

〈一般研究課題〉 内視鏡手術における
術野の3次元復元の精度向上について
助成研究者 愛知工業大学 北坂 孝幸



内視鏡手術における 術野の3次元復元の精度向上について

北坂 孝幸
(愛知工業大学)

Improving the accuracy of 3D reconstruction of the surgical field in endoscopic surgery

Takayuki Kitasaka
(Aichi Institute of Technology)

Abstract :

In this paper, we propose a method for 3D reconstruction of the surgical field in endoscopic surgery. The endoscopic camera image has a narrow field of view since magnifying a view of the affected area, which may cause accidental injury to other organs or tissues by surgical instruments outside the field of view. To this problem, we have developed a new trocar (a cylindrical device that holds an endoscope) with multiple ultra-compact cameras, and have attempted to expand the field of view by combining images. We attempted to estimate the depth from the combined images and to reconstruct the surgical field in three dimensions. A machine learning method called MiDaS is used for depth estimation, and a distance map is generated. Experimental results using four ultra-compact cameras confirmed the feasibility of 3D reconstruction of the surgical field.

1. はじめに

1.1. 研究の背景

近年、さまざまな分野で奥行き推定を試みる研究が活発に行われている。その代表例として、自動車と障害物間の距離を推定し、事故を未然に防ぐ技術や、スマートフォンでの写真撮影時に遠景をぼかす写真加工技術が挙げられる。また、医療分野においては、腹腔鏡下手術で立体腹腔鏡を用いて体内の奥行きを把握する方法や、手術器具と臓器の3次元的な位置関係を提示することで医師

の視野を拡大し、手術の安全性を高める研究が行われている。本研究では、これらの背景を踏まえ、特に腹腔鏡手術に焦点を当てて奥行き推定に取り組む。

腹腔鏡手術では、腹部に直径5～12mm程度の小さな穴を3～5箇所ほど開け、内視鏡や手術用鉗子、ハサミなどを挿入するためのポートを設置する。挿入された内視鏡を通じて撮影された腹腔内の映像がモニターに表示され、医師はそのモニターを見ながら、他のポートを通して挿入した手術器具を操作して手術を行う。

腹腔鏡手術は、開腹手術に比べて患者の体への負担が少なく、傷口が小さいため回復が早いという利点がある。また、肉眼では確認が難しい細部まで映像を通じて把握できる点でも優れている。一方で、モニターに映し出される映像が2次元であるため、奥行き情報を把握しづらいという課題も存在する。このため、モニター映像では判別が難しい箇所では予期しない事態が発生した際、迅速な対応が難しくなる可能性がある。

この課題に対応するため、本研究ではモニター映像を3次元的に復元する手法を提案する。図1に示すように、モニターに表示される映像は2次元であり、奥行きの把握が非常に困難な状態である。もしモニター映像が3次元的に表示されるようになれば、前後左右の位置関係がより明確に理解でき、手術の安全性向上や患者の負担軽減が期待できる。

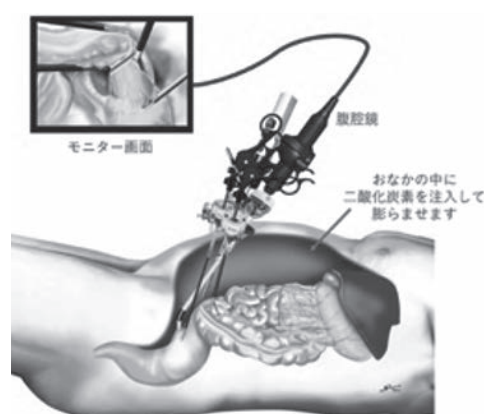


図1 腹腔鏡のイメージ [1]

1.2. 本研究の目的

上述のように、内視鏡手術は、患者への侵襲が低いという利点があるが、モニターを見ながら操作を行う医師の負担は大きい。内視鏡カメラから得られる映像が2次元であるため、奥行き感が欠如し、患部の正確な位置関係を把握することが難しいという課題が存在する。特に、視覚的な奥行き情報が不足することにより、細かい操作や予期せぬ状況への迅速な対応が困難となることがある。本研究では、奥行き推定技術を用いて2次元映像を3次元的に再構築し、手術中の視覚情報を大幅に強化することを目指す。具体的には、機械学習を活用して内視鏡映像から高精度な奥行き情報を取得し、その情報を基に手術視野の3次元的な可視化を実現する。本手法により、医師が患部の位置関係を直感的かつ正確に把握できるようになり、手術の安全性と効率性を向上させるとともに、患者のリスクを低減することが期待される。

2. 奥行き推定及び術野の3次元復元の手法

2.1. 概要

本手法では、AIを活用した深度推定技術を用い、画像や動画から高精度な奥行き情報を取得する。具体的には、単一または複数のカメラで撮影した映像をもとに深度推定を行い、画像のステッチングと3次元復元を通して、術野を3次元的に可視化する。

2.2. AIを活用した深度測定

深度推定には、AIモデル「MiDaS」[2]を採用する。このモデルは事前学習済みの畳み込み

ニューラルネットワーク(CNN)を基盤としており、単一の画像から直接深度情報を推定する。

MiDaSのアルゴリズムは以下の手順で動作する：

1. 入力画像を正規化し、モデルの入力に適したサイズにリサイズする。
2. CNNを通じて各ピクセルの深度値を推定する。この深度値は視差や距離に基づく相対的なスケールで表現される。
3. 推定結果として得られる深度マップを正規化し、色分けによって可視化する。深度マップでは、近い領域が赤、遠い領域が青で表示される。

これにより、奥行き情報を直感的に理解できるだけでなく、術野の三次元構造を把握する基盤を構築する。

2.3. 画像のスティッチング

複数のカメラで取得した画像や動画を結合することにより、広範囲の術野映像を合成する。スティッチング処理には画像統合ソフトウェア「Panorama Stitcher Mini」[3]を使用した。

Panorama Stitcher Miniは以下の手順でパノラマ画像を生成する：

1. 各画像の特徴点を抽出する。特徴点抽出は、画像の輝度やエッジの変化が大きい部分を検出する手法で行われる。これにより、画像内の特徴的な領域が抽出され、後続の処理で画像間の対応を見つける基盤が構築される。
2. 特徴点間の対応を特定し、それらの位置関係を基に画像全体の変換を表す行列を計算する。この行列により、画像の幾何学的な関係を正確に表現することが可能となる。
3. 計算した変換行列を使用して画像を幾何学的に調整し、他の画像と正確に重ね合わせる。この処理により、各画像が統一された視点に整合される。
4. 最終段階では、画像間の境界を滑らかに調整するブレンディング処理を実施し、自然な統合画像を生成する。

この一連の手法により、単一のカメラでは取得不可能な広範囲の視野が確保される。

2.4. 3次元復元

深度測定とスティッチングにより得られたデータを基に、元の画像や動画から3次元復元する。深度マップを用いて奥行き情報を抽出し、元画像の色情報と組み合わせることで三次元点群を生成し、OpenGLを用いて3次元復元する。処理は以下の手順で実施する：

1. 深度マップの正規化

深度マップは相対的な奥行きを示す値として生成されるため、カメラのベースラインや焦点距離を用いて物理的なスケールに変換する。これにより、実際の距離に基づく三次元点群データを生成可能となる。

2. 三次元点群の生成

正規化された深度データをもとに、各ピクセルの(x, y)座標と深度値(z)を算出し、三次元空間上の点群を生成する。また、各点には元画像の色情報を付与し、視覚的にリアルな表現を実現する。

3. 三次元モデルの構築

OpenGLを用いて点群データを視覚化する。点群は黒背景の環境に描画され、各点の位置や色が三次元的に反映される。視点を自由に変更することで、術野の立体的な構造を確認できる。

4. ノイズ除去と視覚化の最適化

点群データには細かなノイズが含まれるため、後処理として統計的手法を用いたフィルタリングを行い、不要なデータを除去する。

本手法により、術野全体を立体的に再構築し、手術支援において有用な3次元情報を提供することが可能となる。

3. 実験結果及び考察

3.1. 実験環境

術野全体をカバーするためにカメラ4台を用いて映像を収集し、MiDaSを用いて出力された値により、3次元復元をする実験を行った。表1に実験環境を示す。

表1 実験環境

CPU	Apple M1
OS	macOS Big Sur version11.2.3
メモリ	16 GB
カメラ	Pixel 6a
使用ソフト	Visual Studio Panorama Stitcher Mini
使用言語	Python (OpenCV , Open3D)

3.2. 実験結果

3.2.1. 4台のカメラからの画像合成

収集された映像(図2)は、ステッチング匠により1つの広範囲な画像に合成される(図3)。

3.2.2. MiDaSを用いた深度測定

合成画像を入力として、MiDaSモデルを用いて深度マップを生成した。生成例を図4に示す。MiDaSは、画像中の奥行き情報を推定し、カメラに近い部分を明るく、遠い部分を暗く表示することで、視覚的に奥行きを表現する。

3.2.3. 距離値の精度測定

深度マップ(図4)は、画像内の各画素に対する相対的な奥行き情報を提供する。これを用いて、カメラから被写体まで距離を20cmとした時に、カメラの高さを0cm～5cmまでの間で1cmずつ高くして精度測定を行った。カメラの高さによる推定精度の違いを表2に示す。

3.2.4. 3次元復元結果

統合画像(図3)および深度マップ(図4)を基に、3次元復元を実施した。図5に示すように、本手法により生成された3次元モデルは、入力された深度マップおよび画像に基づき、対象物の形状や奥行きを再現している。特に、深度マップの解像度やノイズが復元結果に直接的な影響を与えることが確認された。

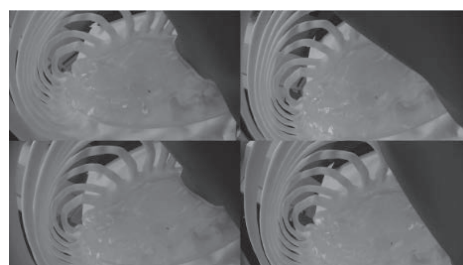


図2 4台のカメラを使用して撮影した画像

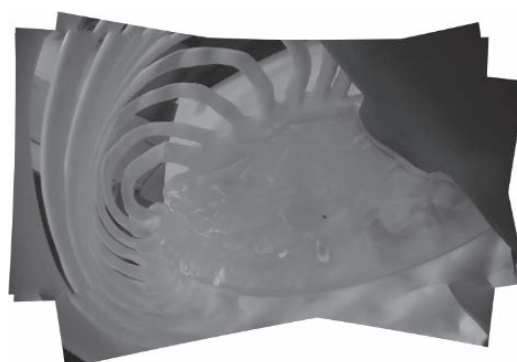


図3 合成結果の例

3.3. 考察

実験の結果、カメラ映像から術野の3次元復元が可能であることを確認した。しかし、復元精度は十分ではなく、いくつかの課題がある。まず、深度マップ生成におけるMiDaSモデルの精度に関してである。生成された深度マップ(図4)には一部にノイズが含まれており、このノイズが3次元復元結果(図5)において細かなゴミとして現れることが確認された。これらの問題は、復元結果の視認性や利用可能性を低下させる要因となっている。さらに、MiDaSの精度評価(表2)では、カメラの高さが増加するにつれて、距離測定の精度が低下する傾向が確認された。これは、被写体がカメラの中心からズレてしまうためであり、カメラの高さが変化することで対象物との相対的な位置関係が変化し、深度推定が不安定になることが原因であると考えられる。この問題を解決するためには、カメラの位置を一定に保つ工夫や、広角レンズや魚眼レンズを用いて視野を拡張する手法が有効であると考えられる。

さらに、3次元復元におけるゴミの浮き出しは、深度マップに含まれるノイズがそのまま点群データに反映されていることが原因と考えられる。この問題を解決するためには、深度マップ生成後に適切なフィルタリング処理を施し、ノイズを除去することが必要である。例えば、ガウシアンフィルタやメディアンフィルタなどの画像処理技術を用いることで、深度マップの滑らかさを向上させることが可能である。



図4 MiDaSを用いて深度測定した結果

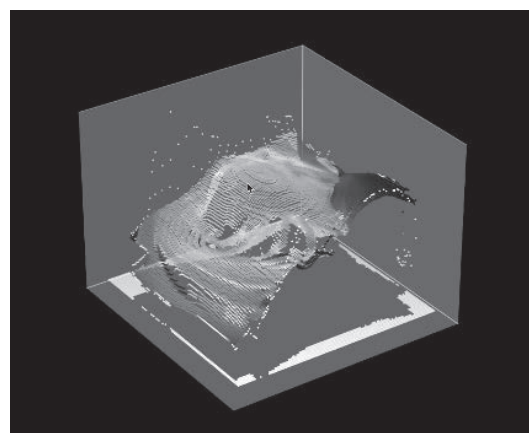


図5 3次元復元結果の例

参考文献

- [1] 大腸の腹腔鏡手術- 医療法人山下病院, https://www.yamashita-hp.jp/suijun/geka/large_intestine.html
- [2] R. Ranftl, K. Lasinger, D. Hafner, K. Schindler, V. Koltun, “Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 44, Issue 3, pp. 1623-1637, 2022
- [3] DG Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, vol. 60, pp.91-110, 2004