

ロボットの的確な判断を可能にする認識技術

Image Recognition Techniques Enabling Precise Understanding for Robots

松村 正文 MATSUMURA Masafumi 伊藤 聡 ITO Satoshi 赤木 琢磨 AKAGI Takuma

近年、AIや深層学習の急速な進歩により、ロボットの進化と普及が進んでいる。一方、国内では少子高齢化や人手不足などの解決といった社会課題に対応するため、ロボットへのニーズや期待が年々増加し適用範囲が拡大している。ロボットの高度化を実現するには、状況を的確に認識する目に相当する機能の役割が重要になっている。

東芝グループは、ピッキングや、荷積み・荷降ろし、自律移動、スポット溶接検査など、多種多様な機能を持つロボットを開発している。今回、ロボットの高度化を可能にする、物体の種類に依存しない物体領域抽出技術や、様々な形状の物体に対応できる3次元姿勢推定技術、マーカーなしで回転誤差1°以下、並進誤差10 cm以下で自己位置を推定する技術、スポット溶接の非破壊検査を効率化する超音波センシング技術などの認識技術を開発した。

With the rapid progress of artificial intelligence (AI) and deep learning, the advancement and dissemination of robots have been accelerating. The range of robot applications has also been expanding, from the handling of repetitious work at manufacturing sites to various industrial spheres including production, logistics, and physical distribution systems in order to address social issues that have arisen in Japan due to the labor shortage accompanying the aging of society and declining birthrate in recent years. Techniques corresponding to human vision that make it possible to precisely understand operating conditions play a critical role in the sophistication of robots.

The Toshiba Group is engaged in the development of robots equipped with a variety of functions including piece picking, loading and unloading, autonomous mobile, and spot-welding validation robots. In order to enhance the sophistication of these robots, we have developed the following image recognition techniques: (1) an instance segmentation technique that is independent of the type of object, (2) a three-dimensional (3D) posture estimation technique that responds to various shapes of objects, (3) a self-localization technique that achieves an identification accuracy of within 1° in rotation error and within 10 cm in translation error, and (4) an ultrasonic sensing technique that makes it possible to evaluate spot welds nondestructively.

1. まえがき

近年、ロボットの進化と普及が加速している。その要因の一つとして、AIや深層学習の急速な進歩による、ロボットを構成する各技術の向上やロボットシステム全体の高度化が挙げられる。一方、国内では少子高齢化や人手不足などの社会的背景により、ロボットに対するニーズや期待が年々増えつつある。これらの技術的進化と社会的ニーズにより、ロボットの応用先は、従来の反復作業の現場から、製造や、流通、物流などの各産業分野へ広がりを見せている。

ロボットを動作させる際は、まず状況をセンシングして認識し、それに基づいて計画を立ててロボット自身を制御するが、そのサイクルは、CPS（サイバーフィジカルシステム）で構成されている。このとき、状況を的確に認識するロボットの“目”の高機能化が、CPS全体の性能向上につながる。このロボットを更に上位の大きなCPSとつなぐことで、認識結果をデータ分析やサービス連携などに活用する新たなCPSを構築することができる。東芝グループは、それらを実

現する認識技術の開発に取り組んでいる。ここでは、ロボットの目を高機能化する各種認識技術について述べる。

2. ロボットの高度化を支える認識技術

東芝グループは、多種多様なロボットを開発しており、その高度化を進める上で重要な、ロボットの目を高機能化する認識技術を開発している。例えば、物流向けのロボットでは、対象物を正しく認識して的確につかむことが重要となるため、物体領域抽出技術や3次元姿勢推定技術を開発している。また、自律移動ロボットでは、自身の現在地を正確に把握する必要があるため自己位置推定技術を、スポット溶接検査ロボットでは、検査結果のばらつきを抑えられる超音波センシング技術を開発している。以下に、これら四つの認識技術について述べる。

2.1 物体領域抽出技術

ロボットにおいては、対象物を正確に認識することが重要である。物体領域抽出技術¹⁾は、入力画像に対し、物体の種類ごとに物体の領域を画素レベルで推定する技術であ

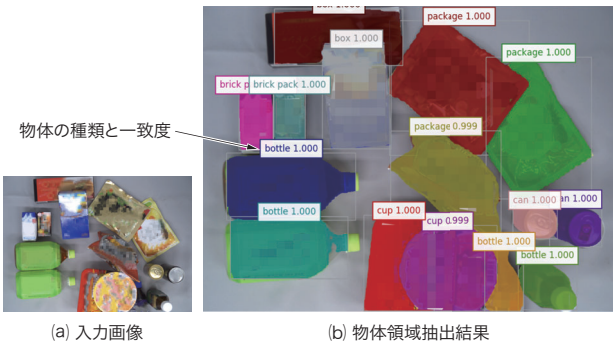


図1. 一般の物体領域抽出技術の例

物体の種類ごとに大量の教示データを用意して学習しておく、入力画像から、物体の種類ごとに各物体の領域を画素レベルで抽出している。

Example of generic instance segmentation technique

る。例えば、図1において、ペットボトル(bottle)という物体の種類に着目すると、全部で四つの領域がそれぞれ画素レベルで抽出されていることが確認できる。この技術は、物体の種類ごとに物体領域を教示した大量の教示データを事前に用意し、それらを学習しておくことで実現されている。したがって、未学習の種類の物体については、領域を抽出することができない。未学習の種類の物体についても、教示データを用意して再度学習を行えば対応できるが、物体の種類を一つ増やす度に数百枚から数千枚という大量の教示データを用意する必要があり、コスト面で現実的ではない。また、処理対象となり得る物体の種類は多岐にわたり、事前に全てを想定することは困難である。そこで、このような未学習の種類の物体にも対応できるように、物体の種類に依存しない物体領域抽出技術を開発した。

一般の物体領域抽出技術は、物体の属する種類を推定するとともに、物体の形状を推定することで物体領域を抽出する。ここで、同じ種類の物体同士の形状は類似していることが多いため、物体の形状は、物体の種類からある程度の類推が可能である。すなわち、物体の種類を区別する一般の物体領域抽出技術では、物体の形状推定が物体の種類に大きく依存するので、形状推定そのものの汎用性が限定的になる。

そこで、今回開発した技術では、物体の種類によらずに多様な形状推定が行えるように、分類を“ペットボトル”や“缶”といったものではなく、それらを総称した“物体”という一つの種類にまとめて、形状推定に特化することにした。この“物体”という種類には多様な形状の物体が含まれるため、汎用性の高い形状推定が可能となり、様々な形状の物体に対して領域抽出できることが期待される(図2)。

図3は、入力画像(a)に対し、物体の種類ごとに学習する

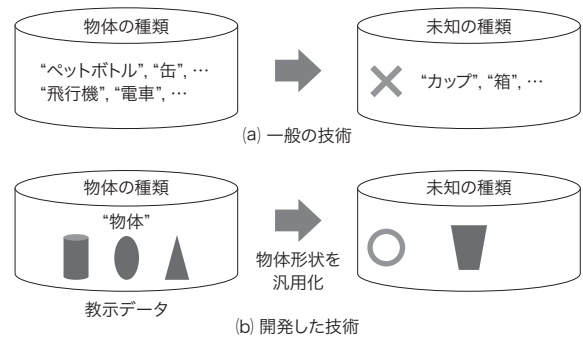


図2. 開発した物体領域抽出技術の概要

物体の種類を区別せずに一つの種類にまとめ、物体の形状推定に特化することで、形状推定の汎用性を高めている。

Outline of generic and newly developed instance segmentation techniques



(a) 入力画像



(b) 一般の物体領域抽出技術
物体の種類を区別して学習



(c) 開発した物体領域抽出技術
物体の種類を区別せずに学習

図3. 開発した物体領域抽出技術の実験結果

開発した技術では、一般の技術よりも多くの物体の領域が抽出できており、検出性能が改善している。

Results of experiments on instance segmentation using generic and newly developed techniques

技術(一般の物体領域抽出技術)の処理結果(b)と、種類を一つの“物体”にまとめた技術(開発した物体領域抽出技術)の処理結果(c)であり、未登録の物体に対する検出性能が改善されていることが分かる。開発した技術と一般の技術で、学習に用いた教示データは同じであり、違いは、物体の種類を一つにまとめるか否かだけである。このように、開発した技術は、より汎用性の高い形状推定ができることが確認された。

2.2 3次元姿勢推定技術

物体をつかむピッキングロボットでは、対象物を的確につかんだり、箱に隙間なく詰めたりするため、物体がどのような姿勢で配置されているかを認識することが重要である。対象物を箱状の定型物に限定すれば、ステレオカメラなどで

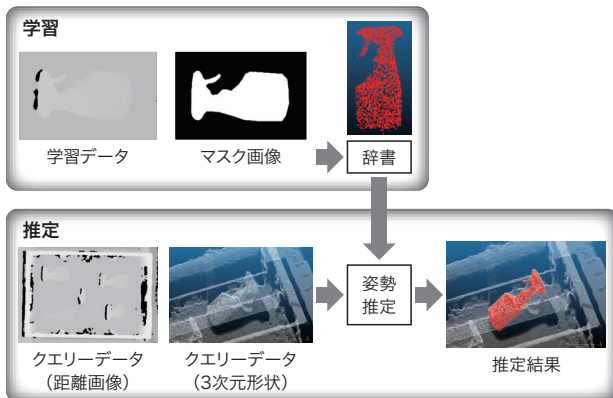


図4. 3次元姿勢推定の流れ

学習データから3次元の形状特徴を辞書登録しておき、箱内を撮影した距離画像から辞書登録済みの物体との一致度が高い箇所を探索することで、相対姿勢を推定する。

Flow of processes for 3D object posture estimation

3次元の姿勢を求められるが、現場では、箱以外の様々な形状に対応してほしいという要求が強い。そこで、多様な形状に対応する3次元姿勢推定技術を開発した。

ピッキングの際は、短時間で多くの物体を処理する必要がある。また、一般に、現場では物体が重なり合っているため、遮蔽物があっても姿勢推定ができることも重要である。そこで、これら二つの要件を満たすため、東芝欧州研究所で開発した技術²⁾を活用した。

この技術では、あらかじめ姿勢を推定したい形状について、3次元的な形状特徴を表現する辞書を作成しておき、姿勢を推定したい物体が含まれるクエリーデータ(距離画像)の中から、辞書と3次元的な一致度が高い箇所を探索することで両者の間の相対姿勢を推定する(図4)。距離画像は、ピクセルに対応する奥行き情報が保存されたものであり、取得方法としては、ステレオカメラを用いる方法のほかに、単眼カラー開口カメラを用いた距離推定技術³⁾もある。

3次元姿勢推定技術は、クエリーデータと辞書との部分的な対応に基づいて姿勢を推定するため、物体がある程度遮蔽されても姿勢推定が可能であることや、動的計画法を用いて探索を効率化しているため、比較的lowコストで処理可能なことなどの特長がある。

2.3 自己位置推定技術

自律移動ロボットが所定位置へ移動するには、ロボット自身が現在地を正確に把握する必要がある。従来の技術では、磁気テープや2次元バーコードなどのマーカーを床面に貼り、マーカーを認識することで自己位置を推定していた⁴⁾。しかし、この手法には、マーカーの敷設が必須であることや、周囲環境の変化に対するロバスト性などの制約があ

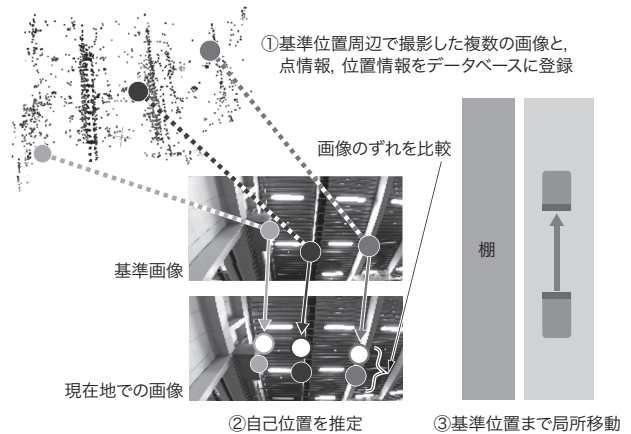


図5. 画像による自己位置推定技術を用いたロボットの移動

自律移動ロボットが、基準画像との画像のずれを認識して自己位置を推定することで、基準位置まで精度良く移動できる。

Self-localization using images for autonomous mobile robot

る。また、タイヤの回転数から自己位置を推定する手法もあるが、カーブや急停止などによるタイヤのスリップで推定位置がずれてしまう問題があった。

そこで、この問題を解決するため、画像を用いて位置ずれを補正する、マーカーが不要な位置推定技術を開発した。この技術では、①基準位置で撮影した画像(以下、基準画像と略記)をデータベースに登録しておき、②その基準画像を基に自己位置を推定し、③自律移動ロボットを基準位置まで移動する(図5)。

①のデータベース登録では、基準位置の周辺でステレオカメラにより天井の画像を複数撮影し、カメラ位置とともにデータベースに登録する。②の自己位置推定では、基準画像と現在位置で撮影した天井画像を比較し、画像に映っている天井位置のずれから、基準位置に対するロボットの相対位置を算出する。天井画像を利用することで、床にある設備の配置換えなどの環境変化に強い位置推定を実現できる。

実際の物流倉庫で精度検証を行い、回転誤差1°以下、並進誤差10 cm以下で自己位置推定できることを確認した。

2.4 超音波センシング技術

スポット溶接は、接合対象の金属板を電極棒で加圧しながら電流を流すことで金属板間を溶解凝固させて接合する手法で、自動車のボディーの組み立てなどに利用されている。スポット溶接が適切に行われたことを確認するために非破壊検査が浸透しつつあり、東芝電力検査サービス(株)は、超音波検査装置Matrixeyeを提供している。検査時には接合部へ超音波プローブを当てるが、検査員が超音波プローブの角度を微調整しながら操作する必要がある。そこで、ロボットアームにMatrixeyeの超音波プローブを備え付

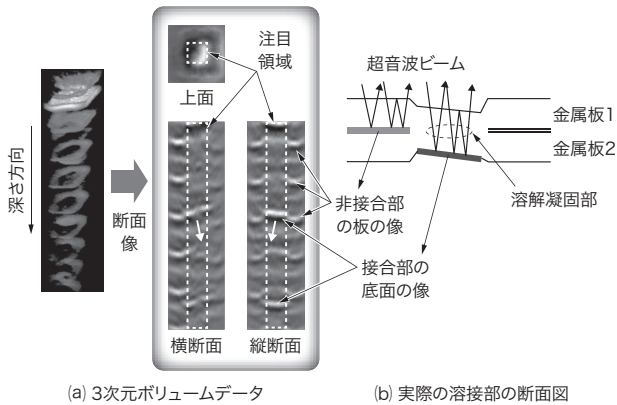


図6. 3次元ボリュームデータを用いた溶接部傾き方向の推定方法
開発した技術は、接合部の底面からの反射像を用いて、超音波プローブに対する溶接部の傾き（白矢印）を算出する。

Estimation of inclination of welded part acquired through 3D volume data

け、角度調整を行いながら検査するスポット溶接検査ロボットを開発した。

検査では、接合部に対して垂直に超音波プローブを当て、接合面積を適切に計測する。しかし、接合部の傾きは、溶解凝固の状態が変わる上、外観からでは判断が困難なため、適切な角度で検査できなかつたり、角度調整回数にばらつきが生じたりして、効率的な検査が行えなかつた。そこで、検査時に得られる超音波画像である3次元ボリュームデータから接合部の傾きの方向を推定する手法を開発した。

3次元ボリュームデータは、接合部では、接合対象である金属板の底面からの反射像が、高い輝度で描出される（図6）。また、超音波の多重反射で、像は繰り返して描出される。開発した手法では、3次元ボリュームデータにおける中央部分の注目領域について、3次元の輝度勾配の平均を算出して反射像の傾きの方向を推定する。このとき、3次元ボリュームデータの深度が深くなると、多重反射で輝度勾配の信頼度が低下することを考慮する。これらにより得られた方向を、接合部の傾きとしてロボットアームに通知し、超音波プローブの角度を調整する。検証した結果、これまで10回の探索を要していた角度調整が、99%の確度で4回以下にて完了することを確認した。

3. あとがき

ここでは、東芝グループが開発した、四つの認識技術について述べた。これらの技術は、ロボットを上位CPSのエッジコンピューティング端末としてつなぐことで、更に新たな価値を創出できる。例えば、サイバー空間において、深層学習で作成したモデルをメンテナンスする技術を活用すること

で、ロボットの目の性能を維持・向上させながら、長期的に安定運用することが可能になる。

また、東芝グループは、目だけでなく、触覚や嗅覚など、ロボットを智能化する上で必要なセンシング技術にも取り組んでいる。そのために必要な次世代センサーデバイスも手掛けており、ロボットの正確な位置・姿勢検知に必須である高感度・低ドリフト・広帯域なジャイロセンサー⁽⁵⁾や、匂いセンサーの基盤として、グラフェンと呼ばれる高感度な炭素材料を用いた高い選択性を持つ小型のバイオセンサー⁽⁶⁾などの開発も進めている。今後、これらの研究開発にも積極的に取り組むことで、ロボットの更なる智能化を目指し、社会課題の解決に貢献していく。

文献

- (1) Pham, Q. V. et al. "BiSeg: Simultaneous Instance Segmentation and Semantic Segmentation with Fully Convolutional Networks". The Proceedings of the British Machine Vision Conference 2017. London, UK, 2017-09, British Machine Vision Association (BMVA). 2017, Tuesday, 5 Sep. 58.
- (2) Zach, C. et al. "A dynamic programming approach for fast and robust object pose recognition from range images". 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015). Boston, MA, 2015-06, IEEE. 2015, p.196-203.
- (3) 三島 直, ほか. 単眼カメラで撮影した1枚の画像から精度良く距離計測できるカラー開口撮像技術. 東芝レビュー. 2018, 73, 1, p.39-43. <https://www.toshiba.co.jp/tech/review/2018/01/73_01pdf/f02.pdf>. (参照 2019-03-25).
- (4) Jeisung, L. et al. A vision-based automated guided vehicle system with marker recognition for indoor use. Sensors. 2013, 13, 8, p.10052-10073. <<https://www.mdpi.com/1424-8220/13/8/10052/pdf>>. (accessed 2019-03-25).
- (5) Gando, R. et al. "A MEMS rate integrating gyroscope based on catch-and-release mechanism for low-noise continuous angle measurement". 2018 IEEE Micro Electro Mechanical Systems (MEMS). Belfast, UK, 2018-01, IEEE. 2018, p.944-947.
- (6) Saito, T. et al. "Amino group termination for wafer-scalable graphene biosensor". Biosensors 2018 (28th Anniversary World Congress on Biosensors). Miami, FL, 2018-06, Elsevier. 2018, P3.133.



松村 正文 MATSUMURA Masafumi
研究開発本部 研究開発センター
メディアAIラボラトリー
Media AI Lab.



伊藤 聡 ITO Satoshi
研究開発本部 研究開発センター
メディアAIラボラトリー
電子情報通信学会会員
Media AI Lab.



赤木 琢磨 AKAGI Takuma
東芝インフラシステムズ(株)
小向事業所 SA設計第三部
電子情報通信学会会員
Toshiba Infrastructure Systems & Solutions Corp.