

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	A Graph Spectral Approach to Human Action Recognition from Depth Maps
著者(和文)	KerolaTommi Matias
Author(English)	Tommi Matias Kerola
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第10383号, 授与年月日:2016年12月31日, 学位の種別:課程博士, 審査員:篠田 浩一,徳永 健伸,小池 英樹,藤井 敦,下坂 正倫
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第10383号, Conferred date:2016/12/31, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	KEROLA Tommi Matias		
論文審査 審査員		氏 名	職 名		氏 名	職 名
	主査	篠田 浩一	教授		下坂 正倫	准教授
	審査員	徳永 健伸	教授	審査員		
		小池 英樹	教授			
		藤井 敦	准教授			

論文審査の要旨 (2000 字程度)

この論文は、“A Graph Spectral Approach to Human Action Recognition from Depth Maps”と題し、英文7章から成っている。

第1章「Introduction」では、研究の背景と目的について述べた上で本論文の構成を示している。まず、近年、通常の映像とともに被写体との距離(深度地図)も観測できる安価な RGB-D カメラが登場し、それを用いて、「右手を振る」、「立ち上がる」などの人間の動作を認識する研究が盛んになっていることを述べている。その上で、本研究の目的は、その認識性能の向上のために深度地図から得られるグラフスペクトルを特徴量として用いる手法を開発することであると述べている。

第2章「3D Human Action Recognition」では、RGB-D カメラなどから得られる3次元データを入力とした動作認識の従来研究を紹介し、そこで用いられる特徴量や識別器について述べている。現在はスパース・コーディングで得られる特徴量を入力としてサポートベクターマシン(SVM)で識別を行う手法が主流であることを述べた上で、特に、RGB-D データから骨格構造を抽出する際に起きる誤りに対して頑健であり、かつ、学習時とは異なる角度から撮影された場合でも認識性能が劣化しない方法が必要とされていることを述べている。

第3章「Graph Signal Processing」は、本研究でグラフスペクトル特徴の抽出のために用いるグラフ信号処理技術についてその概要を述べている。種々のグラフスペクトル特徴の中でも、グラフラプラシアン行列から得られる spectral graph wavelet transform (SGWT) が有効であり、各種の応用で用いられていることを述べている。また、スパース・コーディングと比べ、グラフ構造を予め与える必要がある反面、その計算量はより少ないことを述べている。

第4章「Representing Actions as Graphs」では、ある特定の時刻における動作の状態を表す3次元データをグラフとして表現する方法について議論している。3次元データから抽出された骨格構造を用いる方法と、spatio-temporal key point (SKTP)を用いる方法を提案し、前者は大局的な関節の動きを表現するのに優れ、後者は骨格構造の抽出誤りなどのノイズに対しより頑健で局所的な動きを表現するのに適していると述べている。また、撮像角度の異なりを正規化する手法も提案している。

第5章「Spectral Graph Sequences (SGSs)」では、3次元データを表すグラフの時系列をグラフ構造とみなし、そこから特徴量を抽出し動作を識別する方法を提案している。この方法では、まず、各時刻に対応するグラフから SGWT を抽出し、それに対し時系列方向にピラミッド・プーリングを行い特徴量とする。次に、骨格構造から得られた特徴量と SKTP から得られた特徴量に対し、各々独立に SVM を用いて認識を行い、そのスコアを統合して認識結果を得る。さらに、SGWT の算出に必要な計算量を効果的に削減する方法も提案している。

第6章「Experimental Evaluation」では、第4章と5章で提案した手法の評価を行い、その結果を考察している。研究利用目的のために公開されているデータベースを用いて性能評価を行い、特に撮像角度の異なるデータを収集した N-UCLA Multiview Action 3D Data set を用いた評価で、従来手法のうちもっとも性能の高い手法の認識率が 75.8%のところ、提案手法の認識率は 90.8%と大幅な改善を得たことを述べている。さらに、提案手法を構成する各々の要素について、性能向上への寄与の有無とその程度について解析し、提案手法による改善の理由を考察している。

第7章「Conclusions and Future Work」では本論文で得られた成果をまとめ、将来の研究における課題について述べている。

以上、本論文は、RGB-D カメラを用いた人間の動作認識に対して、グラフ信号処理技術を用いた方法を提案し、その有効性を評価実験により確認している。本論文の成果は、様々な物体の動きを自動認識する技術の発展に貢献するとともに、信号処理と機械学習を用いたパターン認識分野に学術上寄与するところが大きい。よって本論文は博士(学術)の学位論文として十分価値があるものと認める。