

博士論文（要約）

ロボット手術支援のための
術中画像認識に関する研究

中澤 敦士

近年、ロボット技術の進歩によりロボット手術は世界中で普及してきている。ロボット手術では、動作縮小や手の震えのフィルタリングによって高い精度で手技を行うことができるようになり、またロボットの多彩な自由度によって従来人間の手では難しかった場所における治療を容易にした。しかし、現状のロボット手術は医師の手の動作を忠実に模倣してロボット術具が動くため、上手な術具の動かし方だけでなく、下手な動かし方も同様に忠実に再現してしまうといった課題がある。また、動作縮小によって手術が長時間化してしまうといった課題も存在している。

これらの課題を解決するために、近年ロボット手術における支援や自動化の研究が盛んに行われている。支援や自動化を行うことによって、治療成績の向上や、手術時間の短縮、医師の負担の軽減が期待されている。しかし、現状では、自動化に関する研究のほとんどが固定された実験的な環境下で行われたものであり、また失敗するケースもまだまだ見られる。これらは環境認識能力の欠如によるものであり、現在の手術の状況を正確に認識できていないことによって引き起こされる。実際の手術のような動的な環境に対応し、高い成功率を達成するためには、術具や治療対象などの位置・姿勢を正確に認識する技術が必要不可欠である。

術具に関しては、ロボットに取り付けられたエンコーダ情報を用いて機構学計算を行うことによって位置・姿勢を推定することが可能であるが、実際にはたわみや取り付け誤差の影響で計算した位置・姿勢と実際の位置・姿勢との間に誤差がある。この誤差は高い精度を要する手技を支援する上では致命的であり、他のセンサ情報を用いて補う必要がある。また、支援を行う上では、術具だけでなく治療対象などの認識も必要とされる。これまでにマーカやセンサを取り付けて認識する研究も行われているが、マーカやセンサの取り付けには滅菌が必要であり、手術のワークフローを複雑化させてしまう。また、治療対象や縫合針にマーカやセンサを取り付けることは現状困難であり、治療対象などの位置・姿勢を認識することは難しい。そのため、マーカやセンサを用いずに術具などの位置・姿勢を認識する手法として、近年内視鏡画像や顕微鏡画像を用いて認識する手法が提案されている。術中画像を用いた認識は、従来のセットアップを保ったまま認識を行うことが可能であり、追加のデバイスやマーカなどを必要とせずコストも低い。また、近年画像認識技術は飛躍的に進歩してきており、高い精度で認識ができるようになってきている。

そこで、本研究では、ロボット手術における支援技術の開発に向けて術中画像から各物体の位置・姿勢を推定する技術の開発を行う。支援する手技としては、顕微鏡下手術における血管吻合を対象に設定した。血管吻合を支援する上で必要とされる認識対象としては、術具、血管、縫合針、縫合糸が挙げられる。術具や縫合針は剛体であるため、定量的な位置・姿勢の表現が容易であり、特徴的な形状をしていることから物体形状を基に位置・姿勢を推定することが可能である。血管や縫合糸に関しては、術中に変形する物体であるため、定量的に位置・姿勢を表現することが困難であるが、物体形状を抽

出すことができれば、様々な支援手法の開発に役立つと考えられる。そこで、本研究では、まず物体形状を抽出するためにセグメンテーションを行い、その後、手技を支援する上で最も認識が重要とされる術具を対象を絞って位置・姿勢の推定を行った。

物体の位置・姿勢を推定する上では、形状を正確に抽出する必要がある。そこで、物体の境界検出をセグメンテーションと同時にを行うことによって物体形状を正確に抽出することを試みた。提案手法では、まずエンコーダによって入力画像から特徴量を抽出していき、2つのデコーダによって特徴量を統合しながら、セグメンテーションと物体境界検出を同時に行う。提案手法では、セグメンテーションと物体境界検出の双方でエンコーダを共有することによって、エンコーダにおいてより物体境界付近の空間的な情報を抽出することができるようにしている。高次元の特徴量と低次元の特徴量を結合させる際には、**Residual Refinement Module (RRM)**を用いることによって各特徴量を洗練させてから特徴量を統合した。また、物体境界の情報をさらにセグメンテーションに活かすため、デコーダにおいて抽出された物体境界の情報を含有した特徴量をセグメンテーションの特徴量と統合する**Boundary Attention Module (BAM)**を提案した。

人工血管吻合の画像、及びda Vinci Surgical Systemを用いた豚の動物実験の画像に対して提案手法の評価を行ったところ、本手法によって物体形状をより正確に捉えられるようになったことが示された。人工血管吻合の画像においては、影や反射といったノイズに頑健になり、物体同士の境界をより明確に捉えることができるようになった。一方で、縫合針の認識精度が低いことが課題として挙げられた。これは、ネットワークを学習させる際に、全てのピクセルが等しく損失関数に寄与するため、物体の大きさが小さい縫合針はあまりネットワークの学習に寄与しないことが原因であると考えられる。そのため、物体の大きさに影響されずセグメンテーションを行う手法が必要であると考えられる。また、豚の動物実験の画像に対しては、ロボット術具のシャフト、リスト、把持部がより正確に区別できるようになり、生体内の画像に対する有効性が確認された。

次に、縫合針の認識精度を向上するため、物体検出を用いたパノプティックセグメンテーションの手法を提案した。提案手法では、まず前景の物体に対して物体検出、及びインスタンスセグメンテーションを行った。物体検出を用いたインスタンスセグメンテーションでは、物体検出によって検出されたボックスからはみ出ってしまった物体の領域をセグメンテーションできず、また結果の解像度が粗くなってしまうという欠点がある。そこで、本研究では各ピクセルに対して物体の平均的な特徴量との類似度を算出することによって、画像全体に対して解像度の高いインスタンスセグメンテーションを行うことを提案した。セマンティックセグメンテーションに関しては、高解像度のインスタンスセグメンテーションと特徴量を共有しながら行った。そして最後に、インスタンスセグメンテーションとセマンティックセグメンテーションの結果を統合することによって、パノプティックセグメンテーションの出力を得た。

人工血管吻合のデータセットを用いて評価実験を行ったところ、術具、血管、縫合針、

縫合糸を高い精度で認識可能であることが示された。物体検出を用いて各物体を物体の大きさに関わらず等しく学習することによって、大きさの小さい縫合針の認識精度を飛躍的に向上させることができた。また、提案手法により、物体検出の精度が悪い場合でもボックスの外にはみ出てしまった物体の領域に対してインスタンスセグメンテーションを行うことができることが確認でき、物体検出におけるボックスの推定精度に関わらず物体形状を正確に抽出できることを確認した。

最後に、パノプティックセグメンテーションの結果を利用した術具の位置・姿勢の推定手法を提案した。顕微鏡下手術では、術具が頻繁に術野の内外を行き来するため、画像情報と機構学情報を上手く組み合わせながら認識を行う必要がある。そこで、画像情報に基づく類似度と機構学情報に基づく類似度を組み合わせながらパーティクルフィルタによって術具の状態を推定する手法を提案した。画像情報に関する類似度に関しては、あらかじめ術具の各姿勢に対して保存しておいた輪郭モデルを用いて実際の輪郭形状との類似度を算出し、機構学情報に関する類似度に関しては、機構学情報から得られた術具の状態を画像座標系に変換した際の誤差を算出してパーティクルフィルタに利用した。パーティクルフィルタでは、物体検出の結果と機構学情報を照らし合わせながら術具の移動量を決定し、画像情報に基づく類似度と機構学情報に基づく類似度の双方を用いて各パーティクルが保存する術具の状態と実際の術具の状態の類似度を算出した。

人工血管吻合の画像において、提案手法の有効性を評価したところ、高い精度で術具の位置・姿勢を推定できることが示された。頻繁に術具が術野の内外を移動する場合や術具同士の重なりがある場合でも正しく位置・姿勢を推定することが可能であることを示した。

本研究では、ロボット手術における支援技術の開発に向けて、術中画像認識技術の開発を行った。本研究で開発した術中画像認識技術は、手術の様々な状況に対応しながら手技の支援を行う上で非常に重要な技術であり、本研究で開発した術中画像認識技術がロボット手術における支援技術の開発を促進し、医療のさらなる発展に貢献することが期待される。