

# 知識やモデルを利用した顔画像解析

Analysis of Facial Images by Using Knowledge and Model

馬場口 登 山内 智 相原 恒博  
Noboru Babaguchi Satoshi Yamauchi Tsunehiro Aibara

愛媛大学 工学部 電子工学科  
Faculty of Engineering, Ehime University

あらまし 3次元シーンを処理対象とするコンピュータビジョンの分野では、これまで、積木世界などの人工物体を対象とすることが多かった。しかし、3次元ビジョンシステムをより高度化するには、自然物体についても検討する必要がある。そこで、本報告では自然物体の中でも比較的単純な構造を持つと考えられる人の顔画像を対象とし、知識やモデルを利用した画像解析システムを提案する。本システムで用いる知識は、画像処理に関するヒューリスティックな知識、顔の構造に関する事実、見え方の知識等であり、同時に顔の標準的な3次元モデルも利用する。

Abstract It is desired that the three dimensional vision system should possess the ability to deal with the natural scenes. In this paper, we propose an image analysis system for human faces which are regarded as one of the simplest natural objects with respect to their structure. This system is based on three types of knowledge, (1)heuristics about image processing algorithms, (2)facts about the human face, and (3) knowledge about the view world. Also the three dimensional model of human faces is used to analyze the facial images.

## 1. まえがき

3次元世界を表すシーンと呼ばれる2次元画像から、奥行きなどの3次元情報の抽出や画像の3次元処理を目的とするコンピュータビジョンの分野では、これまで積木世界のような拘束条件の強い人工物体を対象とすることが多かった。しかしながら、より高度な画像解析システムを現実のものとするには、処理対象を自然物体にも拡張して種々の知見を得る必要がある。近年では自然シーンについても対象とする範囲や問題を限定したシステムも検討されつつある。

本報告では、知識やモデルを利用する顔画像解析システムを提案する。画像解析を行う際に、知識や対象となるもののモデルを利用

することの重要性は従来よりしばしば指摘されている[1][2]。特に、単一の画像、すなわち単眼画像を前提とする場合には知識やモデルの利用は不可欠である。

さて、処理対象に人間の顔の画像を選んだ理由を以下に列挙する。第1に、人間の顔は自然物体の中でも、比較的単純な構造であり、知識及びモデル利用の手段を決定しやすいと推察されること。第2に、コンピュータで取り扱うことの難しい情緒的な特徴(笑う、泣く、怒るなど)を有しており、興味深い処理対象であること。そして第3に、顔画像は実用面からも個人照合の一手段として有力であることなどである。これまでに、顔画像を対象とした報告はあるが[3]-[5]、本システム

では、3次元モデルの利用により、画像の3次元的取り扱いに重点を置く[6]。本研究のねらいは、目や口などの顔の特徴を表す領域の抽出、顔の向いている角度の推定、正(側)面画像からの側(正)面画像の推定、顔画像ベースとのマッチングによる個人照合、情緒的特徴の抽出等の機能を有する顔画像専用理解システムを構築することであるが、現段階では、目標とする機能の一部しか持っておらず、本稿はその中間報告である。

次章以下、利用する知識、モデルを説明し、本システムの概略を示していく。

## 2. 知識・モデルの利用

画像解析能力を向上させるため、本システムでは以下に述べるような3タイプの知識を利用する。又、与えられた2次元単一画像から3次元情報を得るのは理論的に不可能であるので、顔の表面の形状を表す3次元モデルを導入する。

### 2.1 画像処理に関するヒューリスティックな知識

この知識は画像処理アルゴリズムについてのヒューリスティックスであり、ある特徴を得るにはどのような画像処理が適切かということを示すものである。本システムにおいては、目・鼻・口などの顔の特徴を表す部分(以下、顔の特徴部と呼ぶ)を抽出すると言うゴールを設定したとき、それを得るための画像処理アルゴリズムの組合せにより間接的に表現する。すなわち、アルゴリズムの特性とゴールの特性間の関係に依存するものである。例えば、顔と背景のセグメンテーションにはエッジ抽出(LaplaceやSobelオペレータ)を行い、画素濃度の異なる領域では濃度ヒストグラムによる2値化を用いる。個別の領域を調べる場合にはラベリングを用い、領域の細かな形状を調べる場合には射影や細線化を用いる。又、エッジを抽出する際にも利用するオペレータの特性を考慮しなければならない。ここで利用する画像処理アルゴリズムについては3.2.2.で述べる。

### 2.2 顔の構造に関する事実(fact)

人がいかなる向を向いていたとしても変化

することのない事実であり、顔の特徴部である目、鼻、口、顎の位置関係や各々の属性を表すものである。つまり、目は2つあり、鼻は目の下にあるとか、口は鼻の下に、顎は口の下にあると言ったようなことである。これは、主に処理、推論結果の妥当性を検証するために用いられ、推論のための重要な要素となる。

### 2.3 見え方(view world)の知識

顔の角度(正面を $0^\circ$ 、側面を $90^\circ$ )とそのときの顔の特徴部の見え方を記述したものである。例えば、顔が横を向くに従ってその方向の目は小さく見え、このことを目の領域を示すパラメータで表現する。又、角度の変化に伴う各特徴部の相対的な位置の変化もこの種の知識に含める。

### 2.4 顔の3次元モデル

人間の顔の一般的な表面の形状を表したものであり、阿部らの3次元モデル[4]に類似したものである。顔画像の3次元処理には、種々のパラメータによって変化させることのできるモデルを作成する必要があるとの考察により用意したものである。ここでは、3次元空間上での代表的な座標点の集合により記述する。図1に3次元のモデルを示す。このモデルは正面を向いた状態に設定されていて、顔の幅によってモデルを変化させることができ、顔の特徴部の変化に応じて局所的に変化させることができる。

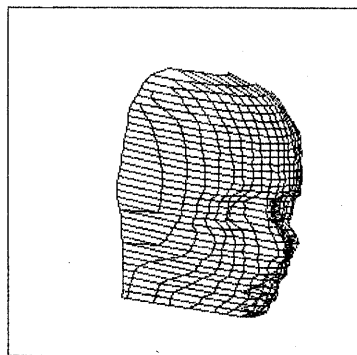


図1 3次元モデル

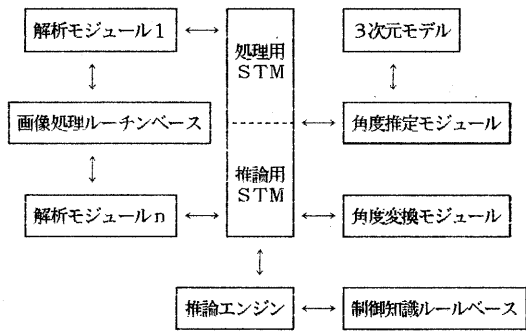


図2 顔画像解析モデル

### 3. システムの構成

本システムの構成を図2に示す。このシステムは、パターン理解システムに適すると考えられるブラックボードモデル[7][8]に類似したもので推論エンジン、制御知識ルールベース及び推論用STM (Short Term Memory) からなる推論機構と処理用STM及び各モジュールからなる処理機構に分けられる。

#### 3.1 推論機構

##### 3.1.1 制御知識ルールベース

制御知識ルールベースは

IF 条件 THEN 動作

のプロダクションルールの集合であり、条件部が満たされたとき、何らかの動作、すなわちモジュールの起動、パラメータの変更などが実行される。

##### 3.1.2 推論用STM

これは推論に直接関係するデータを管理記憶するものである。ここでのデータとは、各モジュールから出力される結果である。

##### 3.1.3 推論エンジン

システム全体の制御、モニタリングを担当する部分である。特に、解析モジュール各々

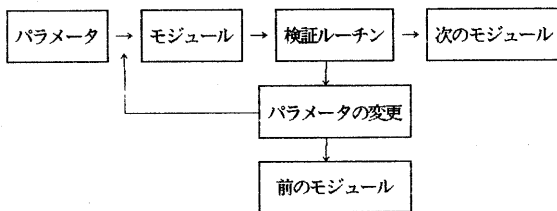


図3 モジュールの動作

については、図3に示すように処理用STMのパラメータを与え、モジュールを起動させる。そして、モジュールによって得られた結果がモジュールごとの検証ルーチンによって結果の妥当性が判定され、矛盾がなければ次のモジュールが起動させる。矛盾が生じれば、パラメータの変更を行った後、そのモジュール自身あるいは前のモジュールを起動させるようになっている。検証ルーチンの動作は、以下の通りである。対応する制御知識ルールベースの各ルールの条件が推論用STM内のデータと照合され、条件に合致したルールが選択される。選択されたルールにより次の行動が実行される。この結果として推論用STMが更新される。

#### 3.2 処理機構

これは推論エンジンからの命令により画像に対して処理を施すものである。また、前述の推論機構に必要な各データを求める。

##### 3.2.1 処理用STM

モジュール間は互いに独立であり、互いに他のモジュールを直接参照することがないので、一時的に記憶するための領域が必要となる。これが処理用STMであり、ここには、原画像、濃淡画像、2値画像、エッジ画像、ラベル画像、モデル及び各種パラメータ等の必要な部分が記録される。

##### 3.2.2 画像処理ルーチンベース

各々の特徴抽出を行うのに必要な2次元画像の基礎的処理ルーチンの集合である。その中にはエッジ抽出 (Sobel, Laplaceオペレータ)、2値化、細線化、濃度ヒストグラム、ラベリング、射影等が含まれており、各解析モジュールによってアクセスされる。

次に、各画像処理アルゴリズムの概略とその特性を示す。

##### <2値化>

濃淡画像やSobel及びLaplaceオペレータを施した画像に対して固定しきい値処理により2値画像を作成する。図4の濃淡画像の2値画像を図5に示す。但し、図4はディザ法により表示したものである。

##### <Sobelオペレータ>

このオペレータは各画素で図6の3×3の



図4 濃淡画像

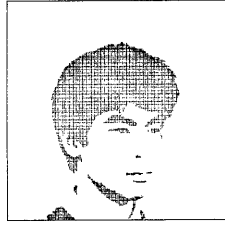


図5 2値画像

$X_1$	$X_2$	$X_3$
$X_4$	$X$	$X_5$
$X_6$	$X_7$	$X_8$

図6 微分マスク



図7 Sobel画像

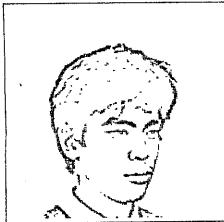


図8 Laplace画像

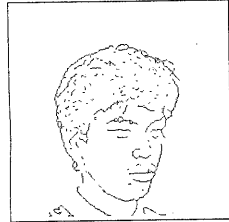


図9 細線画像

マスクにより

$$|X_1 + 2X_2 + X_3 - X_6 - 2X_7 - X_8|$$

$$+ |X_1 + 2X_4 + X_5 - X_3 - 2X_5 - X_8|$$

を計算したものである。本オペレータの特徴は1次微分であるためにエッジが太く抽出され、連続した領域を得ることができる。図7にSobelオペレータの実例を示す。

<Laplaceオペレータ>

このオペレータは各画素で図6の3×3のマスクにより

$$X_2 + X_4 + X_5 + X_7 - 4X$$

を計算したものである。本オペレータの特徴は2次微分であるためにエッジを細く抽出することができるが、連続した部分が途切れる可能性を持っている。図8にLaplaceオペレータの実例を示す。

以上のような特徴よりSobelとLaplaceオペ

レータを使い分ける。

<細線化>

細線化の方法は種々あるが、ここでは、8連結性の保存されるHilditchのアルゴリズムを用いる。図9に細線画像を示す。

<ラベリング>

ラベリングは2値画像に対するもので、8連結領域各々にラベルを与える。これにより、各領域が独立に取り扱うことができる。

<雑音除去>

ラベル付けされた領域の画素数がある一定値以上の領域を削除する。

<濃度ヒストグラム>

濃淡画像のある範囲に対し、その濃度値についてのヒストグラムを取る。2値化におけるしきい値設定の際に用いる。

<射影>

画像内あるいは画像全体のある矩形範囲に対し、垂直、水平両方向の射影を取る。2値画像の場合には範囲内の領域に関する大まかな特徴が得られる。

### 3.2.3 解析モジュール

数個の解析モジュールがあり、顔の特徴部を抽出するものである。モジュール自身には、処理用STMからデータを受け取る入力部とモジュールで得られた結果を処理用STMに引き渡すための出力部及び画像処理ルーチンベースをアクセスするための部分がある。各モジュールは、

入力: 画像の種類

対象とする範囲(座標値)

パラメータ

使用ツール: 画像処理アルゴリズム群

出力: パラメータ

座標値

としてフレーム風に表すことができる[9]。

以下に個々のモジュールの詳細を示す。各モジュールの処理内容は図10に示すが、同図中、実線、破線、一点鎖線、二重線で囲んだものは各々対象画像、画像処理アルゴリズム、各種パラメータ、出力を表す。又、図3に示すように、一つのモジュールが実行された後、推論エンジンにより検証ルーチンに入る。このとき出力結果が所望のものであれば

次のモジュールへ移行するが、そうでないときの検証ルーチンのルール（制御知識ルールベース内にある）も以下に記述する。

<端点モジュール>

入力：濃淡画像

使用ツール：Laplaceオペレータ，射影

出力：頭の頂点，顔の左右端

処理方法：図10（a）

検証ルーチン：

顔の横幅が縦方向の長さよりも長い

=>処理範囲の変更

<目のモジュール>

入力：濃淡画像

頭の頂点，顔の左右端

2値化しきい値

目の範囲の推定領域

使用ツール：2値化，ラベリング，射影，

Sobel, Laplaceオペレータ，

雑音除去

出力：目の数，目を囲む範囲の座標

処理方法：図10（b）

検証ルーチン：

両目とも抽出されなかった

=>2値化しきい値，推定領域の変更

片目が抽出されている

=>推定領域の限定

<鼻のモジュール>

入力：濃淡画像

目を囲む領域の座標

2値化しきい値

使用ツール：Laplaceオペレータ，

2値化，射影

出力：鼻を囲む領域の座標

処理方法：図10（c）

検証ルーチン：

鼻が抽出されなかった

=>2値化しきい値の変更

<口のモジュール>

入力：濃淡画像

鼻を囲む領域の座標

使用ツール：Sobel, Laplaceオペレータ，

細線化，射影

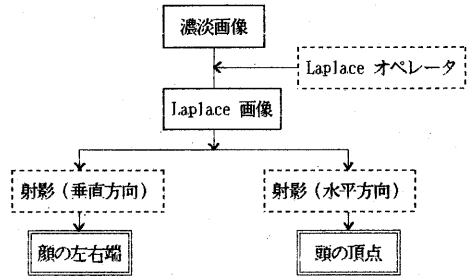
出力：口を囲む領域の座標

処理方法：図10（d）

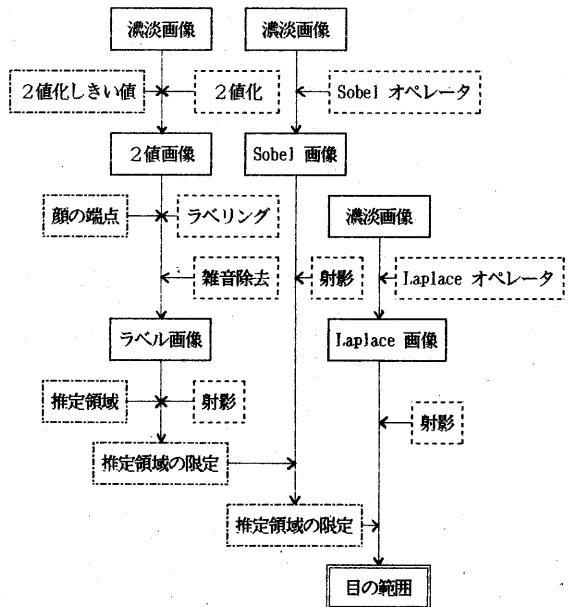
検証ルーチン：

口が抽出されなかった

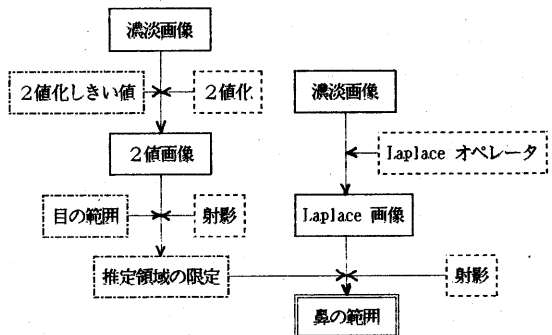
=>領域の限定



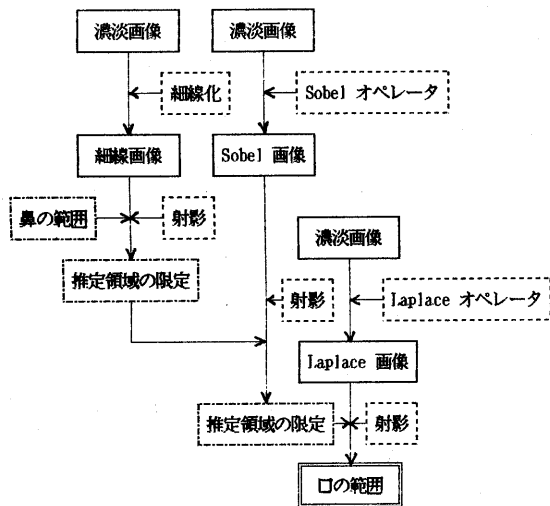
(a) 端点



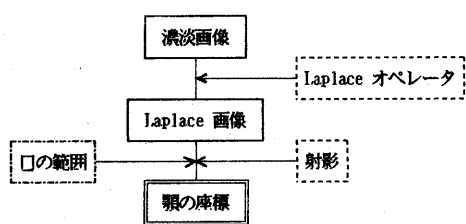
(b) 目



(c) 鼻



(d) 口



(e) 顎

図 10 各モジュールの処理内容

<顎のモジュール>

入力: 濃淡画像

口を囲む領域の座標

使用ツール: Laplaceオペレータ, 射影

出力: 顎の座標

処理方法: 図 10 (e)

検証ルーチン:

顎が抽出されなかった

=>しきい値の変更

図 11 の (a) から (d) は, 各特徴を抽出する際の推定領域を, (e) は, 抽出結果を示している。

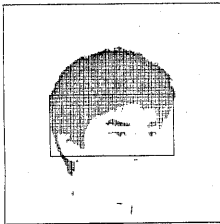
3.3 角度推定モジュール

基本的な考え方は, 3次元モデルと画像との対応を調べることにより角度を推定すると言うものである。但し, モデル全体の対応を

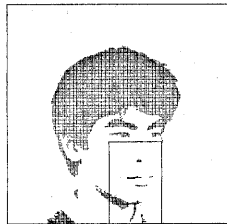
考えるのは, 効率面で得策ではないので, 局部的に目の部分を対象とする。3次元モデルは顔の前半面を想定しているため, 3次元モデルの縁となる部分に対応する画像中の部分を最初に見つける必要がある。

図 12 のように, 両目の midpoint A を通る断面を半円であると仮定し, 顔の両端点の midpoint を O としたとき, 画像内において点 B を含むと考えられる範囲が設定できる。この範囲内において髪と顔の境界点を 3次元モデルに対応する点とする。

次に, 3次元モデルと画像のサイズを一致させるため, 図 12 のように 3次元モデルの X 軸上での仮想的な長さを  $R \times g / m$  とする。そして, 3次元モデルを回転させ, 3次元モデル, 画像上の右目と左目の幅を各々,  $r_{r1},$



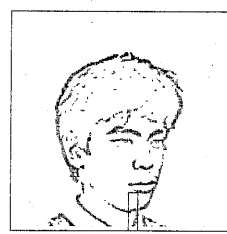
(a) 目の推定領域



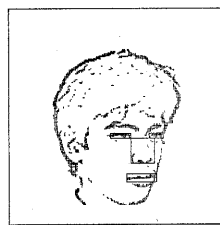
(b) 鼻の推定領域



(c) 口の推定領域



(d) 顎の推定領域



(e) 抽出結果

図 11 特徴部抽出

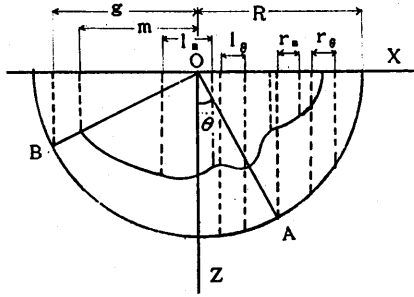


図12 角度の推定

$r_g, l_n, l_g$ としたとき,  $r_g / l_g - r_n / l_n$ の絶対値が最小となる角度 $\theta$ を推定する顔の角度とする。

#### 4. 実験結果とその結果

実験に用いた画像は, CCDカメラから入力した画面サイズが256×256で, 濃淡レベルが0から63の濃淡画像である。実験データの詳細は, 男性8人, 女性2人に対してそれぞれ約0°, 15°, 30°, 45°の角度(但し, 15°~45°に対しては右向き, 左向きの何れか1つであり, 画像に対して右の方向を正とする)で入力した合計40サンプルの画像に対して, 顔の特徴部の抽出, 及び顔の角度の推定を試みた。

実験結果を各々表1, 表2にまとめる。図13, 図14に特徴部抽出例, 誤り抽出であった画像例を示す。実験サンプルのうち, 良好な結果の得られなかったものの理由として, エッジ抽出の不備が挙げられる。SobelやLaplaceオペレータをかけたとき, その後処理として2値化をし, 2値画像を処理対象とする場合が多い。固定しきい値に依っているので所望のエッジが得られないことがある。雑音の多い画像からも良好な特性を持つエッジ抽出アルゴリズムを備えることが必要であろう。また, 2値化処理においても濃度レベルに対応できる可変しきい値を導入することも望まれる。第二の理由としては, 2値画像において目と髪の領域が連結している場合が多いことが挙げられ, 領域形状に基づくセグメンテーションが必要である。第三の理由としては,

顔の実世界座標での傾きがある。ここでは, 理想的な条件下での顔画像しか考察していないのでこれについては今後の重要な課題である。又, システム構成上の問題として, 顔の特徴部の解析モジュールのトリー状の配置がある。これにより, ある1つのモジュールの

表1 特徴抽出の結果

		端点	目	鼻	口	顎
1	0	○	×	×	×	○
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	○	○	○	○
2	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	○	○	○	○
3	0	○	○	○	○	×
	15	○	○	○	○	×
	30	○	○	○	○	×
	45	○	×	×	×	×
4	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	○	○	×	×
5	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	×	×	×	×
6	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	×	×	×	×
	45	○	○	○	○	○
7	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	○	○	○	○
8	0	○	×	×	×	×
	15	○	○	○	○	○
	30	○	×	×	×	×
	45	○	○	○	○	○
9	0	○	×	×	×	×
	15	○	○	○	○	○
	30	○	○	○	○	○
	45	○	○	○	○	○
10	0	○	○	○	○	○
	15	○	○	○	○	○
	30	○	×	×	×	×
	45	○	×	×	×	×
抽出率[%]		100	77.5	77.5	77.0	70.0

表2 角度の抽出結果

	0度	15度	30度	45度
1	3	16	-31	-43
2	2	13	-24	42
3	12	16	-24	43
4	8	16	-33	-41
5	2	-10	20	-34
6	0	-7	34	-45
7	6	-15	-25	-47
8	5	-8	33	44
9	-4	-9	24	30
10	1	11	31	38

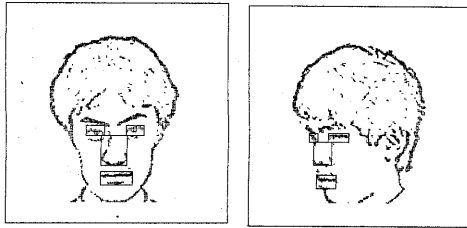


図13 特徴部抽出例

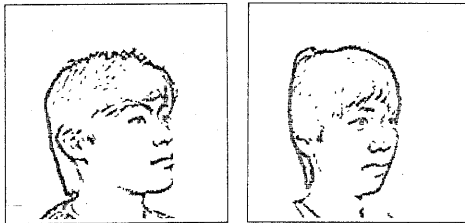


図14 誤り抽出例

影響が他のモジュールに波及することがあるのでネットワーク状にすることも考察する余地がある。

### 5. むすび

本報告では、画像処理に関するヒューリスティックな知識、顔の構造に関する事実、見え方の知識等の知識と顔の表面形状を表す3

次元モデルを利用した顔画像解析システムを提案した。顔の特徴部の抽出と顔の角度の推定に対する実験の結果、比較的良好な結果が得られた。本システムはブラックボード的システムであるので、モジュールの追加、修正等が容易であると言う特徴を持つ。

現段階では、まだ、各モジュールの能力も十分ではない上に目標とする機能も有していない。今後、顔画像の角度の変換、個人照合のための特徴抽出などを検討し、より高度な解析が可能なシステムを目指して行きたい。

### 参考文献

- [1] 白井良明: "コンピュータビジョン", 信学誌, 69, 11, pp. 1152-1156 (1986).
- [2] 松山隆司: "画像処理における知識の利用", 情報処理, 24, 12, pp. 1437-1445 (1983).
- [3] M. Nagao: "Control Strategies in Pattern Analysis", Pattern Recognition, 17, 1, pp. 45-56 (1984).
- [4] 阿部, 梁, 江島, 木村: "顔の3次元モデルに基づく顔画像の自動識別について - B-spline 曲面を用いる場合 -", 昭和61年信学総全大, 1601.
- [5] 梁, 阿部, 江島, 木村: "顔の3次元モデルに基づく顔画像の自動識別について - B-spline 曲線を用いる場合 -", 昭和61年信学総全大, 1602.
- [6] 山内, 馬場口, 相原: "知識・モデルに基づく顔画像解析", 昭和61年電気関係学会四国支部連合大会, pp. 325-326.
- [7] 長尾真: "画像認識論", コロナ社.
- [8] 上野晴樹: "知識工学入門", オーム社.
- [9] 松山, 尾崎: "LLVE: トップダウン・セグメンテーションのための画像処理エキスパートシステム", 情処学論, 27, 2, pp. 191-204 (1986).