

氏 名 高橋 正樹

学位（専攻分野） 博士（情報学）

学位記番号 総研大甲第 1516 号

学位授与の日付 平成 24 年 3 月 23 日

学位授与の要件 複合科学研究科 情報学専攻  
学位規則第 6 条第 1 項該当

学位論文題目 映像解析による人物動作理解に関する研究

論文審査委員 主 査 教授 佐藤 真一  
教授 杉本 晃宏  
准教授 佐藤 いまり  
助教 孟 洋  
教授 佐藤 洋一 東京大学

## 論文内容の要旨

近年では個人が所有するPCやTVにもカメラが搭載されるようになり、映像を用いたコミュニケーションが一般的となっている。またインターネット上には大量の映像が溢れ、様々な映像を誰もが視聴できる環境が整っている。街中にはいたるところに監視カメラが設置され、屋内外の映像を常時記録している。このように、現代では映像が様々な形で日常生活の中に浸透しており、簡単にアクセス可能な身近なメディアとなっている。

本論文はこのような実環境における映像を対象とした、人物動作認識技術の高度化を提案するものである。生活環境への映像の普及に伴い、ユーザ動作による機器操作、特定動作をクエリとした映像検索など、人物動作認識技術に対する期待は高まっており、その需要も多岐にわたる。しかし実環境における映像は撮影条件や被写体動作が多様であり、安定した解析が困難であることが多い。本研究はこれらの課題を確認するとともに、頑健な人物動作認識技術の確立を目指すものである。

第1章では本研究の背景として、日常生活における人物動作認識への社会的ニーズについて述べる。また人物動作の多様性に対処するため、意図の強さに基づく人物動作の分類を行う。分類した各動作の認識技術に対するニーズを確認するとともに、その実現に向けた課題を検討する。最後に、映像解析による一般行動認識技術についてまとめ、関連技術の現状を確認する。

第2章では意思伝達動作であるジェスチャに焦点をあて、その認識手法について検討する。近い将来の大画面・高解像度のTV視聴環境では、リモコンに代わる新たなマンマシインタフェースが求められている。中でも映像解析による人物ジェスチャ認識は接触型デバイスが不要であり、次世代TVの特徴である没入感を損なうことなく操作できるため、新たなインタフェースとして期待を集めている。またジェスチャによる操作は、映像コンテンツ内オブジェクトとの自然なインタラクションを実現するうえでも有効である。本章では、はじめに次世代TV視聴環境でのジェスチャ認識における要件を考察する。続いて、ジェスチャ認識における先行研究を紹介し、本章での研究目的を明確にする。具体的には、次世代TV視聴環境での対話型ジェスチャ認識の実現へ向け、奥行き情報の利用方法について検討する。またユーザの自然な動作の認識を実現するため、長期の時間情報を含む画像特徴を検討する。そしてこれらの検討を踏まえた新たなジェスチャ認識手法を提案する。様々な実験を通して提案手法を評価し、次世代TV視聴環境でのインタフェースとしての有効性を確認するとともに、今後の拡張性を考察する。

第3章では人物の一般行動に焦点をあてる。人混みで混雑した実環境での監視映像を対象とし、混雑映像から一般行動を頑健に認識する手法を提案する。現代では屋内・屋外問わず監視カメラが普及しているが、その映像の確認はほとんどの場合人間によって行われている。膨大な映像量に対する作業者の数は少なく、非効率な監視を余儀なくされている。またそのほとんどは映像の事後確認にとどまり、犯罪の未然防止や直後の検出には活かされていない。そのため、不特定多数の人物行動を自動認識する技術への期待が高まっている。本章ではまず混雑映像の解析における問題点を列挙し、その課題を確認する。続いて一般行動認識に関する先行研究を紹介し、関連技術の現状について述べる。そして広域特徴に基づく手法、局所特徴に基づく2つの手法について検討する。前者は人物領域検

出に基づく手法であり、人物領域の軌跡からその行動を認識する。後者は特徴点軌跡に基づく手法であり、多数の軌跡特徴から映像内で発生した行動を認識する。2手法の比較を通し、実環境で有効に機能する人物行動認識手法を検討する。独自の様々な実験による評価と、映像検索に関する国際的評価型ワークショップ TRECVID Surveillance Event Detection タスクへの参加を通し、提案手法の有効性を確認するとともに、実用化へ向けての課題を確認する。

第4章では、映像を視聴しているユーザの筋運動系情動を計測し、得られた映像特徴からユーザの内部状態（注目度）を推定する手法を検討する。ユーザの内部状態は、情動動作として一部身体に表出すると考えられるが、その動作は微小であり、正確に計測することは難しい。さらに計測した情動動作がそのまま内部状態を表しているとは限らない。たとえば映像コンテンツ内特徴など、外部要因も考慮する必要がある。これらの原因により、可視情報からの内部状態推定は一般に困難であり、本テーマは挑戦的課題と言える。本章では、はじめに脳科学や心理学など、他分野研究を含めた人物の内部状態推定に関する先行研究を紹介する。次にユーザが注目状態にあるときに表出する情動動作を検討し、それら動作と注目度に関する仮説を示す。目視正解データによる仮説の検証の後、各情動特徴を自動取得する手法を提案する。最後に、自動計測した情動特徴から注目度を算出する注目度推定器を作成し、その性能を評価するとともに、提案手法の将来性を考察する。

第5章では、本論文の成果をまとめる。本論文の成果は、実環境における人物動作認識のニーズとその実現へ向けた課題を確認し、各ニーズにおける人物動作認識手法を提案するとともに、その有効性を検証したことにある。さらに人物の意図により動作を分類し、各分類における動作から意図を理解する手法を示したことにある。特に無意図的動作と呼ばれる情動動作から人物の潜在的興味を推定する可能性を示せたことは、大きな成果であると考えられる。従って本研究は、家庭における新たなマンマシンインタフェースや個人プロフィール推定、監視カメラをはじめとする各種映像における人物行動検出など、実生活環境で利用可能な人物動作理解技術の確立へ向け、大きな貢献をしたと考える。

## 博士論文の審査結果の要旨

本論文は、実環境映像によるジェスチャ認識、監視映像解析、人物の意図推定等の実応用への要求を受け、映像による人物動作認識技術の高度化を提案している。テレビ操作のためのジェスチャ認識技術、監視カメラ映像中の人物の一般動作解析技術、映像視聴環境における視聴者の注目度推定技術という具体的な問題に対し、課題の分析並びに解決による実証を行っている。

第1章では本研究の背景として、日常生活における人物動作認識への社会的ニーズの検討を行っている。また人物動作の多様性に対処するため、意図の強さに基づく人物動作の分類を行っている。これらにより、解決すべき課題の具体的な検討を行っている。

第2章では意思伝達動作であるジェスチャに焦点をあて、その認識手法について検討している。特に、次世代TV視聴環境でのジェスチャ認識の実現を目的とし、TOFカメラによる奥行き情報と、ユーザの自然な動作の認識のための長期の時間情報を含む画像特徴との融合により、実践的なジェスチャ認識を実現している。特に、奥行きを含む3次元の軌跡に対するBag of Features法によるジェスチャ認識手法を実現し、その性能評価を行っている。

第3章では、実環境での監視映像を対象とし、一般行動を頑健に認識する手法を提案している。まずは、人物領域の軌跡からその行動を認識する手法、次いで局所特徴量を用いた特徴点軌跡に基づく手法を提案し、これら二手法の比較を通し、実環境で有効に機能する人物行動認識手法を検討した。

第4章では、映像を視聴しているユーザの筋運動系情動を計測し、得られた映像特徴からユーザの内部状態（注目度）を推定する手法を検討している。これは実際には計測不可能な人間の意図を推定しようという、挑戦的な課題である。提案方法では、ユーザが注目状態にあるときに表出する情動動作を検討し、それら動作と注目度に関する仮説を立てている。次いで、各情動特徴を自動取得する手法を提案し、注目度を算出する注目度推定器を構築している。

第5章では、本論文の成果をまとめている。

本論文では、実環境における人物動作認識のニーズとその実現へ向けた課題を確認し、各ニーズにおける人物動作認識手法を提案するとともに、その有効性を検証しており、社会的な要求に的確に答えている。その一方、認知心理的な知見に基づき、人物の意図により動作を分類し、各分類における動作から意図を理解する手法を示しており、特に無意図的動作と呼ばれる情動動作から人物の潜在的興味を推定する可能性を示しており、学術的な意義も高い。従って、家庭における新たなマンマシンインタフェースや個人プロフィール推定、監視カメラをはじめとする各種映像における人物行動検出など、実生活環境で利用可能な人物動作理解技術の確立へ向けた貢献は少なくないと考えられる。

出願者は、以上の研究成果を、査読つき国際雑誌論文2件、査読つき国際会議2件などにまとめている。

本論文の提案内容は、新規性・有効性・信頼性が十分に備わっており、学術的価値

並びに社会的なインパクトも十分に認められる。したがって、本論文が複合科学研究科における博士授与の基準に達していると判断できる。

平成 24 年 1 月 24 日に、博士論文公聴会にて出願者による博士論文の内容の発表、並びに質疑応答が行われた。その公聴会の結果を受けて、非公開にて、審査委員による審査が行われた。

論文内容の発表では、研究の背景と狙い、映像による動作解析のための動作の分類、社会的ニーズの高い人物動作解析の具体的なシナリオに向けた技術の実装と評価、ならびに本研究の結論について報告された。質疑応答に対し、出願者は的確かつ明快に回答した。

その結果、出願者は情報学並びに関連する分野に関し、学位を授与するに十分な学識を有するものと判断し、合格とした。