

3D-2

局所的特徴強調アルゴリズムと
オプティカルフローを用いた人間の顔の認識*

浅見賢 國井利泰†

東京大学理学部情報科学科‡

1 はじめに

コンピュータがする入力・演算・出力の3つの仕事のうち、演算・出力の能力は近年大きく向上したが、それに比べてコンピュータに対して人間が意志を伝える手段に関する研究は遅れていると言わざるを得ない。

人間がコンピュータに意志を伝える手段として従来はキーボードなどの機械的手段が主に用いられてきたが、最近は音声・画像などのより自然な入力方法が研究されるようになってきている。

人間同士の意志疎通にもっともよく使われる手段は音声であるとい一般に思われているが、面と向かいあっている2人の人間の間でやりとりされている情報の半分以上が表情、見ぶりなどの非音声的手段で伝達されているという研究もある。

本研究では人間の顔画像からの情報抽出の方法を提案する。最終的には表情の自動認識をすることを目標とするが、そのためには認識された音声から意味を抽出し、相手の感情を推定することまでが必要であり、現在の技術ではまだ無理である。

2 表情認識に関する過去の研究

人間の顔面には鼻のように起伏が激しく明るさが大きく変化するところ、唇や眉のようにテクスチャが変化するところ、頬のようにどちらともあまり変化しないところが混在している。

明るさやテクスチャが大きく変わるところではエッジ検出のオペレータによってそれを検出することができる[9]。それほど大きくはない明るさやテクスチャの変化はあらかじめそれを強調しておけば同様に検出することができる。

これらのオペレータを作用させ、細線化を行った結果から特徴点を抽出するためには顔についての知識・モデルの利用が不可欠である[1]。筋肉のモデルについては、実際の表情筋の構成は非常に複雑であるため、それをモデル化するのは困難であるので、コンピューターグラフィックスの領域からの簡略化された筋肉モデルの提案[8]などを利用することも考えられる。

頬や額など比較的のっぺりした部分についてはオプティカルフローを利用するのが効果的である。オプティカルフローを顔面の画像処理に用いる試みとしてはあらかじめ決めておいた領域の中のオプティカルフローから表情筋の動きを検出するのに使う例があり[5]、口唇輪郭の抽出の際に特徴点の動きを追跡するための補助的手段として使う例[6]がある。ここでは筋肉の動きを検出するとともにオプティカルフローに注目して領域分けすることによって筋肉の影響範囲の動きを見つげるために使うことを考える。

オプティカルフローを領域分けに用いる試みとしては、顔面の画像処理ではないが、ラプラシアンから得られる零交差画像による領域の

縁に沿ってオプティカルフローを求め、それを利用して認識を行う方法がある[3]。

3 特徴強調

入力された画像には影になっている部分がある。影には面の向きなどの情報が含まれている半面、小さな凹凸の陰影やテクスチャの変化が大きな影の中に吸収されてしまったりすることもある。これはエッジ検出のオペレータにとっては好ましくないことであり、何らかの方法で大きな変化に隠された小さな変化を強調することが必要である。

特徴強調のためのアルゴリズムには画像データによって変換の関数が増えるものとししないもの、また大域的なものや局所的なものなど、さまざまな分類の方法がある[2]。本研究で扱う顔画像のように光の当り方によって明るさが大きく変化する場合に画像データの局所的な性質によって変換関数を決定するものがよい。ここでは輝度の強さの尺度である平均と輝度の散らばりの尺度である標準偏差を使って変換する方法を用いる。

画像をいくつかの長方形の格子に分割し、それぞれの中の輝度データの平均値を求める。それをその長方形中のデータの代表値とし、その値を長方形の中心の値としてそれらの間に双線形パッチ(bilinear patch)を張ることによってある程度滑らかに全体の様子を近似することができる。標準偏差についても同様にして全体の様子を反映するパッチを作る。本来は長方形を1ピクセルずつずらしながら平均と標準偏差を求めてゆくのがよいのであるが、計算量を減らし、かつ同等の効果を得るためにこのような方法をとっている。

これらの値を用いて次の式によって各ピクセルの値を変換する。

$$g(x) = 0.5\alpha + (1-\alpha)\cdot\mu(x) + 0.5\beta\cdot\frac{f(x)-\mu(x)}{\sigma(x)} \quad (1)$$

ただし明るさの値は0と1の間とし、 α と β はどの程度強調するかを決める係数である。均一な背景を無用に強調するのを防ぐために最終項の分母である標準偏差が非常に小さい時(ある閾値を下回る場合)には分母をその閾値で置き換えるものとする。また、 $g(x)$ が0と1の間でない場合には0あるいは1にするものとする。

長方形の大きさを変えることによって強調される特徴の粒度を変えることができる。長方形を小さくすることによって、より細かい特徴を強調するような効果が得られるが、それに従って無視される特徴の大きさもどんどん小さくなるので、あまり長方形を小さくしても意味がない。我々の実験では500ピクセル四方程度の画像では一辺が20~30ピクセルの長方形で平均・標準偏差を計算した時にもっともよい効果が得られた。

4 特徴検出

エッジ検出のオペレータには1次微分に基づくものや2次微分に基づくもの、また微分ではなくパターンマッチングによるものなどが

*Recognition of Human Face by a Local Enhancement Algorithm and Optical Flow

†Satoshi ASAMI, Toshiyasu L. KUNII

‡The University of Tokyo

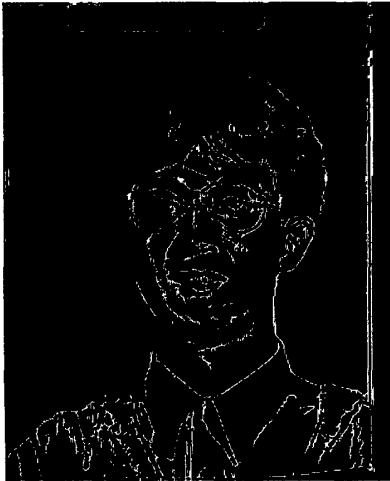


図1 処理例

ある [7]。ここでは Roberts と Sobel のオペレーターを用い、その結果を細線化して特徴点を抽出している。図1に局所の特徴強調した画像に Sobel のオペレータを適用し、その結果を細線化したものを示す。

5 オプティカルフロー

オプティカルフローを計算する方法としては画像フレーム間の特徴点の動きを追跡する方法(マッチング法)と動画像中の輝度の時間微分と空間微分の間になり立つ方程式を使う方法(勾配法)[4]の2つがある。ここでは後者を使う。

人間の表情の変化は表情筋の動きによってつくり出される。表情の変化は皮膚が表情筋に引かれて動くことによって観察される。表情筋が他の多くの筋肉と違う点は他の多くの筋肉は骨と骨を接続してその骨を動かすために用いられているのに対して表情筋は骨(頭蓋骨)と皮膚を接続して皮膚を変形させるのに使われている点である。表情の変化を読みとるためにはその皮膚の領域ごとの動きを検出しなければならぬので、そのためにオプティカルフローを用いる。

時間 t における位置 (x, y) の点の輝度を $E(x, y, t)$ とし、パターンは明るさが変わらずに位置だけ移動すると仮定すると

$$E_x u + E_y v + E_t = 0 \quad (2)$$

が成り立つ。ここで $(u, v) = (dx/dt, dy/dt)$ がオプティカルフローであり、 E_x, E_y, E_t をそれぞれ x, y, t に関する E の偏微分とする。これに制約条件として「オプティカルフローの空間的变化(ここではグラディエント E_x^2 の和)を最小にする」というものをつけることによってオプティカルフローを求めることができる。

実際には雑音や量子化誤差などのため式(2)を厳密に成り立たせることにはあまり意味がないので、左辺を小さくすることを考える。すると雑音とおなじくらいの大さきを持つある係数 α を使い、ラプラスアンのある点の値を近傍の点の平均から引いたもので近似することによって

$$\begin{aligned} (\alpha^2 + E_x^2)u + E_x E_y v &= (\alpha^2 \bar{u} - E_x E_t) \\ E_x E_y u + (\alpha^2 + E_y^2)v &= (\alpha^2 \bar{v} - E_y E_t) \end{aligned} \quad (3)$$

と変形できる。これをガウス-ザイデル法で解くことによってオプティカルフローを求める。

求めたオプティカルフローをあらかじめ入力してある皮膚領域に関する知識に従ってクラスタリングし、その動きと上記の特徴点によって表情を抽出する。

求めたオプティカルフローはまたレームとフレームとの間の特徴点の動きを推定するのに使用することができる。また、特徴点を探す際にオプティカルフローの値の大きいところを重点的に探すことによって特徴点抽出の効率をあげることもできる。オプティカルフローの不連続線によって画像を領域に分ける際には逆に見つかった特徴点の情報を利用することができる。これらの各領域でのオプティカルフローの動きによって表情筋の働きによる顔表面の皮膚の変形を検出するとともに、領域のトポロジーの変化によって表情の変化を検出する。

6 まとめ

局所の特徴強調アルゴリズムとオプティカルフローによる顔画像の解析手法を提案した。局所の特徴強調によって大きな画面の明暗に隠された細かい特徴を抽出することができる。オプティカルフローによって頬などの特徴点が見つみにくい部位の動きを検出するとともに領域分けをすることができる。

今後の課題としては特異点理論を応用することによって顔の凹凸を検出し、それを利用して表情認識をすること、また特徴強調、特徴抽出、オプティカルフローの処理などの本質的に並列性を多く含む手続きを並列化するとともにそれらを統合した並列システムを構築することなどがあり、検討を進めている。

参考文献

- [1] 馬場口 登, 山内 智, 相原 恒博. 知識やモデルを利用した顔画像解析. 情処研報 (CV), Vol. 46, No. 1, pp. 1-8, January 1987.
- [2] Rafael C. Gonzalez and Paul Wintz. *Digital Image Processing*. Addison-Wesley, 1987.
- [3] Ellen C. Hildreth. *The Measurement of Visual Motion*. The MIT Press, 1984.
- [4] Berthold K. P. Horn and Brian G. Schunck. Determining Optical Flow. *Artif. Intell.*, Vol. 17, pp. 185-203, 1981.
- [5] 間瀬 健二. オプティカルフロー抽出による表情筋の動作検出. 信学技報, Vol. 89, No. 469 (PRU89-128), pp. 17-24, March 1990.
- [6] 光本 浩士, 岡崎 耕三, 梶見 直樹, 田村 進一. エネルギー関数とオプティカルフローを用いた口唇輪郭の抽出・補間と追跡. 情報処理学会論文誌, Vol. 31, No. 3, pp. 444-453, March 1990.
- [7] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Press, 1982.
- [8] Keith Waters. A Muscle Model for Animating Three-Dimensional Facial Expression. *ACM SIGGRAPH*, Vol. 21, No. 4, pp. 17-24, July 1987.
- [9] 柳 修, 若山 忠雄, 宮田 務. 人物像等に適したエッジ画像の算出とそのエッジ線検出. 情処研報 (CV), Vol. 59, No. 4, pp. 1-8, March 1984.