

人間行動の認識と遂行 のための実世界知能情報学

國吉康夫 大津展之
情報理工学系研究科知能機械情報学専攻

概要

人間の振舞いを認識・理解し、それに呼応してタイミングよく適切に情報提供したり手助けする機能は、知能情報システムの一つの理想形であろう。その実現のためには、複雑・不確実な実世界中でも信頼性よく行動の認識・遂行ができる手法の確立が必要である。本研究では、統計的パターン認識、ニューラルネット、ヒューマノイド実験などのアプローチにより、新しい適応的行動認識・遂行手法の構築を進めている。

1 はじめに

行動認識機能の研究の歴史は決して新しいものではないが、近年、マルチモーダル対話システムの発展形として、人間の振舞いを視覚等で自動的に認識し適切な応答を行う Perceptual User Interface の研究や、セキュリティへのニーズの高まりに伴う知的監視システムの実用化などにおいて、急激にその重要性が高まっている。

一方、ヒューマノイドロボットの開発が急激な進歩を遂げ、実世界応用の研究も進展している。この中で、人間協調タスクは非常に重要な位置をしめており、人間の行動理解とそれに基づいてどのように協調、手助け行動を生成するかという問題は、重要性と現実性を増している。

本研究では、実世界の中での行動認識と行動生成を、複雑、不確実な状況下で、動的な対象について信頼性よく安定に意味を抽出したり実現したりする問題として捉え、統計的パターン認識、適応ニューラルネット、ヒューマノイド全身行動実験などのアプローチから、新たな原理と手法の確立を目指している。

2 実世界情報システムの数理的基礎

実世界の情報は時空間的に分布するパターン情報であり、曖昧さや不確かさ、また時間的な動きや変容を特徴とする。従ってこれらを取り扱う上で、パターン認識と学習、特に時間を含む多変

量解析的手法や、確率統計的手法（ベイズ推定、ベイジアンネット）といった情報数理的手法がキーとなる。これら理論的基礎の上に、認識、理解、推論、学習機能の研究を進めている。具体的には、情報統合と学習・自己組織化、マルチモーダル学習に基づく概念獲得、実環境での人の仕草や表情の認識・理解、状況認識と適応行動などに取り組んでいる。

2.1 汎用的特徴抽出手法としての高次局所自己相関（HLAC）法の数理的拡張と整備

静止画像（2D空間）認識のための特徴抽出法としてこれまで開発した高次局所自己相関（HLAC）を、時間を扱えるように拡張整備した。これによって、時系列（1D時間）の場合に自己回帰モデルとの整合性が良くダイナミックスの同定が可能となること、また立体HLAC（2D空間+1D時間）により動画像の効率の良い認識が可能となる知見が得られた。一般に立体HLACは、3次元の形で得られる任意のデータ（3-way data）からの汎用的な特徴抽出手法となるものであり、実世界情報を扱う強力な数理的手法が得られたことになる。

2.2 動画像認識への応用

動画像は2D（空間）+1D（時間）の立体（3-way）データと考えられる。これに立体HLACを適用することにより、その位置不変性と対象加法性を生かし、動物体の位置に依らない（分節を要しない）認識、また複数動物体の同時認識が可能となることを実データにおいて示し、動画像認識手法としての有効性を示した。具体的には、5人の4行動（左右方向の歩く・走る）に対して99.8%の認識率で個人の同定と行動(gait)の同定が得られている。

また、空間2D（画像）のHLAC特徴抽出

を動画像に適用すると時系列データが得られるが、それらを自己回帰分析と組み合わせることで、より複雑なジェスチャなど、人間の動作の認識が可能となる。具体的には、HMM 識別を用いて、48名の17種のジェスチャデータに対して95.7%の高い認識率が得られている。

さらに、移動物体の追跡や異常行動の検出について、新しい有効な手法を開発した。

2.3 顔認識への応用

社会生活において顔は重要な役割を占めている。顔を見ることで、個人、性別、年齢、表情などを知ることができる。こうした顔情報を認識する機能をコンピュータに持たせれば、人と情報システム（ロボットを含む）とのより円滑で親しい対話が可能となる。

既に筆者らは HLAC 特徴抽出を顔認識（個人の同定）に応用し高い認識率を得ているが、さらに表情認識へ応用を拡大した。まず、因子分析と合わせることで、位置あわせと顔の切り出しを要しない高速な個人と表情の同時認識が可能となることを確認した。具体的には、9人、7種の表情の認識率は、各々、98%、80.8%、同時 79.7%であった。表情認識は人間にとっても微妙であるので、これらは高い認識率と言える。

さらに、表情など、場所によって重要度が異なる認識課題に対応するため、HLAC 特徴を場所に応じて重み付けし、それらの重みを学習適応的に求める方式を開発した。数理的には、これは従来の Eigen-Face や Fisher-Face 手法を特別の場合として一般化した手法となっていて、認識率が向上することを確認した。

2.4 マルチモーダル学習を用いた概念獲得

幼児は音声と画像を統合することで物事概念を獲得する。この過程を、認知心理的知見も踏まえつつ、正準相関分析などの多変量解析手法や情報理論的な基準を用いてモデル化し、その有効性を確認した。

2.5 移動ロボットにおける随伴行動の適応的獲得

移動ロボットの物体追従問題を具体例に、マルチモーダルなセンサ情報の統合学習を行うことで、既定の安定な行動をする中で適応的に新しい行動を獲得することができるシステムのモデルを考案し、トイロボットで有効性を確認した。これにより、言語などで教えられる固定的な行動ル

ールだけでは厳密に対応できないような環境におかれたとしても、ロバストにタスクを続行することが可能となる（行動の汎化）。

3 人間型身体性の情報構造解明と全身ダイナミック動作への応用

ヒューマノイドロボットの行動実現の研究は急速に進展しており、二足歩行にとどまらず様々な作業状況での行動実験がなされている。しかし、従来の行動生成法は、立位のバランスを崩さないような安定な状態を接続して作られる、いわば準静的な軌道である。このため、動作全体の速度が制限される、瞬間的に大きな力を発生することが難しいなどの問題がある。

一方、人間は身体や操作物体のダイナミクスを活用した行動スキルを有しており、不安定な領域も積極的に利用しつつ、高速な運動や、大きな力を要するタスクをこなすことができる。

本研究の目的の一つは、そのようなスキルをヒューマノイドロボット上に実現し、その行動能力を拡大しようとするものである。

身体や操作対象物のダイナミクスを利用するという事は、それを認識・同定し、それに合わせた適切な制御則を適用することを意味する。実世界中のタスクは複雑で、身体や物体の状態は刻々と変化し、それらのダイナミクスも変化する。従って、ダイナミクスの切り替わり、すなわち境界とその次に存在するダイナミクスの同定が必要となる。また、系の状態は不確定性を有するので、適切な制御則とは、従来型のように系の状態軌道を完全に規定するのではなく、系本来のダイナミクスを破壊せずその安定性や発散的方向性を利用するため、必要以外の時には系に介入しないことを意味する。

このことから、系への介入は疎であり、ダイナミクス境界における状態制御（新たなダイナミクス成立のため）や制御則切り替えを主眼とすることになる。このような制御戦略は大域的には安定であり、多少の身体・環境状態の変動に対してロバストな、記号的性質を有すると考えられる。これはいわば、「コツ」に相当する概念であり、個体を超えて伝達可能な意味の情報といえる。従って、これはコミュニケーションや協調・介助において、行為概念記号や他者行動への介入点という、要になる情報構造を規定すると考えられる。

これらの内容を具体的に明らかにすることが本研究のもうひとつの目的である。

3.1 全身行動実験用ヒューマノイドロボット

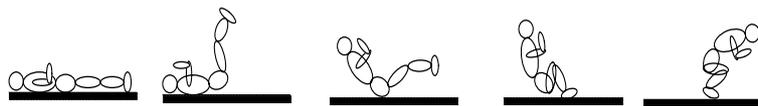
この目的のために、任意姿勢全身ダイナミック運動可能な等身大人間型ロボット（名称：R. Daneel Study 1、右図）を完成した。身長 150cm、体重 70kg、46 自由度。人間に非常に近い関節可動範囲と、瞬発力と素早さを重視し、滑らかな外形を持つ。このため、任意姿勢でダイナミックに自重を操る能力を有し、従来の二足歩行ロボットとは一線を画す。このような全身での任意ダイナミック動作が可能で等身大ヒューマノイドは世界でも他に例がなく、本研究における身体性解明実験のプラットフォームとして非常に強力な役割を果たす。



3.2 ダイナミック起き上がり行動

人間型身体の特徴を活用した行動の例として、ダイナミックな起き上がり行動を取り上げ、人間の行動戦略の分析と上記ロボットでの行動実験を行った。

これは、床に仰向けに寝ている状態から、両脚を振り上げて、その反動を使って一気に立ち上がる行動である（上図）。



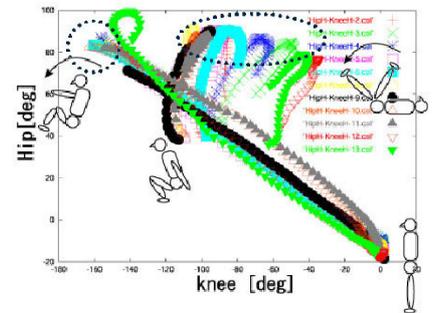
る行動である（上図）。

これは、運動能力の優れた人以外にとっては、多少の技を要する動作である。しかし、一度コツをつかめば、そう難しいものではない。ところが、この運動を厳密に記述する方程式を立て、解くのは難しい。なぜなら、背面の様々な部分で床に接触しつつ転がるため、床からの反力や摩擦力の影響が極めて複雑・不確定だからである。また、これは非周期的な、発散的な（安定状態に収束していない）運動を複数つないだものになっている。

3.3 人間の行動戦略の分析、コツと身体性

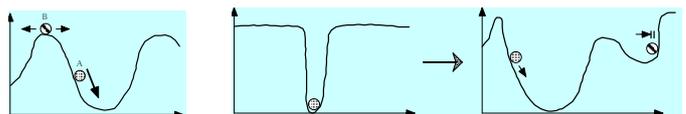
このような動作を人間はどうやって達成しているのか？ 人体の運動をモーションキャプチャ装置でリアルタイム計測してみると、興味深い事

実が判明した。被験者と同じ動作を何回も繰り返して実行して、そのデータを位相空間（各関節角度を座標軸にとった空間）にプロットしてみると、動作のあるフェーズでは軌道が同じ位置に重ね書きされ、別のフェーズでは毎回異なる位置にばらけたのである（右図）。つまり、人間の制御戦略は、動作中の各時点ごとに非一様であり、きちんと制御している部分と適当に流している部分があるかのように見えた。

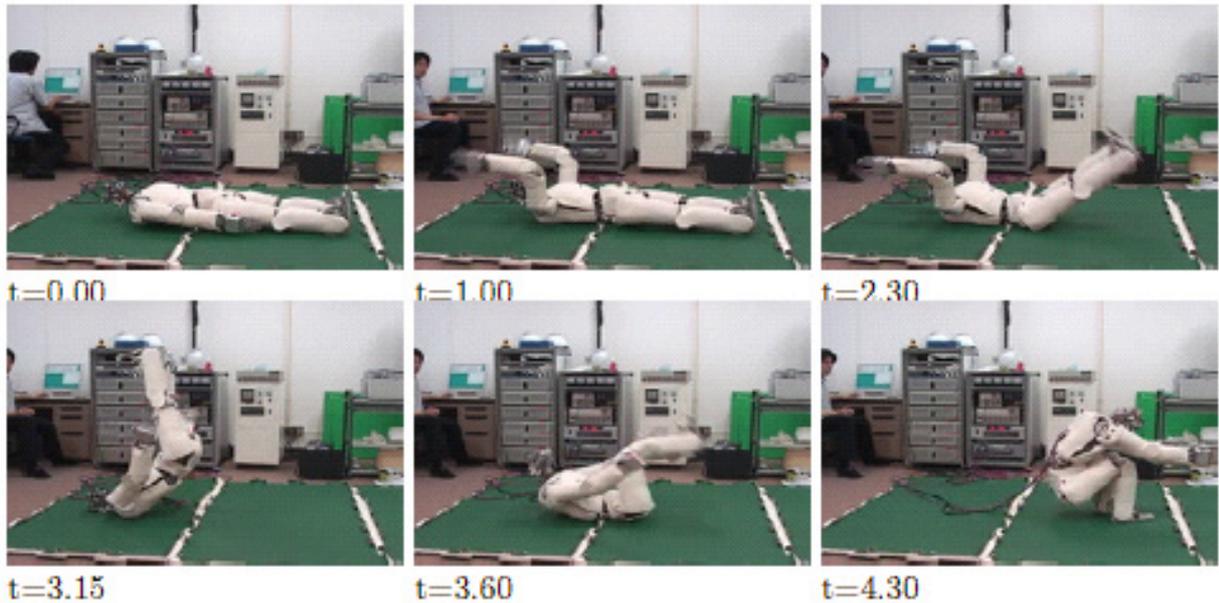


次に、人間の軌道が収束した部分、特に、着地点が腰から足に遷移する時点のまわりに着目して、この運動の力学モデルを解析してみると、確かに、その瞬間にある厳しい条件を満たしていないと起き上がりが成功しないことが分かった。また、その直前の部分では、この厳しい条件を満たすために脚の動かし方やタイミングに一定の調整幅が許されることも分かった。つまり、この起き上がり動作の位相空間中には、ごく疎らに、非常に条件の厳しい状態があり、そこをきちんと通過しないと動作全体が成功しないが、それ以外の部分では、状況に応じて変化させて構わない、すなわち、状態軌道はばらけてよい、ということが分かった。

この動作の力学的性質を簡単に模式化した図を示す。足を振り下げながら転がっている状態は、ポテンシャルの谷の斜面を転がり落ちている玉に相当する。この状態では、玉に多少の力を加えても、その運動を左右することは難しい。逆に、多少の外乱や不確実性があっても、状態変化の傾向は変わらない。従って、この状況では自然の力学に身を任せた方が賢明だ。一方、玉が峰



の頂上にいるときは、わずかな努力で玉をその場で維持したり、思った方に転がしたりすることができる。そして、その効果は大きい。つまり、隣合った力学構造の境界付近が、制御をかけるべきポイントといえる。さらに、このポテンシャルの



谷の形は、身体の形によって決まるので、玉の移動に合わせてタイミングよく姿勢を変化させることで力学構造自体を変化させて玉を制御できる。平らに寝ている状態は深く細い井戸、脚を振り上げ背中を丸めた瞬間になだらかな広い谷に変化し、勢いあまって足の上を超えて前のめりに倒れこんでも、手を突き出しておけば壁となって玉が止まる。

このように、多様なダイナミクスを含む全身運動では、疎らな制御ポイントが現れ、それらは状態軌道としては収束する部分であるらしいことが分かった。これらは、われわれが「コツ」と呼ぶものに相当するのではないかと考えている。そして、これはまさに身体性の一つの具体例である。つまり、身体の物理特性や制約条件ゆえに発生する情動的構造と言える。

3.4 実機ヒューマノイドによるダイナミック起き上がり実験

以上の原理を適用して、実機ヒューマノイドでの起き上がり実験に成功した。身長 150cm、体重 70kg という、小柄で小太りな体格のマシンであるが、身体各部に大きな瞬発力を有し、脚の振り上げ開始からわずか2秒という素早い起き上がり動作が世界で始めて実現された。(写真)。この成果は、国内外のメディア (NHK、民放各社、新聞各紙、CNN, BBC 等) で広く報道されると共に、権威ある国際会議での発表3件並びに国内学会発表2件により学術的に高く評価された。この実験結果とシステム性能、および国内外の反響は、

当初想定した範囲を超えた。

4. おわりに

実世界中での人間と情報システムのインタラクションを支える新たな手法を目指して、時々刻々変化する複雑な実世界情報から本質的な情報を抽出し、制御するための理論基盤、認識手法、ロボット行動制御に関する研究を進め、人間の動作認識、カテゴリ学習、全身ダイナミック動作などの実験に成功した。