

Web カメラ映像を用いた確率ネットワークによる姿勢判別手法

Posture Classification Method by Probabilistic Networks based on Images recorded by Web Camera

岡本 勝^{*1}, 赤井 悠子^{*1}, 松原 行宏^{*1}

Masaru OKAMOTO^{*1}, Yuko AKAI^{*1}, Yukihiro MATSUBARA^{*1}

^{*1}広島市立大学大学院情報科学研究科

^{*1}Graduate School of Information Sciences, Hiroshima City University

Email: okamoto@hiroshima-cu.ac.jp

あらまし：本稿では、Web カメラを用いて身体動作を計測し、取得情報から作業分類を行う手法を提案する。本手法では、Web カメラで撮影した映像から複数の手法・機器を用いて関節などの位置を算出するため、計測時のノイズや計測環境での遮蔽物の影響による判別精度の低減を避けられる。さらに、確率的に作業姿勢に分類することによって、わずかな姿勢の違いや計測ノイズによる曖昧さの影響を受けないユーザの姿勢判別を行える。

キーワード：姿勢判別, Bayesian Network, Web カメラ, スキル学習

1. はじめに

音楽やスポーツなどのスキルを学ぶ学習や、労働作業における作業負担軽減などを目的とした指導などにおいて、適切に指導を行うためには対象者の身体姿勢情報の計測が重要となる。上述した学習・指導において、学習者に対して指導情報の提示を行い、正しい状態へ身体位置、関節角度を誘導することが基本的な指導内容になるが、適切なタイミングで指導内容を学習者へフィードバックするためには、ユーザの姿勢を正確に判別する必要がある。

多くのスキル学習の研究において様々な計測・状態判別が行われているが（例：文献[1]）、一般的には事前に解析した情報に基づいたヒューリスティックなアプローチが用いられている。そのため、対象となるデータや指導内容が追加されるたびに、再度解析を行い、学習支援システムに実装する必要がある。一方、藤澤らは工場作業や看護作業の状態判別を自動的に行う手法を提案した[2]。この手法では赤外線式モーションキャプチャシステムを用いて作業者の姿勢を計測し、ファジィ推論によって自動的に判別を行えるが、ファジィルール作成の自動化を実現できていないため、データの解析に時間がかかると問題が残る。

そこで本研究では、確率的に動作判別する手法を提案する。本手法では、動作判別にベイジアンネットワーク（以下、BN と略記）を用いることにより、自動的に学習を行え、計測値に欠損が含まれる場合でも与えられた情報から事後確率を導出できる。また、複数の計測手法による計測値ごとに算出した事後確率を計測環境や事後確率のあいまいさを考慮して統合し、統合後の事後確率をもとに姿勢判別を行うことにより、単一計測手法を用いた結果よりも精度の高い姿勢判別を目指す。

2. 提案手法

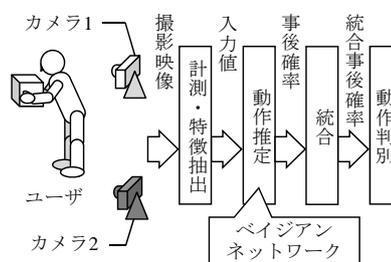


図1 システム構成図

図1に提案手法の構成図を示す。本手法では、二計測手法を用いてユーザを計測し、構築したネットワークを用いて判別対象となる姿勢に対応する事後確率を導出し、事後確率を統合して姿勢判別を行う。

計測・特徴抽出部では、二つの撮影機器・手法（計測手法1, 2）によってユーザを撮影し、取得した映像から画像処理技術を用いて関節などの部位の位置情報を取得する。本稿では、計測手法1では、ARToolkit ライブラリで認識できるマーカをユーザの体の各計測箇所に装着し Web カメラを用いて計測する。計測手法2では、民生品である赤外線カメラ（Xbox360 Kinect センサー、Microsoft 社）を用いて被験者の各関節部位の位置情報を取得する。さらに、計測時のノイズや姿勢のわずかなずれなどによる影響を低減させるために、フィルタ処理や、計測信号の正規化を行う。

動作推定部では、事前に計測データと対応する姿勢情報をもとに構築した BN を用いて入力値に対する各姿勢の事後確率を導出する。

統合部および動作判別部では、二つの計測手法による動作の各事後確率を統合しユーザの姿勢を判別する。計測手法 m ($m=1,2$) のデータ番号 n における信号が姿勢 c ($c=1,2,\dots,C$) に対応する事後確率を $O_{c,m}(n)$ とすると、統合後の事後確率 $IO_c(n)$ は次式より求める。

$$IO_c(n) = \frac{\sum_{m=1}^2 w_m(n) O_{c,m}(n)}{\sum_{m=1}^2 w_m(n)} \quad (1)$$

ここで、各計測手法の重み $w_m(n)$ を次式より求める。

$$w_m(n) = \varepsilon_m (1 - H_m(n)) \quad (2)$$

ε_m は計測環境の安定性などを考慮した一定の重みである。 $H_m(n)$ は計測手法 m ($m=1, 2$) における動作の事後確率に対するあいまいさを表すエントロピーであり、次式より求める。

$$H_m(n) = \frac{-\sum_{c=1}^C O_{c,m}(n) \log O_{c,m}(n)}{\log C} \quad (3)$$

統合した動作の事後確率の中で最も確率が高くなる動作をユーザの行った動作として判別する。このとき、あいまいな識別結果による誤識別を防ぐためエントロピーを利用し、その値が一定値を超えた場合は識別を保留する[3]。

3. 評価実験

提案手法の有効性の検証を行うために単一の計測手法での判別結果の比較と、BN 以外の識別手法による判別結果との比較を行った。比較手法では判別を行う識別子として、(1)単一の BN (計測手法 1 または 2 を用いた計測データを判別)、(2)階層型ニューラルネットワーク (以下、NN と略記)、(3)確率型 NN (LLGMN [3]) を用いた。比較手法(2)、(3)での入力信号として、計測手法 1 を用いて計測したデータを用いた。また、計測時に計測データの欠損が発生した場合は比較手法では識別を行えないため、比較手法での識別が誤った場合と同様に扱い、計測欠損によって判別を失敗したと想定した。

判別対象姿勢は工場での実作業の基本となる 10 姿勢[2]とした (表 1 参照)。各計測手法につき学習用データとして姿勢ごとに 400 データを取得し、BN を構築した。また、判別用データとして姿勢ごとに 100 データを用意した。また、統合した事後確率のエントロピーが 0.7 を超えた場合は判別を保留した。

表 2 に提案手法と比較手法での全姿勢の識別率を示す。提案手法の識別率は、92.3%となった。一方、単一の BN で判別を行った場合には、計測手法 1, 2 を用いた判別において、それぞれ 88.6%、64.4%の精度となった。これらの結果より、識別結果を統合することによって識別精度の向上が確認でき、一方の BN による判別精度が低下する場合でも、統合後の判別精度の低下が発生しないことを示せた。

次に、比較手法を用いた場合の実験結果と比較すると、階層型 NN を用いた場合、判別精度が大きく低下することが確認できた。これは、階層型 NN の学習に膨大なデータを要することと、表 1 の姿勢ごとの入力信号パターンの違いがあいまいな組み合わせが存在するため、NN が正確に学習できなかったと考えられる。一方、確率 NN を用いた場合でも、提案手法および計測手法 1 で計測したデータを単一

の BN で判別した結果よりも判別精度が低下していることが確認できた。この結果は、欠損を含むデータを用いた姿勢判別を確率 NN が行えないことが原因と考えられる。欠損が発生しなかったデータのみを判別した場合、確率 NN は単一の BN と同様の精度で判別できていたため、確率的に判別を行うことの有効性が確認できたと考えられ、信号の一部が計測できなかった場合でも判別できる点から BN を用いた提案手法の有効性が示せた。

表 1 判別対象姿勢 [2]

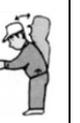
番号	1	2	3	4	5
姿勢例					
番号	6	7	8	9	10
姿勢例					

表 2 判別実験結果

手法	識別率
提案手法	92.3%
単一 BN (計測手法 1)	88.6%
単一 BN (計測手法 2)	64.4%
階層型 NN	48.6%
確率 NN [3]	83.8%

4. おわりに

本稿では、映像を用いた二つの計測手法によって身体計測を行い、各計測情報から BN を用いて導出した事後確率をもとに動作判別を行う手法を提案した。評価実験では、各計測手法による事後確率を統合することで、計測手法単体や従来手法での姿勢判別と比較して判別精度が向上することを確認した。

今後は、実際の教育分野での利用に適した計測手法や他作業への適用可能性を検討する。

参考文献

- (1) 曾我真人, 松田憲幸, 瀧寛和, “デッサン描画中に描画領域に依存したアドバイスを提示するデッサン学習支援環境”, 人工知能学会論文誌, Vol. 23, No. 3, pp. 96-104 (2008)
- (2) 藤澤一暁, 松原行宏, 岩根典之, 真嶋由貴恵: “作業姿勢分析における関節角度に基づくファジィ推論を用いた姿勢自動分類手法”, バイオメディカル・ファジィ・システム学会誌, Vol. 10, No. 1, pp. 126-129 (2008)
- (3) T. Tsuji, O. Fukuda, H. Ichinobe and M. Kaneko: “A Log-Linearized Gaussian Mixture Network and Its Application to EEG Pattern Classification,” IEEE Trans. on Sys. Man and Cyb. Part-C, Vol. 29, No. 1, pp. 60-72 (1999)