

報告番号	甲	第	号
------	---	---	---

主 論 文 の 要 旨

論文題目 形状モデルフィッティングによる顔および人体の高精度な姿勢推定
氏 名 木下 航一

論 文 内 容 の 要 旨

現在、我々の生活の多くの活動は機械によって支えられている。家電製品、自動車、医療機器、交通システム等の公共インフラ、工場における生産ロボットなど、その例は枚挙に暇がない。機械の高性能化に伴い、我々の生活の中で機械が果たす役割はますます大きいものとなってきており、人と機械が円滑にコミュニケーションすることの重要性は日々増加している。これまで一般的に、機械は人の状況に関わらず画一的なサービス・機能・情報等を利用者に提供し、人がそれに合わせる必要があった。しかしながら、人が発信するさまざまな情報を機械がセンシングすることができれば、機械が人の状態を理解し、それに対して最適なサービス・機能・情報等を提供することが可能になると考えられる。例えば自動車の例では、運転者の顔向きや視線を検知して、危険な状況にあれば警報を発する、速度を緩めるなどのサービスが考えられる。また生産ロボットの例では、周囲の人の姿勢を検知することで、その人の状態に合わせて作業速度を調整したり、部材を保持する、部品を手渡すといった共同作業が可能となる。

本論文ではこのような人と機械が協調した社会の実現を目指し、機械が人の行動を認識、理解することを目的として、人の状態を推定する技術に関する検討を行った。具体的には対象を顔および人体とし、これらの姿勢を高精度に推定することを目指した。画像による顔および人体の姿勢推定を実現するためには、部位の位置を高精度に検出することが必要になる。ここで部位とは、顔の場合は目、鼻、口等であり、人体の場合は頭、腕、胴体などを指す。これは“画像中から対象物（顔もしくは人体）を検出する”、“検出された領域内で部位の詳細な位置を検出し、姿勢を推定する”、の2段階の処理によって行われることが一般的である。

本論文では後者の部位位置検出に関して、一般的な顔や人体の形状を表す“形状モデル”を画像上に適合（フィッティング）させる新たな手法を提案する。ただし、

顔と人体とでは、形状変化の仕方や見え方に大きな差があることから、全く同一の手法で両者に対応することは難しい。本論文では、顔および人体に対して、それぞれ適した特徴量サンプリング手法を提案し、さらに人体に対しては、大きな姿勢変化に対しても精度よくフィッティングするための手法を提案する。

また人体に関しては、前者の検出技術についても課題が多い。特に、従来技術では体の一部が隠れた場合に、検出性能が大きく低下するという問題がある。そこで、本論文では、人体検出の性能向上についても検討を行い、一部の部位が隠れた際の性能低下を抑えるための、新たな手法を提案する。

本論文は5章で構成される。第1章は序章であり、研究背景と目的、意義について述べる。第2章では、摂動特徴量を用いた新たなフィッティング手法を提案し、その原理および顔画像への適用方法について述べる。第3章では、一部の部位が隠れた状況でも高い精度を維持するための新たな人体検出手法を提案する。第4章では、第2章で導入したフィッティング手法を人体に対して適用するための方法について述べる。また、人体特有の課題に対応しフィッティングを高精度化するための改良手法についても説明する。第6章では、本研究を総括し、今後の課題および展望について述べる。以下、第2~4章の概要を述べる。

第2章 摂動特徴量を用いた形状モデルフィッティングによる高速・高精度な顔特徴点検出：

顔画像に対するモデルフィッティングは影や表情変化などのノイズの影響を受けやすく、高精度な結果が得られにくいという問題があった。また多数の繰り返し演算を必要とするため、実時間処理の実現は難しかった。本章ではこれらの問題を解決するため(1)形状モデル上でのサンプリング点の構造的配置(2)特徴量による形状パラメータの摂動量推定、を行うフィッティング手法を提案する。目や口など顔特徴点周辺の局所特徴量は、影や表情変化などに対して比較的ロバストであり、フィッティングにおける重要な情報となる。しかし局所的な情報だけに着目すると、正解位置からの少しのずれに対してもフィッティングが困難になりやすく、また、周辺情報を用いないため局所解に陥りやすいという課題がある。(1)のサンプリング手法は、形状モデル全域を疎に覆いながら、各特徴点周辺に関しては密なサンプリングを行う。これにより影や表情変化などのノイズに対してロバストなフィッティングが実現できる。また(2)により、正解位置からずれた状態の形状モデルの特徴量から、そのモデルのずれ量(摂動量)が推定できる。これによって特徴量から一度の行列演算で形状修正することが可能となり、処理時間を大幅に削減できる。

性能評価には、公共のデータベースに加え、さまざまな表情変化や実環境画像を含む多様な条件のデータベースを使用した。評価の結果、提案手法は照明変化や表

情変化などさまざまなノイズに対して、従来手法と比較してロバストな性能を示した。1 画像あたりの処理時間は従来手法が 0.264sec であったのに対して、提案手法は 0.017sec であり、約 15 倍の高速化を実現した。このことにより、実時間処理が可能な、高速・高精度な顔姿勢推定技術を実現した。

第 3 章 部位スコアマップ多段統合による隠れにロバストな人体検出：

静止画からの人体検出技術は、不審者検知や危険行動検知、車載機器による歩行者検知など、幅広い用途での活用が期待され、盛んに研究が行われている。しかしながら従来の人体検出技術では、人体の一部に隠れが生じた際に検出能力が大きく低下することが一般的である。人体検出が必要となる多くの用途において、いつでも全身が隠れなく見えているという状況を想定することは困難であり、実用上大きな課題となっている。

本章ではこの問題に対応するため、部位スコアマップの情報を多段的に統合する人体検出手法を提案する。

スコアマップとは、検出器の出力をマップ状に記録したものである。検出器は画像の各位置で識別処理を行い、その結果を数値として出力する。一般に 0 より大きい値はその場所に物体が存在することを示し、値が大きいほど、その可能性が高いことを意味する。この数値を各画素ごとに記録したものがスコアマップである。

提案手法はスコアマップを検出器への入力として用いることで、人体画像から生じる検出パターンとそれ以外の画像から生じる検出パターンをより明確に区別することができる。さらに多段統合の仕組みを提案する。これは、1 つの部位に隠れが生じても大きな影響を受けないような統合検出器を複数学習し、これらの出力をさらに統合して最終的な検出結果とする手法である。これにより、一部の部位が隠された場合と人体以外の画像とを、より明確に区別することが可能である。

性能評価には、一般的に人体検出技術の性能評価に用いられる公共の画像データベースを使用した。この時、同データベース画像を加工することで一部の部位に隠れを生じさせたデータセットを構築し、これによって一部の部位が隠れた状態についても検出性能の評価を行った。

評価の結果、提案手法は従来技術と比較して、隠れない画像に対する検出率（誤検出 10% 時）が 88.1% から 96.0% に向上し、さらに肩の隠れに対して 50.5% から 78.2%、左半身の隠れに対して 35.6% から 83.2%、それぞれ検出率が向上した。以上により、従来技術と比較して隠れに対してロバストな、高精度な人体検出技術を実現した。

第 4 章 摂動特徴量を用いた形状モデルフィッティングによる高速・高精度な人体姿勢推定：

顔と比較して、人体は形状変化の自由度が高いことに加え、衣服等による見えの変化が激しく、これらによる特徴量変化が大きい。そのため人体に対して高速・高精度なモデルフィッティングを実現することは、非常に困難な問題である。

本章ではこの問題を解決するため、人体画像に対して、撮動特徴量を用いた形状モデルフィッティング手法を適用することを提案する。しかし人体の持つ上記のような特性により、顔に対して有効であった手法が人体にもそのまま適用できるとは限らない。本章ではまず、顔に対して有効であった特徴量サンプリング手法は人体に対しては不適切であることを示す。その上で、近年人体検出に用いられる HOG 特徴量を、人体モデルフィッティングに適用することを検討する。その結果、HOG 特徴量を形状モデルの各ノード上でサンプリングする手法によって、人体に対するモデルフィッティングが実現できることを確認した。さらに人体検出による位置ばらつきや、大きな姿勢変動がある状況でも高精度なフィッティングを実現するため、Coarse-to-Fine の考えに基づく段階的フィッティング手法を提案する。

性能評価の結果、提案手法は従来技術に比べて高い性能を示し、従来技術と比較して、フィッティング成功率が 50.0% から 74.0% に向上し、段階的フィッティングのない場合と比較しておよそ 9% の性能向上が得られた。処理時間は従来手法が 21.5sec、提案手法は 0.28sec であり、大幅な高速化を実現した。

以上のように、本論文では顔および人体に対して、撮動特徴量を用いた形状モデルフィッティング手法を提案し、高速・高精度な部位位置検出を実現した。また人体に関しては、部位スコアマップの多段統合によって、従来手法と比較して隠れにロバストな、高精度な人体検出を実現した。これらにより、顔および人体に関して高精度な姿勢推定を実現した。