

表情変化における顔の部位の二次元・三次元移動量の比較*

4B-2-02

岩田 満 田野 俊一†

電気通信大学大学院情報システム学研究科‡

1. はじめに

近年, 我々の周囲にコンピュータが急速に浸透してきている. ハードウェア技術の進歩により, コンピュータ自体の性能は格段に向上しているが, 人間との接点である入力インタフェースについてはキーボードやマウスによる入力デバイスを用いた計算機中心のインタフェースが未だに広く利用され続けている.

人間同士の Face to Face コミュニケーションは, 言葉以外に顔の表情や身振り, 手振り, 声色などのノンバーバル (非言語) 情報を用い, 円滑に行われている. したがって, コンピュータがノンバーバル情報を人間と同程度に理解・表現することが可能となれば, 人間とコンピュータとの共同作業におけるコミュニケーションはよりスムーズになると考えられ, 擬人化エージェントに関する研究[1]が数多く行われている.

表情はノンバーバル情報の中でも重要な情報であるといわれている. そのため表情認識に関する研究が盛んに行われており, 主な研究例として, 顔の各部位の特徴点を利用した研究[2]や, Ekman が提唱した Action Unit を用いた研究[3], 表情筋の動きを考慮した研究[4]がある. これらの多くは, 二次元の顔画像からの表情認識を行っている. 顔の三次元形状モデルを利用した研究[5-7]もあるが, これらは二次元の入力画像と三次元顔モデルとのマッチングを行い, 表情筋の三次元の動き情報を抽出している.

半導体技術の進歩により, 距離画像をこれまで困難であったビデオレート (30 フレーム/秒) で得ることが最近になって可能となった. これにより, 顔の表情を連続的に三次元画像として得ることができる. 本論文では, 距離画像の利用が表情認識に有効であるかどうかを検討する.

2. 二次元線画の表情認識

前研究として, 二次元の線画の顔からの表情認識システムを構築した[8]. システムの評価実験から, システムが推論した感情と, 被験者が感じ取った感情とはほぼ合っているという評価が得られたが, 「恐怖」や「嫌悪」の認識精度が他の表情よりも低いという結果も得られた.

実画像を用いた表情認識においても同様の傾向が見られ, 例えば文献[7]では, 喜び, 悲しみ, 怒り, 驚きの 4 表情では平均 87% の認識率であるのに対して, これに恐怖と嫌悪を加えた 6 表情では誤認識が多く存在し, 平均認識率は 62% まで下がっている.

本研究では, 従来の表情認識では, 顔の三次元形状が二次元平面に投影された画像を用いているために認識率が低いと考え, 表情の三次元データが表情認識において有効であるかを確認するために, 二次元画像から得られた顔の部位の移動量と, 三次元画像から得られた移動量との比較を行う.

3. 顔の距離情報収集

三次元表情データを収集するために, 被験者の表情変化を 3D カメラで撮影する. 被験者は, 17 歳～31 歳の 13 名 (男性: 9 名, 女性: 4 名) である.

3.1. 装置

表情撮影には, 図 1 に示す 9 個のカメラを備えた高速ステレオビジョンを用いる. このシステムは, 9 個のカメラから入力された画像をステレオ処理し, 実時間 (ビデオレート) で距離画像を生成する. 図 2 に高速ステレオビジョンシステムで撮影したセンタRGB画像と, 被写体の距離が疑似カラーで表

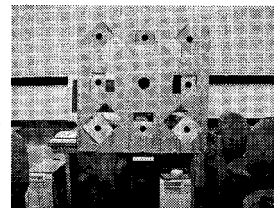


図 1: ステレオビジョンシステム(9 眼カメラ)

*Comparison of Two- and Three-Dimensional Facial Movements for Recognition of Facial Expression

†Mitsuru Iwata and Shun'ichi Tano

‡Graduate School of Information Systems, University of Electro-Communications

現された距離画像の例を示す。

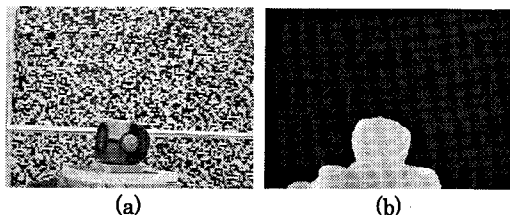


図2：センターRGB画像(a)と距離画像(b)

3.2. 撮影方法

撮影する表情は、表情認識の研究で広く用いられている Ekman の感情分類である喜び、悲しみ、驚き、恐れ、怒り、嫌悪の6種類である。各被験者に、無表情の状態からそれぞれの表情を表出してもらい、その変化過程をカメラで撮影する。

4. 顔の部位の移動量

4.1. オプティカルフローの計算

顔の各部位の移動方向と移動量を抽出するため、ブロックマッチング法（相関法[9]）により動画像からオプティカルフローを計算する。ブロックマッチング法は、図3に示すように2枚の連続したフレームにおいて、第1フレームの任意の画素 $P(x, y)$ に対する次フレームでの対応点 $Q(x', y')$ を探索する手法である。第1フレームの点 P を中心とする領域をテンプレートブロックとし、第2フレームの注目ブロックとの輝度差分評価関数値が最小となる点を対応点とする。第1フレームの注目点 $P(x, y)$ から第2フレームの対応点 $Q(x', y')$ へのベクトルが、注目点におけるフローベクトルとなる。この処理をフレーム全ての画素について行う。

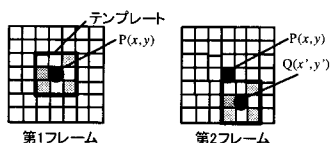


図3：ブロックマッチング

輝度差分評価関数 $E_t(x, y, x', y')$ は次式(1)で定義される。

$$E_t(x, y, x', y') = \sum_{m=-L}^L \sum_{n=-L}^L |I_C(x+m, y+n, t) - I_C(x'+m, y'+n, t+1)| \quad (1)$$

ここで、 $I_C(x, y, t)$ はセンター画像の第 t フレームの点 (x, y) の輝度値、 $I_C(x', y', t+1)$ は第 $t+1$ フレームの点 (x', y') の輝度値であり、テンプレートブロック

の大きさは幅と高さが共に $2L+1$ ピクセルである。

本研究ではテンプレートブロックの大きさを5ピクセル ($L=2$) とし、次フレーム中の探索範囲を点 $P(x, y)$ を中心とする 22×22 ピクセルのブロック内に限定して計算時間を短縮している。

4.2. 結果

オプティカルフローの計算から得られた、二次元画像上のフローベクトルの一例を図4(a)に示す。これは強い喜びの表情についてフローベクトルを計算した結果である。また、距離方向の移動ベクトルを距離画像から算出した結果を図4(b)に示す。図4(b)の左下方向の矢印は手前方向への移動を表し、右上方向の矢印は奥方向への移動を表す。図5は二次元方向の総移動量と距離方向の総移動量を示したものである。図5(a)の濃く塗りつぶされた箇所が二次元移動量の特に多い部分である。また、図5(b)の濃く塗りつぶされた箇所は手前方向への移動が特に多い部分、薄く塗りつぶされた箇所は奥方向への移動が特に多い部分を表す。

図4(a)および図5(a)から、喜びの表情では口角から頬の領域が外側上方向に大きく移動していることがわかる。また、図4(b)および図5(b)からは、口角から頬の領域がへこんでいることがわかる。

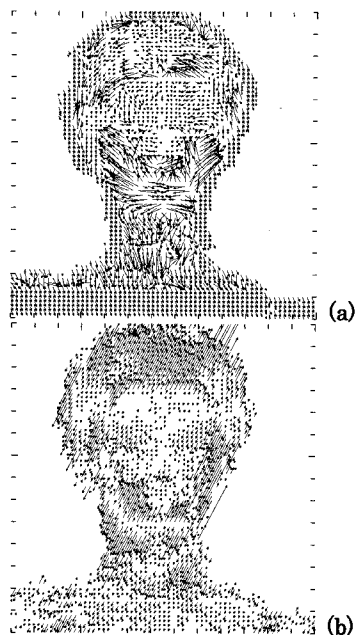


図4：二次元方向のフローベクトル(a)と距離方向のフローベクトル(b)

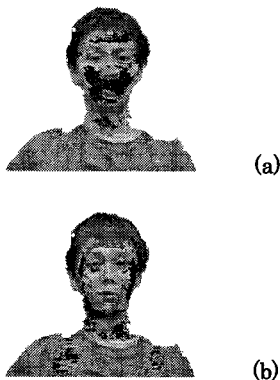


図 5: 二次元方向の総移動量(a)と
距離方向の総移動量(b)

これらの結果から、各表情について以下の傾向が見られた。

喜び

- ・口角から頬にかけて外側上方向へ動き、上唇、鼻唇溝も上がる。
- ・口角から頬、および顎にかけてへこんでいる。
- ・口の周囲で、二次元の移動量と距離方向の移動量の間に違いがみられる。距離方向の移動量については、唇ではあまりみられず、むしろ口角から頬にかけての箇所で見られている。

悲しみ

- ・目の周囲が特に下がり、口の周辺も下がる。
- ・うつむき加減になり、顔の上側が前に出る。
- ・(特に目の周辺では) 距離方向の移動は二次元方向の移動ほどは変化しない。

驚き

- ・全体的に上方向へ動いている。口角は上へ動くが、下唇と顎は下がる。
- ・目や眉(眉間)、頬がへこむ被験者が多い。
- ・頬と目の下の部分で、二次元方向の移動量と距離方向の移動量の間に違いがみられる。

怒り

- ・目と眉が二次元方向に大きく動いている。
- ・目や眉の周辺はふくらみ、口の周辺はへこむ。顔全体が前へ出る被験者が多い。

恐れ

- ・口角と目が下がり、眉が上がる。頬が動く。
- ・口の周りがへこむ(顔の下半分が後ろへさがる)。

目、眉、頬もへこむ。

嫌悪

- ・頬、口が大きく動く。口は下がり、目は外側下方向へ動く。
- ・顔全体が後ろへ下がる。口や頬の周辺が特に大きくへこむ。

喜びや驚きの表情では、二次元方向の移動量と距離方向の移動量との間に違いが見られたが、恐れや嫌悪の表情では大きな違いは確認できなかった。

5. 三次元形状を考慮したオプティカルフロー

画像の輝度情報のみを用いて計算されたオプティカルフローでは、恐れや嫌悪などの表情に違いが見られなかったため、本研究では輝度情報の他にカメラからの距離情報も用い、三次元形状を考慮したオプティカルフローを計算し、顔の各部位の移動ベクトルの違いを見る。

従来手法との違いは、輝度差分評価関数の代わりに、輝度・距離差分評価関数を定義している点である。輝度・距離差分評価関数 $E_D(x, y, x', y')$ は以下の式(2)で定義する。

$$E_D(x, y, x', y') = \sum_{m=-L}^L \sum_{n=-L}^L [w_1 \cdot |I_C(x+m, y+n, t) - I_C(x'+m, y'+n, t+1)| + w_2 \cdot |I_D(x+m, y+n, t) - I_D(x, y, t) - \{I_D(x'+m, y'+n, t+1) - I_D(x', y', t+1)\}|] \quad (2)$$

ここで、 $I_D(x, y, t)$ は距離画像の第 t フレーム上の点 (x, y) の輝度値であり、カメラからの距離に相当する。また、 $I_D(x', y', t+1)$ は第 $t+1$ フレーム上の点 (x', y') の輝度値である。 w_1 と w_2 はそれぞれ、輝度情報と距離情報を重視する度合いである。

5.1. 結果

恐れ表情についてオプティカルフローを計算した結果を図 6, 7 に示す。図 6 は $w_1=1, w_2=0$ の時の二次元方向のフローベクトルと距離方向のフローベクトルであり、形状情報を利用しない場合である。図 7 は $w_1=1, w_2=1$ のフローベクトルであり、形状情報も利用した場合である。

三次元形状情報を用いることで、オプティカルフローの精度は多少向上しているが、図 6 と図 7 との間に大きな違いは見られないため、大幅な認識率向上は期待できないと考えられる。

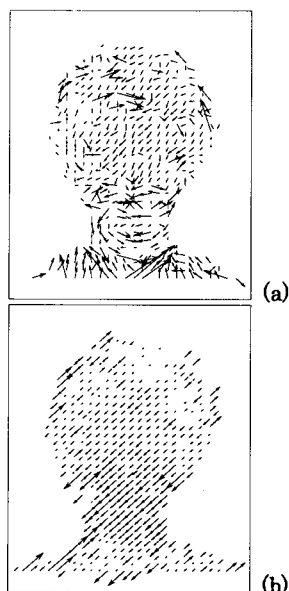


図6：形状情報を利用しない場合の二次元方向のフローベクトル(a)と距離方向のフローベクトル(b)

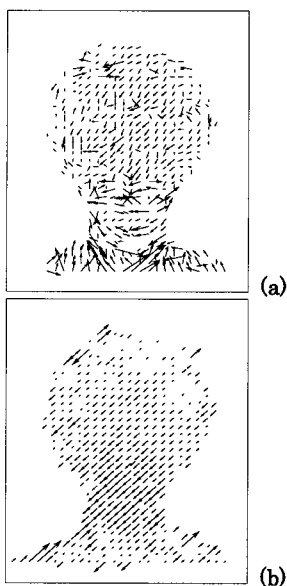


図7：形状情報を利用した場合の二次元方向のフローベクトル(a)と距離方向のフローベクトル(b)

6. まとめ

本論文では、従来の二次元表情認識モデルの認識精度向上のために距離データの利用を検討した。三次元画像が表情認識において有効であるかを確認するために、二次元画像から得られた顔の部位の移動量と、三次元画像から得られた移動量との比較を行った。喜びや驚きなど違いがはっきり分かる表情

では違いがみられたが、恐れや嫌悪などのあいまいな表情については明確な違いがみられず、表情認識においては、二次元情報だけでも十分であることがわかった。違いがわかりにくい表情の認識についてはユーザの状況を考慮するなど、画像以外の情報も必要になる。

三次元データを用いることで、顔を背景から切り出しやすくなり、また、鼻の位置が特定しやすいため、顔の各部位の特徴点位置が検出しやすくなる。そのため、顔の部位の切り出しの際は距離情報が有効利用できると考えられる。

謝辞

本研究の実施において、高速ステレオビジョンを快くお貸し頂きましたコマツ研究本部基礎技術研究部に深く感謝いたします。

参考文献

- [1] 速水悟, 竹澤寿幸: マルチモーダル情報統合システムの研究動向, 人工知能学会誌, Vol. 13, No. 2, pp. 206-211, 1998.
- [2] Mei Wang, Yoshio Iwai, Masahiko Yachida: Expression Recognition from Time-Sequential Facial Images by use of Expression Change Model, Proc. of the Third International Conference on Automatic Face and Gesture Recognition, pp. 324-329, Nara, Japan, 1998.
- [3] Gianluca Donato, Marian Stewart Bartlett, Joseph C. Hager, Paul Ekman, and Terrence J. Sejnowski: Classifying Facial Actions, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 10, pp. 974-989, 1999.
- [4] 矢崎和彦, 石川貴博, 森島繁生: オプティカルフローを用いた正面顔画像からの顔面筋パラメータの自動推定, 電子情報通信学会技術研究報告, PRMU, パターン認識・メディア理解, Vol. 97, No. 596, pp. 121-128, 1998.
- [5] 金子正秀, 羽鳥好律, 小池淳: 形状変化の検出と3次元形状モデルに基づく顔動画の符号化, 電子情報通信学会論文誌 B, Vol. J71-B, No. 12, pp. 1554-1563, 1988.
- [6] 崔昌石, 原島博, 武部幹: 顔の3次元モデルを用いた顔面表情の分析, 電子情報通信学会論文誌 D-II, Vol. J74-D-II, No. 6, pp. 766-777, 1991.
- [7] 太田寛志: 顔面構造に基づいた顔動作解析と表情認識に関する研究, 静岡大学大学院理工学研究科博士論文, 1998.12.
- [8] 岩田満, 鬼沢武久: 絵のつながりを考慮した絵情報の言語的表現, 電子情報通信学会論文誌, Vol. J84-D-II, No. 2, pp. 337-350, 2001.
- [9] 安居院猛, 長尾智晴共著: 画像の処理と認識, 昭晃堂, 東京, 1992.