

# 確率的因果構造を考慮した日常生活行動の画像認識

Image Recognition of Behavior in Daily Life under the notion of Probabilistic Causal Structure

河田 諭志\*<sup>1</sup>      本村 陽一\*<sup>2</sup>      西田 佳史\*<sup>2</sup>      田中 和之\*<sup>1</sup>  
 KAWATA Satoshi      Yoichi MOTOMURA      Yoshifumi Nishida      Kazuyuki Tanaka

\*<sup>1</sup>東北大学 大学院情報科学研究科

Graduate School of Information Sciences, TOHOKU University

\*<sup>2</sup>産業技術研究所 デジタルヒューマン研究センター

Digital Human Research Center, National Institute of Advanced Industrial Science and Technology

日常生活における人間の行動を観察して認識・理解することは、人工知能の研究における大変重要な基本問題であるとともに、知的情報サービスの実現にも必要な技術である。本研究では日常生活の中での人間の行動を観測した動画等のセンサデータから行動を理解するための人間行動モデルの統計的学習を行う。モデルはセンサにより観測された信号と因果的な関連情報を入力として確率推論を実行し、行動を推定する。本手法は、データマイニングによるセンサデータからの行動検出、機械学習による行動の識別、ベイジアンネットワークを用いた隠れ変数や欠損の推定、確率推論という流れで構成される。

## 1. はじめに

近年、各種センサ技術や情報処理技術の発達により、我々人間の日常生活において多くの電子的情報が発生している。このような実生活における行動の断片として得られた電子的情報を単なる信号のレベルから一歩踏み込んで、意味のある情報として意味づけを行うことが人工知能研究の上で大変重要な課題となっている。また、人間の行動を観察してそれを認識・理解することは、様々な工業製品の開発や乳幼児などの社会的弱者の危険回避を考える上でも大きな関心を呼んでいる。[1]

本研究では、実生活行動において観測されたセンサデータを加工していくボトムアップの処理と、モデルを生成してベイジアンネットワークによる確率推論を行うトップダウンの処理の融合により、人間の日常生活行動から発生した情報から確率的な因果構造を抽出する。その上でこれを、日常生活における実アプリケーションへ適用することを検討する。

## 2. グラフ構造を用いたベイズ推定

### 2.1 ベイズ推定

複数のクラスラベルを  $C_i$  とし、信号パターン  $x$  に対する尤度を  $\Pr(x | C_i)$  とする。また、データが生成される空間に特有な制約や発生頻度の偏りを表す事前知識の確率分布を  $\Pr(C_i)$  とする。これらを用いて、事後確率

$$\Pr(C_i | x) = \frac{\Pr(x | C_i) \Pr(C_i)}{\sum_j \Pr(x | C_j) \Pr(C_j)} \quad (1)$$

を最大化するようなクラスラベル  $C_i$  を決定することが、パターン認識問題におけるベイズ推定である。

期待予測誤差を最小にするようにベイズ推定の性能を上げるためには、事前分布を積極的に利用することが必要であり、データの発生源となる実世界の問題領域の事前知識を反映した事前分布をモデル化することが重要になる。従って事前分布については、データが生成される過程における因果的な構造や状態空間の制約、頻度の偏りなどの良い近似となっていることが重要である。[2]

近年の計算機能力の向上から、複雑な構造を反映できるグラフィカルモデルの適用が十分現実的となってきており、実生活空間に内在する構造や事前知識をうまく扱える可能性が生まれている。中でもベイジアンネットワークは他のグラフィカルモデルと比べてモデルに因果構造を持ち込みやすく、意味的な解釈が容易であることなどの理由から人間が問題固有の事前知識を導入しやすい。最近では、ベイジアンネットワークを用いた応用事例も多数具体的に示されてきている。[3]

### 2.2 ベイジアンネットワーク

ベイジアンネットワークは、様々な事象をノードとして表し、そのノード間の確率的な因果関係をリンクとして表した確率モデルである。各リンクにはそれぞれ、親ノードを条件とした条件付確率が割り当てられている。このような構造を用いることで、ある事象が観測された際のその原因を推定するという確率推論を行うことが可能となる。

### 2.3 事前分布のモデル化

真の事前分布が本来特定の状況 ( $S$ ) において条件付けられた  $\Pr(C | S)$  のようなものであれば、状況をパラメータとして持たない確率分布  $\Pr(C)$  を用いたベイズ推定は、例えデータフィッティングが最適であっても確率的な近似精度は低下してしまう。条件付確率  $\Pr(C_i | S)$  を表すためには適切な説明変数  $S$  を見つけ、それによる依存関係をモデル化する必要がある。

そのため、状況依存性のある事前確率として、重要な説明変数を関係の強さに応じてグラフ状に結合したベイジアンネットワークにより表すという解決方法が考えられる。事前確率をベイジアンネットワークの確率推論から計算し、それと最尤推定に基づく識別器の出力とを統合したベイズ解を求めれば、推定性能の向上が期待できる。また、実際の問題領域でデータが発生する因果的な構造を反映して事前分布をモデル化する、つまりベイジアンネットワークのグラフ構造を問題内の因果的な構造に忠実なモデルとして決定することが重要となる。

## 3. 実験の概要

### 3.1 実験環境

現在、産業技術総合研究所内の実験ルームにおいて人の日常的な生活行動を観測して実験データを収集している。この実験

連絡先: 河田 諭志 (skawata@smapi.is.tohoku.ac.jp)

東北大学 大学院情報科学研究科 応用情報科学専攻

ルームでは、部屋の中の人や物体に超音波センサをつけることで、人や物体の各時刻における位置情報を  $x, y, z$  の座標データとして取得できる。また部屋の天井部分にカメラが設置されており、部屋の中の様子を動画で撮影できるようになっている。さらに、撮影された動画を基に専門のスタッフが手動で、部屋の中の人や物体の行動に対して 1 秒ごとにラベル付けを行っている。例えば対象となる人が歩いている、座っている、立っているといった行動ラベルの情報が利用可能となっている。

これらのデータに対し、どのようにして意味づけを行っていくかを考える。

### 3.2 実験の流れ

センサにより観測された情報から抽出された特徴量を  $x$  とする。この  $x$  と、その時刻における行動ラベル  $C_i$  とのセットを学習して識別器を構成する。理想的な学習後の識別器は近似的に式 (1) の事後確率を最大にするベイズ識別を行うものと見なせる。しかし、クラス間の発生頻度が観測時間や観測場所に依存しているような場合、事前分布  $\Pr(C_i)$  は状況  $S$  に依存したものになっている。そこでこれを条件付確率  $\Pr(C_i | S)$  として考え、ベイジアンネットワークを用いて学習する。これと通常の最尤解を学習した識別器から得られる尤度  $\Pr(x | C_i)$  とを結合した式 (2) の事後確率を最大とするクラスを識別結果とする。

$$\Pr(C_i | x) = \frac{\Pr(x | C_i) \Pr(C_i | S)}{\sum_j \Pr(x | C_j) \Pr(C_j | S)} \quad (2)$$

Intel が公開しているフリーのコンピュータビジョンライブラリ “OpenCV” [4] を用い入力画像  $G$  からオプティカルフローを計算するなどして、いくつかの特徴量  $x$  を求めた。教師信号となる行動ラベルは、センサデータに対して手動で割り当てた。入力画像  $G$  から抽出した特徴量  $x$  を介して、自動的にラベル情報を得る識別器は OpenCV 内のサポートベクターマシン (CvSVM) を用いて構成する。構成された識別器に対して未知の画像  $G$  から抽出した特徴量  $x$  を入力し、割り振られた行動ラベルに対する尤度を計算したものを式 (2) の第 1 因数  $\Pr(x | C_i)$  として用いる。

また、式 (2) 右辺分子の第 2 因数  $\Pr(C_i | S)$  は、ラベル空間における状況  $S$  の下での行動ラベルの事前確率である。今回はこれを計算するために、ラベル間の確率的因果構造を考慮することにする (図 1)。すなわち、場所や時系列を考慮してラベ

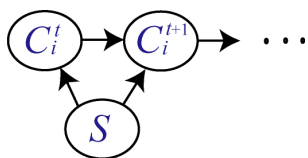


図 1: ラベル間の確率的因果構造

ル情報のベイジアンネットワークを生成して推論を実行することにより、例えば「状況  $S$  で時刻  $t$  に  $C_i^t$  という行動が起きたら、次の時刻  $t+1$  に  $C_i^{t+1}$  という行動が起きやすい」といった因果構造の形で事前知識を導入して  $\Pr(C_i | S)$  を求めることになる。

行動ラベルの実験ルーム内における場所依存性を確認するため、実験ルーム内の幼児の行動に対して手動で割り振られたラベルの、実験ルーム内に定義した  $xy$  座標平面へのマッピングを行った (図 2)。各点の色と行動ラベルは以下の通りと

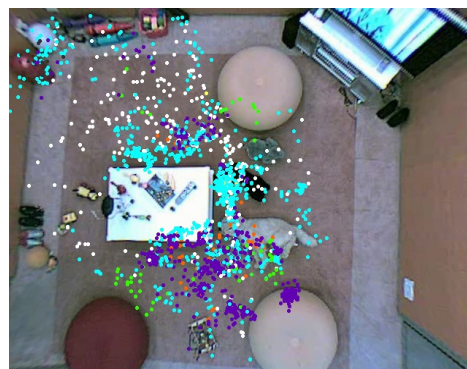


図 2: 行動ラベルのマッピング

なっている。

水色:「立っている」、白色:「歩いている」、紫色:「座っている」、緑色:「よじ登っている」、青色:「仰向け」、桃色:「寝返り」、橙色:「這い這い」、黄色:「うつぶせ」

図 2 から、例えば右下のクッション付近には「座っている」というラベルが多く振られているのが分かる。この行動 (サンプル数 1797) について、画像を  $8 \times 6$  の小領域に区切りその領域情報のうち行動ラベルが観測されたもの 30 を選んだ。これを  $S = (s_1, s_2, \dots, s_{30})$  とし、ある領域において時刻  $t$  に行動  $C_i^t$  が起こった条件の下での時刻  $t+1$  の行動  $C_i^{t+1}$  を集計する。この集計表から 3 つの条件付確率表  $\Pr(C_i^t | S)$ ,  $\Pr(C_i^{t+1} | C_i^t, S)$ ,  $\Pr(S)$  を求めた。この条件付確率表と図 1 の構造に従いベイジアンネットワークを構築した。

このモデル上で確率推論を実行することで、人が領域  $S$  に入った時の行動の確率を  $\Pr(C_i^{t+1} | C_i^t, S)$  として推定することができる。それを式 (2) に代入することで画像からの行動推定精度を向上させることができると期待できる。

## 4. まとめと今後

本稿では、確率的因果構造により日常生活行動を画像から推定するために、実際の行動観測から確率的な因果構造を抽出し、これをベイジアンネットワークによりモデル化した。今後は、OpenCV を用いて画像から抽出した特徴量から、尤度の高い特徴量を調査していく。また今後、行動ラベル群から生成したベイジアンネットワークの評価も行う。それらを用いて行動ラベルの事後確率を計算し、提案手法の有効性の評価を口頭発表時に行う予定である。

## 参考文献

- [1] 本村陽一, 西田佳史: 因果確率構造モデリングによる行動理解と日常生活支援, ニューロコンピューティング研究会, 信学技報, NC2005-46, (2005) .
- [2] 本村陽一, 西田佳史: ベイズ推定における事前分布のグラフ構造モデリングと実生活行動理解, 情報処理学会論文誌コンピュータビジョンとイメージメディア, 第 18 号 (2007) .
- [3] 本村陽一, 岩崎弘利: ベイジアンネットワーク技術, 東京電機大学出版局, (2006) .
- [4] Intel Open Source Computer Vision Library (<http://www.intel.com/technology/computing/opencv/>)