

H-066

カメラによる体動観測に基づく心理状態推定 Estimation of State of Mind Base on the Motion of Pose Using Cameras

ブンヨンスクサコント 青木 恭太
Souksakhone Bounyong Kyota Aoki

1. はじめに

近年、カメラやセンサー、画像処理技術の進歩により、私たちの生活のさまざまな場面で非常に役に立っている。例えば、教育やセキュリティ、医療、障害者の支援などがある。だが、教師は学生たちがどれだけ真剣に授業を聞いているか分からない。また、言葉ができない障害者の心理を、手話なしで理解することはとても困難だ。そこで、人々が対面で会話をしなくても、コンピュータシステムを利用してお互いの気持ちを理解することができればどれほど便利になるだろうかと考えた。

最近では特に、表情認識といって、表情に着目し感情を推定する研究が盛んに行なわれている。人間の身体の状態は、心理的な状態の影響を受けて変化する。身近な例で言えば、映画を見る場合、感動的なシーンや怖いシーン、どきどきするシーンなど、視聴者の感情が表情に表れたり、体動などに現れたりすることがある。このような目に見える情報を使用し、目に見えない人間の感情や心理状態を推定して分析することにより、情報提供システムのパフォーマンスを向上させることができると考えられる。

現在、すでに表情認識の技術を用いて実用化に向けて研究が行なわれている。例えば、観客の表情より競技の試合に対する満足度を推定する研究[1]や、顔画像から授業の主観的難易度を推定する研究[2]などがある。また、表情だけでなく人間の眼球運動の変化に着目し心理状態を推定する研究もある[3]。

表情や眼球を観測することが困難な状況もある。体動の観測はより容易である。心理状態が姿勢や体動の変化を引き起こす。本研究では、体動を観測し心理状態をリアルタイムで推定するシステムを開発することを目的としている。実験方法は、2台のカメラを使用し、被験者の正面と側面の2方向から上半身を観測する。体動の平均動き量、及び各周波数成分の振幅の2つを用いて、心理状態を推定する。

2. 実験

2.1 体動観測

観測システムでは2台のUSBカメラを使用する。2台のカメラを使用することにより、体動の3次元データを推定することができる。上半身の動きを捉えるため、被験者の正面と側面にカメラを設置し、録画する。そして被験者に、モニター上で、難しく面白くなさそうな英語小説を10分読んで貰う。また、被験者が面白いと思うであろう「漫画」を10分読んで貰う。観測の間、被験者が自由に動けるよう、画面操作はゲームパッドを使用する。英語小説と漫画を読んだ後、読書中の興味や難しさについて心理状態に関するアンケートを実施する。合計32回実施し、被験者は6

名だった。10分間で平均1万3千フレームが得られた。

2.2 画像処理

データ処理の手順を図1の流れで行なった。特徴点の検出には、Lucas&Kanadeの手法を使用した(図2)。体に集まっている特徴点だけを追跡するために、K-meanアルゴリズムを用いて、連続した2フレームの、各特徴点の移動距離を測定し、移動距離と方向に基づいて、特徴点を5つのクラスターに分類したクラスタリングを行なった。そして、最大移動距離が含まれ、かつ、特徴点の分散が小さいクラスターを選択することにより、体の領域に集まっている特徴点を特定できた(図3)。

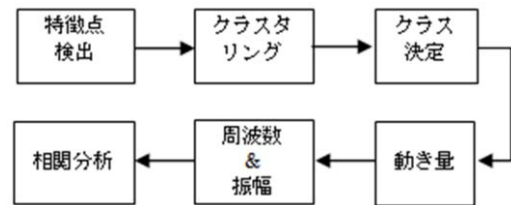


図1 データ処理の手順



図2 検出されたコーナー特徴点

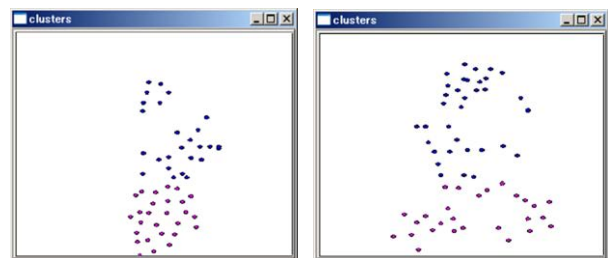


図3 クラスタリングによって得られた体の領域

ここでの動き量とは、クラス決定の段階で得られた特徴点の位置の2フレーム間の差である。また、平均動き量は、10分間で得られた全フレームの動き量の絶対値の平均値である。

各周波数成分の振幅を求めるには、10分間のデータを0.1秒間隔で動き量を算出し、フーリエ変換を行う。

2.3 心理状態推定結果

実験で用いたアンケートは以下の通りで、5段階評価の形式をとった。

- Q1. 面白かったか
- Q2. 難しかったか
- Q3. 集中したか
- Q4. 理解したか
- Q5. 続きを読むみたいか

2.3.1 平均動き量による推定

アンケートとの相関係数を図4(a)と4(b)に示す。縦軸は相関係数であり、横軸はアンケートの質問項目である。

図4より、両カメラで観測された上下の平均動き量はアンケートの各質問項目と相関が見られる。また、正面カメラより観測された上下の平均動き量はQ1及びQ5と、強い負の相関が見られ、Q2とは正の相関が見られる。側面カメラより観測された上下の平均動き量も同様だが、前後の平均動き量はほとんど相関が見られなかった。

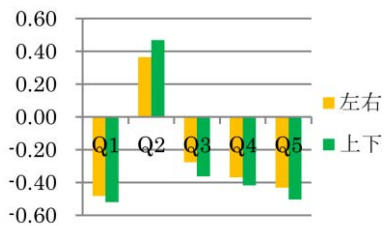


図4(a) 正面カメラの結果

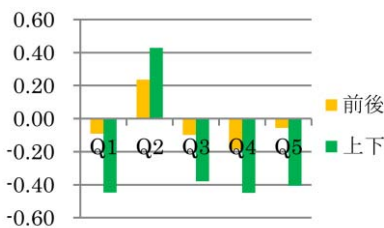


図4(b) 側面カメラの結果

2.3.2 各周波数成分の振幅による推定

周波数成分の振幅とアンケートの相関を求めるには、周波数帯を分割し、それぞれの範囲内の振幅の平均を求め、アンケートと相関を計算する。結果を表1に示す。表1では、高い相関が見られるパラメータのみ表示する。大文字のアルファベットは周波数帯を表す。Bは1~1.4[Hz]、Dは2~2.8[Hz]、Eは2.8~4[Hz]である。また、FxとFyはそれぞれ正面カメラから観測されたx軸とy軸の振幅で、Syは側面カメラから観測されたy軸の振幅である。

表1 アンケートと周波数の振幅との相関係数

質問	周波数範囲：相関係数		
	Fx	Fy	Sy
Q1	D: -0.41		B: -0.40
Q2	E: +0.51	E: +0.40	B: +0.55
Q3	E: -0.41		B: -0.46
Q4	D: -0.41		B: -0.44
Q5	D: -0.43		B: -0.40

Fxにおいて、周波数が2から2.8[Hz]の時の振幅はQ1、Q4、Q5と負の相関が見られる。また、周波数が2.8から4[Hz]の時の振幅はQ2と正の相関が見られ、Q3とは負の相関が見られる。

Fyにおいて、周波数が2.8から4[Hz]の時の振幅Q2と正の相関が見られるが、それ以外は見られなかった。

Syにおいて、周波数が1から1.4[Hz]の時の振幅はQ2以外、各質問項目と負の相関が見られる。

3. むすび

平均動き量を用いた場合、両カメラより観測された上下の動きは「難しかった」と正の相関が見られた。同様に、各周波数成分の振幅を用いた場合においても、周波数が2.8から4[Hz]の時、上下の振幅が「難しかった」と正の相関が見られた。また、周波数成分の振幅を用いた場合は、周波数が2.8から4[Hz]の時、左右の振幅も「難しかった」と高い相関が見られた。つまり、「難しい」と感じる心理状態は、周波数2.8から4[Hz]の時の上下、または左右の振幅に比例するのではないかと考えられる。図4及び表1から分かるように、側面カメラより、上下の平均動き、及び周波数が1から1.4[Hz]の時の振幅は各アンケート項目と相関が見られた。つまり、この周波数帯において、側面から観測された上下の振幅は心理状態を推定することが可能ではないかと考えられる。

正面から観測した上下の平均動き量とアンケートの各質問項目と相関が見られたが、各周波数成分の振幅の場合では、「難しかった」状態以外、相関が見られなかった。

実験で得られた全ての相関計数において、絶対値は高々0.55であるため、決して強い相関関係とは言えない。実際に心理状態を推定するために、より強い相関関係の下で行うべきである。今回の実験では6名の被験者のデータを区別せずに、分析を行なった。個人の癖によって、同じ心理状態であっても体動が異なるため、強い相関関係が見られない原因ではないかと考えられる。また、検討した周波数の範囲は適切でないということも考えられる。今後はそれらの解決も含めて、より適切なデータを得るため、実験方法を改善していく。

参考文献

- [1] 松川徹, 日高章理, “複数人物の表情認識による観客の満足度の自動定量評価”, 筑波大学大学院システム情報工学研究科システム開発型研究プロジェクト(2008).
- [2] 中村和晃, 角所考, 美濃導彦, “e-learning環境における学習者の観測に基づく主観的難易度の推定”, MIRU2007, PP.1029-1034, 2007-7
- [3] 平山雄介, “眼球運動の変化に基づく心理状態の推定”, 電気通信大学大学院(2003)