

## 3D-1 FFTを用いた正面顔画像認識の検討

佐々木 努 赤松 茂 末永康仁

NTT ヒューマンインタフェース研究所

## 1 まえがき

人が顔画像により行なっている情報のやりとりを機械上で実現させる技術はマンマシンインタフェースを向上させるために必要な技術である。この研究の一環として、我々は、カラー画像の色空間におけるクラスタリングに基づく領域分割とラベリング法[1]を顔画像に適用して目や口などの造作を抽出し、これを顔の位置合わせに用いる顔認識システムについて検討を進めている[2][3]。

ここでは、我々が目指す顔認識実験システムの分類に使用する照合時の特徴としてフーリエ変換像のワースペクトルを用いた場合に関して比較的規模の大きな顔画像データベースを用いたシミュレーション実験による評価を示す。

## 2 顔画像の分類

## 2.1 パタンの正規化

照合用特徴は、位置および濃淡値に対して正規化を施した顔画像から抽出した。

顔パタンの正規化と照合領域の設定に用いる基準点は、領域分割・ラベリング法で抽出の可能性が高いものとして、右目中心点  $E_r$ 、左目中心点  $E_l$ 、口点  $M$  の3点を選んだ。なお、今回の実験では、各造作の位置を手操作によって抽出するが、このとき、直接左右目の中心点、口点を直接選ぶことは困難なため、実際には左右の目頭、目尻、口端の計6点から基準点を求めた。

また、濃淡値の正規化については、各顔画像の特徴抽出用顔画像領域内の輝度値の分散を一定値に揃えることで行なった。

## 2.2 フーリエ変換像による照合用特徴

照合に用いる特徴にはテクスチャ解析などで用いられるフーリエ変換像のワースペクトルを用いた。フーリエ変換像は像の水平方向の移動に無関係という性質を持っており、位置ずれなどが問題になる顔画像の照合に適していると思われる。

対象画像を  $f(x, y)$ 、これのフーリエ変換像を  $F(u, v)$  とすると、 $f$  の濃度値変化が大雑把で少ないほど  $F$  の原点近くにおける値が大きく、逆に、 $f$  の濃度値変化が細かく微妙になるに従い、 $F$  上の大きな値の部分は原点から遠ざかる。このような特性から、フーリエ変換像は、通常、半径方向と角度方向で評価される。ここで、

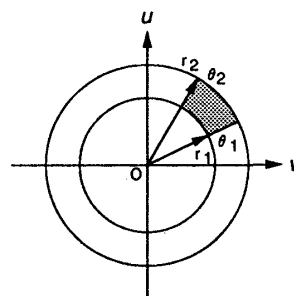


図1: FFT ワースペクトル  
半径方向ワースペクトル分布:

$$F_1(r) = \int_0^{2\pi} |F(r, \theta)|^2 d\theta$$

角度方向ワースペクトル分布:

$$F_2(\theta) = \int_0^{\infty} |F(r, \theta)|^2 dr$$

とすると、 $F_1(r)$  はテクスチャの細かさを抽出(コントラスト、均一性、エントロピーに対応)し、 $F_2(\theta)$  はテクスチャの方向性を抽出(相関に関与)するといえる。

本報告では上記の特性を反映するために、フーリエ変換像を角度方向と半径方向に分割することで得られるセクタごとのワースペクトルを求め、これを照合用特徴として用いた。例えば、図1に示される  $r_1, r_2, \theta_1, \theta_2$  で囲まれるセクタのワースペクトルは、次式のように表される。

$$\phi_{r_1 r_2 \theta_1 \theta_2} = \sum |F(u, v)|^2$$

なお、入力パターンと標準パタンの類似性の尺度には、マハラノビス距離を用いた。

## 3 実験

実験に用いた顔画像は正面平常顔とした。

学習用データ: 人物 269 人(男 257 人, 女 12 人(眼鏡使用者: 男 125 人, 女 1 人)) について 1 サンプル/人の顔画像データを用意し、これを標準パターンとして認識辞書作成に用いた。

テスト用データ: 学習用データに含まれる 20 人(男 19 人(眼鏡使用者: 2 人), 女性 1 人) について間隔をあけて 4.5 サンプル/人を収集し、全部で 94 パターンを用意した。

Experiments on Front-view Face Image Classification Using FFT Power Spectrum

Tsutomu SASAKI, Shigeru AKAMATSU and Yasuhito SUENAGA

NTT Human Interface Laboratories

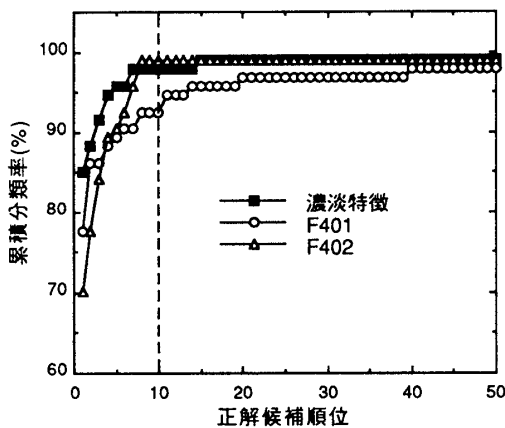


図 2: 累積分類率

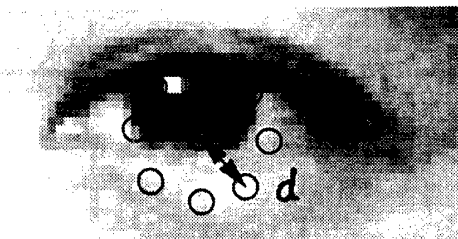


図 3: 基準点に人為的変動を与えた例

3.1 実験 I: 分類能力の評価

カテゴリ内の変動を含むテストデータに対する候補選択能力という観点から累積分類率による評価を行なった。また、比較のために 128×128 画素の原画像を 16×16 画素の縮小画像に変換(以後、濃淡特徴と呼ぶ)した後ユークリッド距離で分類する手法についても実験した。

なお、照合の際に用いる FFT パワースペクトルを得るための FFT 像の分割は等間隔におこない、以下の組合せについて実験した。ここで、下表中の半径方向分割数の ( ) 内の数字は、照合用特徴として用いた原点からのセクタ数を示している。

照合特徴名	F401	F402
角度方向分割数	64	64
半径方向分割数	8 (4)	16 (8)

実験結果を図 2 に示す。結果から、第一位認識率という点では濃淡特徴を用いた手法の方が優れている。一方、フーリエ特徴を用いた方は半径方向の分割数を多くした方がよい。

また、上位 10 位以内に安定に正解候補を挙げるといふ面では、濃淡特徴を用いた場合と、半径方向の分割数を多くしたフーリエ特徴を用いた場合の結果はほぼ同じであり、累積分類率は 97% 以上の結果であった。

3.2 実験 II: 位置ずれ変動に対する耐性評価

本実験では基準点の抽出を手操作により行なっているが、これを自動化した場合、基準点検出位置の位置ずれ変動は避けられず、これに対する認識系の耐性が求められる。

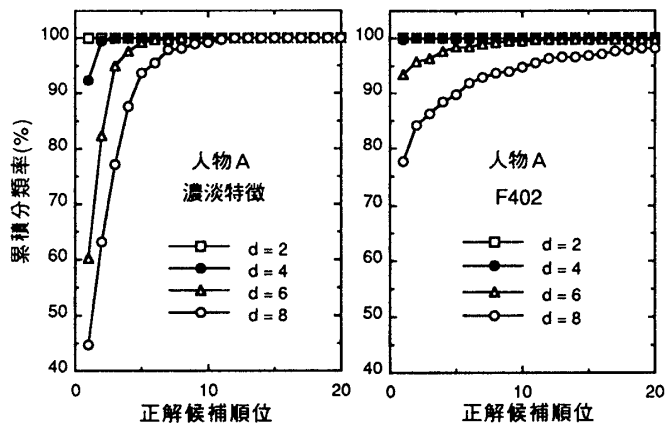


図 4: 位置ずれ変動に対する耐性評価

そこで、ある人物の標準顔画像パタンの 3 つの基準点に対して図 3 に示すような放射状方向に独立に人為的な変動  $d$  を与え、変形パタンを発生させた。この変形パタンをテストデータとした認識実験を行なった。実験結果を図 4 に示す。

図 4 より、フーリエ特徴を用いた場合の方が第一位認識率の落ち込みは少なく全体的に良い結果になっている。これはフーリエ特徴が水平変動に対して強いという性質が現れているためと思われる。

また、フーリエ特徴を用いる場合においては、2, 4 画素の変動に対して影響はないが、6 画素以上では影響が現れる。しかし、6 画素の変動であっても 10 位以内に正しい候補者を選び出せることが分かる。この結果から、フーリエ特徴を用いる認識系では、基準特徴点を抽出する際の精度は 6 画素程度にした方がよいと考えられる。

4 まとめ

人物の分類・識別の際の、照合特徴にフーリエ変換像のパワースペクトルを用いた場合について検討した。その結果、基準点の変動に対して高い耐性能力があることが分かった。またこの時、269 名の登録人物の中から正解候補をほぼ確実に 10 位以内に絞れることが確認された。

今後は、フーリエ特徴のチューニングを行ない、より優れた認識系を実現する。

謝辞

日頃御指導いただき視覚情報研究部小林幸雄部長ならびに、御討論頂いた視覚情報研究部の諸兄に感謝致します。

参考文献

- [1] ワレス, 赤松, 末永「Face Image Labelling By Probabilistic Interpretation Trees」, 信学全大秋, D-410, (1990-10)
- [2] 佐々木, 赤松, 末永「パタン照合法にもとづく正面顔画像の分類能力評価」, 信学全大秋, D-412, (1990-10)
- [3] 赤松, 佐々木, 末永「正面顔画像による個人識別の基本検討」, テレビジョン学会技術報告, Vol.14, No.54, pp.7-12 (1990-10)