

氏名（本籍） 齋藤裕佑（東京都）
学位の種類 博士（工学）
学位記番号 乙第992号
学位授与の日付 2018年3月19日
学位授与の要件 学位規則第4条第2項該当
学位論文題目 深度映像による時系列的な位置情報を用いた部位分割及び部位着目による人物動作認識手法の提案

論文審査委員 （主査）教授 西山 裕之
教授 大和田勇人 教授 滝本 宗宏
准教授 原田 拓 准教授 竹村 裕

論文内容の要旨

人物動作認識技術とは、人間の現実世界での”動作”というアナログな信号を、コンピュータ上における一種のデジタル信号に置き換えたのち、学習及び認識する技術のことである。動作認識技術は1980年代から研究され、在宅健康管理システムや危機監視システムなどへの応用が期待されている。人間の動作を詳細に認識することができれば、より高度なサービスの展開が可能になるだろう。例えば、家庭内において”水を飲む”という動作や、”胸を抑えて倒れ込む”という動作を識別し認識できれば、脱水症状に陥る前に別居している家族に警告を伝えたり、身体的な異常が生じてから救急が駆けつけるまでのタイムラグを短縮させることが出来る。

本論文では、このような人物動作認識技術において、特に屋内における実生活環境に対応し、なおかつ認識する動作のラベルをより細かく詳細に生成し、さらに効率よく認識を行うことのできる動作認識手法を提案することを目的とする。屋内における実生活環境では、照明環境、撮影角度や被写体の角度、またカメラと被写体の間の障害物などによる閉塞(オクルージョン)が生じるなどの問題が発生する。そこで、照明環境、撮影角度や被写体の角度に依存しづらい、深度映像による動作認識技術が研究されてきた。深度映像とは、赤外線などによって撮影範囲内の1ピクセルごとに、カメラと物体との距離情報を内包した時系列データを指す。

ここで、既存の深度映像による動作認識技術は体全体の動作を一つのラベルとしているものが多く、十分に閉塞を考慮できているとは言い難い。そこで本研究ではまず、身体の

一部が閉塞によって塞がれている場合でも可視可能な部位の認識に影響を与えることがなく、また部位ごとのより細かな情報の提供が可能となるような特徴量生成の手法を提案する。骨格認識技術により関節位置などを取得した上で、部位ごとに基準となる関節である、静的な関節を設定する。その関節の方向情報からいくつかの、階級という空間を分割した領域を決定する。このように、静的な関節とその関節ごとに応じた空間を設定することで、互いの部位が独立したヒストグラムを生成することができる。例えば、"投げる"という行動において、例えば上半身の中央部が傾いたときでも、腕の動きが同じであれば、腕の部位におけるヒストグラムは上半身の傾きに影響されない。最後のステップとして、静的な関節から続く全ての関節の位置情報の推移を、分割した空間の領域にカウントしていくことで、ヒストグラムを生成する。また、既存研究[36]と同様な分類の行動ラベルについて、閉塞の有無や姿勢など、それぞれの条件を設定した上でデータを取得し、ヒストグラムを生成し評価するための、行動のデータセットを作成する。そして、閉塞によってヒストグラムがどのように影響されるか、また被験者によってどのように変化するかを、距離尺度である EMD(Earth Mover's Distance)[7] や Histogram[8] という指標を用いて評価を行う。

一方、咀嚼と健康には関連があることがよく知られている。咀嚼動作は、顔という部位の中でも口における関節運動の一つと捉えることができる。那須ら[48]によれば、咀嚼能力と健康余命の間には相関があり、十分な咀嚼能力を持っていることが長く日常生活を送ることができると言える。ここで健康余命とは、不健康な状態に陥り、日常生活に支障をきたすまでの年齢としている。不健康な状態とは、ADL(日常生活動作)に関する7つの質問項目(入浴、更衣、食事、起床、歩行、外出、排泄)と IADL(手段の日常生活動作)の7項目(食事の準備、買い物、金銭管理、電話、軽い家事、交通機関の利用、服薬管理)の合計14項目のうち、いずれか1項目以上で「非常にむずかしい」あるいは「できない」と回答した場合を指す。身体全体としての動作だけではなく、咀嚼動作のような、より細かい部分に着目した動作も含めて認識することができれば、見守り対象者の健康を管理することに役立つだろう。そこで本研究では、身体全体のうちの一部として顔の骨格動作に着目し、顔の表出情報を用いて、咀嚼動作を認識する手法を提案する。顔の表出情報とは、顔の輪郭や目や鼻などの特徴点となる位置情報および、目が空いている、笑っている、または舌を出しているなどの状態を含めた情報である。咀嚼している時の顎の動きに着目し、咀嚼している時の顎の動きの繰り返し性、すなわち定常性を考慮した手法を提案する。その際、矩形波窓[10]によるノイズ除去を行った上で、自己相関関数による定常性を特徴として用いる。また、咀嚼している時の顔の表出情報を記録したデータセットを作成し、咀嚼しているかどうかだけではなく、何を咀嚼しているか、あるいは誰が咀嚼しているかについて評価実験を行う。

さらに、身体の動作と顔の動作を包括して認識する手法を提案する。その際、身体の動作から顔の動作に認識の焦点を移動できるかどうかを検討する。時系列的な動作認識においてだけでなく、様々な非時系列的な環境認識などにおいても、センサーに映る情報を全て処理しているのでは計算リソースを無駄にってしまう。特に、時系列的なセンサー

情報は容量が大きくなる傾向にあり、処理すべき計算量も非時系列的な情報と比べて大きくなる傾向にある。本研究においては、まず第 2 章で述べる身体全体の部位の認識手法が実行され、その際、手と顔の距離やあるいはその時系列的ヒストグラムから、顔の動作認識を行う手法を提案する。また、食べる、飲む、携帯端末で話す、の 3 つの動作について、身体の骨格情報と顔の骨格情報を同時に取得できる環境にて記録し、データセットを作成した。そして、身体の動作と顔の動作が連動しているかどうかを確認し、認識の焦点を変更することが可能かどうか検証を行う。

論文審査の結果の要旨

本論文は深度映像情報を用いて、屋内実環境における閉塞（オクルージョン）を考慮した人物の動作認識手法を提案している。従来の深度情報を用いた人物動作認識の研究では、特定の実験環境内で全身に対する深度情報を用いた動作内容の分類器を生成する。そのため、机や椅子などの家具が混在する屋内実環境等、閉塞が生じる環境では人物に対する深度情報が正確に収集できないため、分類器の判定精度の低下が問題となり、実生活環境におけるユーザへのサービスへ適用することは困難であった。

本研究では、実生活環境における閉塞を考慮するために深度画像情報から得られる身体骨格における関節群の位置情報を 6 つの部位に分け、その部位ごとに独立した特徴量を生成することで、機械学習器を用いた部位ごとの身体動作に対する分類器の生成を可能にしている。そして、生成した 6 つの分類器を用いたアンサンブル推定を行うことで、身体動作の判定を実現している。これは、従来の研究にはない独創的な人物の動作に対する認識手法であり、本手法を用いることにより、家具などにより局所的な閉塞が生じたとしても、可視部位に着目した動作認識が可能になる。本研究では、機械学習器として、ラベルごとの確率推算値を算出して判定するランダムフォレストを用いている。そして 6 つの分類器によるアンサンブル推定の重みに、部位ごとの確率推算値を適用することで、閉塞が存在する空間における身体動作の判定精度の低下を最小限に留めることに成功している。これにより、深度映像情報から得られた身体骨格における関節群情報を用いた屋内実環境における人物の身体動作認識が可能となる。

次に、身体骨格における一部分として顔を一つの部位と見做し、顔における口や顎関節などの動きによる動作認識にも着目し、顔部位に関わる動作（食事、会話）を推定する手法を提案した。具体的には、深度情報を用いて顔部位に関わる動作に対する特徴量を生成し、機械学習器を用いた分類器の生成を行い、その動作の判定を可能にしている。本研究では、分類器の作成において、口の動きの定常性を特徴量として加えることで、その精度の向上に成功している。これにより、人物の身体動作認識だけでなく、顔部位を含む人物動作認識が可能となる。

以上により、本研究では顔部位における動作内容の認識を含む人物動作認識を、深度映

像から得られた情報に基づき実現するための手法を提案し、その有効性を示した。さらに、本提案内容を屋内の実生活環境におけるユーザへのサービスへ適用するための考察も行っている。本考察では、単純に身体動作認識と顔動作の認識の組み合わせを行うだけでなく、身体の 6 部位の分類器に顔部位の分類器を加えてアンサンブル推定することにより、人間の複合的な動作の認識を可能にし、実生活環境におけるサービスの実現を目指している。これは本研究の特徴である、人物部位の分割及び着目を活用した考え方であり、今後、顔以外の部位を新たに特徴量として加えた場合にも、拡張可能な手法であると考えられる。

以上の様に、本論文は屋内実環境における閉塞（オクルージョン）を考慮した人物の動作認識に対する情報工学的手法として独創的かつ先端的な研究であり、今後の高齢化および単身者の増加に伴う健康管理サービスや様々なユーザ支援サービスへの貢献が期待される。よって、本論文は博士（工学）の学位論文として十分に価値あるものと認められる。