

## 相関演算に基づく顔の動き推定法

4M-3

塚本 明利† 李 七雨† 辻 三郎‡  
† イメージ情報科学研究所 ‡ 大阪大学基礎工学部

### 1 はじめに

顔画像を利用した新しいMMIの実現のため、表情認識や動作理解などの研究が盛んに行なわれている。しかし、人間の顔は個人差や表情変化などにより様々であるため、安定した情報抽出は困難である。

筆者らは以前、テキストチャマッピング手法を用いていろいろな方向を向く顔の暫定的なモデル画像を作成し、これを用いた顔の動き追跡法を示した[1]。本稿では入力画像から獲得した様々な方向を向く顔をモデル画像としてメモリ上に蓄積してゆき、これらとの相関値により顔の向きを推定する手法について述べる。

### 2 正面顔の探索法

正面顔の探索は、画像における輝度およびエッジ要素の情報に基づいて行なう。しかし、それらの値は照明条件や個人性などの影響により不安定であるため、ここでは画像上の任意の領域を  $N$  個のブロックに分割し、その各ブロックにおける以下の2つのパラメータを用いる。

$$\text{Darkness } d_i = N(d_i)/N(i) \quad (1)$$

$$\text{Plainness } p_i = N(p_i)/N(i) \quad i = 1 \dots N \quad (2)$$

ここで、 $N(i)$  はその領域の第  $i$  ブロックにおける画素数、 $N(d_i)$  はそのうち輝度が領域内の全画素の平均輝度よりも低い画素の数、 $N(p_i)$  はエッジ要素が領域内の全画素の平均値よりも低い画素の数である。エッジ要素は Sobel Operator により得た。

正面顔の領域は、入力画像における様々な領域において、以下のマハラノビス距離[2]が最小となるものとして探索する。

$$D^2 = A\Sigma^{-1}A^T \quad (3)$$

$$A = [d_1 - \bar{d}_1, \dots, d_N - \bar{d}_N, p_1 - \bar{p}_1, \dots, p_N - \bar{p}_N]$$

Motion Estimation of Human Face by Image Correlation  
Akitoshi TSUKAMOTO†, Chil-Woo LEE†, Saburo TSUJII‡

† Laboratories of Image Information Science and Technology

‡ Faculty of Engineering Science, Osaka University

$$\Sigma = [\sigma_{ij}]_{i,j=1 \dots 2n}$$

ここで  $\bar{d}_i, \bar{p}_i$  はそれぞれサンプル画像の顔領域における  $d_i, p_i$  の平均値であり、 $\Sigma$  は共分散行列である。

### 3 他の方向を向く顔の探索法

以降のフレームにおいては、獲得したモデル画像によるテンプレートマッチングを繰り返すことにより顔領域の追跡を行なう。すなわちモデル画像との相違度が最小となる領域の探索を繰り返す行なう。

モデル画像  $M$  と入力画像上の領域  $A(x,y)$  との相違度は以下の式で与えられる。

$$D(A(x,y), M) = \sum_x \sum_y |A(x+i, y+j) - M(i,j)| \quad (4)$$

ここで、相違度に対する閾値  $th$  を十分小さく設定し、相違度がこの閾値以下で最小の領域を、モデル画像  $M$  の顔の向き  $\vec{M} = (\phi, \psi, \theta)$  にほぼ等しい方向を向く顔の領域であると仮定する。モデル画像  $M$  で探索可能な顔の方向の変動許容範囲  $\Delta \vec{M} = (\Delta\phi, \Delta\psi, \Delta\theta)$  は次のようにして得る。

$$\text{dist}(M_{\Delta\phi, \Delta\psi, \Delta\theta}, M) \leq th \quad (5)$$

ここで  $M_{\Delta\phi, \Delta\psi, \Delta\theta}$  は  $M$  をテキストチャとして顔の3次元モデルにマッピングし、これを  $(\Delta\phi, \Delta\psi, \Delta\theta)$  だけ回転させて得られた合成顔画像である。また、 $\text{dist}(M_{\Delta\phi, \Delta\psi, \Delta\theta}, M)$  は2つの画像  $M_{\Delta\phi, \Delta\psi, \Delta\theta}$  と  $M$  の間の相違度である。

入力画像における顔の向きが変わり、モデル画像  $M$  との相違度が閾値  $th$  を越えたときには、上記変動許容範囲内で向きを変えた顔の暫定モデル画像をいくつか作成し、それらを用いて向きが変わった顔を探索する。すなわち、各暫定モデルと以前の顔の位置の近傍領域とのマッチングを行なった結果、領域  $A$  と方向  $\vec{M}' = (\phi', \psi', \theta')$  を向く暫定モデル  $M'$  が最小の相違度を与えたとすると、そのときの近傍領域  $A$  を  $\vec{M}'$  を向く顔の領域として得る。この領域の画像は新しいモデル画像  $M'$  としてメモリ上に蓄積し、以降の入力画像列における顔の動き推定において用いる。

### 4 角度推定法

入力画像における顔の追跡処理において、モデル画像  $M$ ,  $M'$  との相違度がともに閾値  $th$  以下の領域  $N$  は、それぞれの相関値を用いて以下の式により領域  $N$  における顔の方向  $\vec{N}$  を推定する。

$$\vec{N} = (dist(n, M) - err)\vec{M}' + (dist(n, M') - err)\vec{M} \tag{6}$$

ここで  $\vec{M}, \vec{M}'$  はそれぞれ領域  $M, M'$  に存在する顔の向きの推定角度である。また  $err$  はテキストチャマッピング処理の誤差による相違度の増加分を考慮したパラメータである。

### 5 処理実験結果

#### 5.1 正面顔の探索処理結果

まず顔探索処理実験の結果について述べる。正面顔のサンプル画像として、予め 22 人分 152 枚の画像を準備し、式 (3) における  $A, \Sigma$  を求めた。顔領域は Fig.1 のように 35 ブロックに分割した。探索実験においてはサンプル画像とは別人の 28 人分 37 枚のテスト画像を用いた。サンプル画像、テスト画像ともに顔の大きさや表情、背景などは任意である。

Table.1 に本手法による実験結果を示す。表中探索正解率は各テスト画像において式 (3) の  $D^2$  が最小となった領域が目視検査の結果実際に顔であったテスト画像の割合、また候補含有率は  $D^2$  の小さい上位 20 位までの候補領域の中に実際顔を示すものが含まれていたテスト画像の割合である。

探索正解率	56.7%	(21/37)
正解含有率	86.5%	(33/37)

Table 1: 顔探索実験結果

Table.1 より、撮影環境の違う画像において別人の顔を探る実験において、約半数以上の顔領域を探ることができ、本手法の有効性を確認した。なお、探索に失敗したテスト画像においては髪型の違いが顕著であったため、統計的性質がサンプル画像のものと異なり、探索できなかったものと思われる。

#### 5.2 顔の動き推定結果

次に実画像における顔の動き推定処理の結果を示す。Fig.2 は最初のモデル画像及びそれから作成された暫定モデル画像である。Fig.3 に約 250 フレーム間における顔の方向の推定角度を、その時の動作とともに示す。この実験においては相違度の閾値  $th$  は 15/画素とし、また相違度の推定増加分  $err$  は 25/画素とした。同図より、ほとんど全ての顔の動きが得られていることがわかる。

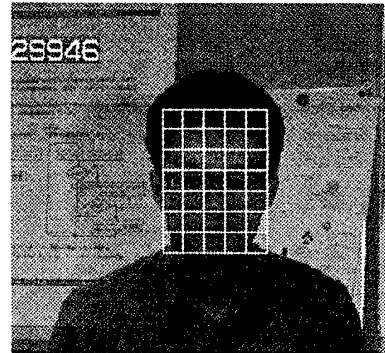


Figure 1: 顔領域ブロック



Figure 2: 最初のモデル画像と暫定モデル画像

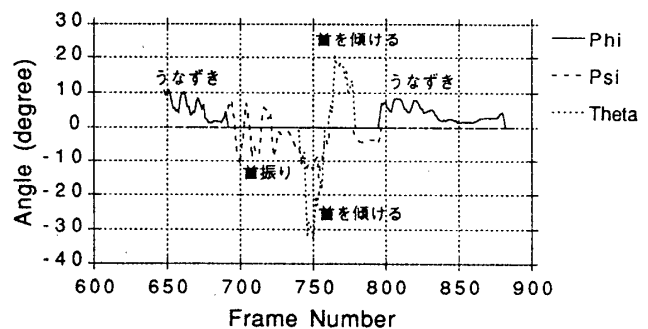


Figure 3: 顔の動き推定結果

### 6 おわりに

実験結果から、単純な相関演算だけで顔の動きが推定できることが確認できた。今後はリアルタイムに顔の動きを得ることのできるシステムの構築を目指してゆきたい。

#### 参考文献

[1] 塚本, 李, 辻, "3次元モデルとテキストチャマッピングによる顔の追跡", 信学技報, PRU93-96.  
 [2] 田中, 脇本, "多変量統計解析法", 現代数学社 1983.