

## 審査の結果の要旨

氏名 キタニ クリス マコト

本論文は「Modeling and Recognizing Human Activities from Video」（映像にもとづく人物行動のモデリングと認識）と題し、動画からの人物動作の学習、行動文法の学習、行動認識のための推論という人物行動解析における主要な問題に対して解決のための手法を提案し、人物行動解析のための一貫した枠組みを示したものであり、全体で6章で構成されている。

第1章「Introduction」（はじめに）では、本研究の背景と目的について論じた後、提案される手法の適用例を明確に説明し、論文で提案される3つの手法の概要を述べている。

第2章「Preliminaries」（準備）では、当該研究分野における本研究の位置付けについて整理した上で、本研究における重要な3つの概念である Event（イベント）、Activity（行動）、Action（動作）を定義し、心理学分野における知見から行動を構文的なモデルを用いて表現するという本研究の基本アプローチの妥当性を論じている。

第3章「Learning action primitives」（プリミティブ動作の学習）では、関連研究に共通する問題点、具体的には、映像にもとづく動作カテゴリの学習手法は動き特徴のみを用いており、動作に関連している物体や背景のアピランス特徴を用いていないことを指摘した上で、人物の動きのみならず、視覚的文脈をも考慮した人物動作カテゴリの教師無し学習手法を提案している。具体的には、動作カテゴリを学習するために(1) 動き特徴と視覚的文脈の二つを考慮した生成モデルと(2) 大量のデータを処理するためのクラスタリング手法を提案している。複雑な背景のもとでさまざまな動作を行った場面を記録した映像を用いた実験により、視覚的文脈を用いることによって動作カテゴリがより精度良く獲得できることを示し、プリミティブ動作の学習について本手法の有用性を示している。

第4章「Learning the structure of activities」（人物行動の構造学習）では、自然言語の構文解析に用いられている確率文脈自由文法を映像からの人物行動解析へ適用した先行研究を紹介したのちに、それらに共通する問題点を整理している。本研究では、映像からの人物行動学習では文章の場合と異なり記号列に多くのノイズが含まれるために文法をうまく学習できないという課題に対し、最小記述長原理にもとづき、ノイズを除外した終端記号集合とそれに伴う文法を獲得する手法を提案している。具体的には、終端記号の全組合せを評価し、各々の部分集合の下で得られた文法の複雑さと観測データの記号列尤度とのバランスにもとづき、評価値の高い終端記号集合と文法の候補を特定することにより、記号列に含まれるノイズを除去しつつ行動文法の基本構造を獲得することを可能としている。シミュレーションデータと実データを用いた実験により、提案手法の有効性を示している。

第5章「Recognizing structured human activities」（構造的な人物行動の認識）では、構文モデルを用いて人物の行動を認識する手法を提案している。具体的には、画像処理による基本動作要素検

出部分の不確実性に対応するために、文脈自由文法を階層的ベイジアンネットワークへと変換したのちに動作記号列を解析することを提案している。更に、自然言語処理で用いられている削除補間法を利用し、重なっている行動にも対応可能とすることを提案している。サーベイランスビデオを基に生成された動作記号列から重なった行動を認識することにより、本手法の有用性を示している。

第 6 章「Conclusion」（まとめ）では、本研究の新規性と技術的貢献をまとめた上で、今後取り組むべき課題を述べている。

以上これを要するに、本論文では、人物動作の学習、行動文法の学習、行動認識のための推論という行動解析に関する 3 つの重要な課題に対して、動きと視覚的文脈の両方を考慮した動作カテゴリの教師無し学習手法、ノイズを考慮した行動文法学習手法、階層的ベイジアンネットワーク削除補間法を用いた行動認識手法を提案し、シミュレーションおよび実データを用いた実験により各手法の有効性を評価したものであり、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。