

人間の動作認識による能動的な人間支援システムの開発

横山 法子 山口 友之 橋本 周司

早稲田大学理工学部応用物理学科[†]

1. はじめに

近年、人の行動を見守りさりげなく支援できるロボットや環境の開発が盛んに行われている[1]. このような能動的支援を行うにあたり、人の行動からその人の内面状態を読み取ることができれば、より幅広く、その状況に合った支援が行えると考えられる. 内面状態を推定する研究としては、運転中の居眠りを検知するシステム[2]などがあるが、これらはあくまで注意・警告が主目的である. そこで、我々はコミュニケーションのきっかけにもなるような支援システムの一例として、人間の机上動作に注目し、励ましなどの話しかけを行う学習支援システムの開発を試みた. 本稿では、画像を用いたユーザの姿勢判定からその人の意識状態を推定し、話しかけを行うシステムの概要を述べ、実験結果を報告する.

2. システム概要

まず、パソコン上に取り付けた USB カメラからユーザの画像を取得し、姿勢を判定する. そして姿勢の時系列データをもとにニューラルネットワークにより意識状態の推定を行う. さらに、意識状態の時系列データをもとにベイジアンネットワークを使って、話しかけを行う(図 1).

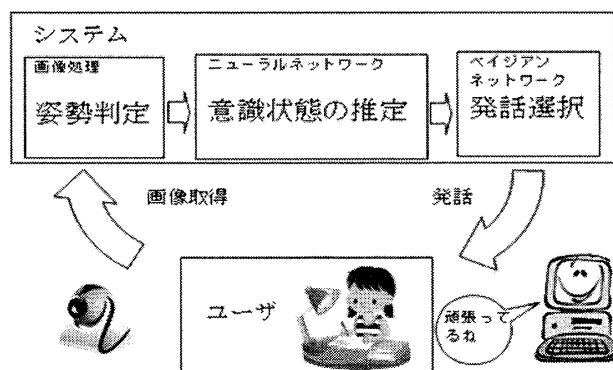


図 1 システム概略図

2.1. 姿勢判定

まず、グレースケール画像の背景差分を閾値処理することにより、シルエットを抽出する. 次に、事前に作成したテンプレートのシルエットとの距離を計算し、姿勢判定を行う. シルエット間の距離は画素ごとの画素値差の絶対値とした. 判定姿勢の種類としては、「正常姿勢」、「傾いた姿勢」、「うつぶせの姿勢」、「かなり下を向いた姿勢」の 4 種類を用意した.

2.2. 意識状態の推定

前述の姿勢判定は 2fps の速度で行い、5 分間のデータをもとに意識状態の推定を行う. 推定には、バックプロパゲーション法による 3 層構造のニューラルネットワークを用いる(図 2). 入力信号には、5 分間のデータ(長さ 600)を各姿勢の累計回数、各姿勢間の遷移回数の計 10 個のパラメータに変換したものを用いた. 出力信号となる意識状態としては、「寝てる」「眠い」「やる気がない」「集中している」の 4 状態を用意した. この 4 状態は、人間の意識状態を示す一つの指標であるフェーズ理論[3]をもとに選出した. 学習では、実際に取得した時系列データに手動で状態を付与することで教師データを作成し、各意識状態 100 データずつ用いて 10000 回行った.

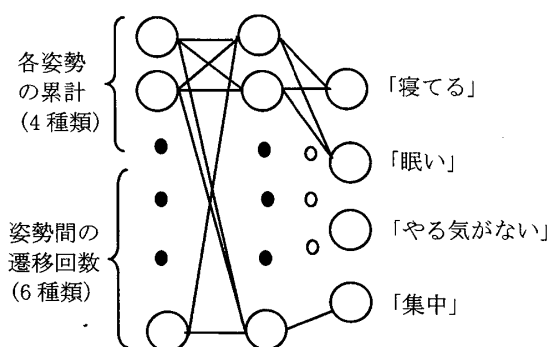


図 2 意識状態の推定

2.3. ユーザへの発話

意識状態の時系列データをもとに話しかけの内容を選択する. 学習者の意識状態等の状況ノードとコンピュータが話しかける発話ノードとの

An active support system based on recognition of human motion during deskwork

[†]Noriko Yokoyamao, Tomoyuki Yamaguchi, Shuji Hashimoto
• Department of Applied physics, Waseda University

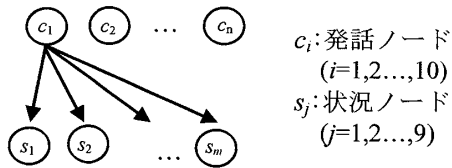


図3 ユーザへの話しかけ

関係を図3のようにモデル化し、ベイジアンネットワーク[4]を用いて学習、選択を行う。状況ノード \$S\$ は9種類、発話ノード \$C\$ は10種類である。学習時には、実際にある発話 \$c_i\$ が観測された時の状況ノード \$s_j\$ の値を \$t_k\$ とし、その回数を \$n\$ として、条件付確率 \$P_{s_j|c_i}(t)\$ を式(1)のように決定した。正規分布 \$f_k(t)\$ の分散 \$\sigma^2\$ は経験的に 100 とした。

$$f_k(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-t_k)^2}{2\sigma^2}\right)$$

$$P_{s_j|c_i}(t) = \frac{\sum_{k=1}^n f_k(t)}{n} \quad (1)$$

選択時には、ある状況 \$S^i=(s_1^i, s_2^i, \dots, s_9^i)\$ のときの各発話ノードの確信度を式(2)で計算し、確信度が閾値以上であれば発話する。

$$BEL(c_i) = P(S^i | c_i) = \prod_{j=1}^m P_{s_j|c_i}(s_j^i) \quad (2)$$

3. 評価実験

被験者の一日の学習中の状況に応じて教師信号となる発話を与えて、ベイジアンネットワークで学習させる。次にシステムを稼働させながら被験者に机上で勉強を行ってもらい、システムからの発話を観察した。パソコン上に取り付けた USB カメラの画像をシステムへの入力とした。ベイジアンネットワークの出力値である発話ノードとして「頑張っているね」などの励ましやいたわり、ほめる言葉 10 種類を用意し、状況ノードとして表1の9種類を用意した。

実験の結果、意識状態に応じてシステムからの発話が得られた(表2)。寝ていれば心配し、集中していればほめるといったように、妥当な発話が確認できた。また、この集中時の発話「頑張っているね」の確信度の推移は図4であった。集中する時間が長くなるにつれ、発話「頑張っているね」の確信度が徐々に上がり、25分、60分で閾値(0.3 × 10⁻¹³)を超えると発話。発話直後は、確信度が減少していることがわかる。

表1 状況ノード

ノードの意味	ノード数
各内面状態の一日の累計時間	4
各内面状態の継続時間	4
直前の発話が行われてからの経過時間	1

表2 意識状態(自己評価)と発話内容の例

意識状態	システムからの発話内容
寝てる	寝ていて大丈夫?
集中	頑張っているね

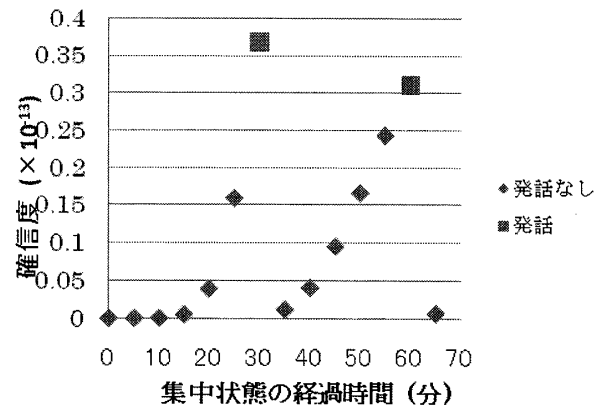


図4 集中時「頑張っているね」の確信度の推移

4. おわりに

システムから人間への能動的な支援の一つとして学習支援システムを作成し、相手の様子から気づきを得て、能動的に話しかける機能を持たせた。今回は机上姿勢からの意識状態の推定であったが、今後は他の動作にも注目し、感情などのより深い内面状態を推定できるようにしたい。

謝辞

本研究の一部は、早稲田大学ヒューマノイド研究所、グローバル COE プログラム「グローバル ロボット アカデミア」、科学技術振興機構 CREST 研究「人を引き込む身体的メディア場の生成・制御技術」の研究助成を受けて行われた。

参考文献

- [1] 佐藤知正, “生活を支えるネットワーク知能機械—ロボティック・ルームの試み”, 東芝レビュー, vol. 56, no. 9, pp. 28-32, 2001.
- [2] 足立和正ら, “画像処理によるドライバの運転状態の監視”, 電気学会論文誌 E, vol. 126, No2, pp. 31-37, 2006.
- [3] 橋本邦衛, “安全人間工学”, 中央労働災害防止協会, 1984.
- [4] 須崎弘樹, “ベイジアンネットワーク入門(1)”, 日本医画像工学会誌, vol. 21, no. 4, pp. 315-318, 2003.