

## 顔と手の認識による視覚インタフェース技術の構築

Construction of Vision-based Interface Technology using Face and Hand Recognition

渡辺博己† 本郷仁志† 安本 護† 丹羽義典† 山本和彦‡

†(財)ソフピアジャパン地域結集型共同研究推進室、科学技術振興事業団 ‡岐阜大学工学部

Hiroki Watanabe† Hitoshi Hongo† Mamoru Yasumoto†

Yoshinori Niwa† Kazuhiko Yamamoto‡

†Softopia Japan Foundation, Regional Intensive Research Project, JST

‡Gifu University

## 1. はじめに

人間と機械が円滑なコミュニケーションを行うためには、ノンバーバル情報を無視することはできない。特に、人間の意図を機械に伝えるためには顔や手の認識による視覚インタフェース技術が重要である。そこで、我々は、顔やハンドサインの識別技術によるユーザフレンドリーなインタフェースの実現を目指している。本研究では、顔やハンドサインの認識を行い、判別空間の構成の違いによる、認識率の変化について考察する。

## 2. 顔・手領域の抽出

図1がカメラから得られた画像である。まず、最も大きい肌色領域を顔領域として抽出する。図1中の白枠の矩形で囲まれた部分が顔領域として抽出された部分である。次に、顔領域の周囲に存在する肌色領域を手領域として抽出する。手領域は2つ存在するが、人間が対象に向かってサインを出すときには、顔に近い位置で出す傾向があるため、ここでは顔領域に最も近い手領域を1つ抽出する。図1中の黒枠の矩形で囲まれた部分が手領域として抽出された部分である。

肌色領域は、人間の知覚色差が幾何学的に近似できるとされているLUV表色系を用いて決定した。肌色の取り得るUV値の範囲内で最も画素数の多いUV値を基準肌色とし、基準肌色UV値と入力画像の色差ヒストグラムに判別分析法を適用することで、肌色領域を抽出する[1]。

## 3. 特徴量の抽出

前節で得られた顔、手領域に対し、4方向面特徴[2]を用いてそれぞれの特徴量を抽出する。4方向面特徴は文字認識の分野において開発され、有用性が証明された手法の1つである。特徴抽出方法は、入力画像からPrewittオペレータを用いて、水平、垂直、右上がり、右下がりの4つの方向面を抽出する。そして、これら4つのエッジ方向面をそれぞれ正規化、低解像度化し、低解像度化された方向面の画素の濃淡値を特徴量として抽出する。抽出された特徴面の一例を図2に示す。この特徴量は4つの方向面に分けてから解像度を低くするため、入力画像の解像度を直接低くする場合よりも、高解像度でのエッジ情報が保持される。また、特徴面の濃淡値を特徴量としているので、空間配置情報を特徴量として保持している。

## 4. パターン認識

パターン認識には線形判別分析[3][4]を用いた。線形判別分析は、より低い次元で、しかも、クラス間の分離を最大限強調する線形写像を構成する多変量

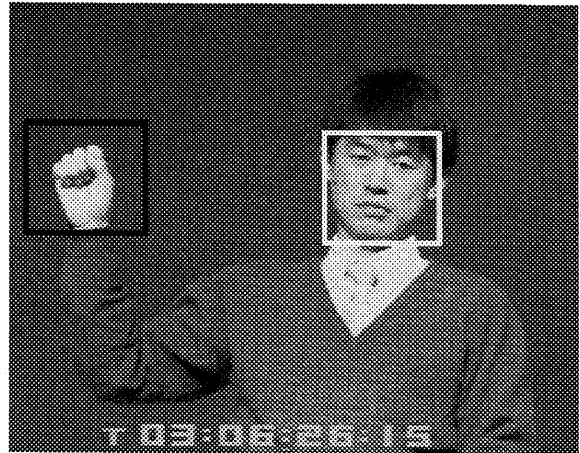
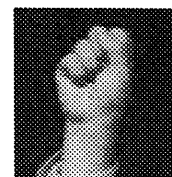


図1 抽出結果



(a)入力画像



(b)水平面 (c)右上がり面 (d)右下がり面 (e)垂直面

図2 4方向面特徴

解析手法である。

学習の際には、画像データから得られた特徴量を線形判別分析することにより、係数行列と各クラスの平均ベクトルを生成し判別空間を構成した。しかし、様々な特徴量をもつものを同一判別空間内に分布させると、局所的に分布が集中したり、個々の分布状態が拡散したりするため、よい判別空間を構成することは困難であると予想される。そこで、適度な分布状態を形成するため、認識対象物をカテゴリに分類し、その中で判別空間を構成するように階層化を行った。

認識実験では、未知の画像データから特徴量を抽出し、係数行列を用いた写像変換より生成されるベクトルとのユークリッド距離を測ることでパターン認識を行った。この時、階層化された判別空間による認識の際には2回判別式を計算することになる。

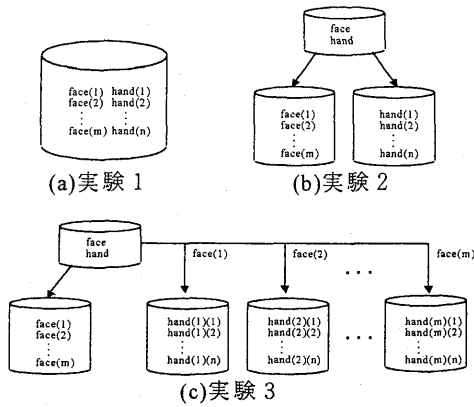


図3 判別空間の構成

5. 認識実験

顔は、カメラを中心に9つの指標(3x3, 15度間隔)を設置し、各指標を見ているときの顔をそれぞれ100枚、計900フレーム分を1セットとして5人分を学習した(計4,500フレーム)。ハンドサインは、11種のサインをカメラに対して正面に向け、100フレーム分を5人について学習した(計5,500フレーム)。これらの画像について4方向面特徴を用いた特徴量を計算し、判別空間の構成の違い(図3)による3つの実験を行った。

5.1 実験1

顔とハンドサインを全て1つにまとめた、16クラスの判別空間を作成し実験を行った。結果を表1、2に示す。表1は学習データと未知データに対する顔の認識率の平均である。表2はそれぞれのハンドサインの未知データに対する認識率の平均である。ハンドサインに関しては、学習データに被験者を含む場合と含まない場合の結果も示す。

5.2 実験2

対象のカテゴリ(顔、ハンドサイン)に応じて、大分類のクラスを作り、同じカテゴリの中で小分類のクラスを作るといった2段階に階層化された判別空間を作成し実験を行った。ここで、大分類、小分類の手法を用いたのは、顔や手といった明らかに特徴が違うもので同一の判別空間を構成した場合よりも、異なる判別空間を構成した場合の方が、認識率が向上することを期待したためである。結果を表1、2、3に示す。表3は大分類での顔とハンドサインの未知データに対する認識率の平均である。ここでも、学習データに被験者を含む場合と含まない場合の結果を示す。

5.3 実験3

実験2と同様に階層化された判別空間を構成するとともに、ユーザ毎にハンドサインの小分類のクラスを作るといった多面化された判別空間を作成し実

験を行った。ここで、ユーザ毎の小分類の手法を用いたのは、ユーザ毎に微妙な形状、姿勢の違いをもつハンドサインで同一の判別空間を構成した場合よりも、異なる判別空間を構成した場合の方が、認識率が向上することを期待したためである。結果を表2に示す。

表1 顔の認識率

|     | 学習データ(%) | 未知データ(%) |
|-----|----------|----------|
| 実験1 | 99.6     | 94.0     |
| 実験2 | 100      | 96.6     |

表3 大分類の認識率

| 実験2     | 顔(%) | ハンドサイン(%) |
|---------|------|-----------|
| 被験者含む   | 100  | 100       |
| 被験者含まない | 99.0 | 92.0      |

5.4 考察

実験により、判別空間の階層化は認識率を向上させることが示された。実験3で行った多面化の平均認識率は低下しているが、これは、ユーザによって姿勢の変化を受けやすいサインの認識率が極端に下がったためである。それを除外すると、全体的に認識率は上がっていた。また、学習データが少なかったことも認識率の低下の要因と考えられる。よって、ユーザ間で姿勢が異なるハンドサインにおいては、個別の判別空間の構成は有効であると考えられる。

6. まとめ

顔やハンドサインの識別技術による視覚インタフェースの実現に向けて、判別空間の構成の違いによる認識率の変化について考察し、判別空間の構成方法が認識率に影響を及ぼすことを示した。

今後、これらの結果を基に認識率の向上を目指し、システムの安定性を高くする方法を検討していく予定である。また、複雑背景や複数人が存在する環境におけるシステムの実現に向けて研究を行っていく予定である。

参考文献

[1]村田, 羽飼, 本郷, 加藤, 山本, “形状変化に強い目・口領域の抽出方法の提案”, 信学技法, PRMU97-159, pp.55-60, 1997  
 [2]K.Yamamoto, “Present State of Recognition Method on Consideration of Neighbor Points and Its Ability in Common Database”, IEICE transactions, Vol.E79-D, No.5, pp.417-422, 1996  
 [3]大津, “パターン認識における特徴抽出に関する数値的研究”, 電総研研究報告, 第811号, 1981  
 [4]栗田, “柔らかな情報処理のための統計的手法の応用に関する研究”, 電総研研究報告, 第957号, 1993

表2 ハンドサインの認識率

| 被験者含む   | 手1   | 手2   | 手3   | 手4   | 手5   | 手6   | 手7   | 手8   | 手9   | 手10  | 手11  | 平均   |
|---------|------|------|------|------|------|------|------|------|------|------|------|------|
| 実験1     | 100  | 71.7 | 92.3 | 97.0 | 100  | 100  | 100  | 88.7 | 81.0 | 86.7 | 99.3 | 92.4 |
| 実験2     | 100  | 71.0 | 96.0 | 85.0 | 100  | 99.3 | 100  | 94.0 | 98.7 | 99.0 | 87.0 | 93.6 |
| 実験3     | 93.7 | 84.3 | 98.3 | 79.3 | 99.3 | 100  | 97.3 | 72.0 | 91.3 | 100  | 97.7 | 92.1 |
| 被験者含まない | 手1   | 手2   | 手3   | 手4   | 手5   | 手6   | 手7   | 手8   | 手9   | 手10  | 手11  | 平均   |
| 実験1     | 81.7 | 65.7 | 14.3 | 90.0 | 40.7 | 41.3 | 95.3 | 61.3 | 90.0 | 32.0 | 9.7  | 56.5 |
| 実験2     | 68.0 | 63.3 | 20.3 | 89.7 | 50.0 | 38.0 | 99.7 | 58.0 | 84.7 | 39.0 | 3.0  | 55.8 |