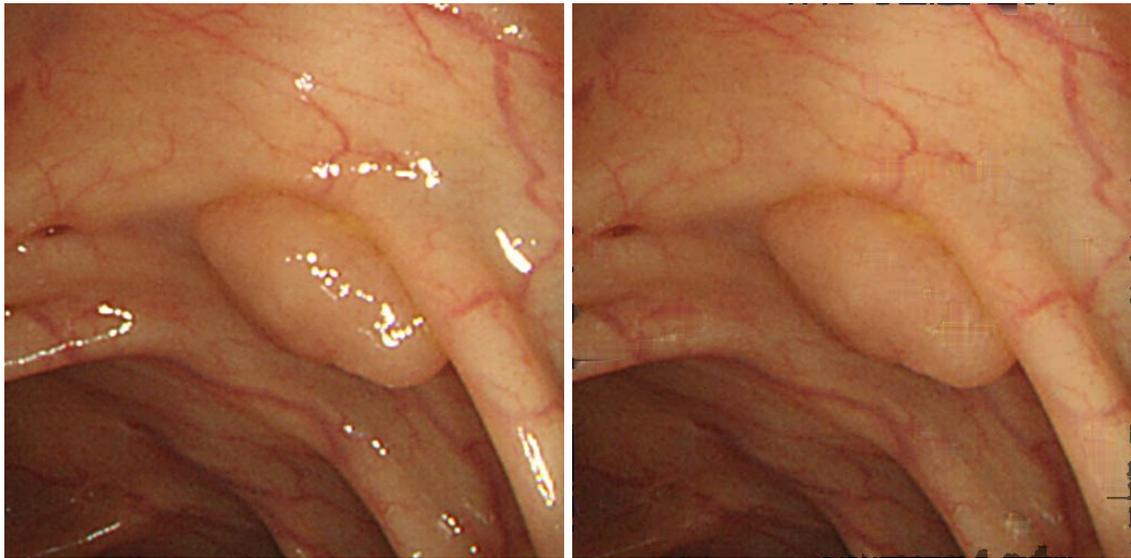
Lambert 画像生成手法として、論文 [35] では複数の視点から対象物を撮影することで鏡面反射成分が除去された Lambert 画像を生成している. また論文 [36] では、物体の回転を利用して元画像を Lambert 画像に変換する手法を提案している. しかし、内視鏡環境では視点の異なる複数の画像間において、カメラの運動と物体の蠕動運動の両方が含まれてしまうため、上記手法を用いることができない. そこで、1 枚の内視鏡画像から Lambert 画像を生成する論文 [37] が提案されている. この手法では、画像の RGB 情報を用いて反射率のクラスタリングを行い、各クラスタにおける相対的な反射率を求める. そして、相対的な反射率の差を吸収し、均一な反射

率を割り当てることで、元の異なる色の反射率の領域に対してランバート反射率に近い画像を生成する。またこの手法で用いられている Mean-Shift クラスタリングは初期点がランダム性を持つことから安定した結果を得ることができないという問題があるため、一様化の処理を複数回行い結果の平均を取ることで結果の安定性を向上させているが、これにより計算コストが高くなり処理時間が長くなってしまっている。

そのため本論文では、短い処理時間で内視鏡画像から反射率が一様な Lambert 画像を生成する手法 [38] を用いている。この手法では、鏡面反射領域を検出するために高速フーリエ変換による領域検出とカラー成分の閾値による領域検出を使用している。高速フーリエ変換による領域検出では、RGB 成分のうち画像中で最も少ない成分を求め、その成分に対して高周波成分のみが残るように高速フーリエ変換を行う。これは、鏡面反射領域の RGB 値は画像全体のカラー成分に対して全て高い傾向にあり、鏡面反射領域付近で高周波成分が得られやすいためである。これにより鏡面反射領域の輪郭となる部分を得る。またカラー成分の閾値による領域検出では、青色または緑色に閾値を与えることで鏡面反射成分を持つ画素を検出する。内視鏡画像には赤色成分が広い分布で多く含まれており、閾値で分けることが困難なため赤色は使用していない。求めた2つの鏡面反射領域に対して膨張処理をかけて論理和をとることでマスク画像を生成し、生成したマスク画像をもとに鏡面成分領域を補間する。補間処理は注目点の8近傍に対してエッジ情報をできるだけ残すように行う。それは距離の重みに応じて係数を分布させたフィルタを適用し、8近傍の画像の和をとることで補間を行っており、その結果を図3.5に示す。補間された画像の RGB 空間を HSV 空間へと変換することで、色相 H のヒストグラムによる反射係数ごとのクラスタリング及び明度 V による反射率一様化処理を実現している。これにより、血管領域と腸内肉壁領域に対して一様な反射率を割り当てることができ、色が異なる反射率領域に対して反射率の差が吸収された Lambert 反射に近い画像を生成できる。一様な反射率処理の手順を以下に示す。

Step 1. HSV 表色系の色相 H のヒストグラムによる分類を行う。

Step 2. 分類された最大クラスと注目クラスの境界で2点を取り、明度 V の比率を



(a) 入力画像

(b) 除去及びその補間した画像

図 3.5: 鏡面反射領域の除去及びその補間

算出する.

Step 3. 反射率が一様な物体表面において隣接する2点間の明度 V はほぼ等しいと仮定できるため, 注目点の明度 V の比率を最短距離の最大クラスに属する点の比率と等しくなるようにして, 最大のクラスへ吸収させる.

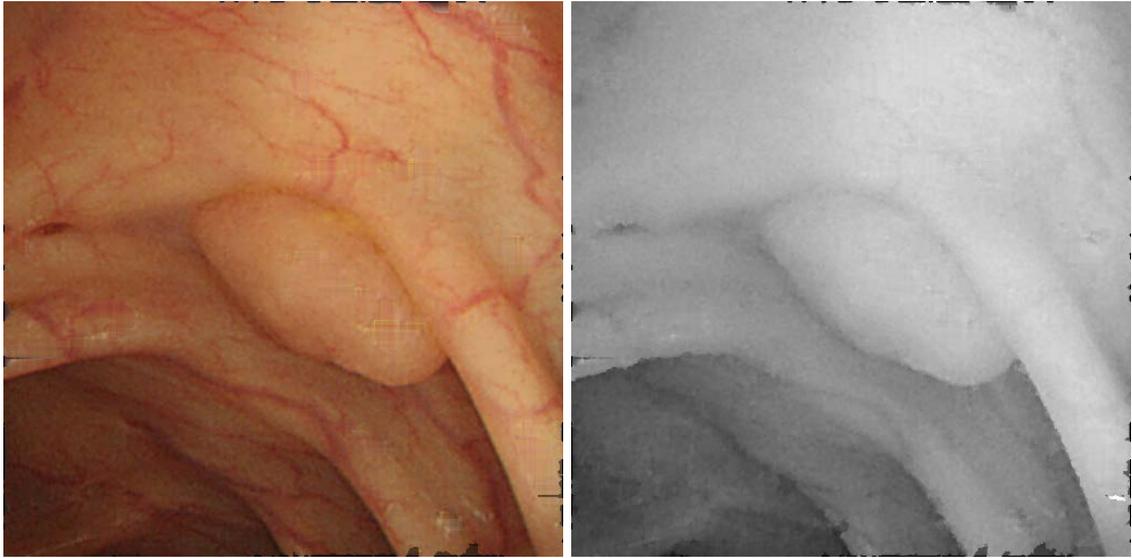
Step 4. 最大クラスから色相 H の値に近い順に全てのクラスに対して同じ処理を行うことで均一な反射率にする.

その結果を図 3.6 に示す.

3.5 三次元形状復元

ランバート反射の仮定で観測される輝度値 E は, 奥行き Z , 表面勾配パラメータ (p, q) , 反射係数 C で表される. この関係から, 奥行き Z を他のパラメータで表現するために, 式 (3.16) で表される光学的制約式が導かれる. E は輝度値, x, y は観測画像座標, f はレンズの焦点距離である.

$$Z_1 = \sqrt{\frac{CV(-px - qy + f)}{E(p^2 + q^2 + 1)^{\frac{1}{2}}}} \quad (3.16)$$



(a) 入力画像

(b) 一様な反射率処理した画像

図 3.6: ランバート画像の生成

ただし、式 (3.16) を簡略化するために V として式 (3.17) を定義する.

$$V = \frac{f^2}{(x^2 + y^2 + f^2)^{\frac{3}{2}}} \quad (3.17)$$

また、近傍点での Z と (p, q) の関係は類似しており、 (p, q) を用いた積分関係で奥行き Z を表現すると、式 (3.18) の幾何学的制約式が得られる. t (trial) は画像上の Z が未知の注目点, k (known) は Z が既知の近傍点を表す.

$$Z_2 = \frac{Z_k(f - px_k - qy_k)}{(f - px_t - qy_t)} \quad (3.18)$$

式 (3.16), 式 (3.18) から得られる奥行き Z は各点において等しくなければならないため、2項を等しくして並べ替えると、Eikonal 方程式の形になる. ここでは、FMM (Fast Marching Method) の枠組みを用いて、Eikonal 方程式を解く.

ここで、 Z_1 と Z_2 は等しいはずで、この2つの式から注目点での奥行き Z を求めることができる. そこから、 $Z_1 = Z_2$ となり、式 (3.19) が成立する. また、式 (3.18) は既知の p, q を用いているため、式 (3.19) では $(p, q) = (p_k, q_k)$ とする.

$$\sqrt{\frac{CV(-px - qy + f)}{E(p^2 + q^2 + 1)^{\frac{1}{2}}}} = \frac{Z_k(f - p_k x_k - q_k y_k)}{(f - p_k x_t - q_k y_t)} \quad (3.19)$$

$(f - px - qy) \approx (f - p_k x - q_k y)$ という近似を用いて式 (3.19) を展開すると、式 (3.20) の形になる。

$$\sqrt{p^2 + q^2} = \sqrt{\left(\frac{C}{E}\right)^2 A - 1} \quad (3.20)$$

ただし、式 (3.20) を簡略化するために A として式 (3.21) を定義する。

$$A = \frac{V^2(f - p_k x_t - q_k y_t)^6}{Z_k^4(f - p_k x_k - q_k y_k)^4} \quad (3.21)$$

さらに、式 (3.20) を Z の x , y 方向それぞれにおける数値差分から求まる式 (3.22) に置き換える。

$$\sqrt{\max(p_-, -p_+, 0)^2 + \max(q_-, -q_+, 0)^2} = f_{ij} \quad (3.22)$$

$$\begin{aligned} p_- &= \frac{Z_{i,j} - Z_{i-1,j}}{X_{i,j} - X_{i-1,j}} \\ p_+ &= \frac{Z_{i+1,j} - Z_{i,j}}{X_{i+1,j} - X_{i,j}} \\ q_- &= \frac{Z_{i,j} - Z_{i,j-1}}{Y_{i,j} - Y_{i,j-1}} \\ q_+ &= \frac{Z_{i,j+1} - Z_{i,j}}{Y_{i,j+1} - Y_{i,j}} \end{aligned} \quad (3.23)$$

ここで、 f_{ij} は式 (3.20) の右項、 i, j は注目点の画素座標である。このとき、式 (3.22) を奥行き Z について解くと、式 (3.24) となる。

$$Z_{ij} = \begin{cases} \frac{Z_a + Z_b + \sqrt{2f_{ij}^2 - (Z_a - Z_b)^2}}{2} & (|Z_a - Z_b| < f_{ij}) \\ \min(Z_a, Z_b) + f_{ij} & (|Z_a - Z_b| \geq f_{ij}) \end{cases} \quad (3.24)$$

$$Z_a = \min(Z_{i-1,j}, Z_{i+1,j}) \quad (3.25)$$

$$Z_b = \min(Z_{i,j-1}, Z_{i,j+1})$$

Eikonal 方程式に基づくポリープ形状復元の手順を以下に示す。

Step 1. 最大輝度値を持つ画素を初期点とし、勾配パラメータを $(p, q) = (-x/f, -y/f)$ で設定し、式 (3.16) を用いて奥行き Z を算出することで既知の点とする。初期点以外の点については $Z = \infty$ とする。

Step 2. 初期点を中心とした4近傍を注目点とし、初期点における (p, q) を用いて、この4近傍における Z を式 (3.24) により計算する。

Step 3. Step2で算出した4近傍における Z の x , y 方向それぞれにおける数値差分から求まる式(3.23)を用いて, 4近傍の (p, q) を更新する.

Step 4. Step3で更新された (p, q) を用いて式(3.24)から, 隣接する注目点の Z を更新する. ただし, 注目点の X, Y は $(X, Y) = (Zx/f, Zy/f)$ を用いたワールド座標である.

Step 5. 求めた注目点 Z のうち, 最小値となる画素を既知とする.

Step 6. すべての画素の Z が計算されるまで, Step2~Step5を繰り返す.

3.6 反射係数の推定

3.5節の形状復元手法では反射係数 C は既知として行われるため, ここではその正しい反射係数の推定方法について説明する.

本論文の形状復元手法では, 任意の反射係数 C を与えて形状復元結果を得ることができる. しかし, 光学的制約式(3.16)からもわかるように反射係数 C が大きいほど奥行き Z の値も大きくなる. さらに, 全画素の奥行き Z を推定したあと, 透視投影における三次元座標 X, Y と画像面座標 x, y の関係式 $(X, Y) = (Zx/f, Zy/f)$ によって各画素の X, Y を計算しているため, その幅も同様に大きくなる. つまり, 反射係数 C が大きいほど復元結果の幅が大きくなり, 反射係数 C が小さいほど復元結果の幅が小さくなる. そのため, 正しい反射係数 C を求めなければポリープの絶対的な大きさを推定することはできない. そこで, 血管領域の幅を基準として復元対象全体の幅をミリメートル単位で定義することで, その幅と形状復元結果が一致するように反射係数 C を最適化により推定する. なお, 実行時間を削減するために最適化には形状復元で用いる画像から解像度を下げた圧縮画像を用いる. まず, 大きさの基準となる血管領域の幅をピクセル単位とミリメートル単位で取得し, そこからピクセルとミリメートルの変換基準を求める. ピクセル単位での幅取得には3.3節の手法を用いて, ミリメートル単位での幅は本論文では既知と仮定しているためその情報を用いる. そして, 得られた変換基準により形状復元したい画像全体の幅をpxからmmに変換できる. 反射係数推定の手順を以下に示す.

Step 1. 反射係数 C を任意の値に決定し、三次元形状復元を行う。

Step 2. 幅の誤差を式 (3.26) のように定義する。なお、予測値の幅は式 (3.27)、測定値の幅は式 (3.28) を用いて求める。ここでいう X 座標は Step1 で行なった形状復元結果を用いる。

$$\text{誤差} = \text{予測値} - \text{測定値} \quad (3.26)$$

$$\text{予測値 [mm]} = \text{復元対象の幅 [px]} \times \text{変換基準 [mm/px]} \quad (3.27)$$

$$\text{測定値 [mm]} = X \text{ 座標の最大値 [mm]} - X \text{ 座標の最小値 [mm]} \quad (3.28)$$

このとき、式 (3.29) のように幅の誤差が正のとき反射係数を正の方向に動かし、幅の誤差が負のとき反射係数を負の方向に動かす。

$$C_{n+1} = \begin{cases} C_n + \gamma & (Error > 0) \\ C_n - \gamma & (Error < 0) \end{cases} \quad (3.29)$$

ここで、 γ は反射係数を動かすときのステップサイズである。

Step 3. 誤差の符号が入れ替わるとき動かすステップ幅を減少させる。

$$\gamma = \gamma\alpha \quad (3.30)$$

ここで、 α はステップサイズの減少率である。

Step 4. 幅の誤差が 0.01mm 未満になるまで Step1 から Step3 を繰り返す。

第4章 実験

提案手法の有効性を確認するため、ポリープのシミュレーション画像を用いて実験を行うことで精度を評価した。また、実画像に対しての有効性を確認するために内視鏡画像を用いて実験を行った。

4.1 シミュレーションによる実験

点光源・透視投影条件のシミュレーション画像を作成し、三次元形状復元を行うことで精度を評価した。ここでは、ポリープを半球と見立ててシミュレーション球画像を作成する。

4.1.1 ポリープシミュレーションの作成

本論文では Lambert 画像から FMM の手法を用いて三次元形状復元を行うため、実験に使用するシミュレーション球画像は点光源・透視投影の環境を想定した反射率が一律な Lambert 画像を作成する必要がある。さらに、その精度評価のため形状を表す三次元座標も同時に用意する必要がある。そのため、作成される Lambert 画像の各画素 x, y が三次元空間のどこに位置するかを計算する。内視鏡環境下における透視投影モデルが図 2.1 のように表されるとき、半径が R 、中心が $(0, 0, Z_0)$ の球を想定する。このとき、透視投影における視線ベクトルと球との交点を計算することで、各画素 x, y から奥行き Z を求めることができる。なお、透視投影における視線ベクトルと球との交点を計算には式 (4.1) を用いる。

$$\begin{cases} X^2 + Y^2 + (Z - Z_0)^2 = R^2 \\ X = \frac{x}{f}Z, Y = \frac{y}{f}Z \end{cases} \quad (4.1)$$

ただし、 (x, y) は投影される画像の二次元座標であり、 (X, Y, Z) は球の三次元座標である。式 (4.1) を二次方程式の解の公式を用いて Z について解くと、式 (4.2) が得

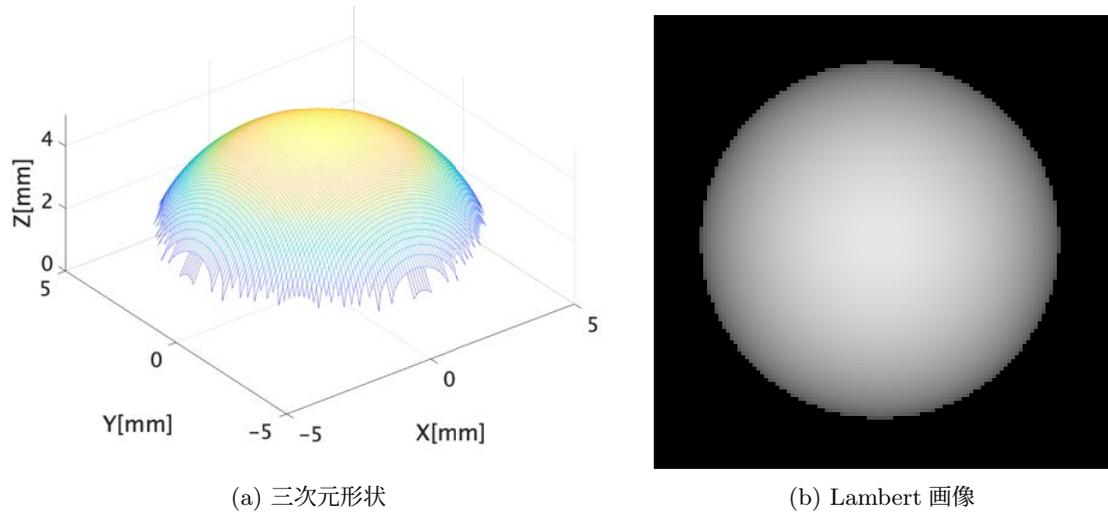


図 4.1: シミュレーションポリープ

られる．解の公式では解が2つ得られるが，1つは視線には写らない視線ベクトルと球の奥側の交点なため，マイナスの解を選択する．

$$Z = \frac{f^2 Z_0 - \sqrt{f^4 Z_0^2 - f^2(x^2 + y^2 + f^2)(Z_0 - R^2)}}{x^2 + y^2 + f^2} \quad (4.2)$$

さらに，式(4.2)は奥行きを表すものなので，式(4.3)を用いて球の中心を原点とした高さの式に変換する．

$$Z = Z_0 - \frac{f^2 Z_0 - \sqrt{f^4 Z_0^2 - f^2(x^2 + y^2 + f^2)(Z_0 - R^2)}}{x^2 + y^2 + f^2} \quad (4.3)$$

また，Lambert 画像の輝度値計算には光学的制約式(3.16)を輝度値 E について解いた式(4.4)を使用する．

$$E = C \frac{f^2(-px - qy + f)}{(x^2 + y^2 + f^2)^{\frac{3}{2}} Z^2 (p^2 + q^2 + 1)^{\frac{1}{2}}} \quad (4.4)$$

ここで，球の半径 $R = 5[\text{mm}]$ ，中心座標 $Z_0 = 15$ ，焦点距離 $f = 10[\text{mm}]$ ，反射係数 $C = 90$ ，Lambert 画像サイズ $120 \times 120[\text{px}]$ として，式(4.3)，式(4.4)を用いてシミュレーション画像を作成する．作成したシミュレーション球の三次元形状を図4.1a，ランバート画像を図4.1bに示す．

4.1.2 血管領域シミュレーションの作成

ポリープの大きさを推定するために、参照物体として用いる血管領域がカメラに対して水平な場合と傾いている場合の2つのシミュレーション画像を作成する。血管領域の幅は実際には0.2[mm]程度と言われているが、シミュレーションでは低い解像度で実験を行うため、見やすくするために1[mm]程度として作成する。なお、三次元形状の計算はポリープシミュレーションの作成と同様の手順で求める。血管領域が位置する形状は平面だと考えて、点 $P = (X_0, Y_0, Z_0)$ を通り法線ベクトル $\mathbf{n} = (n_x, n_y, n_z)$ である平面の方程式と視線ベクトルとの交点を式 (4.5) から求める。

$$\begin{cases} n_x(X - X_0) + n_y(Y - Y_0) + n_z(Z - Z_0) = 0 \\ X = \frac{x}{f}Z, Y = \frac{y}{f}Z \end{cases} \quad (4.5)$$

式 (4.5) を Z について解き、その後高さの式に変換した結果を式 (4.6) に示す。

$$Z = Z_0 - \frac{f(n_x X_0 + n_y Y_0 + n_z Z_0)}{n_x x + n_y y + n_z f} \quad (4.6)$$

また、血管領域シミュレーションでは Lambert 画像は作成せず、二値化画像を作成する。内視鏡画像の場合は U-Net を用いて血管検出を行うが、シミュレーションでは既に血管領域が検出されたものとして扱う。

ここで、点 $P = (0, 0, 13)$ 、 $\mathbf{n} = (0, 0, 1)$ 、 $Z_0 = 15$ 、焦点距離 $f = 10$ [mm]、画像サイズ 120×120 [px] として、式 (4.6) を用いて水平な場合のシミュレーション画像を作成する。平面の高さは図 4.1a においてシミュレーション球の底面が高さ $Z = 2$ に位置するためそれに合わせるように作成する。作成した三次元形状を図 4.2a、血管領域画像を図 4.2b に示す。シミュレーション上での血管領域の幅は 0.975[mm] としている。さらに、点 $P = (5, 0, 8)$ 、 $\mathbf{n} = (1/2, 0, \sqrt{3}/2)$ 、 $Z_0 = 15$ 、焦点距離 $f = 10$ [mm]、画像サイズ 120×120 [px] として、傾いている場合のシミュレーション画像も同様に作成する。この平面の方程式は Y 軸周りに 30 度傾いた平面である。作成した三次元形状を図 4.3a、血管領域画像を図 4.3b に示す。シミュレーション上での血管領域の幅は 0.8728[mm] としている。

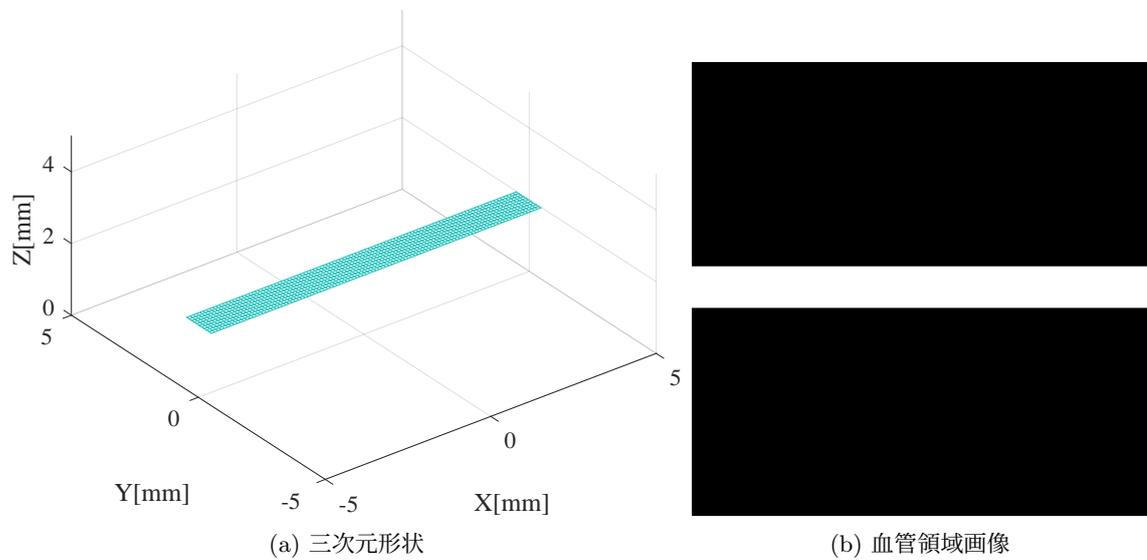


図 4.2: シミュレーション血管領域（傾きなし）

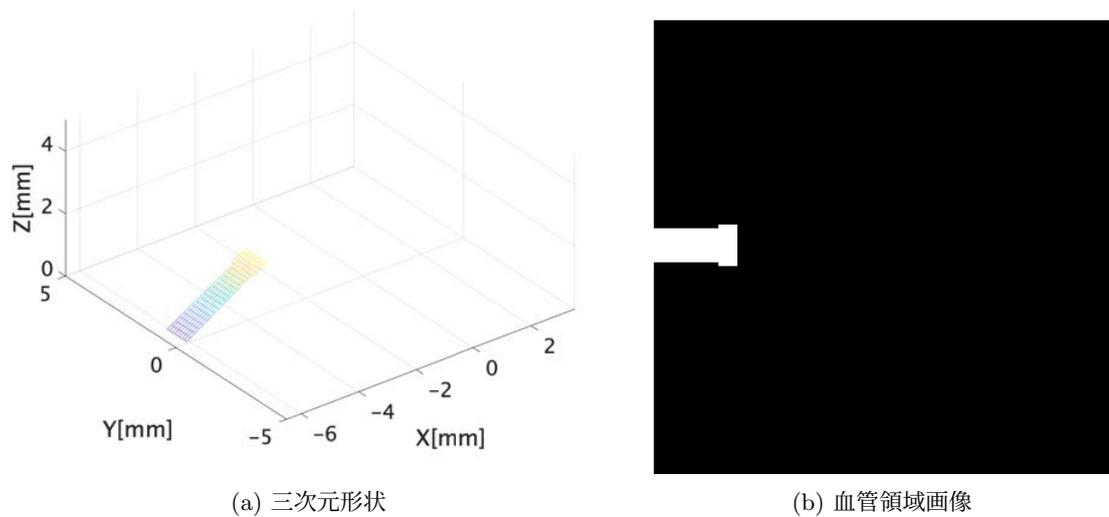


図 4.3: シミュレーション血管領域（傾きあり）

4.1.3 血管領域の幅取得

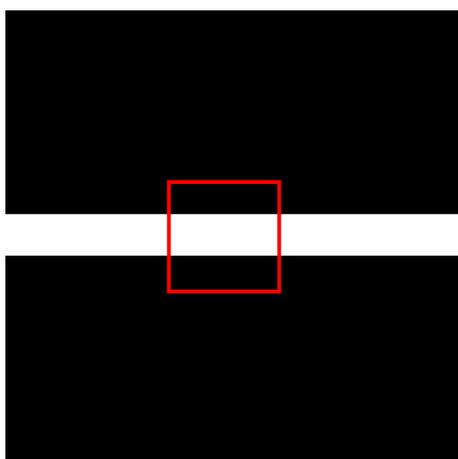
血管領域の傾き補正の有効性を示すため、ここでは傾き補正しない場合も検証する。傾き補正を行わずに血管領域画像から直接血管領域の幅をピクセル値で取得する場合、図 4.3b のどの箇所の幅を計測するかによって結果が異なる。本論文ではポリープ底面の高さ $Z = 2$ を基準としているため、血管領域の中でも高さが $Z = 2$ に

近い図 4.3b の右側の太い箇所から取得する。また、比較のため図 4.3b 全体からも取得する。図 4.4a, 図 4.5a, 図 4.5c に血管領域の幅取得にしようとした範囲を示す。さらにこれらを用いて、血管領域を線形回帰させた結果を図 4.4b, 図 4.5b, 図 4.5d に示す。

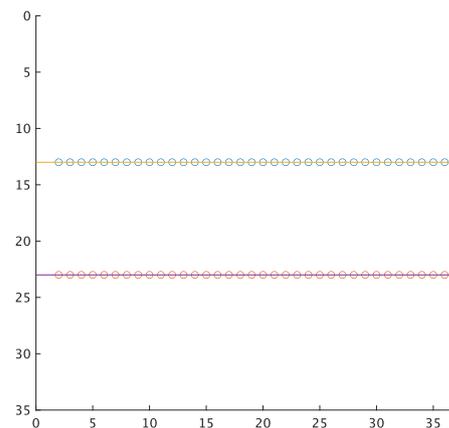
次に、傾き補正を行なった後に血管領域の幅をピクセル値で取得する。任意の反射係数 C を与えて、傾きあり血管領域の Lambert 画像 4.6a から FMM による形状復元を行なった結果を図 4.6b に示す。ここでは、反射係数 $C = 300$ としている。また、式 (3.4) を用いて平面近似した結果を図 4.6c に示す。その結果に対して、図 4.3b をマスク画像として血管領域の形状を抽出した結果を図 4.6d に示す。血管領域の三次元形状の各点の傾き p, q から式 (3.3) を用いて天頂角 $\alpha = 142.5^\circ$ 、方位角 $\beta = 0.0134^\circ$ を取得する。そこからアフィン変換による傾き補正を行なった結果を図 4.6e, 線形回帰の結果を図 4.6f に示す。最後に、式 (3.10) を用いて画像面上の血管領域の幅をピクセル単位に変換する。取得した血管領域幅の結果を表 4.1 にまとめる。

表 4.1: ピクセル単位の血管領域の幅

傾きなし (図 4.4b)	傾きあり		
	領域一部 (図 4.5b)	領域全体 (図 4.5d)	アフィン変換 (図 4.6e)
10.00[px]	10.00[px]	8.381[px]	9.154[px]

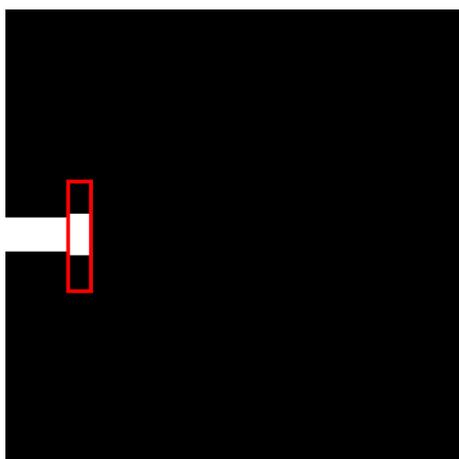


(a) クロップ箇所

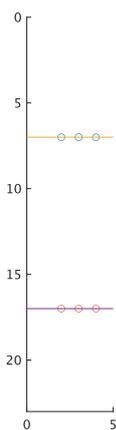


(b) 線形回帰画像

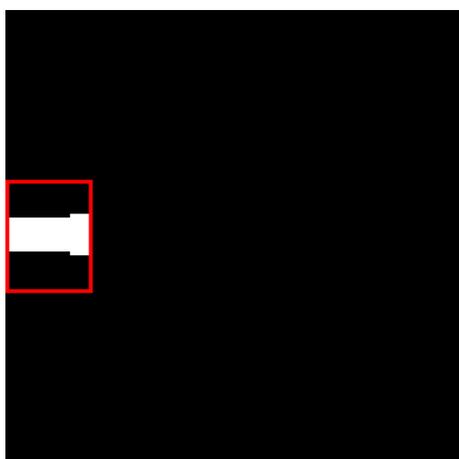
図 4.4: シミュレーション血管領域 (傾きなし)



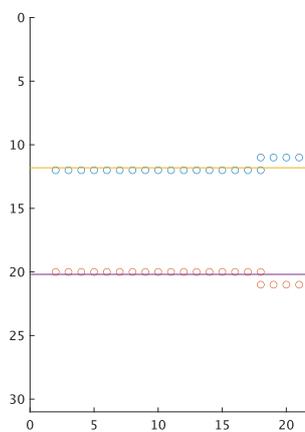
(a) クロップ箇所



(b) 線形回帰画像

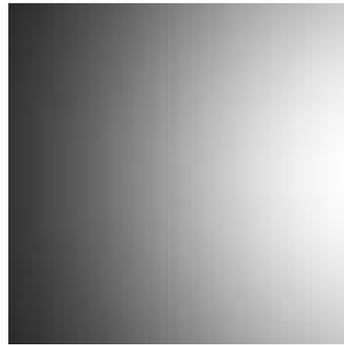


(c) クロップ箇所

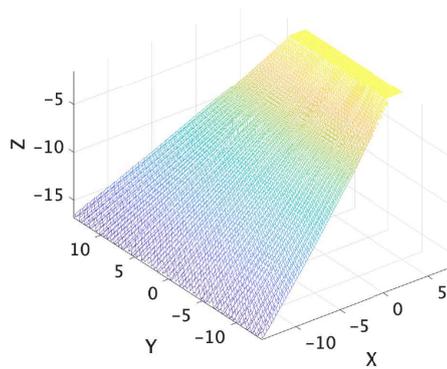


(d) 線形回帰画像

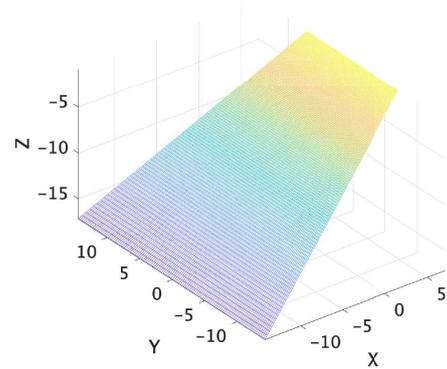
図 4.5: シミュレーション血管領域 (傾きあり)



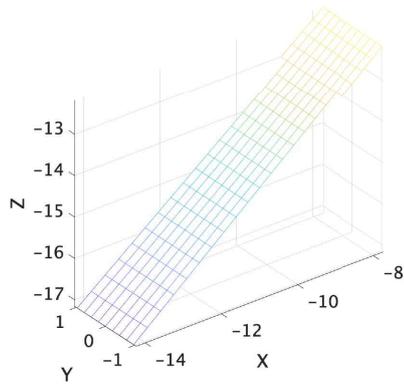
(a) ランバート画像



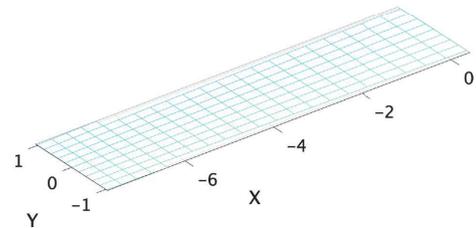
(b) 三次元形状復元



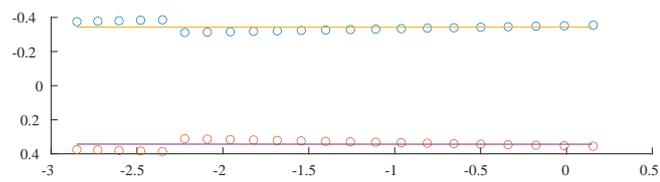
(c) 平面近似



(d) 血管領域形状



(e) 傾き補正後形状



(f) 線形回帰画像

図 4.6: シミュレーション血管領域 (傾き補正)

4.1.4 反射係数の推定及び形状復元

図 4.1b から円の直径を測ると、復元対象の幅は 94[px] とわかる。既知とする血管領域の幅 [mm] と計測した血管領域の幅 [px] から単位の変換基準を計算し、式 (3.27) を用いて形状復元される大きさの予測値 [mm] を計算する。その情報をもとに反射係数 C の推定及び形状復元を行う。形状復元の結果を図 4.7 に示す。また、最適化パラメータの初期値を表 4.2、最適化の結果を表 4.3 に示す。なお、X、Y 軸方向の幅はそれぞれの軸方向の端中央からの 5 点平均の差から計算し、Z 軸方向の高さは半球の頂点の高さと半球の底面となる円周上の点の高さ平均との差から計算する。

血管領域がカメラに対して傾いていない場合において、反射係数の誤差は 1.07%、幅の誤差は X 軸、Y 軸でそれぞれ 0.163%、0.958% となった。このことから、反射係数、形状の幅はともに真値に近い数値となっており、提案手法の有効性が確認できた。しかし、Z 軸方向の高さは約 0.4[mm] 程度の誤差が生じてしまっている。これは、FMM は復元初期点となる半球の頂点から全体を形状復元するため、初期点から離れた点ほど徐々に誤差が蓄積してしまうことによるものだと考えられる。また、血管領域が傾いている場合における反射係数、形状の幅の 3 つの結果を比較すると、アフィン変換の結果が最も真値に近い数値になっており、反射係数の誤差は 4.01%、幅の誤差は X 軸、Y 軸でそれぞれ 2.42%、0.272% となった。これにより、アフィン変換による血管領域の傾き補正の有効性が確認できた。アフィン変換補正の結果が傾きなしの結果と比較して、反射係数、形状の幅の精度が下がっているのは血管領域の三次元形状復元に誤差があるためだと考えられる。一方で、Z 軸方向の高さは領域一部の結果が最も誤差が少ない。これは先ほど述べた FMM の誤差と大きさが小さく復元されたことによる誤差が相殺して誤差が少なくなったものだと考えられる。

表 4.2: パラメータの初期値

パラメータ	値
反射係数 C	150
ステップサイズ γ	25
ステップサイズ減少率 α	0.8
誤差の許容値 [mm]	0.01

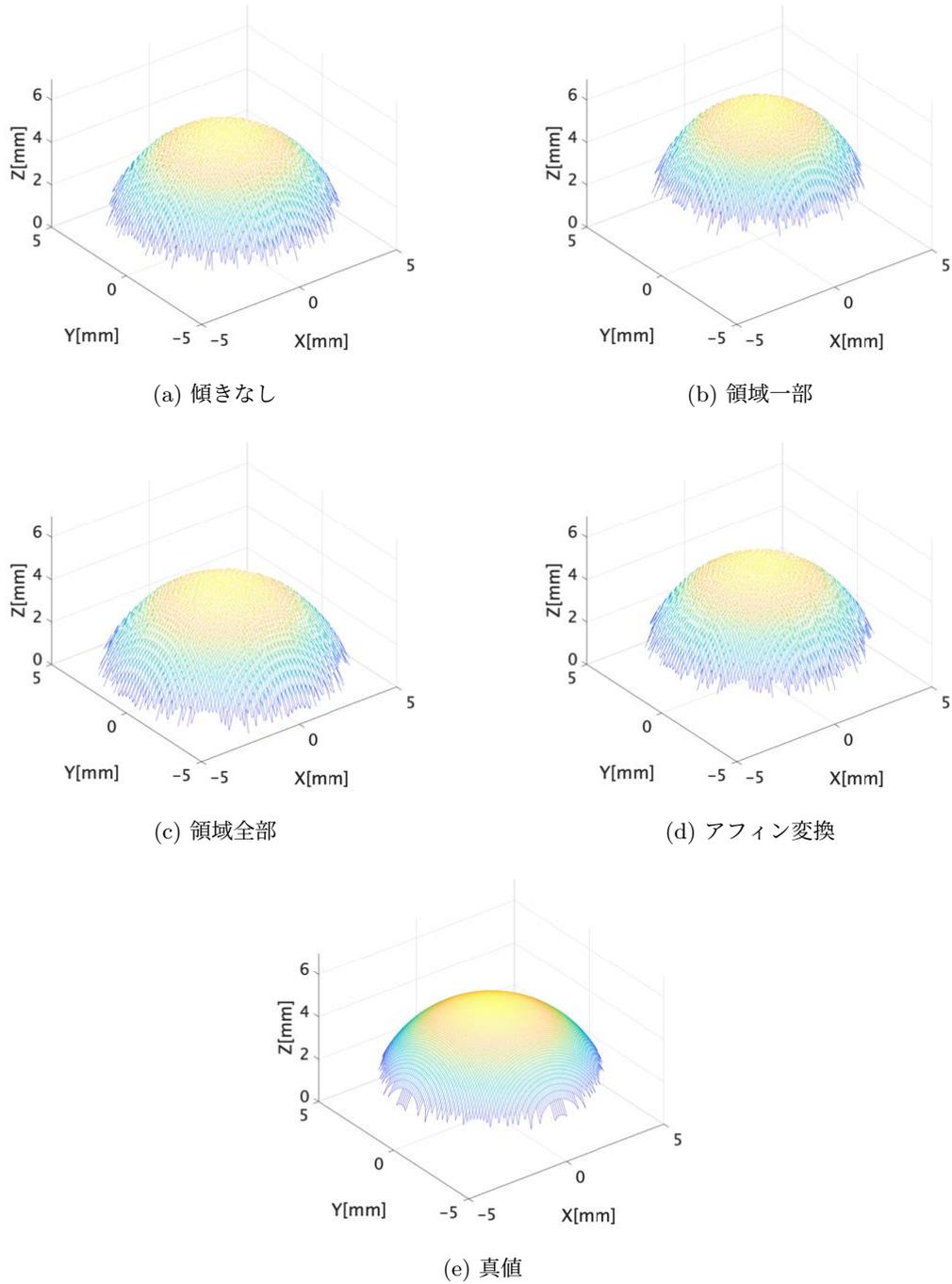


図 4.7: 形状復元結果

表 4.3: 反射係数推定及び形状復元結果

	傾きなし	傾きあり			真値
		領域一部	領域全体	アフィン変換	
血管領域の幅 [mm]	0.9750	0.8728	0.8728	0.8728	-
血管領域の幅 [px]	10.00	10.00	8.381	9.154	-
復元幅の予測値 [mm]	9.165	8.204	9.789	8.962	-
反射係数 C	90.96	71.05	104.8	86.39	90.00
X 軸方向の幅 [mm]	9.168	8.206	9.786	8.961	9.183
Y 軸方向の幅 [mm]	9.271	8.188	9.787	9.208	9.183
Z 軸方向の高さ [mm]	3.129	2.880	3.292	3.070	2.690

4.2 実画像による実験

内視鏡画像を用いて実験を行う。実験で用いる内視鏡画像は図 4.8a, 図 4.9a である。撮影環境は点光源, 透視投影であり, 画像サイズは 863×863 [px] である。

4.2.1 内視鏡画像の前処理

形状復元を行う前処理として, 入力する内視鏡画像から血管領域抽出, ランバート画像生成を行う。図 4.8a に対して U-Net による血管検出を行った結果が図 4.8b である。血管領域がグレー画像で抽出できていることがわかるが, 血管抽出画像から明確な血管領域を抽出するため, 大津の二値化を適応する。その結果を図 4.8c に示す。ここで, 図 4.8d の赤枠で示す箇所のように可能な限りポリープに近い血管領域をクロップする。可能な限りポリープに近い血管領域を選択するのは, ポリープが位置する底面を高さのベースとしており, 参照物体も同じ高さに存在するものを使用することが望ましいためである。クロップした結果を図 4.8e に示す。

次に, 論文 [38] の手法を用いて鏡面反射除去とランバート画像の作成を行う。図 4.8a に対して鏡面反射除去を行なった結果が図 4.8f であり, その後ランバート画像を生成した結果が図 4.8g である。ここから先ほど選択した血管領域を囲む範囲で血管領域をクロップする。クロップ箇所を図 4.8h, クロップした結果を図 4.8i に示す。なお, クロップする画像サイズは 120×120 [px] である。ここまでの処理を同様に図 4.9a にも行う。その結果を図 4.9 に示す。

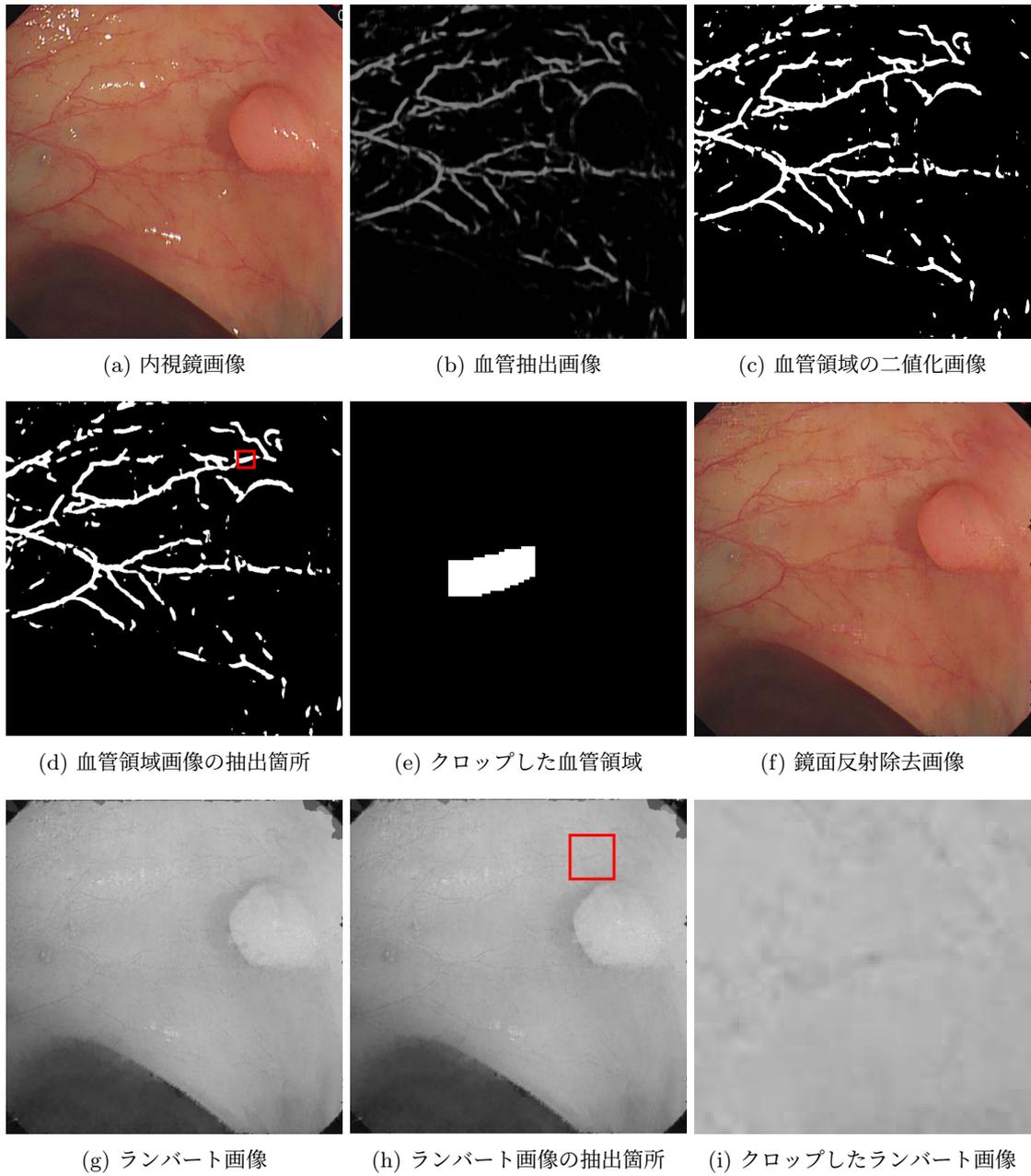


図 4.8: 形状復元の前処理 1

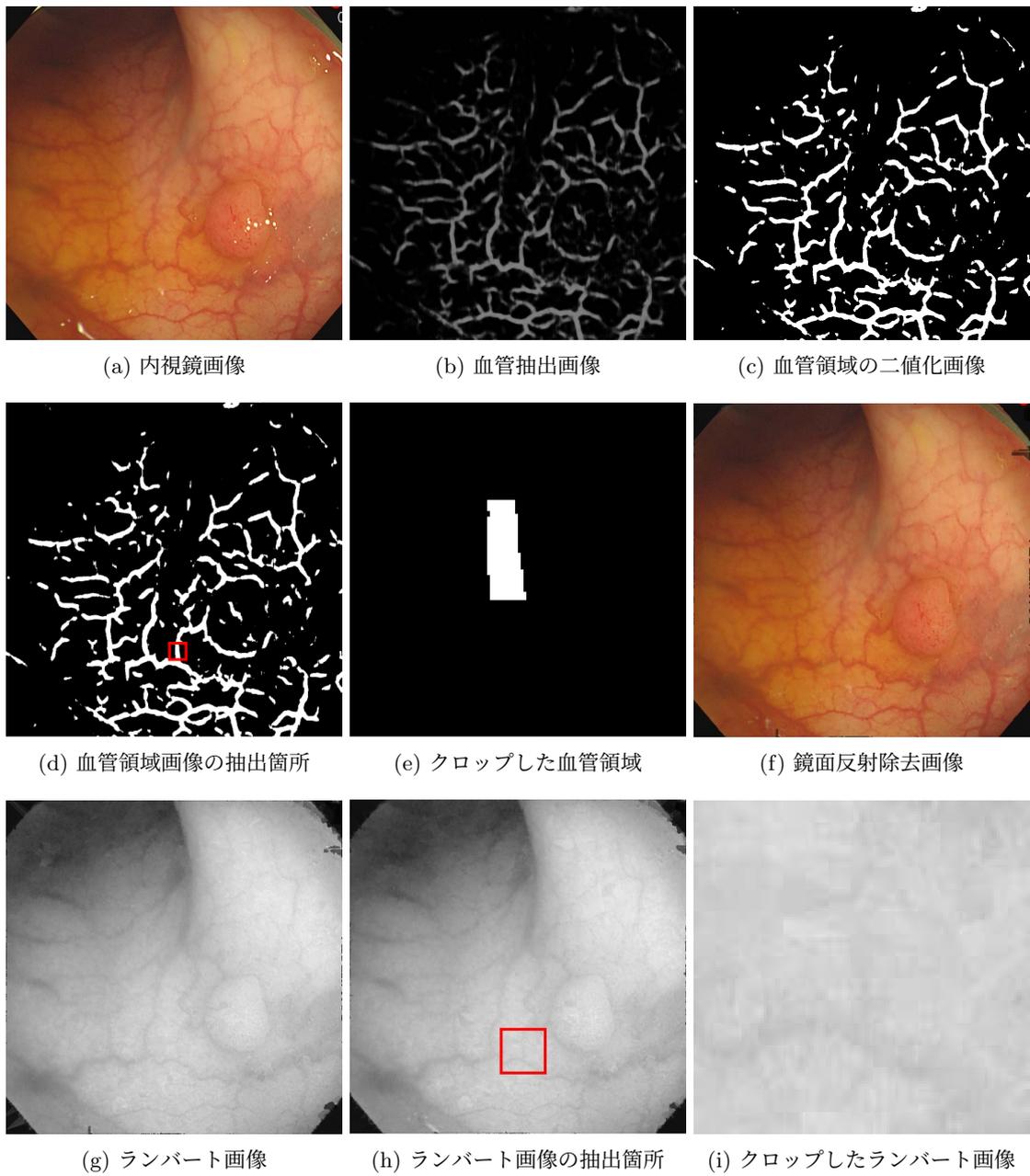


図 4.9: 形状復元の前処理 2

4.2.2 血管領域の幅取得

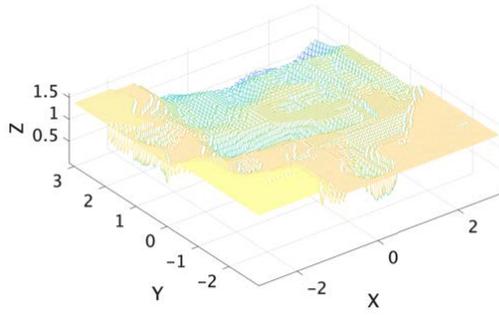
図 4.8e, 図 4.8i を用いて血管領域の傾きをカメラに対して水平に位置するように補正し, そこから血管領域の幅を取得する. 任意の反射係数 C を与えて, 図 4.8i を形状復元した結果を図 4.10a, さらに式 (3.4) を用いて平面近似を行なった結果を図 4.10b に示す. ここでは, 反射係数 $C = 300$ としている. その結果に対して, 図 4.8e をマスク画像として血管領域の形状を抽出した結果を図 4.10c に示す. 血管領域の三次元形状の各点の傾き p , q から式 (3.3) を用いて天頂角 α , 方位角 β を取得する. そこからアフィン変換による傾き補正を行なった結果を図 4.10d, 線形回帰の結果を図 4.10e に示す. 最後に, 式 (3.10) を用いて画像面上の血管領域の幅をピクセル単位に変換する. ここまでの処理を同様に図 4.9e, 図 4.9i に対しても行う. その結果を図 4.11 に示す. また, ここまでのパラメータの結果を表 4.4 にまとめる.

表 4.4: 血管領域の幅取得におけるパラメータ

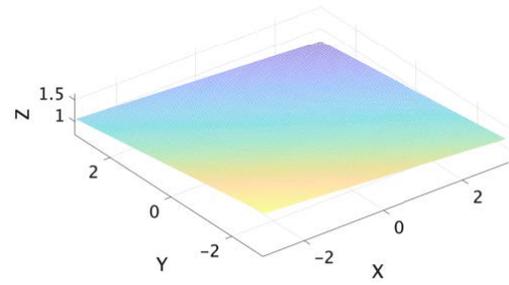
	内視鏡画像 1	内視鏡画像 2
X 軸方向の傾き p [度]	3.565	0.062
Y 軸方向の傾き q [度]	5.372	3.595
天頂角 α [度]	6.420	3.591
方位角 β [度]	56.43	89.01
血管領域の幅 [px]	11.84	10.42

4.2.3 反射係数の推定及び形状復元

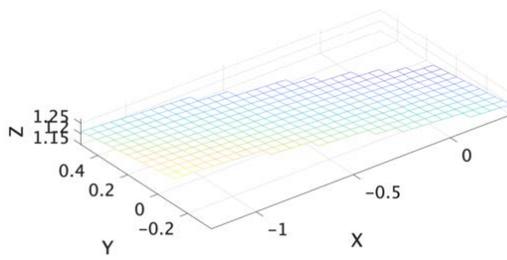
図 4.8g, 図 4.9g からポリープ部分をクロップし, 三次元形状復元を行う. クロップした画像は図 4.12a, 図 4.13a であり, サイズはそれぞれ 250×250 [px], 220×220 [px] である. 既知とする血管領域の幅 [mm] と計測した血管領域の幅 [px] から単位の変換基準を計算し, 式 (3.27) を用いて形状復元される大きさの予測値 [mm] を計算する. ここで, 最適化の計算時間を削減するために入力画像から解像度を 120×120 [px] に圧縮した画像を生成する. この圧縮画像を用いて先ほど定義した予測値と一致するように反射係数 C を推定する. 形状復元の結果を図 4.12, 図 4.13 に示す. また, 表 4.2 の初期パラメータを用いて最適化した結果を表 4.5 に示す. この結果から, 図 4.12a, 図 4.13a のポリープサイズはともに約 4 [mm] と得ることができた.



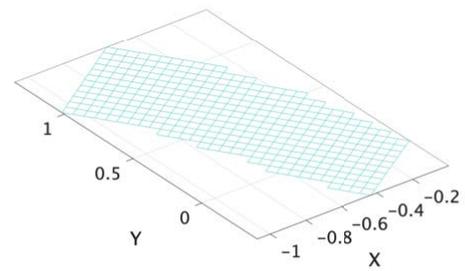
(a) 三次元形状復元



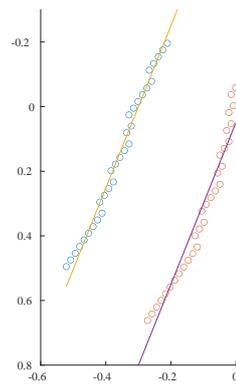
(b) 平面近似



(c) 血管領域形状

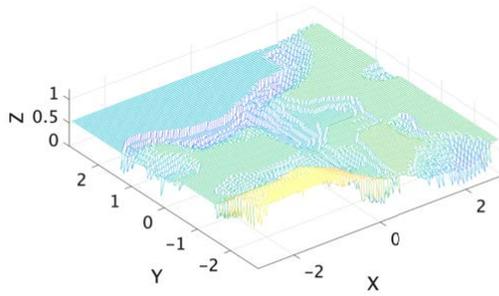


(d) 傾き補正後形状

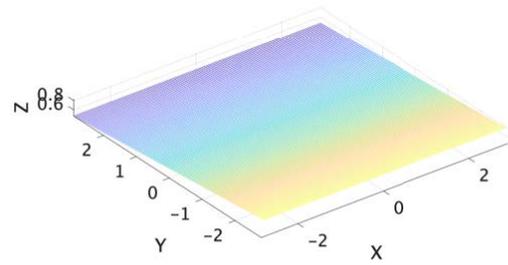


(e) 線形回帰画像

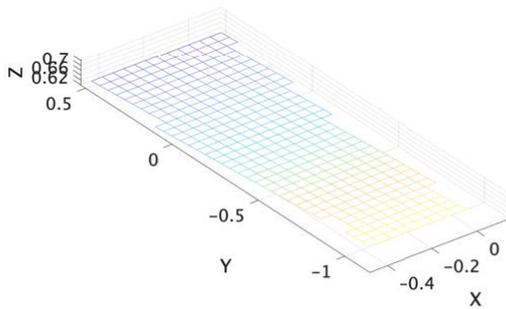
図 4.10: 実画像 1 の血管領域補正



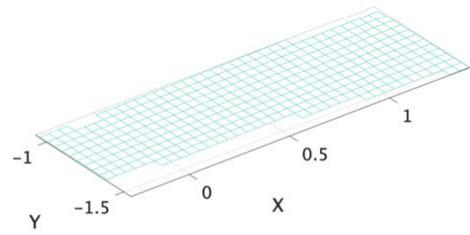
(a) 三次元形状復元



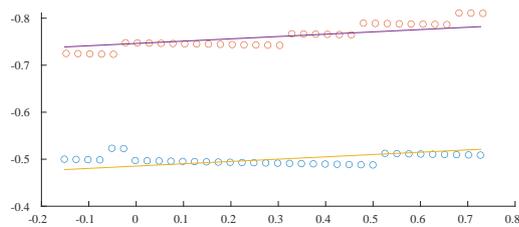
(b) 平面近似



(c) 血管領域形状



(d) 傾き補正後形状



(e) 線形回帰画像

図 4.11: 実画像 2 の血管領域補正

表 4.5: 反射係数推定及び形状復元結果

	内視鏡画像 1	内視鏡画像 2
血管領域の幅 [mm]	0.200	0.200
血管領域の幅 [px]	11.84	10.42
復元幅の予測値 [mm]	4.206	4.202
反射係数 C	148.0	165.6

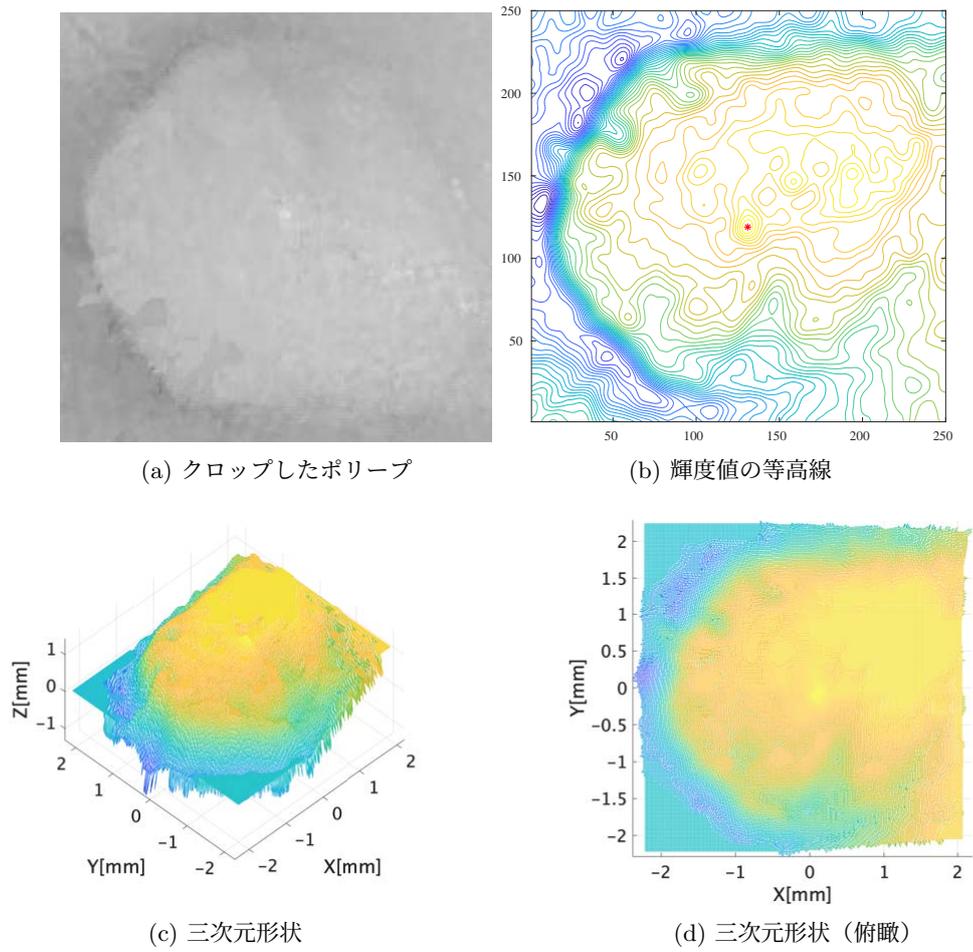


図 4.12: 実画像 1 の三次元形状復元復元

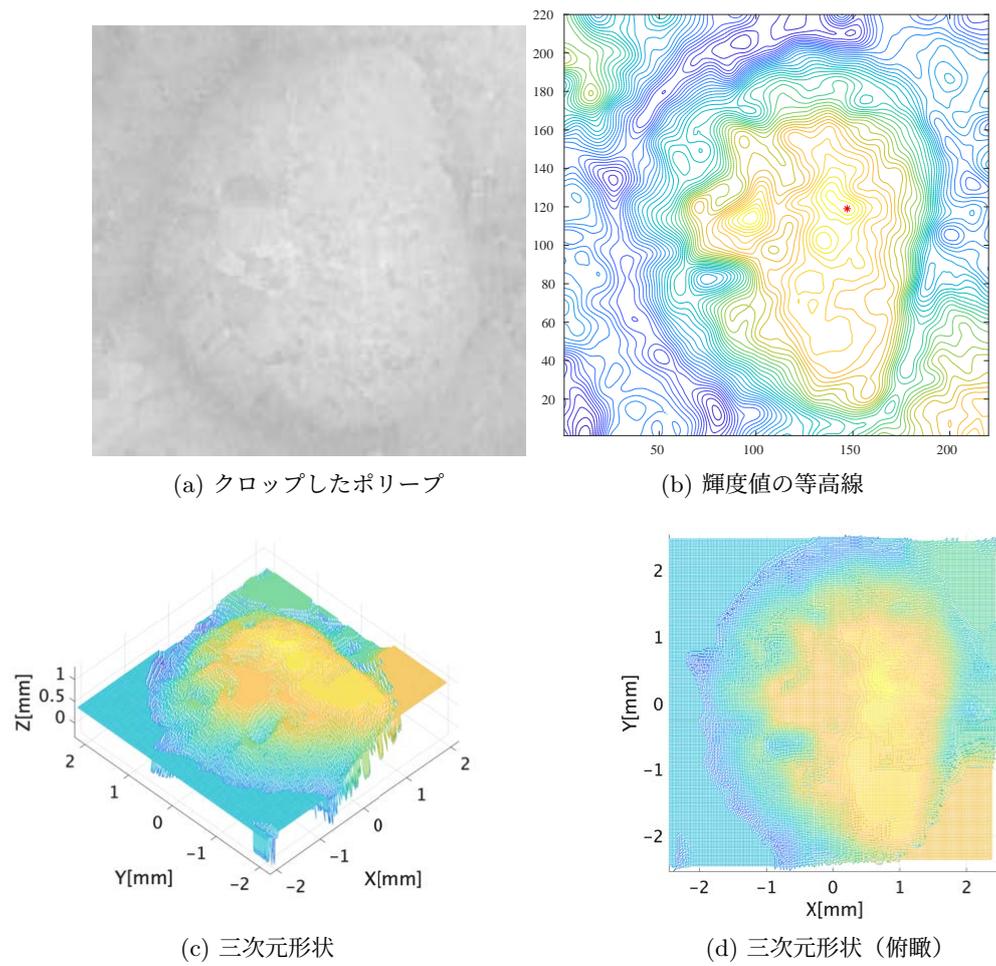
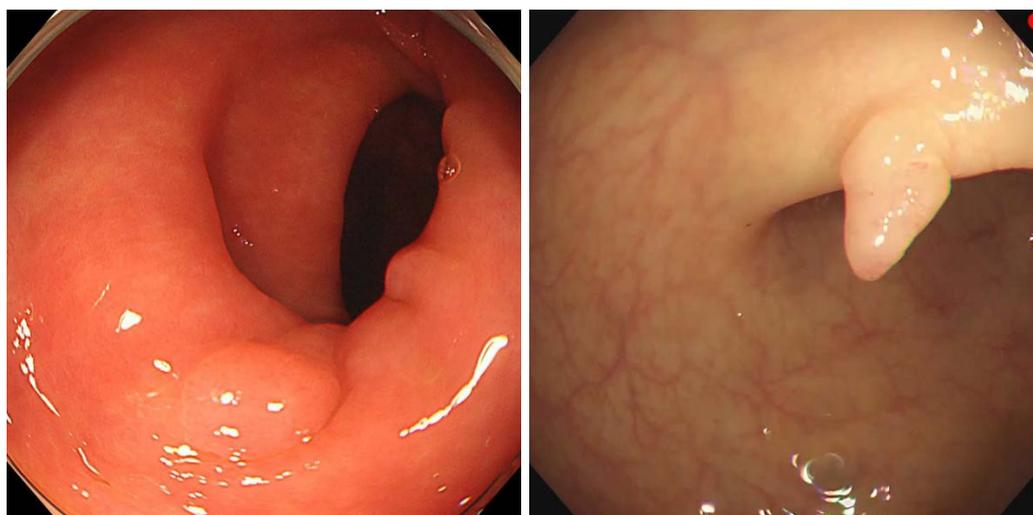


図 4.13: 実画像 2 の三次元形状復元復元

第5章 むすび

本論文では、1枚の内視鏡画像から高速な形状復元手法を用いたポリープ形状復元及び絶対的な大きさ推定手法を提案した。これにより、内視鏡画像内にカメラに対して水平な血管領域が存在しない状況であっても、1枚の内視鏡画像から物体表面の反射係数が推定できる。シミュレーション実験では、アフィン変換による傾き補正を行った反射係数の誤差が4.01%、幅の誤差がX軸、Y軸でそれぞれ2.42%、0.272%となっていることから、血管領域の傾き補正、反射係数推定手法の有効性を示した。また実画像実験において、ポリープの形状復元及び絶対的な大きさ推定ができていることから、提案手法の有効性を示した。

今後の課題として、本手法が適応できない例について述べる。内視鏡画像5.1aでは、大きさの指標となる血管が写っていないため、ポリープの絶対的な大きさを求めることは難しい。また内視鏡画像5.1bでは、血管が画像奥にしか存在せず、ポリ-



(a) 内視鏡画像 1

(b) 内視鏡画像 2

図 5.1: 提案手法が適応できない例

プが位置する周囲には血管が存在しない。本手法ではポリープが位置する底面を高さのベースとしており、そこから遠く離れた参照物体を利用すると大きさ推定の精度が下がってしまうと考えられる。

今後の展望としては、様々なシーンでの追加検証を行い実画像実験による精度評価を行うことが挙げられる。また、本手法では血管領域の幅を医師の経験的な判断により決定しているが、その幅仮定の妥当性も検証していきたい。

謝辞

本研究を行うにあたり、終始熱心にご指導を頂いた名古屋工業大学准教授 舟橋健司先生に深く感謝いたします。また、日頃よりご指導いただきました中部大学教授 岩堀 祐之先生に感謝いたします。

なお、本研究において、内視鏡画像の提供をはじめ、ご討論、ご協力を頂きました愛知医科大学教授 春日井 邦夫先生と特任教授 小笠原 尚高先生に感謝いたします。

さらに、本研究を進めるにあたり多くの助言を頂き、御協力いただいた名古屋工業大学 舟橋研究室の皆様、および中部大学 岩堀研究室の皆様に深く感謝し、厚く御礼申し上げます。

参考文献

- [1] H. Itoh, M. Oda, K. Jiang, Y. Mori, M. Misawa, S. Kudo, K. Imai, S. Ito, K. Hotta, K. Mori , “Binary polyp-size classification based on deep-learned spatial information”, *International Journal of Computer Assisted Radiology and Surgery*, vol.16 , pp.1817–1828, 2021.
- [2] C. Chen, Y. F. Zheng, “Passive and active stereo vision for smooth surface detection of deformed plates”, *IEEE Transactions on Industrial Electronics*, vol.42, no.3, pp.300-306, 1995.
- [3] R. Munguia, A. Grau, “Monocular SLAM for Visual Odometry: A Full Approach to the Delayed Inverse-Depth Feature Initialization Method”, *Mathematical Problems in Engineering*, vol.2012, 2012.
- [4] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, J. D. Tardós, “ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM”, *IEEE Transactions on Robotics*, vol.37, no.6, pp.1874-1890, 2021.
- [5] S. Sumikura, M. Shibuya, K. Sakurada, “OpenVSLAM: A versatile visual SLAM framework”, *Proceedings of the 27th ACM International Conference on Multimedia*, pp.2292–2295, 2019.
- [6] M. Labbe, and F. Michaud, “RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation”, *Journal of Field Robotics*, vol.36, no.2, pp.416-446, 2018.
- [7] E. Rosten, T. Drummond, “Machine learning for high-speed corner detection”, *European Conference on Computer Vision*, vol.1, pp.430–443, 2006.

- [8] D. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, vol.60, no.2, pp.91–110, 2004.
- [9] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, “SURF: Speeded up robust features”, *Computer Vision and Image Understanding*, vol.110, no.3, pp.346–359, 2008.
- [10] M. Muja, D. G. Lowe, “Fast approximate nearest neighbors with automatic algorithm configuration”, *VISAPP International Conference on Computer Vision Theory and Applications*, pp.331–340, 2009.
- [11] M. Calonder, V. Lepetit, C. Strecha, P. Fua, “BRIEF: Binary robust independent elementary features”, *European Conference on Computer Vision*, pp.778–792, 2010.
- [12] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, “ORB: An efficient alternative to SIFT or SURF”, *IEEE International Conference on Computer Vision*, pp.2564–2571, 2011.
- [13] S. Leutenegger, M. Chli, R. Y. Siegwart, “BRISK: Binary robust invariant scalable keypoints”, *IEEE International Conference on Computer Vision*, pp.2548–2555, 2011.
- [14] A. Alahi, R. Ortiz, P. Vandergheynst, “FREAK : Fast Retina Keypoint”, *IEEE International Conference on Computer Vision and Pattern Recognition*, pp.510–517, 2012.
- [15] A. Gionis, P. Indyk, R. Motwani, “Similarity search in high dimensions via hashing”, *Proceedings of International Conference on Very Large Data Bases*, pp.518–529, 1999.
- [16] J. Engel, T. Schops, D. Cremers, “LSD-SLAM: Large-Scale Direct Monocular SLAM”, *European Conference on Computer Vision*, pp.834–849, 2014.
- [17] B. K. P. Horn, “Obtaining shape from shading information”, *MIT Press*, pp.123–171, 1989.

- [18] P. W. Verbeek, B. J. H. Verwer, “Shading from shape, the eikonal equation solved by gray-weighted distance transform”, *Pattern Recognition Letters*, vol.63, pp.681–690, 1990.
- [19] J. A. Sethian, “A Fast Marching Level Set Method for Monotonically Advancing Fronts”, *Proceedings of the National Academy of Sciences of the United States of America*, vol.93, no.4, pp.1591–1593, 1996.
- [20] R. Kimmel, J. A. Sethian, “Optimal Algorithm for Shape from Shading and Path Planning”, *Journal of Mathematical Imaging and Vision*, vol.14, no.3, pp.237–244, 2001.
- [21] A. Tankus, N. Sochen, Y. Yeshurun, “Shape-from-Shading Under Perspective Projection”, *International Journal of Computer Vision*, vol.63, no.1, pp.21-43, 2005.
- [22] S. Y. Yuen, Y. Y. Tsui, C. K. Chow, “A Fast Marching Formulation of Perspective Shape from Shading under Frontal Illumination”, *Pattern Recognition Letters*, vol.28, no.7, pp.806-824, 2007.
- [23] Y. Iwahori, K. Iwai, R. J. Woodham, H. Kawanaka, S. Fukui, K. Kasugai, “Extending Fast Marching Method under Point Light Source Illumination and Perspective Projection”, *ICPR2010*, pp.1650-1653, 2010.
- [24] K. Tatematsu, Y. Iwahori, T. Nakamura, S. Fukui, R. J. Woodham, K. Kasugai, “Shape from Endoscope Image based on Photometric and Geometric Constraints”, *Procedia Computer Science*, vol.22, pp.1285-1293, 2013.
- [25] H. Usami, Y. Iwahori, M. K. Bhuyan, A. Wang, N. Ogasawara, K. Kasugai, “Recovering Polyp Shape from an Endoscope Image Using Two Light Sources”, *International Journal of Software Innovation*, vol.5, no,2, pp.33-54, 2017.

- [26] Y. Iwahori, S. Miyazaki, H. Usami, Y. Hanai, B. Kijirikul, K. Kasugai, “Automatic Detection of LST-Type Polyp by CNN Using Depth Map”, Handbook of Artificial Intelligence in Healthcare, pp.177-196, 2021.
- [27] Y. Iwahori, D. Yamaguchi, T. Nakamura, B. Kijirikul, M. K. Bhuyan, K. Kasugai, “Estimating Reflectance Parameter of Polyp Using Medical Suture Information in Endoscope Image”, ICPRAM 2016, pp.1-7, 2016.
- [28] Y. Iwahori, T. Suda, K. Funahashi, H. Usami, A. Wang, M. K. Bhuyan, K. Kasugai, “Shape Recovery of Polyp from Endoscope Image Using Blood Vessel Information”, Computational Science/Intelligence and Applied Informatics, pp.165-184, 2017.
- [29] S. Emoto, Y. Iwahori, K. Funahashi, H. Usami, N. Ogasawara, K. Kasugai, “Graph Matching Approach between Endoscope Images for Non-Rigid Motion using Blood Vessel Structure”, Procedia Computer Science, vol.176, pp.1754-1762, 2020.
- [30] H. Usami, Y. Iwahori, M. K. Bhuyan, A. Wang, N. Ogasawara, K. Kasugai, “Polyp Shape Recovery using Vascular Border from Single Colonoscopy Image”, BIOIMAGING 2019, pp.104-111, 2019.
- [31] A. F. Frangi, W.J. Niessen, K. L. Vincken, M. A. Viergever, “Multiscale vessel enhancement filtering”, Medical Image Computing and Computer-Assisted Intervention, pp.130-137, 1998.
- [32] M. Golhar, Y. Iwahori, K. Funahashi, M. K. Bhuyan, K. Kasugai, “A Robust Method for Blood Vessel Extraction in Endoscopic Images with SVM-based Scene Classification”, ICPRAM 2017, pp.1-10, 2017.
- [33] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, MICCAI 2015, pp.234-241, 2015.

- [34] orbix Srl, “Retina blood vessel segmentation with a convolutional neural network”, <https://github.com/orobix/retina-unet>, 2016.
- [35] K. Schluns, A. Koschan, “Global And Local Highlight Analysis In Color Images”, 2000.
- [36] Y. Iwahori, T. Nakagawa, R. J Woodham, S. Fukui, H. Kawanaka, “Shape from Self-Calibration and Fast Marching Method”, ICPR 2008, pp.1-4, 2008.
- [37] Y. Shimasaki, Y. Iwahori, D. R. Neog, R. J. Woodham, M. K. Bhuyan, “Generating Lambertian Image with Uniform Reflectance for Endoscope Image”, IWAIT2013, pp.1-6, 2013.
- [38] N. Ikeda, H. Usami, Y. Iwahori, B. Kijirikul, K. Kasugai, “Generating Lambertian Image by Removing Specular Reflection Component and Difference of Reflectance Factor Using HSV”, ITC-CSCC 2016, pp.547-550, 2016.

発表論文リスト

投稿論文

1. S. Emoto, Y. Iwahori, K. Funahashi, H. Usami, N. Ogasawara, K. Kasugai, “Graph Matching Approach between Endoscope Images for Non-Rigid Motion using Blood Vessel Structure”, *Procedia Computer Science*, vol.176, pp.1754-1762, 2020.
2. 江本 峻, 岩堀 祐之, 舟橋 健司, 宇佐 美裕康, 小笠原 尚高, 春日井 邦夫, “血管構造に基づく非剛体運動のための画像マッチング手法”, *MIRU 2020*, IS2-2-22, 2020.

口頭発表

1. 江本 峻, 岩堀 祐之, 舟橋 健司, 宇佐 美裕康, 小笠原 尚高, 春日井 邦夫, “内視鏡画像の血管構造に着目したグラフマッチング手法”, *WiNF 2020*, S-1A-2, 2020.