

## エージェントネットワーク Socia の能動的視覚機能

4P-2

田中吾一 石黒浩 石田亨  
京都大学工学部情報工学教室

## 1 はじめに

計算機ネットワーク利用者を支援するエージェントに視覚機能を持たせ、エージェントに利用者の挙動を把握させることを試みる。電子会合支援を例にとると、視覚機能を有するエージェント（以下、視覚エージェントと呼ぶ）は、把握した利用者の存在や挙動を基に、会合への参加・不参加、また参加中の会合の中断や終了を管理する。

本論文では、人物を対象とする視覚機能として、顔領域の探索・追跡および動きの判定などの画像処理技術を提案する。図1に視覚機能の構成要素間の情報の流れを示す。顔領域の探索により、存在確認を行ない、追跡処理に移る。追跡処理からは、動作理解を行なう。各処理に失敗すれば、前段階の処理に戻る。

顔領域の探索手法は、多数報告されているが、本システムでは、定性的特徴モデルとのマッチング<sup>[1]</sup>を基本とした手法を用いる。定性的特徴モデルに基づく手法では、以下のような手順で顔を検出する。まず、サンプル顔領域を縦10横7ブロックに分割し、これらのブロック毎に次の特徴量  $L_{face}(i)$ ,  $E_{face}(i)$  を求める。

$$L_{face}(i) = NL_{face}(i)/m \quad (\text{輝度成分}) \quad (1)$$

$$E_{face}(i) = NE_{face}(i)/m \quad (\text{エッジ成分}) \quad (2)$$

ここで、 $NL_{face}(i)$ ,  $EL_{face}(i)$  はそれぞれ領域中の平均輝度成分、エッジ成分より大きい画素数、 $i$  はブロック番号、 $m$  は1ブロックの全画素数である。エッジ成分はソーベルのオペレータを用いて求める。この操作を複数のサンプル顔画像に対して行ない、それらの平均を特徴モデルとする。入力画像中の顔領域の探索においては、様々な大きさの領域を入力画像に対し走査させ、モデル作成時と同様の特徴量を求める。その特徴量をモデルの分布とマッチングさせることにより、その領域の“顔らしさ”を計算し、この値が比較的高い領域を、顔候補領域とする。

Using Active Vision for Interfacing Humans and Socia Agent Network

Goichi Tanaka Hiroshi Ishiguro Toru Ishida

Department of Information Science, Kyoto University

以上の定性的特徴モデルによる探索は、表情や撮影条件の変化に対して頑強であるが、物体の形状に関する解析を行なっておらず、顔検出の安定性に欠ける。また、会合システムのように限られた使用状況のもとでは、動きなどの情報を基に認識率を向上することができる。そこで、本稿では、定性的特徴モデルに基づく手法に以下の2点の改良を加える。

- i) エッジ形状による判定
- ii) 顔の動きによる判定

さらに、顔領域の追跡を行ない、動作理解の機能を実現し、会合システムを支援する。

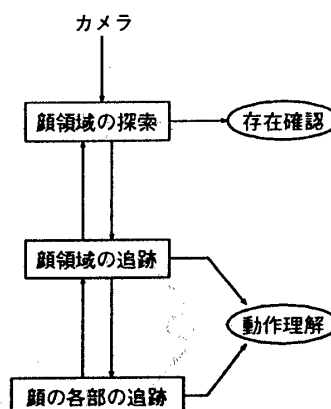


図1: 視覚機能のモジュール構成

## 2 エッジ形状による顔領域の判定

定性的特徴モデルに基づく手法では、エッジによる物体の形状を十分に分析しておらず、顔以外を検出することが多い。そこで、マッチングにより得られた候補領域の中から、以下の手順によりエッジ形状を調べ、さらに候補を絞り込む。

- i) モデル作成時と同様に、エッジの縦・横成分各々について(2)式により計算し、縦・横成分の差が比較的大きなブロックを検出してモデル化する。X成分>Y成分では‘1’、X成分<Y成分では‘-’としてイメージ表示したものを図2に示す。顔の特徴を充分に残し

ていることが分かる。

- ii) テンプレートマッチングによる探索処理で得られた数箇所の候補領域において、各ブロックのエッジ成分の方向の大小を調べ、このモデルとマッチングすることにより最終判定を行なう。



図 2: エッジ形状のイメージ表示

この手法により、9割以上の探索成功率が得られた。

### 3 動きによる顔領域の判定

以上の様な方法で得られた候補領域を一旦顔領域として扱い、以降の追跡処理に移る。追跡処理において、その領域がほとんど移動しないなら、人物の顔である可能性は少なく、もう一度、探索処理に戻るといった処理が必要となる。このシステムが使用される室内環境を考慮すると、移動物体は人物のみである可能性が高いので、この処理は有効である。

### 4 動作理解

探索によって得られた顔領域を追跡し、動きを判定する。追跡には、画像間の相違度を実時間で計算するハードウェアを用いる。確実な追跡のためには、顔領域全体の追跡が有効だが、詳細な顔の動き判定のためには、顔の各部の追跡が必要となる。追跡の確実性の向上および背景の混入を防ぐため、追跡すべき参照画像はなるべく単色部分は避け、かつ顔の中心部に近い必要がある。本実験では両目及び鼻部分を参照画像とした。しかし、各部の追跡は全体を追跡する場合に比べ不安定なため、追跡に失敗した場合は顔領域全体の追跡に処理を戻す。

顔領域の追跡は正面から始めるが、顔の向きが変化すると、正面顔を参照画像とした追跡では正確な追跡は不可能である。そこで、向きがある程度変化した場合、テンプレートを更新する必要がある。図4はテンプレートの例である。左の二つは正面、右の二つは右に傾いた場合の両目のテンプレートである。顔の向き判定は、各部分の位置関係により求める。以下に、判定の基準を示す。

- i) 左右の傾き … 両目の高さの差

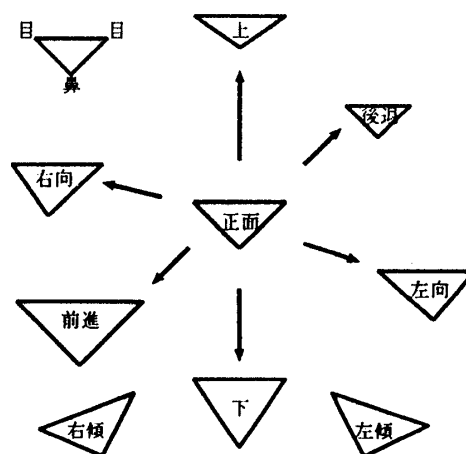


図 3: 顔の動き判定



図 4: テンプレートの例

- ii) 左向き、右向き… 両目間の中心線と鼻の位置  
iii) 上向き、下向き… 両目を結ぶ線と鼻の高さ  
iv) 前後の移動 … 両目間の距離。

図3は、両目と鼻をそれぞれ三角形の頂点に見たて、位置関係を表したものである。この関係により、ある程度顔の向きを推定できるので、マッチングを高速化できる。i), ii) の実際の動きに対しては8割以上、iii), iv) はおよそ6割の成功率を確認している。iii), iv) の成功率は若干低いですが、実際の人間の顔と比較して、図3の三角形が極端に崩れることのないよう、追跡処理に拘束条件を付加することで、成功率は上がるものと考えている。

### 5 おわりに

現在、視覚エージェントの初期バージョンの実装を完了し、顔領域の探索プログラムを電子会合支援システムのインタフェースに組み込み中である。また、顔の向き変化や移動から人物の動作を推定する行動認識モデル、その動作により会合を制御する会合モデルの二つのモデルを現在製作中である。

### 参考文献

- [1] 塚本明利, 季七雨, 辻三郎, “複数のモデル画像による顔の動き推定,” 電子情報通信学会論文誌 Vol. J77-D-II, No.8, pp.1582-1590, 1994.