

## 論文の内容の要旨

論文題目 Action Recognition Based on Discriminative Classification Techniques and Kernels  
( 識別的推論とカーネルに基づく動作認識 )

氏 名 下 坂 正 倫

人の動作の認識は、知的な機械システムが、人と親和に居住空間で活動するための重要な技術の一つである。本研究では、動作認識を統計的推論の問題として捉え、動作に関する人の知識と観測された動作データを、柔軟かつ合理的に統合する手法を構築し、その有効性を検証することを目的とする。

従来、動作の認識は、直感的な計算機対話技術などを実現するための基盤技術として、熱心に研究されてきた分野である。実際、手話、ジェスチャなどを想定したものが数多く研究されてきた。一方、情報通信技術の発達と普及に伴い、計算機や知能ロボットを生活空間に組み込むことで生活者の見守りや支援を可能とするシステムに関する研究が盛んに行われる現在、日常生活の行動に現れる動作を理解するための方法論の構築は重要である。よって、本研究は認識対象の動作として、座っている、立っている、腕を組んでいる等の生活行動に現れる動作を対象とする。

本研究では、従来行われていた多くの動作認識手法の研究と比較して、以下の特徴を有する。まず一つに、認識対象となる動作の共起性、例えば、腕を組んで座っている、などの状況を陽に考慮し認識処理する点である。従来のジェスチャを想定した認識手法では共起性への対応は困難であった。次に、本研究は一貫して識別的な推論のメカニズムに基づいて認識手法を構築している点である。これは従来の多くの研究で採用されている、動的ベイジアンネットワークとも呼ばれる生成的な枠組に基づく方法論と異なり、人が運動を観察し認識した結果との整合性を基準とするものである。さらなる特徴は、人の運動の時系列性や非線形性の扱いを容易にし、統一的な定式化を提供しうるカーネル法に基づくというものである。カーネル法とは2つのデータ間の類似計量に基づく機械学習法の総称で、1990年代以降盛んに研究されているものである。

機械学習の観点で俯瞰すれば、本研究は、カーネルに基づく時系列データ解析問題として捉えることができる。本研究の重要な課題は、従来のカーネル法における問題点を明らかにし新たなカーネルに基づく認識手法の提案と検証である。動作認識を考える上で指摘しうる従来のカーネル法の問題点として、人の運動の時系列特性を陽に考慮する効率的な類似度計算法の研究が不十分であることや、ノンパラメトリックなモデリングに依存する、実用の観点で重要な認識および学習におけるコストが増大してしまうことなどが挙げられる。本研究では、このような問題に対して解決を行う。具体的には、認識性能の高さおよび認識器構築にかかるコストの低減、認識器設計の容易化を実現するために、動作認識というある種の複雑な時系列データ処理の問題を 1) 動作認識の最終的な出力の整合性を考慮する大域的識別、2) 下記の運動間類似度に基づき大域的識別の手がかりとなる情報を提供する、認識対象の各動作についての局所的な識別、3) 運動特徴間の類似度に相当するカーネル計算、の3つの処理に明示的に分割し、各々の要素について新たな計算法を構築する。本研究は、モーションキャプチャデータを利用した性能検証を行い、従来の動作認識法と比較して設計の自由度の高さ、実用上の扱いやすさ、認識の頑健さの観点でバランスよい性能を有することを示す。

以下、論文で提案されている手法とその有効性についてまとめる。

第1章 “Introduction”では、本研究の背景、目的および本論文の構成について述べる。

第2章 “Formulation of Action Recognition as Discriminative Statistical Inference Problems”では、本研究で対象とする動作認識の問題設定を述べ、それらの知的構造化としてグラフィカルモデル上の統計推論として捉え定式化し、その解法として提案手法の概要を述べる。本論文の手法とは、動作の共起性・時間依存性を考慮するグラフィカルモデル上の指数分布族として定式化し、入力時系列運動に対する最終的な認識結果の推論を行う大域的な識別と、その構成要素として大域的な識別の手がかりを提供する、認識対象の各動作の局所的な識別からなるものとする。局所的な識別においてカーネルを利用することで、幅広い運動特性を扱わなくてはならない状況においても、統一的な定式化の実現が可能となることを示す。あわせて、従来の関連手法の特性を列挙し、本研究の手法の妥当性を定性的に示す。

第3章 “Local Action Classification with Kernels, and its Performance Adjustment by Maximum Margin Learning”では、第2章で述べた認識のフレームワークの一要素である、カーネルに基づく動作の局所的な識別について、その定義と最大マージン学習に基づく構築法を提案し、モーションキャプチャデータを用いた性能検証実験により有効性を検証する。

第 4 章 “Margin-Based Query Learning Leveraging Global Similarity of Motion for Local Action Classification”では、動作の局所的な識別の構築プロセスにかかる人的コストを低減することを目的に、問い合わせ学習法と呼ばれる考え方を援用した学習法を提案する。ここでは、第 3 章で提案した学習法の最適化基準を手がかりに、そのわずかな修正のみで問い合わせ学習法に拡張することで、少ないヒューリスティックでの実現を狙う。

第 5 章 “Local Action Classification Encoding Qualitative Prior Knowledge”では、第 3 章、第 4 章で提案した学習法における問題点を明記し、その問題点を解決する。ここでの問題点とは、動作認識における運動データの取得にかかるコストや、非日常的な運動の取得が困難であるために生じるデータの偏在性に起因する、新奇データに対する出力の不安定性である。これが、ノンパラメトリックな表現による識別の定式化と、ガウシアンカーネルに代表される特徴空間内で局所的な作用をもたらすカーネルに起因することを示す。本章では、人の定性的な事前知識を利用し、特徴空間全体に識別則が作用する情報を組み込むことで、それへの解決を狙う。ここではガウシアンプロセスを事前分布とする確率的な枠組で局所動作識別を定式化し、動作の事前知識の組み込みの統一的な扱いを実現し、その検証として、非日常データに対する挙動を通じて提案する計算法の有効性を明らかにする。

第 6 章 “Kernel Design and Feature Discovery Method for Pose Classification”では、運動間の類似計量であるカーネルの設計法について述べる。この章では、立っている、腕を組んでいる、等の姿勢の認識の話題に注力し、姿勢の認識で重要な、着目すべき運動部位の選出と選出した運動部位に基づく運動特徴の設計について、人の事前知識に基づき問題を簡便化する方法を述べる。また、それに対する補完的な位置づけで、大量データを利用して着目すべき運動部位を自動発見する手法について述べる。この自動発見法は、パターン認識問題における特徴抽出の話題と関連するものであるが、本章で提案する手法は、第 3 章と第 5 章で述べた動作の局所的な識別の汎化性能とカーネルとの関連性に基づき、認識に有効に機能するカーネルは、認識に重要な運動特徴によって構成されるという考えから、自動発見を通じて、識別性能の向上も狙った方法論である。

第 7 章 “Marginalized Bags of Vectors Kernels for Dynamic Action Classification”では第 6 章とは対照的に、歩行などの動的な動作に特化した、すなわち運動の時系列性を陽に考慮するカーネルの構築を行っている。この章では人の運動を確率的な時系列モデル、特に力学特性を考慮したモデルで表現し、そのモデルを巧みに組み込むことで、2 つの時系列運動間の類似性を計算する手法を提案する。この章で着目するモデルはスイッチングリニアダイナミクスと呼ばれる、人の運動表現に優れる時系列モデルである。本章で提案する計算法は、このモデルの持つ統計的性質と動作認識の観点で扱いやすいカーネルの性質を整理することで、既存のカーネル計算法と比較し、計算量の観点や導出の安定性の双方で優れるものである。特に、既存のカーネルでは困難であったオンラインでの処理が可能となることは、動作認識のみならず機械学習一般に対する貢献である。この章では、脚運動を例にとり、本章で提案するカーネルの有効性を検証する。

第 8 章 “Multi-Label Action Segmentation with Multi-Task Conditional Random Fields”では、大域的な動作の識別の実現として動作の共起性や時間的な連続性を考慮し、第 3 章から第 7 章まで述べた計算法を統合するモデルを提案する。第 2 章で述べた定式化の詳細にふれ、動作の共起性を考慮した効率的な動作のセマンティクスの符号化を行う。以上の定式化の下で、重要な問題となるのが、効率的な動作の推論法である。動作の共起性を考慮した場合、解析的な方法で推論を行うことは不可能であるためである。本章では、計算効率の高さと近似精度の高さに着目し、変分法に基づく推論手法を新たに提案する。推論法の効率性の観点と、動作の共起性の考慮の有無による頑健性を、人工データとモーションキャプチャデータを利用し有効性を評価する。

第 9 章 “Conclusion”で本論文の結論としてこれまで各章で述べた内容を整理し、識別的な枠組とカーネルに基づく動作認識手法の発展について記す。