

## 審査の結果の要旨

氏 名 八木 拓真

本論文は「Hand-Object Interaction Mining from First-Person Videos」（一人称視点映像からの手-物体インタラクションマイニング）と題し、ユーザが物体を手に取り行動する様子を捉えた映像から、手と物体間のインタラクションの理解を目的に、手-物体インタラクションマイニングと呼ばれる概念を提案し、要素技術と応用システムの両面から手-物体インタラクションマイニングについて論じたものであり、英文で記され全体で5章により構成される。

第1章「Introduction」（はじめに）では、本研究の背景として実環境において手-物体インタラクションの認識が難しいことを述べ、それを克服するための包括的な概念として「手-物体インタラクションマイニング」の概念を提案している。本概念は「ラベルなしデータからの学習」「カテゴリ非依存の定式化」「ユーザへの最小限の介入」の3つの要素からなり、これらの要素に基づいて手法を設計することで実世界において動作する手法が作れること、上記の概念に基づいて3つの個別の取り組みが具現化されることを述べている。

第2章「Hand-Object Contact Prediction」（手-物体間の接触状態予測）では、手と物体が互いに接触している否かを映像から予測する手法を提案している。具体的には、訓練データのアノテーションコストが高いことを背景として、できる限り少ない量のアノテーションから実環境に汎化するモデルを構築するための（i）運動情報を用いた疑似ラベル収集（ii）誘導付き逐次ラベル訂正（gPLC）の2つの技術に基づく弱教師あり学習フレームワークを提案している。実環境における提案手法の有効性を評価するため、新たな評価用ベンチマークが併せて提案され、実験を通じて、提案手法が既存の物体検出ベースの手法群に対して有利であること、手動で収集されたラベルのみを用いて訓練した場合と比べ高い性能で実現可能であることが示されている。

第3章「Category-Agnostic Object Instance Identification」（カテゴリ非依存の物体インスタンス識別）では、映像中に出現する特定の物体を識別する問題が取り組まれている。本問題は従前より取り組まれているが、本研究ではユーザが行動を行う実環境において本問題が従来想定されてきたものより遥かに難しくなる点が指摘され、実際的な状況で出現する多数の未知物体を事前知識なしで識別する問題に取り組んでいる。具体的には、1500インスタンスを超える大規模データセットが構築され、その分析を通

じて物体および背景の見えが動的に変化することで発生する本質的な課題が見出された。

第4章「Assisting Users for Finding Lost Objects」（無くし物の発見支援）では、第2章および第3章で開発された技術を背景に、ユーザが無くし物をした際にその発見を支援する問題が取り組まれている。従来の支援システムでは、探したい物体にセンサを装着するまたは探したい物体の見た目を事前に登録することが要求され、突発的な無くし物に対応できないという問題があった。本章では、首に装着したウェアラブルカメラより撮影した映像から手で操作した物体を自動的に検出・追跡し、それらが最後に出現した瞬間をユーザに提示することによって想起を促す事前登録不要のもの探しシステム（G0-Finder）が提案されている。ユーザは、スマートフォン上のインターフェースから探したい物体の画像を検索し選択することで、その物体を最後に見た瞬間を得ることができる。実験室環境における模擬物探しタスクおよびより多数の物体が出現する実際的なシナリオにおける検索タスクの2つのユーザ実験が行われ、ユーザが少ない精神的負荷でより正確にもの探しが行えること、100物体以上の多数の物体が登場する場合においても短時間で探したい物体の検索が行えることが確認された。

第5章「Conclusion and Future Work」（結論と将来展望）では、本論文の成果をまとめるとともに、その将来的な展望が述べられている。

以上これを要するに、本論文は、「ラベルなしデータからの学習」「カテゴリ非依存の定式化」「ユーザへの最小限の介入」の3要素からなる手-物体インタラクションマイニングの概念を2つの基礎問題および1つの応用問題において包括的に具現化したものであり、一般的な環境下における手-物体インタラクション認識のための基盤技術を提供するものとして、電子情報学上貢献するところが少なくない。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。