

Title	高次局所自己相関特徴による高速画像認識モジュールの開発と自律走行型ロボットへの応用
Author(s)	中川, 弘隆
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/349">http://hdl.handle.net/10119/349</a>
Rights	
Description	Supervisor: 藤波 努, 知識科学研究科, 修士

修 士 論 文

高次局所自己相関特徴による高速画像認識  
モジュールの開発と自律走行型ロボットへの応用

北陸先端科学技術大学院大学  
知識科学研究科知識社会システム学専攻

中川 弘隆

2002年3月

# 修士論文

## 高次局所自己相関特徴による高速画像認識 モジュールの開発と自律走行型ロボットへの応用

指導教官 藤波 努 助教授

北陸先端科学技術大学院大学  
知識科学研究科知識社会システム学専攻

950064 中川 弘隆

審査委員主査 藤波 努 助教授  
審査委員 中森 義輝 教授  
審査委員 林 幸雄 助教授  
審査委員 佐藤 賢二 助教授

提出年月: 2002 年 2 月

# 目次

第1章	序論	1
1.1	背景と研究目的	1
1.2	本稿の構成	1
第2章	高速画像認識モジュール	2
2.1	高速画像認識モジュールの目標要件	2
2.2	画像認識手法の選択	2
2.3	肌色領域の抽出	3
2.3.1	しきい値法による肌色領域抽出	3
2.3.2	肌色領域の冗長	5
2.4	領域のラベリング	6
2.5	高次局所自己相関特徴	7
2.5.1	自己相関関数	8
2.5.2	高次局所自己相関特徴	8
2.6	線形判別分析を用いた識別	9
第3章	ロボットへ搭載する方法	12
3.1	高速画像認識モジュールの搭載方法	12
3.2	実時間画像認識アプリケーション”Smart-EYE”	13
3.2.1	Smart-EYE のシステム構成	16
第4章	実験 1:静止画を用いた顔・手の識別	17
4.1	実験目的	17
4.2	実験内容	17
4.3	実験結果	19
4.4	実験結果の考察	19
第5章	実験 2:ビデオ映像による顔の識別	24
5.1	実験目的	24
5.2	実験内容	24
5.3	実験結果	26

5.4	実験結果の考察	26
第6章	高速画像認識モジュールを搭載したロボット	27
6.1	実装した技術について	27
6.2	プラットフォーム	28
6.3	障害物回避機能	31
6.4	対象物との距離を一定に保つアルゴリズム	32
6.5	JAIST Navigator	35
第7章	結論	37
第8章	今後の課題	38

# 第1章 序論

## 1.1 背景と研究目的

1990年代後半から、日本ではロボットブームが起こり、現在もその熱気は続いている。その流れのきっかけとなったのが、SonyのAIBOなどのエンターテインメントロボットの一般家庭への普及である。家庭用ロボットの多くは、音声や特定のリモコンにより、人間とのインタラクションを行なっている。今後、より自然なインタラクションを実現するための方法として、画像認識による個人の識別やジェスチャ認識の開発が注目を浴びている。今まで提案された画像認識は、顔画像の識別やジェスチャ認識を十分に行なえ、実際に、固定式ロボットや産業用アームロボットに実装されている。しかし、そのシステムは複雑で、特殊なハードウェアを一般に必要とする。家庭に普及させるロボットに組み込もうとすると、ロボットの提供価格が高価すぎるものとなる。こうした理由で、家庭用ロボットに搭載する画像認識に、システムの簡略化や廉価な画像装置を用いた画像認識の処理速度や識別率の向上が強く求められている。

そこで、本研究では、廉価な画像装置によって、顔や手の形状を認識する自律走行型ロボットのための高速画像認識モジュールを開発することを目的とする。また、高速画像認識モジュールを自律走行型ロボットGAIA-2へ搭載し、個人を識別してその人にあった行動をおこすシステムを構築する。

## 1.2 本稿の構成

本稿は、第2章で本研究で用いる画像認識手法の検討と選択した画像認識手法について詳しく述べ、第3章で、高速画像認識モジュールを自律走行型ロボットへ搭載するための手法と、そのために作成した画像認識アプリケーション”Smart-EYE”について述べる第4,5章で、本研究の画像認識手法およびSmart-EYEの評価実験方法とその実験結果、結果の考察を述べる。第6章で画像認識モジュールを搭載した自律走行型ロボットについて述べ、第7章で画像認識モジュールの今後の課題を挙げて、第8章で結論を述べる。

## 第2章 高速画像認識モジュール

### 2.1 高速画像認識モジュールの目標要件

本研究での高速画像認識は、移動型ロボットに搭載することと、個人の顔を識別し、人間とロボットのインタラクションを確立することを前提にしている。そこで、構築する高速画像認識には、以下の目標要件が挙げられる。

- (a) 20 fps(frames per second) 程の高速実時間で識別しなければならない。
- (b) 背景(場所) に依存されない。
- (c) 画面内の対象物の座標を得る必要がある。
- (d) 識別率が高い。
- (e) 廉価な画像装置をもちいて実現可能である。

(a) については、認識物体によって行動を決定するようなロボットにおいて低速画像認識は致命的なものとなるためである。(b) は、移動するロボットから得られる背景画像が常に一定ではないためである。また、(c) においては、画面内の対象物の座標が分からなければ、カメラを支える雲台を制御することが困難になり、ロボットの行動決定に支障が生じるためである。

### 2.2 画像認識手法の選択

(a), (b), (c) の条件を満たす画像認識には、福井らによる顔特徴点抽出を利用したものの [1] や、栗田らによる PARCOR 画像と高次局所自己相関特徴を用いたもの [2] がある。

前者は、瞳及び鼻孔を特徴点として捉え、2次元アフィン変換により顔領域を一定の大きさ、向きに正規化し正規化パターンを抽出し照合する方法である。顔画像認識のみならこの方法でも十分だが、手の形状を識別する際に特徴点を一意に定めることが困難となるので、本研究の画像認識手法には不向きである。

後者は、画像認識に大津、栗田 [3] が提案した高次局所自己相関特徴を用いている。高次局所自己相関特徴は、同背景内で約 100% の識別率を収め、容易なアルゴリズムで構成されているため、高速実時間内処理が可能である。しかし、異なる背景での認識は、高次

局所自己相関特徴が繊細に画像の特徴を捉えるため、識別率はかなり低下してしまう。そこで栗田は複数台のカメラより得られるステレオ動画像から動きに関する情報を抽出した PARCOR 画像を用いて背景と対象物の領域を分離し、抽出した領域に対して高次局所自己相関特徴を求めた。この方法で栗田は、異なる背景下で識別率 100%を実現した。

栗田が提案した手法は、本研究で用いる画像認識手法に最適であると思われる。しかし、PARCOR 画像の生成に少々問題が生じる。PARCOR 画像を生成するには複数台のカメラの同期を取りステレオ動画像を作成する必要がある。ステレオ動画像の生成を高速実時間内で処理するためには、特別なハードウェア環境が必要となる。本研究では容易に組める高速画像認識モジュールの開発を目指しているため、経済面も考慮に入れなければならない。そのため、背景と対象物体の領域を分離する方法に PARCOR 画像以外のものを用いなければならない。

そこで、背景と顔の領域を分離する方法に、肌色領域の抽出を行なった。次節で詳しく述べるが、人間の肌色には特徴があるため、色情報によりある程度の領域を抽出することが可能である。移動中のロボットに搭載されたカメラに写る肌色領域は、顔だけとは限らない。人間の肌色の部分を考えると、顔の他に手が写り込む可能性は大きい。そこで、本研究の画像認識は、顔だけではなく、手の平を識別する。手の平といっても様々な形状があるので、グー、チョキ、パーと個人の顔を識別することを目的とする。

ここで図 2.1 に本研究で構築する高速画像認識モジュールの画像処理手順を示す。高速画像認識を目標としているので、すべての処理は簡潔で精度の良い方法を用いなければならない。本研究では画像全体に対して一回弱の処理を施すだけで、対象領域を抽出することが可能である。次節より肌色領域の抽出方法、領域のラベリング、高次局所自己相関特徴、線形判別分析の順に詳しく説明する。

## 2.3 肌色領域の抽出

本研究での肌色領域抽出は、表色系 YUV, YIQ, RGB のしきい値法による抽出と、エッジ画像を用いた肌色領域の冗長の 2 つの処理で構成されている。次節からこの 2 つの処理について述べる。

### 2.3.1 しきい値法による肌色領域抽出

人間の肌色領域は、しきい値法によりある程度抽出できる [4]。肌色は、表色系 YUV や YIQ の U と I に特徴を持ち、人種によりある程度同じしきい値を持っていることが知られている。本研究では、データ数の関係から黄色人種の肌色を抽出するしきい値を手動で求めた。手動で求めた直感的なしきい値の評価は、同時に濃淡画像の最小 2 乗近似の意味で最適な 2 値化を与えることが証明されている [5]。

本研究では、肌色のしきい値を選定するため、被験者 10 人、総画像数 1000 枚のデータを採取した。各画像について肌色のみを抽出したしきい値を表色系 RGB, YUV, YIQ につい

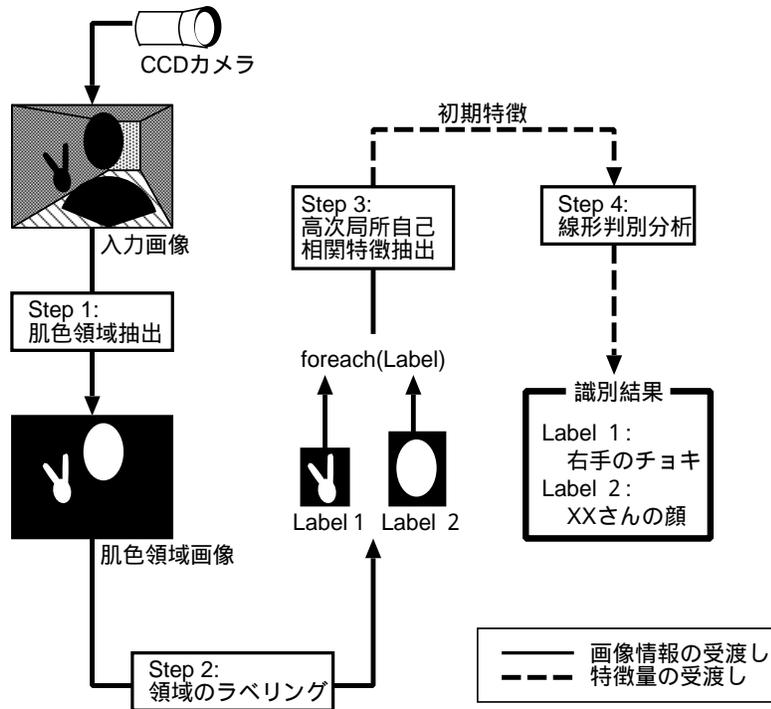


図 2.1: 高速画像認識モジュールの構成

て求めた．表 2.1 に各表色系の内容示す．また，RGB から YUV，YIQ への変換に以下のものを用いた．

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.256 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & 0.322 \\ 0.211 & -0.522 & -0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

選定したしきい値を表 2.2 に示す．本研究では，各表色系のしきい値を満たす画素を肌色領域とした．肌色領域の抽出結果を図 2.2 に示す．

表色系	内容	
RGB	R	赤の輝度
	G	緑の輝度
	B	青の輝度
YUV	Y	輝度
	U	青み成分
	V	赤み成分
YIQ	Y	輝度
	I	肌色を含む, オレンジ~シアン
	Q	I以外の色調

表 2.1: 表色系 RGB, YUV, YIQ の表記内容

表色系	信号	しきい値
RGB	R	[80,248]
	G	[64,224]
	B	[56,187]
YUV	Y	[64,240]
	U	[94,125]
	V	[131,255]
YIQ	Y	[50,240]
	I	[128,166]
	Q	[46,240]

表 2.2: 選定したしきい値

### 2.3.2 肌色領域の冗長

しきい値法を利用して抽出した領域には、画素単位のノイズや、対象領域の欠落が多く見られる。そこで、エッジ画像を用いた領域の冗長を行い、ノイズの削除のため平滑化を行う。

エッジ画像は、YUV の輝度を示す Y をもとに表 2.3 に示す 2 次元エッジ検出オペレータを用いて検出し、ラスタ追跡 [6] にてエッジの追跡・連結を行なったものを使用した。

-1	-1	-1
-1	8	-1
-1	-1	-1

表 2.3: 2次元エッジ検出オペレータ

以下に領域冗長の手順を示す。

- 手順1 現在の画素  $p(x, y)$  がエッジならば手順2へ、そうでなければ手順4へ移る。
- 手順2  $p(x, y)$  を中心に  $(-5, -5)$  から  $(5, 5)$  の 25 画素について、しきい値法で求めた領域である画素をカウントする。カウント数を  $count$  とおく。
- 手順3  $count > 3$  のとき、 $p(x, y)$  にフラグをたてる。
- 手順4 上記の操作を画像の左上から右下に向かってすべての画素について行なう。

フラグの立っている画素を冗長画素とする。最後に全画素について平滑化を行ないノイズを除去する。図 2.3 に肌色領域の冗長結果を示す。

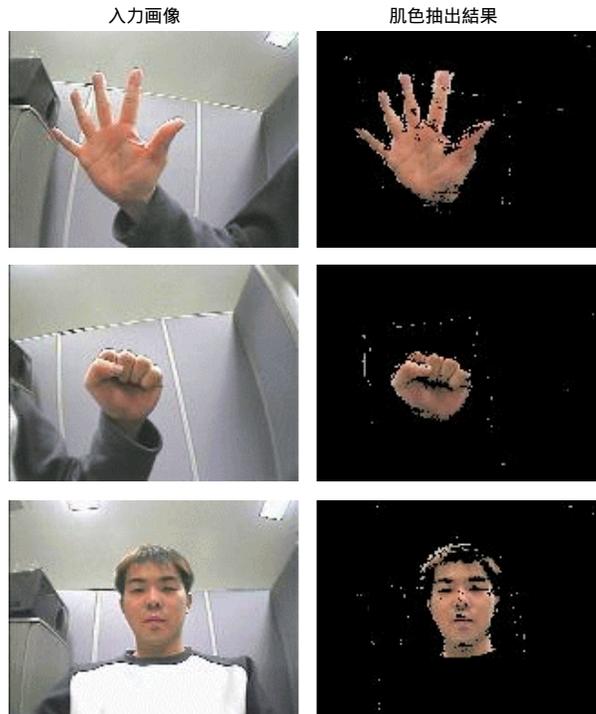


図 2.2: しきい値法による肌色領域の抽出

## 2.4 領域のラベリング

入力画像内に肌色の領域が1つだとは限らない．そこで，認識に入る前に，複数の肌色領域を分類する必要がある．本研究では，分類方法にラベリング手法 [7] を用いた．ラベリングとは，連結している領域ごとにラベル (番号) を与える処理を示す．本研究では，連結の肯否に，8 連結 (8 近傍) を調べる手法をとった．

肌色領域として抽出された画素を  $-1$ ，背景部分は  $0$  で初期化されているとする．ラベリングは，図 2.4 の太枠でできたマスクを画像の左上から右下に走査させていく．今，画素  $A$  を検査しているとする． $C, D, E$  はすでにラベル付けされていることになる．以下に図 2.4 を参考にして本研究で用いたラベリング手法を示す．

- 1  $A$  が  $0$  であれば，3 へ進む．
- 2.1  $A$  が  $-1$  であり，かつ  $B, C, D, E$  が  $0$  であれば， $A$  に新しいラベルを付与して 3 へ進む．
- 2.2  $A$  が  $-1$  であり，かつ  $B, C, D, E$  のいずれかにラベルが振られていれば， $A$  はその割り振られているラベルを付与される．
- 2.3  $A$  が  $-1$  であり，かつ  $B, C, D, E$  にことなるラベルが付与されていれば，最小のラベルを  $A$  に付与し，最小ラベル以外のラベルが付与されているものに，最小ラベルと同じ領域であることを記憶させる．
- 3 上記の操作を画像の左上から右下に向かってすべての画素について行なう．

図 2.5 にラベリング手法で分類された肌色の領域 (赤・緑の枠) を示す．

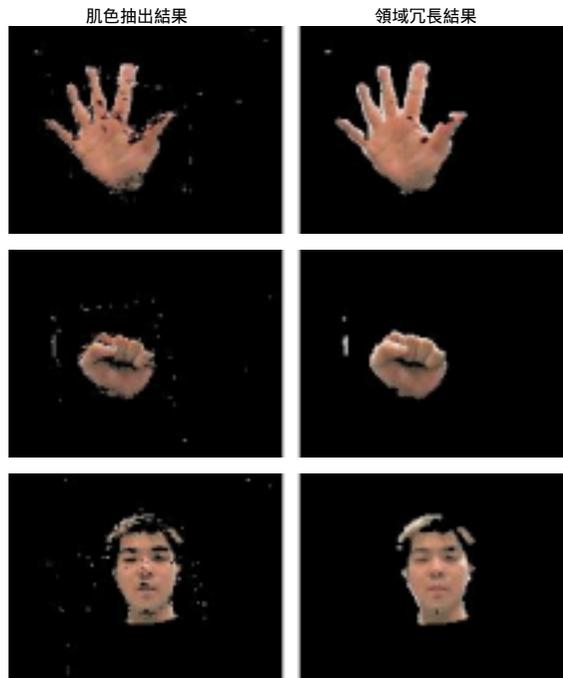


図 2.3: 肌色領域の冗長結果

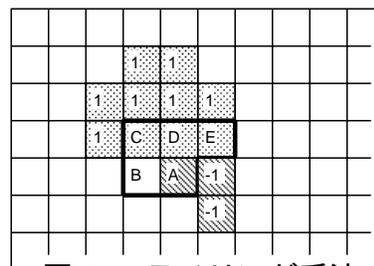


図 2.4: ラベリング手法

以上が本研究での認識対象領域の抽出方法である。

## 2.5 高次局所自己相関特徴

抽出した肌色領域を識別する方法に，大津，栗田が提案した高次局所自己相関特徴を用いる．この節では，まず自己相関関数について述べ，次に高次局所自己相関特徴について述べた後，その有用性について述べる．

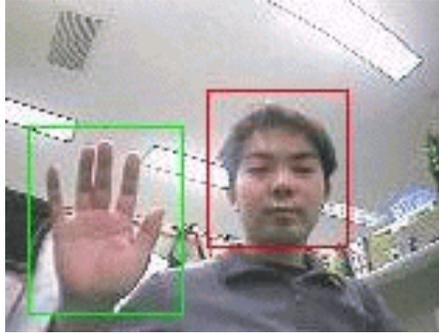


図 2.5: ラベリング手法による肌色領域の分類

### 2.5.1 自己相関関数

平均が 0 の定常時系列データ  $x(t)|t = 0, \dots, N - 1$  が与えられた場合，自己相関関数は，

$$r(l) = \frac{1}{N} \sum_{t=0}^{N-l-1} x(t+l)x(t) \quad (2.1)$$

で定義される．自己相関関数のフーリエ変換は，ピリオドグラム (パワースペクトル密度関数) と呼ばれ，

$$p(f) = \sum_{t=-(N-1)}^{N-1} \exp(-i2\pi fl)r(l) \quad (2.2)$$

のように定義され，時系列信号のフーリエ変換  $X(f)$  と

$$p(f) = \frac{1}{N}|X(f)|^2 \quad (2.3)$$

のような関係が成り立つ．

### 2.5.2 高次局所自己相関特徴

自己相関関数はパワースペクトルと密接な関係があり，定常時系列データの有効な特徴の一つである．大津らは，画像の認識や計測のために有効で基本的な画像特徴として，自己相関特徴を高次へ拡張した高次局所自己相関関数を提案した．

参照点  $r$  での対象画像の輝度を  $I(r)$  とすると， $N$  次自己相関関数は，変位方向  $(a_1, a_2, \dots, a_N)$  に対して，

$$x^N(a_1, a_2, \dots, a_N) = \int I(r)I(r+a_1)\cdots I(r+a_N)dr \quad (2.4)$$

で定義される．従って，高次自己相関関数は，次数や変位方向  $(a_1, \dots, a_N)$  の取り方により，無数に考えられる．しかし，画面に関する加法性の条件を満たすためには，変位方向は参照点  $r$  の局所領域に限定されなければならない．

ここでは、簡単のため高次自己相関係数の次数  $N$  を高々2までとする。また変位方向を参照点  $r$  の周りの局所的な  $3 \times 3$  画素の領域に限定する (図 2.6 参照)。

この場合には、平行移動により等価な特徴を除くと、特徴の数は全部で 25 個になる (図 2.7 参照)。

各特徴の計算は、局所パターンに対応する画素の輝度の積を全画像に対して足し合わせれば良い。こうして計算された特徴は、明らかに、対象の位置に関して不変でしかも画面に対する加法性を満たす。

画像から抽出される高次局所自己相関特徴は、次のような性質を持っている。

- C1 位置に関する不変性: 対象が画像枠内の何処にあっても特徴値はかわらない。
- C2 画面に関する加法性: 画像枠内に複数の対象があれば、全体の特徴値はそれぞれの対象の特徴値の和になる。
- C3 学習による適応性: 例からの学習により適応的に種々の目的に利用できること。
- C4 ノイズに対してある程度、ロバストである。

C3 はシステムの汎用性を確保するものである。

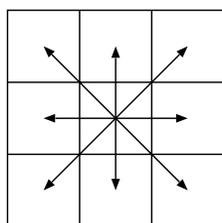


図 2.6:  $3 \times 3$  の参照領域

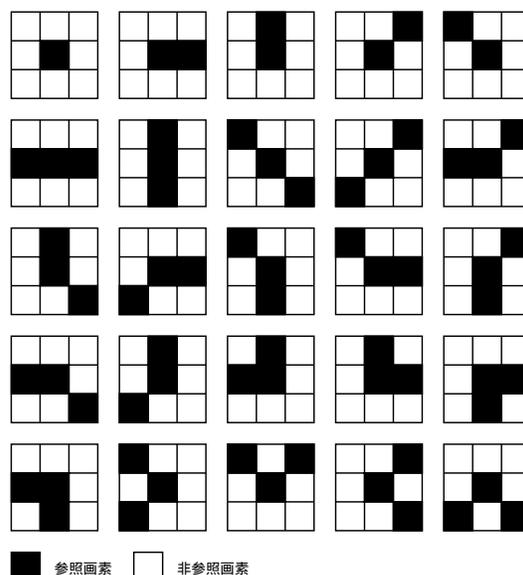


図 2.7: 2 次までの局所パターン

## 2.6 線形判別分析を用いた識別

入力画像より抽出した個々の高次局所自己相関特徴は、認識課題に依存しない一般的で基本的な特徴であり、全体として与えられた認識課題を達成するために必要な情報がある程度抽出していると考えられる。そこで、与えられた認識課題に対して有効な新しい特徴

を，高次局所自己相関特徴を組み合わせて構成する．ここでは，線形判別分析を利用する．新しい特徴  $y = (y_1, \dots, y_M)^T$  は，初期特徴 (高次局所自己相関特徴)  $x = (x_1, \dots, x_M)^T$  の線形写像

$$y = A^T x + b \quad (2.5)$$

によって抽出する．ここで， $A = [a_{ij}]$  および  $b = (b_1, \dots, b_M)^T$  は，それぞれ，線形結合係数行列および定数である．また， $M$  は初期特徴の個数 (ここでは 25 個) を表す．

$K$  クラスのサンプル集合  $\{C_k\}_{k=1}^K$  が与えられると，各クラスの平均，共分散行列，総平均は，平均操作 ( $\sum_{i=1}^L$ ) を  $E_L$  で表すと，それぞれ，

$$\bar{x}_k = E_{C_k} x, \quad (2.6)$$

$$\Sigma_k = E_{C_k} (x - \bar{x}_k)(x - \bar{x}_k)^T, \quad (2.7)$$

$$\bar{x}_T = E_C x = \sum_{k=1}^K \omega_k \bar{x}_k \quad (2.8)$$

となる．

このときの初期特徴のクラス内共分散行列およびクラス間共分散行列は，それぞれ，

$$\Sigma_W = \sum_{k=1}^K \omega_k \Sigma_k, \quad (2.9)$$

$$\Sigma_B = \sum_{k=1}^K \omega_k (\bar{x}_k - \bar{x}_T)(\bar{x}_k - \bar{x}_T)^T \quad (2.10)$$

となる．ここで  $\omega_k$  は，クラス  $C_k$  の先験確率 (通常  $1/K$  とする) である．

判別空間での新特徴  $y$  の各クラスの平均，共分散行列，総平均は，それぞれ，

$$\bar{y}_k = E_{C_k} y = A^T \bar{x}_k, \quad (2.11)$$

$$\hat{\Sigma}_k = E_{C_k} (y - \bar{y}_k)(y - \bar{y}_k)^T = A^T \Sigma_k A, \quad (2.12)$$

$$\bar{y}_T = E_C y = A^T \bar{x}_T \quad (2.13)$$

のように表される．従って，新特徴でのクラス内分散，クラス間分散は，

$$\hat{\Sigma}_W = \sum_{k=1}^K \omega_k \hat{\Sigma}_k = A^T \Sigma_W A, \quad (2.14)$$

$$\hat{\Sigma}_B = \sum_{k=1}^K \omega_k (\bar{y}_k - \bar{y}_T)(\bar{y}_k - \bar{y}_T)^T = A^T \Sigma_B A \quad (2.15)$$

となる．

係数行列  $A$  を決定するための判別空間の良さ (判別力) の評価には，判別基準  $J = \text{tr}(\hat{\Sigma}_W^{-1} \hat{\Sigma}_B)$  が使われる．判別基準  $J$  を最大とする最適な係数行列  $A$  は，Lagrange 乗数

行列を  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$  として,

$$Q(A) = \text{tr}(\hat{\Sigma}_B) - \text{tr}[(\hat{\Sigma}_W - I)\Lambda], \quad (2.16)$$

$$= \text{tr}(A^T \Sigma_B A) - \text{tr}[(A^T \Sigma_W A - I)\Lambda] \quad (2.17)$$

を最大とすることにより求められる．これを  $A$  で偏微分して 0 とおくと，固有値問題

$$\Sigma_B A = \Sigma A \Lambda, \quad (2.18)$$

$$A^T \Sigma_W A = I \quad (2.19)$$

が得られる．これを解くことにより，最適な係数行列  $A$  が求まる．ここで， $\Lambda = \text{diag}(\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N > 0)$  は，固有値を要素とする対角行列であり， $I$  は単位行列である．ここで，判別空間の次元  $N$  は，高々  $\min(K - 1, M)$  で抑えられる．

未知画像が与えられた場合の識別は，その画像から高次局所自己相関特徴  $x$  を求め，判別空間での値  $y$  を計算し，それに最も近い各クラスの平均ベクトル  $\bar{y}_k$  を持つクラスに決定する．ここで， $y$  と  $\bar{y}_k$  との距離  $|\bar{y}_k - y|^2$  は，多クラスの分布間の平均マハラノビス汎距離  $(x - \bar{x}_k)^T \Sigma_W^{-1} (x - \bar{x}_k)$  と密接に関係していることが知られている [8]．つまり，判別分析の場合には，平均クラス内分散の逆  $\Sigma_k^{-1}$  で重み付けたベクトル間の距離を近似的に計算していることに対応する．

以上が本研究で用いる画像認識手法である．

## 第3章 ロボットへ搭載する方法

本節では、高速画像認識モジュールを自律走行するロボットへ搭載するための方法と高速画像認識アプリケーションを作成した。以下これについて詳しく述べる。

### 3.1 高速画像認識モジュールの搭載方法

画像認識モジュールを既存のロボットへ搭載するには、ロボット内部へモジュールを組み込むことが一番良い方法である。しかし、すべてのロボットが画像処理装置を搭載できるスペースや機能を保持しているとは限らない。

そこで本研究では、画像認識装置が搭載されていないロボットを対象とし、さらに、小型 CCD カメラを載せることが可能な大きさの AIBO や Khepera<sup>1</sup>などの小型ロボットやノートパソコンが搭載可能な中型ロボットを対象を絞る。

AIBO や Khepera には、小型 CCD カメラを搭載するのが精一杯である。よって、AIBO や Khepera に画像認識を行なわせる場合、外部パソコンに画像処理・認識を委ねなければならない。十分な画像処理装置を搭載していない中型ロボットについても同様である。本研究では、以下の二つの条件を満たすロボットについて図 3.1 のようなシステムを構築した。

- ロボット本体が何らかの通信機能を持っている。
- 小型 CCD カメラを搭載できる。(有線の場合は、ノートパソコンが搭載可能であるのが望ましい)

CCD カメラの有線、無線の選択やパソコンの種類(固定 PC または、ノートパソコン)の選択によって、画像認識モジュールを搭載するロボットに表 3.1 のような行動制限が生じる。

自律走行型ロボットは、走行可能な場所の制限が少ないものが理想的である。よって、表 3.1 が示すように、

- 無線 CCD カメラ & ノートパソコン
- 有線 CCD カメラ & ノートパソコン(ノートパソコン搭載可能ロボットの場合)

---

<sup>1</sup>スイス・ローザンヌ連邦工科大学 (EPFL) のマイクロコンピュータ・インタフェース研究所 (lami) で開発された研究開発用小型実験ロボットシステム

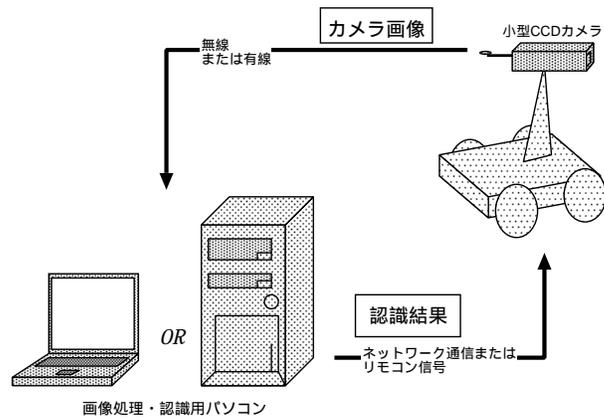


図 3.1: 画像処理システムの構造

	固定 PC	ノートパソコン
有線 CCD カメラ	Kepera などの卓上で動くロボット	中型ロボット
無線 CCD カメラ	無線が届く範囲での行動	ロボットが走行可能な場所

表 3.1: カメラ, パソコンの選択によるロボットの行動制限

- 無線 CCD カメラ & 固定 PC

の組合せが良い。そこで、本研究では、画像処理・認識を行なうパソコンに、ロボットの行動範囲制限が少なくなるようにノートパソコンを用いた。次節に構築した実時間画像認識アプリケーションについて詳しく説明する。

### 3.2 実時間画像認識アプリケーション”Smart-EYE”

構築する実時間画像認識アプリケーションは、容易に組み、また経済的に優位なものではない。そのため、使用 OS に Linux を選択した。

本研究で構築した実時間画像認識アプリケーションを”Smart-EYE”と名付けた。以下 Smart-EYE の仕様を示し、本節のそれ以降のパラグラフで、使用 OS, 画像取り込み, 画像表示について説明する。

- 外部仕様

- Video Capture adapter PC Card REX-9590(RATOC System Inc.) .
- カメラ
  - \* GAIA-2 搭載 CCD カラーカメラ (CHANNEL VISION 6005-B) .
  - \* デジタルビデオカメラ (Sony Digital Handycam DCR-TRV10)

- ノートパソコン Celeron 400MHz Memory 196MB .
- 内部仕様
  - OS:RTLinux-2.2(base: VineLinux 2.1.5 Kernel2.2.18) .
  - ISCC ドライバ+Video4Linux API による画像入力 (解像度 160x120) .
  - C 言語, Gtk+/GDK による画像表示 [9] .
- 主な機能
  - 第 2 章で説明した画像認識手法により, 肌色領域の識別を行なう .
  - 画面内に検出領域を表示する .
  - 識別結果を RS-232C, または TCP/IP ネットワーク通信を用いて送信する .
  - 識別結果を terminal へ出力 .

使用 OS について 高速実時間処理を実現するため, OS に RTLinux<sup>2</sup>を用いた . RTLinux には以下の利点があり, ロボットの制御など行なう際に適した OS である .

- $10^{-9}$ second(1ns) の時間制約が可能となり, 処理開始タイミングと処理終了までの時間制限を保証される .
- Pthread<sup>3</sup>により, リアルタイム処理のプログラムをモジュールの形で作成できる .
- 共有メモリ (mbuff) により, モジュール, プログラム間のデータの受渡しが容易になる .
- RS-232C(serial port) などのハードウェアデバイスドライバがリアルタイムモジュールにより容易に組める .

画像の取り込みについて Linux は ISCC<sup>4</sup>ドライバと Video4Linux<sup>5</sup>を用いて, ビデオ画像をメモリにマップすることが可能である . 本研究で用いた Video Capture card は IBM の Smart Capture Card 互換であるので, ISCC ドライバにより Linux でも使用できる . 高速画像認識を実現するためには, 画像取り込み速度が十分速くなければならない . 上記仕様での画像取り込み速度は, 最大解像度  $320 \times 240$  で約 12 fps, 解像度  $160 \times 120$  で約 20 fps であった . 取り込みに時間をかけることが出来ないので, 本研究では, 解像度  $160 \times 120$  を用いる .

---

<sup>2</sup>標準 Linux にハードウェアリアルタイム機能を拡張機能として付加した Linux . Mexico Institute of Technology で, Dr.Victor Yodaiken や Michael Barabanov らにより開発され, 現在 Finite State Machine Labs, Inc で管理されている .

<sup>3</sup>POSIX 規格のマルチスレッド

<sup>4</sup>IBM Smart Capture Card の略

<sup>5</sup>Video for Linux . URL <http://roadrunner.swansea.uk.linux.org/v4l>

画像表示について 識別した対象物や入力画像などをリアルタイムで確認できるように，画像処理ライブラリ Imlib と GUI ツールキットである GTK+ を利用して，X Window System 上に処理画像を表示した．Imlib や GTK+ の説明は，本稿の主旨から外れるので，文献 [9] を参考にされたい．

### 3.2.1 Smart-EYE のシステム構成

Smart-EYE は、Video4Linux による入力画像の取り込み、GTK+ と Imlib による画像表示、TCP/IP ネットワーク通信による識別結果の出力を行なうユーザプログラムと複数のリアルタイムモジュールによって構成されている。図 3.2 に高速画像認識アプリケーションのシステム構成を示す。

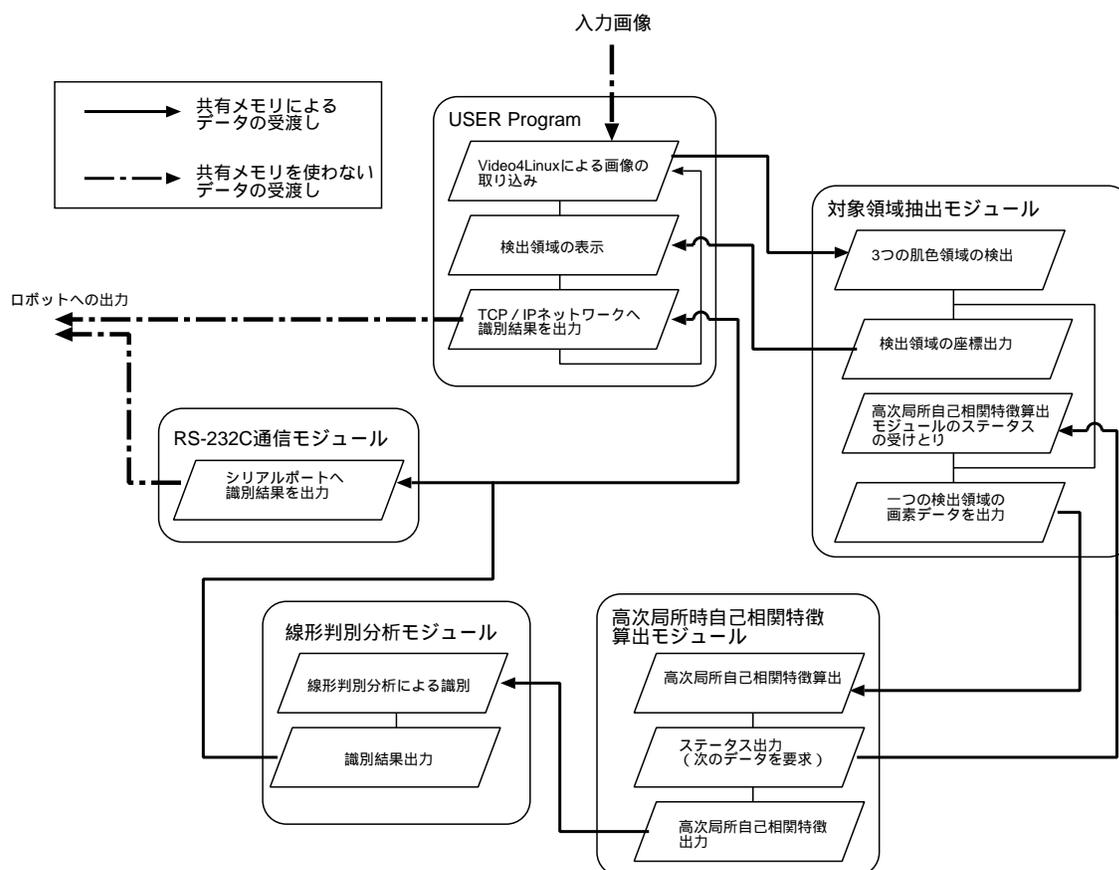


図 3.2: Smart-EYE のシステム構成

線形判別分析で用いる判別空間と代表ベクトルは、静止画を用いてあらかじめ作成しておく必要がある。

# 第4章 実験1:静止画を用いた顔・手の識別

## 4.1 実験目的

本実験の目的は、本研究で用いた画像認識手法の識別率を静止画像を用いて評価することと、対象領域の画像の種類(輝度画像, エッジ画像など)によって識別率がどのように変化するか調べ、考察することである。

## 4.2 実験内容

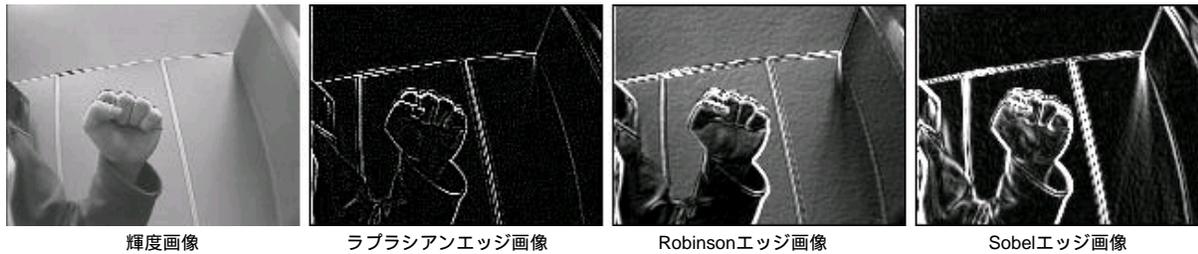
本実験では、識別率の評価に leave-one-out 法を用いる。leave-one-out 法は、全サンプルの中から1サンプルを取り除き未知画像し、残ったサンプルから構成された判別空間を用いて、未知画像を識別する方法である。

画像サンプルに、異なる場所(3箇所)で撮影した、解像度  $160 \times 120$ 、1枚の画像に対して識別対象が1つ撮影されているものを画像総数 1 000 枚(10クラス)用いる。対象領域の画像の種類に以下のものを選び、それぞれについて leave-one-out 法を用いて識別率を求める。

- 画像の種類(図 4.1 参照)
  1. 輝度画像(YUVのY)
  2. ラプラシアンエッジ検出オペレータによるエッジ画像
  3. Robinson エッジ検出オペレータによるエッジ画像
  4. Sobel エッジ検出オペレータによるエッジ画像

- サンプル画像内容(図 4.2 参照)

class 1	右手のグー	class 2	左手のグー
class 3	右手のパー	class 4	左手のパー
class 5	右手のチョキ	class 6	左手のチョキ
class 7	顔 A	class 8	顔 B
class 9	顔 C	class 10	顔 D
			各 100 枚



輝度画像

ラプラシアンエッジ画像

Robinsonエッジ画像

Sobelエッジ画像

図 4.1: 画像の種類



クラス 1 右手のゲー

クラス 2 左手のゲー

クラス 3 右手のパー

クラス 4 左手のパー

クラス 5 右手のチョキ

クラス 6 左手のチョキ

クラス 7 顔 A

クラス 8 顔 B

クラス 9 顔 C

クラス 10 顔 D

図 4.2: 各クラスの静止画像サンプル

### 4.3 実験結果

表 4.1 に各画像の種類を用いたときの識別率を示し，表 4.2-4.5 により詳細な識別率を示す．

画像の種類	認識数	誤認識数	識別率
輝度画像	574	426	57.4 %
エッジ画像 (ラプラシアン)	930	70	93.0 %
エッジ画像 (Robinson)	747	253	74.7 %
エッジ画像 (Sobel)	687	313	68.7 %

表 4.1: 各画像種類の識別率

異なる背景からの識別率は，最大で 93% となった．また，本実験の結果，ラプラシアンのエッジ画像，Robinson のエッジ画像，Sobel のエッジ画像，輝度画像の順に識別率が高いことが得られた．

### 4.4 実験結果の考察

異なる背景での識別率が 93% というのは，本研究で目指す画像認識システムにおいて，十分な結果である．

ここでは，栗田らが輝度画像を用いて 100% 近くの識別率をあげているのに対して，本研究では，輝度画像を用いた際に，57.4% の識別率しか得られなかった理由について述べる．

輝度画像の識別率が低い理由に，肌色領域の抽出方法が密接に関係していると考えられる．本研究では，肌色領域の抽出方法に，精度よりも処理速度を優先して考案されている．そのため，図 2.3 の右図に見られるように，領域の過膨張や過短縮が生じる．冗長が不完全であるため，情報の欠落やノイズが生じ，情報量が多い輝度画像を用いた識別が低下してしまう．エッジ画像を用いた識別は，冗長失敗によるノイズや情報の欠落にロバストであるゆえ，高い識別率が得られたと考えられる．

未知画像 (右手のゲー)

識別したもの	識別率
顔 C	51.0 %
ゲー (LEFT)	25.0 %
ゲー (RIGHT)	13.0 %
顔 B	11.0 %

未知画像 (左手のゲー)

識別したもの	識別率
ゲー (LEFT)	69.0 %
ゲー (RIGHT)	23.0 %
顔 C	5.0 %
顔 B	3.0 %

未知画像 (右手のチョコキ)

識別したもの	識別率
チョコキ (RIGHT)	91.0 %
チョコキ (LEFT)	9.0 %

未知画像 (左手のチョコキ)

識別したもの	識別率
チョコキ (LEFT)	70.0 %
チョコキ (RIGHT)	30.0 %

未知画像 (右手のパー)

識別したもの	識別率
パー (RIGHT)	70.0 %
顔 B	25.0 %
パー (LEFT)	4.0 %
顔 A	1.0 %

未知画像 (左手のパー)

識別したもの	識別率
パー (LEFT)	83.0 %
パー (RIGHT)	17.0 %

未知画像 (顔 A)

識別したもの	識別率
顔 B	49.0 %
パー (LEFT)	18.0 %
パー (RIGHT)	15.0 %
顔 A	8.0 %
顔 D	7.0 %
ゲー (LEFT)	3.0 %

未知画像 (顔 B)

識別したもの	識別率
顔 B	55.0 %
顔 A	25.0 %
顔 D	20.0 %

未知画像 (顔 C)

識別したもの	識別率
顔 C	100.0 %

未知画像 (顔 D)

識別したもの	識別率
顔 A	46.0 %
顔 B	39.0 %
顔 D	15.0 %

表 4.2: 輝度画像を用いた識別率の詳細

未知画像 (右手のグー)

識別したもの	識別率
グー (RIGHT)	85.0 %
兵藤氏の顔	10.0 %
近藤氏の顔	4.0 %
自分の顔	1.0 %

未知画像 (左手のグー)

識別したもの	識別率
グー (LEFT)	94.0 %
グー (RIGHT)	3.0 %
チョキ (LEFT)	1.0 %
パー (LEFT)	1.0 %
近藤氏の顔	1.0 %

未知画像 (右手のチョキ)

識別したもの	識別率
チョキ (RIGHT)	93.0 %
パー (RIGHT)	6.0 %
自分の顔	1.0 %

未知画像 (左手のチョキ)

識別したもの	識別率
チョキ (LEFT)	94.0 %
グー (LEFT)	5.0 %
パー (LEFT)	1.0 %

未知画像 (右手のパー)

識別したもの	識別率
パー (RIGHT)	87.0 %
自分の顔	12.0 %
グー (RIGHT)	1.0 %

未知画像 (左手のパー)

識別したもの	識別率
パー (LEFT)	88.0 %
チョキ (LEFT)	10.0 %
兵藤氏の顔	2.0 %

未知画像 (顔 A)

識別したもの	識別率
自分の顔	89.0 %
近藤氏の顔	6.0 %
パー (RIGHT)	5.0 %

未知画像 (顔 B)

識別したもの	識別率
兵藤氏の顔	100.0 %

未知画像 (顔 C)

識別したもの	識別率
近藤氏の顔	100.0 %

未知画像 (顔 D)

識別したもの	識別率
松久保氏の顔	100.0 %

表 4.3: エッジ画像 (ラブラシアン) を用いた識別率の詳細

未知画像 (右手のゲー)

識別したもの	識別率
ゲー (RIGHT)	43.0 %
ゲー (LEFT)	40.0 %
自分の顔	9.0 %
パー (LEFT)	7.0 %
兵藤氏の顔	1.0 %

未知画像 (左手のゲー)

識別したもの	識別率
ゲー (LEFT)	89.0 %
ゲー (RIGHT)	11.0 %

未知画像 (右手のチョコキ)

識別したもの	識別率
チョコキ (RIGHT)	75.0 %
自分の顔	12.0 %
パー (RIGHT)	8.0 %
兵藤氏の顔	2.0 %
松久保氏の顔	2.0 %
パー (LEFT)	1.0 %

未知画像 (左手のチョコキ)

識別したもの	識別率
チョコキ (LEFT)	61.0 %
パー (LEFT)	28.0 %
松久保氏の顔	6.0 %
ゲー (RIGHT)	1.0 %
自分の顔	2.0 %
兵藤氏の顔	2.0 %

未知画像 (右手のパー)

識別したもの	識別率
パー (RIGHT)	100.0 %

未知画像 (左手のパー)

識別したもの	識別率
パー (LEFT)	80.0 %
自分の顔	14.0 %
ゲー (RIGHT)	6.0 %

未知画像 (顔 A)

識別したもの	識別率
兵藤氏の顔	35.0 %
パー (LEFT)	28.0 %
自分の顔	26.0 %
ゲー (RIGHT)	8.0 %
チョコキ (LEFT)	2.0 %
ゲー (LEFT)	1.0 %

未知画像 (顔 B)

識別したもの	識別率
兵藤氏の顔	73.0 %
自分の顔	21.0 %
チョコキ (LEFT)	6.0 %

未知画像 (顔 C)

識別したもの	識別率
近藤氏の顔	100.0 %

未知画像 (顔 D)

識別したもの	識別率
松久保氏の顔	100.0 %

表 4.4: エッジ画像 (Robinson) を用いた識別率の詳細

未知画像 (右手のグー)

識別したもの	識別率
グー (RIGHT)	43.0 %
グー (LEFT)	26.0 %
顔 A	25.0 %
顔 C	6.0 %

未知画像 (左手のグー)

識別したもの	識別率
グー (LEFT)	56.0 %
グー (RIGHT)	36.0 %
顔 A	8.0 %

未知画像 (右手のチョキ)

識別したもの	識別率
チョキ (RIGHT)	83.0 %
パー (RIGHT)	10.0 %
顔 B	7.0 %

未知画像 (左手のチョキ)

識別したもの	識別率
チョキ (LEFT)	48.0 %
パー (RIGHT)	37.0 %
パー (LEFT)	7.0 %
グー (RIGHT)	5.0 %
顔 A	3.0 %

未知画像 (右手のパー)

識別したもの	識別率
パー (RIGHT)	78.0 %
チョキ (LEFT)	15.0 %
顔 B	7.0 %

未知画像 (左手のパー)

識別したもの	識別率
パー (LEFT)	76.0 %
顔 A	15.0 %
チョキ (LEFT)	6.0 %
グー (RIGHT)	3.0 %

未知画像 (顔 A)

識別したもの	識別率
顔 A	27.0 %
パー (LEFT)	21.0 %
グー (RIGHT)	18.0 %
チョキ (LEFT)	18.0 %
グー (LEFT)	15.0 %
顔 C	1.0 %

未知画像 (顔 B)

識別したもの	識別率
顔 B	77.0 %
顔 D	18.0 %
チョキ (RIGHT)	5.0 %

未知画像 (顔 C)

識別したもの	識別率
顔 C	100.0 %

未知画像 (顔 D)

識別したもの	識別率
顔 D	99.0 %
顔 B	1.0 %

表 4.5: エッジ画像 (Sobel) を用いた識別率の詳細

# 第5章 実験2:ビデオ映像による顔の識別

## 5.1 実験目的

本実験の目的は、第2章で説明した高速画像認識手法の実時間処理速度、識別率をビデオ映像を用いて計測し、考察することである。

## 5.2 実験内容

本実験の実験方法を以下に示す。

1. JAIST 校内で被験者 3 人をデジタルビデオカメラ (Sony Digital Handycam DCR-TRV10) で撮影する。
2. 撮影したビデオ映像から、各被験者の顔が出現した時間を記録し、タイムテーブルを作成する。
3. 撮影したビデオ映像を実時間画像認識アプリケーション Smart-EYE へ入力し、各被験者の顔が識別された時間を記録してタイムテーブルを作成する。
4. 2, 3. で作成したタイムテーブルを比較する。

サンプルビデオ映像の内容を以下に示す。

- I. 被験者とカメラの距離を一定に保ち、異なる背景で撮影したもの  
撮影場所。
  - ロボット実験室内。
  - 本棚の前。
  - セミナー室。
- II. 同じ背景内でカメラから 1 ~ 2m の間を移動する 3 人の被験者。

図 5.1 に線形判別空間を作成する際に用いたサンプル画像を示し、図 5.2 にビデオ映像のサンプルを示す。



被験者 A                      被験者 B                      被験者 C

図 5.1: 線形判別空間を作成する際に用いたサンプル画像



距離一定・ロボット実験室



距離一定・本棚の前



距離一定・セミナー室



距離変動・背景同じ

図 5.2: ビデオ映像のサンプル

## 5.3 実験結果

表 5.1, 5.2 に実験結果を示す .

撮影場所	識別率 (%)	処理速度 (frames/s)
ロボット実験室	92.4	19
本棚の前	19.2	5
セミナー室	84.0	16

表 5.1: 距離一定・異なる背景での識別率と処理速度

対象者	識別率 (%)
被験者 A	16.3
被験者 B	19.2
被験者 C	22.4

表 5.2: 距離変動・同じ背景での識別率

距離変動・同じ背景での識別において、平均速度は、21 frames/s である .

## 5.4 実験結果の考察

線形判別空間を作成したサンプル画像と同じ背景での認識は、かなり高くなっている . また、セミナー室で撮影したビデオ映像を用いたときも、良い結果を得た . しかし、本棚の前では、非常に悪い結果となった .

本棚の前で、識別率が落ちた原因に、肌色領域の抽出の失敗が挙げられる . 今回撮影した本棚は、肌色のしきい値範囲に近い色で構成されているため、本棚全体を肌色領域として抽出してしまう . よって、識別率が低くなる . また、本棚の領域が、入力画像の大半を占めているため、高次局所自己相関特徴の抽出速度が遅くなる .

また、距離変動に非常に弱いことが実験結果に見られ、この対処法を考案することが今後の課題の一つとなる .

# 第6章 高速画像認識モジュールを搭載したロボット

本章では，高速画像認識モジュールを実際に搭載した自律走行型ロボットについて述べる．

## 6.1 実装した技術について

本研究では，以下の要素技術を実装して，特定の人間から逃げ回るロボットシステムを作成した．

- 高速画像認識モジュールによる個人の識別．
- Sonar による障害物回避．
- 対象者との距離を一定に保つ．

本研究で構築した画像認識モジュールは，ある距離に居る人物を識別することが可能であり，その画像内の座標を得ることが出来る．この性質を利用して，特定の人を識別できたら，その人の名を呼び，その人物との距離を一定に保つロボットシステムを構築した．以下に，構築したロボットのプラットフォームと画像認識モジュールを組み込んだ GAIA-2 について説明し，障害物回避，対象物との距離を保つ方法について述べる．

## 6.2 プラットフォーム

まず、本研究で用いる自律走行型ロボット GAIA-2 の仕様と製作会社があげた特徴を説明する。

- 仕様

外観	GAIA-2 鳥瞰図 
サイズ	長さ 49cm, 幅 53cm, 高さ 26.5cm
自重	35Kg
最大荷重	20Kg
継続運航時間	5 時間
最大走行速度	80cm/秒 (3.6Km/h), 速度は連続的に変化可能.
モータ出力	140 ワット (×4)
処理系	モトローラ社製 MC68332, 32bit プロセッサ PC104 . 周辺プロセッサとしてカメラの雲台を制御する PIC を追加 .
標準入出力端子	プログラム・ダウンロードおよびリモコン制御用シリアルポート, ユーザ用シリアルポート (2), デジタル出力 (7), デジタル入力 (6) 端子
基本センサ	超音波センサ (7)、毎秒 40KHz のキャリア波を 20Hz で変調して発射、反射波の強度を検知
CCD カメラ	CCD カラーカメラ (CHANNEL VISION 6005-B)
無線 LAN	BreezeCOM SA-10 PRO.11 Station Adapter
画像処理	AAI 社製 V3 Vision Board

- 特徴

- 行動型技術による屋外・屋内共用の知的ロボット研究開発用プラットフォーム .
- 車体下部、車軸等に防水構造を採用し、最深 10cm の水中でも動作が可能な設計, 15cm の高さの障害物を乗り越えることが可能 .
- 四輪独立差動型の駆動方式を採用したため、小さい回転半径で方向変換が可能 .

- 障害物検出のための防水型超音波センサを車体周辺に装備し、動作環境中の障害物等を検出し回避行動をとることが可能。
- 各種追加センサやアクチュエータなどのオプションの接続も可能。
- モトローラ社製 68332 標準プロセッサ (ベスタボード) の他に、PC104 プロセッサ及びケペラロボットの関数群がそのまま使える 68376 プロセッサ (カメレオン・ボード) もオプションとして搭載可能。
- PC, ノートパソコン, ワークステーション等からのリモート制御の他に, これらのコンピュータで開発したプログラムをダウンロードし, 完全自律運転可能。

#### ● ソフトウェア開発環境

- GNU-C 言語あるいは他の言語で, ロボットの走行に必要な行動群を作成。
- 完成した行動群は GAIA-2 上の RAM にダウンロードし, 機上プロセッサが実行。
- 付属のライブラリ関数を使って, ロボット上のセンサの読み取り, モータ制御信号の設定が可能。
- アルゴリズムの選択, ソフトウェアのモジュール設計, 製作等は研究開発者の自由。

上記に示したように, GAIA-2 は, 研究用に設計された自律走行型ロボットである。

GAIA-2 に高速画像認識アプリケーション”Smart-EYE”を搭載した A4 型ノートパソコンを搭載する方法を図 6.1 に示す。

A4 型ノートパソコンは, GAIA-2 内部にある BreezeCOM SA-10 Pro をネットワーク HUB の代わりに使用して, GAIA-2 の内部パソコンと TCP/IP ネットワーク接続を行なうことができる。また, Video Capture PC Card REX-9590(RATOC System Inc.) により, GAIA-2 付属の CCD カラーカメラからの画像をノートパソコンへ送った。ノートパソコン内蔵スピーカでは, 最大音量が小さいので, アンプ内蔵ステレオスピーカを接続した。

GAIA-2 自体の改造は, CCD カメラと雲台のマウント位置をノートパソコンに隠れない位置に高台を用いて設置した。また, 拡張や配線接続の組み替えが行ない易いように, GAIA-2 外部の周辺機器との接続ケーブルをすべてソケット形式に交換した。図 6.2 に改造・拡張した GAIA-2 を記載する。

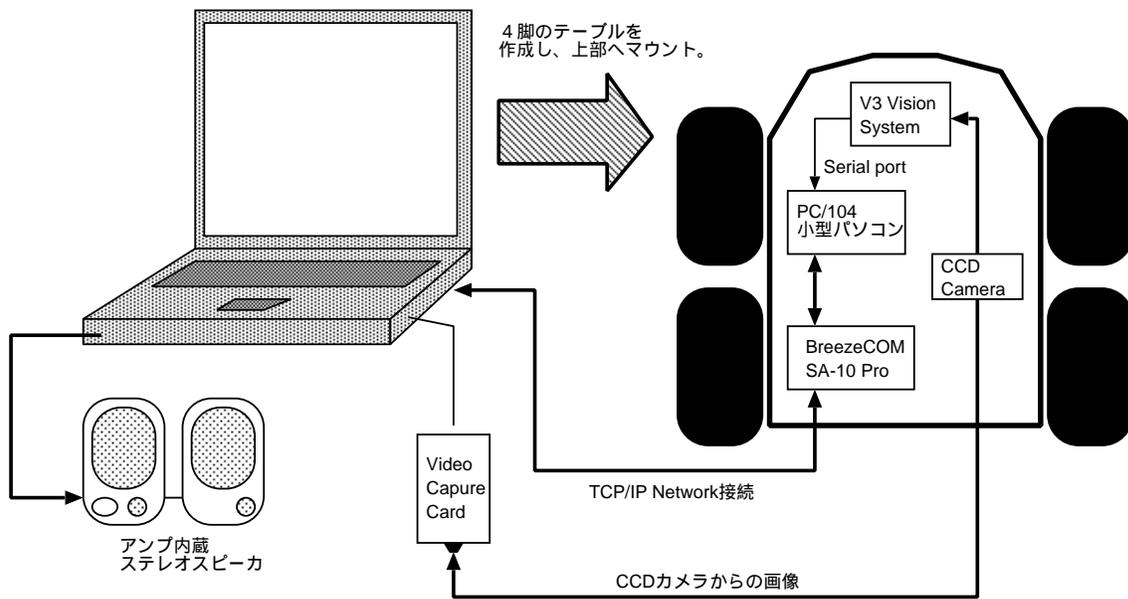


図 6.1: A4 型ノートパソコンと GAIA-2 の接続環境



図 6.2: 改造・拡張した GAIA-2

## 6.3 障害物回避機能

障害物回避は，GAIA-2 に搭載されている Sonar によって行なう．図 6.3 に Sonar の配置を示す．

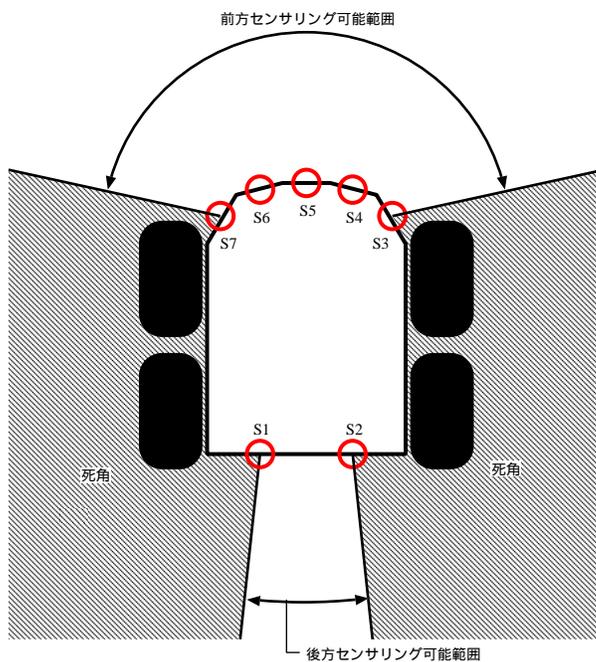


図 6.3: Sonar 配置図

以下に障害物回避アルゴリズムを説明する．

- まず，Sonar を配置状態により 4 つのグループに分ける．ここで用いる変数は，図 6.3 にある英数字に対応しており，各々の Sonar の値を表す．
  - Group 1(後方):  $(S1 + S2)/2$  .
  - Group 2(右斜め前):  $S3 + S4$  .
  - Group 3(前方):  $S5$  .
  - Group 4(左斜め前):  $S6 + S7$  .
- Group の値をもとに，次の手順を繰り返す．

```
if(Group3 <12){
  if(Group4 > others)
    TurnLeft(X); X のモータ回転で左へ旋回
  else if(Group2 > others)
    TurnRight(X); X のモータ回転で右へ旋回
  else if(Group1 > others)
    Back(X/2); X のモータ回転で後退する
} else {
```

```

if(S4 < 15 || S3 < 11)
    TurnRight(X);
else if(S6 < 15 || S7 < 11)
    TurnLeft(X);
else {
    if(Group3 > 30)
        Forward(X*2); X*2 のモータ回転で前進
    else
        Forward(X)
}
}
}

```

このアルゴリズムを用いると，GAIA-2は，Sonar の反応が大きい(障害物が遠い) 方へ機体に向け，なるべく前へ進もうとする．

## 6.4 対象物との距離を一定に保つアルゴリズム

ここでは，対象者の顔をボールに置き換えて説明する．

以下に追従アルゴリズムについて，図 6.4-6.6 を用いて説明する．カメラを支える雲台

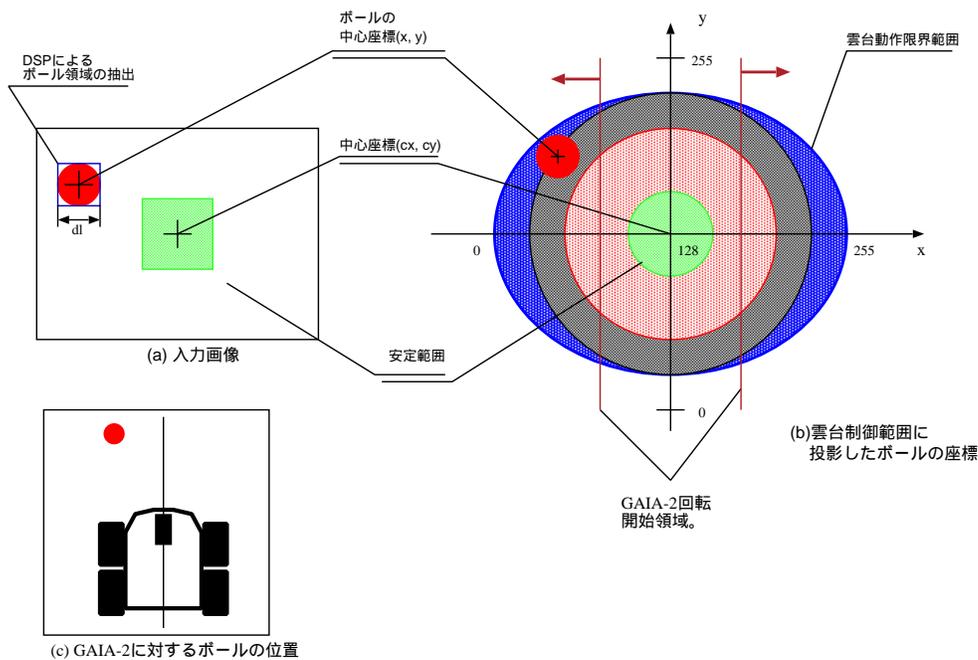


図 6.4: 初期状態

は，縦横に 180 度の回転運動(首振り)が可能である<sup>1</sup>．カメラの制御信号は，ASCII 文字

<sup>1</sup>縦はカメラが雲台に接触するので，実際は 100 度ほどである

一文字であり，整数値に収めると  $[0,255]$  の範囲である ( $[0, 255] = [0 \text{ 度}, 180 \text{ 度}]$ )．入力画像の座標は，雲台の制御空間に投射され，図 6.4(b) のように表すことができる．

これより，図 6.4 を初期状態として，雲台，機体の動きを説明する．雲台は，常に図 6.4(a),(b) にある緑色の領域に，ボールを捉えようとする．その動きを以下に示す．

1. 入力画像において，画像の中心座標  $(cx, cy)$  とボールの中心座標  $(x, y)$  の差分を求める．その値を  $(dx, dy)$  とする．
2.  $(dx, dy)$  を雲台制御範囲に投射する．その値を  $(f(dx), f(dy))$  とする．
3.  $(f(dx), f(dy))$  と雲台制御範囲の緑色の範囲の中心座標との距離  $l$  が，緑色の範囲に収まっていれば終了．そうでない場合は，次に進む．
4.  $l$  が赤い範囲にある場合は，雲台を  $(mx, my)$  動かす．動かす方向は， $(f(dx), f(dy))$  の各符号による．また， $l$  が灰色の範囲にあれば， $(nx, my)$  動かす．ここで  $m < n$  である．

上記のものを約 20 回/秒で実行する．これを繰り返しているうちに，図 6.5 のようにボールを緑色の範囲に捕らえる．

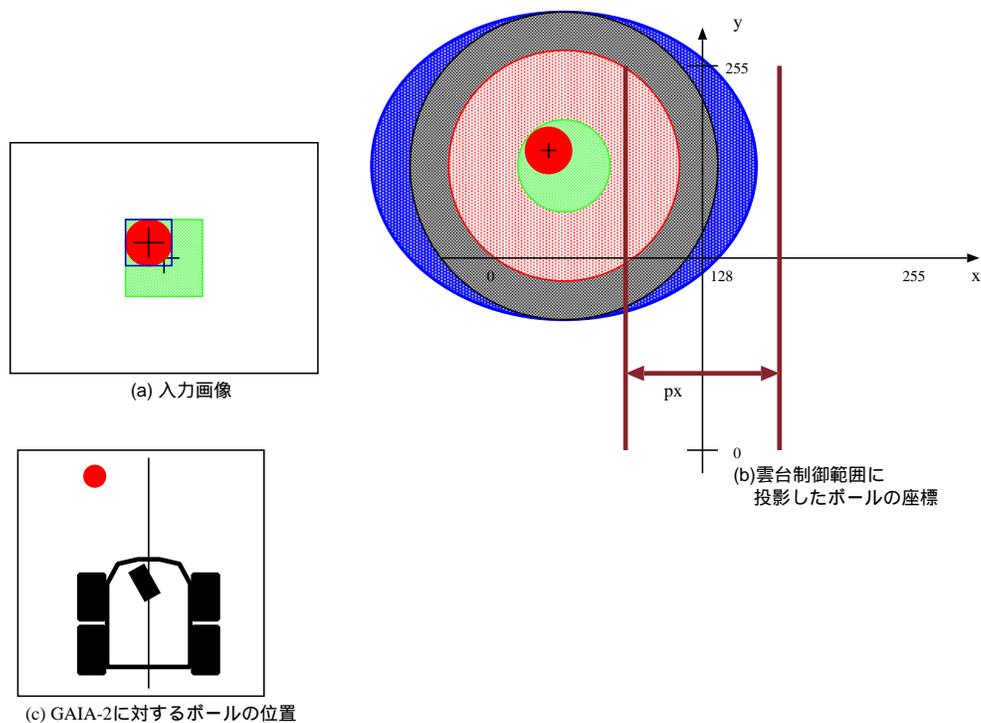


図 6.5: 雲台制御により，ボールを捕らえた状態

雲台制御により，ボールを入力画像の中心付近に捕らえたら，機体の制御に移る．

まず回転運動の制御について説明する．機体の回転運動は，図 6.5(b) での  $x$  軸の動きを示す．次に行なうのは，緑色の領域を茶色の範囲内に収めることである．ここで，茶色の範囲の幅を  $px$  とし，緑色の中心座標を  $(gx, gy)$  とすると，回転運動は， $-px/2 < gx < px/2$  となるまで続く，この時の回転方向は， $gx - 128$  の符号で回転方向を決め， $gx - 128$  の値で速度を決める．この回転運動の際も，上記に示した雲台の制御は行なわれている．回転運動により  $-px/2 < gx < px/2$  となった状態を図 6.6 に示す．

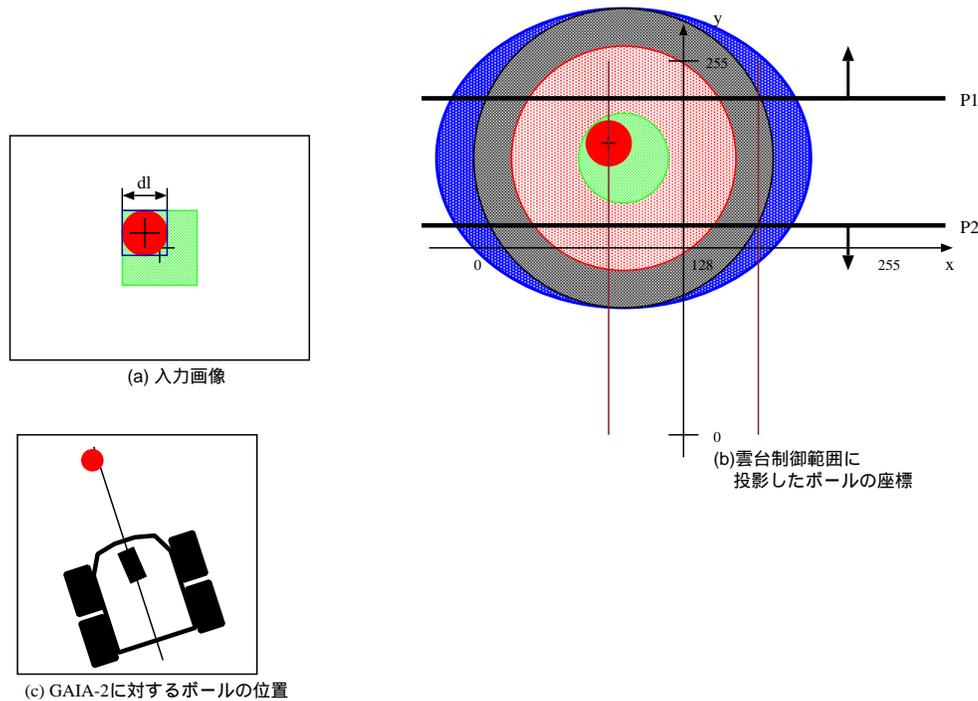


図 6.6: 機体の回転運動でボールを捕らえた状態

次に，機体の前後運動について述べる．前後運動は，ボールと機体の距離を一定に保つために行なう．一定の距離に保つために，抽出したボールの領域の直径  $dl$  を参考にする． $dl$  がしきい値  $cl$  よりも小さい場合は，前進を行ない， $dl > cl$  の場合は，後退する．このとき，しきい値  $cl$  にある程度の許容範囲  $\pm cl$  をもたせることにより，前後運動による機体の振動 (停止しない状態) を回避している．

上記の前後運動の制御だけでは，加速した機体がボールを追い越してしまうなどの問題が生じる．そこで，雲台の  $y$  軸方向の値に  $P1, P2$  のしきい値を持たせ，この範囲に収まるように前後運動を行なうようにした．雲台の  $y$  軸方向の値が， $P1$  よりも大きい，または  $P2$  よりも小さい場合には，後退させることにより， $P1$  より小さくかつ  $P2$  よりも大きい範囲に収めることが出来る．この  $P1, P2$  のしきい値は，ボールの高さによって変化する．例えば，ボールが人間の顔の高さにあり，カメラの設置位置が腰よりも低いとき， $P1, P2$  を限りなく 255 にすると，機体と人間の距離は最も近くなり，128 に近づけると最も遠い場所で距離を保つようになる．

## 6.5 JAIST Navigator

本研究で構築した対象人物の距離を一定に保つ自律走行型ロボットは，人を目的場所まで案内するロボットとして応用することが出来る．そこで，JAIST 案内ロボット”JAIST Navigator”を提案する．図 6.7 を用いてその大まかなタスクを説明する．

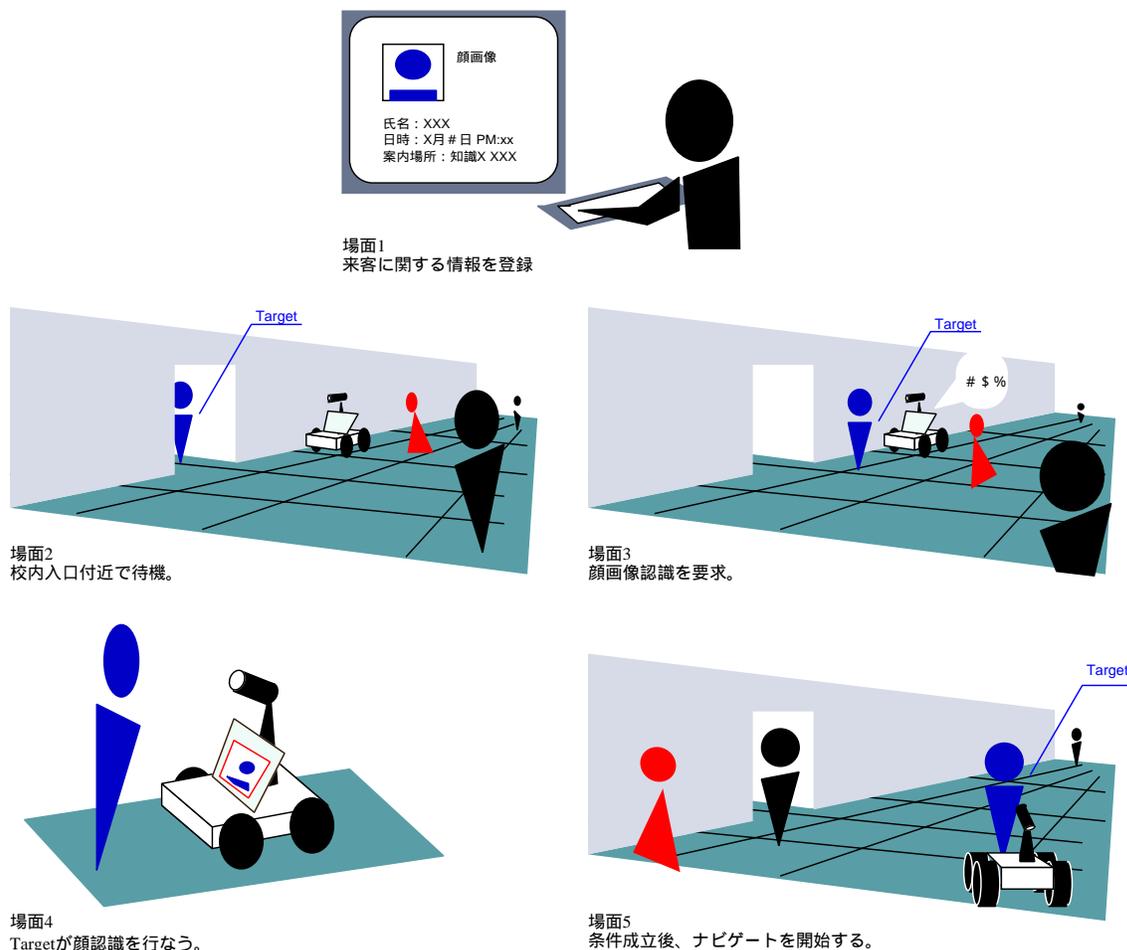


図 6.7: JAIST Navigator のシステム概要

場面 1 来客が訪れる日時，案内先，来客の顔画像をあらかじめ JAIST Navigator へ登録する．

場面 2 校内入口付近に待機し，自機付近に通行人が来るまで待機する．

場面 3 通行人がそばを通り，後部のセンサが反応すると，音声により顔認識を行なうように要求する．

場面 4 音声による指示に従い，Target が顔認識を行なう．

場面 5 Target が登録しておいた来客と判断すると、目的地まで誘導することを音声で告げ、ナビゲーションを開始する。

この JAIST Navigator は、構築したロボットシステムに、GPS などを用いた自機の位置確認や、建物の内部マップの記憶などの機能を拡張することにより実現可能である。

## 第7章 結論

本研究で開発した高速画像認識モジュールは、第6, 7章の実験で高速実時間での識別が可能であり、抽出した肌色領域の中心座標を得ることができ、背景にあまり依存しない高識別率の画像認識モジュールであることが言える。これらは、第2章で挙げた自律走行型ロボットのための画像認識モジュールの目標要件のうち、(a)20 fps 程の高速実時間で識別しなければならない、(b) 背景 (場所) に依存されない、(c) 画像内の対象物の座標を得ることが出来る、(d) 識別率が高い、の4つの条件を満たしている。また、画像認識に必要なものは、CCD カメラが一台と Video Capture Card が一枚だけであり、目標要件にある (e) 廉価な画像装置を用いて実現可能であることを満たす。また、開発した高速画像認識モジュールは、対象物の大きさの変化に弱いという制約があるが、第6章で構築したロボットシステムのようなものには、十分利用価値があることが分かった。

よって、本研究で開発した高速画像認識モジュールは、自律走行型ロボットのためのものとしては、十分なものであると言える。

第6章で、高速画像認識モジュールと自律走行型ロボットを容易に連結することが出来ることを証明することができ、今後、画像認識モジュールを搭載した自律走行型ロボットの活躍に期待が持てる。そこで、本稿では付録に、本研究で構築したロボットシステムの有効活用の例を挙げている。

## 第8章 今後の課題

本研究で考案した肌色領域抽出方法は，処理速度を重視したものであるため，検出した肌色領域に過膨張や過伸縮が良く見られる(図 2.3 参照)．これが，領域のラベリングを行なう際に，間違った領域の結合や分類を引き起こす原因となり，対象領域の誤認識へと繋がる．よって，処理速度を落さず，本研究で考案した肌色領域抽出方法よりも精度の高いものを構築することが今後の課題である．

また，本研究で用いた高次局所自己相関特徴と線形判別分析による識別は，対象物の大きさの変化に弱いことが知られており，その解決法として，対象領域の画像を Log-Polar 画像へ変換した後，高次局所自己相関特徴を用いることにより，大きさに不変な画像認識を実現する方法 [18] がある．Log-polar 画像への変換には，時間が掛かるため，廉価な画像装置を用いた画像認識にとっては不向きである．今後，その Log-Polar 画像への変換処理時間の向上が今後の課題である．

# 謝辞

本研究を行なうにあたり，数々の知識を提供して頂きました櫻井彰人教授，ならびに藤波努助教授に深く御礼申し上げます．また質問に詳細に答えて頂いた荒木修助手に御礼を申し上げます．

GAIA-2 に関する知識を提供して下さった Applied AI System, Inc. の主任技術者井出浩一氏に感謝致します．

最後に，同じ研究室内でともに励んだ同輩，先輩に感謝致します．

# 参考文献

- [1] 福井和宏, 制約相互部分空間法を用いた環境変化にロバストな顔画像認識-照明変動を抑える制約部分空間の学習. 電子情報通信学会論文誌 (D-II),J82-D-II,4,1997,pp.2170-2177.
- [2] 栗田多喜夫, PARCOR 画像の高次局所自己相関特徴を用いた背景変化および平行移動に強いジェスチャ認識, 信学技法,PRMU96-213,1997,pp.159-164.
- [3] N.Otsu and T.Kurita, "A new scheme for practical, flexible and inteligent vision systems", Proc.IAPR Workshop on Computer Vision,1988,pp431-435.
- [4] 今川 和幸ほか, 肌色領域により隠れて見える場合を考慮した手話動画像からの手の実時間追跡, 電子情報通信学会論文誌, Vol.J81-D-II No.8,1998,pp.1787-1795 .
- [5] 大津展之, 判別および最小二乗基準に基づく自動しきい値選定法 信学論, vol.J63-D,no.4,April,1980.
- [6] A.Rosenfeld and A.C.Kak(長尾 真監訳), デジタル画像処理, 近代科学社,1978,p.329.
- [7] 八木 伸幸 ほか, ハイパーリンクブック デジタル映像処理, オーム社,2000,p.62.
- [8] 大津展之, パターン認識における特徴抽出に関する数理的研究, 電子技術総合研究所報告,Vol.818, 1981.
- [9] 飯尾 淳, Linux による画像処理プログラミング, オーム社,2000.
- [10] 小俣光之, C for UNIX システムコールの基礎と応用, 株式会社 秀和システム, 2000 .
- [11] FEST Project 編集委員会 新実践画像処理 Image Processing with HALCON, 株式会社リンクス, 2001 .
- [12] Brad Nichols ほか, 榊 正憲 訳, Pthreads プログラミング, 株式会社 オライリー・ジャパン, 1998 .
- [13] Jeff Tranter 著, 山形 浩生 訳, Linux マルチメディアガイド, 株式会社 オライリー・ジャパン, 1997 .

- [14] Alessandro Rubini 著, 山崎 康宏 訳, Linux デバイスドライバ, 株式会社 オライリー・ジャパン, 1999 .
- [15] 佐藤 淳, コンピュータビジョン -視覚の幾何学-, 株式会社 コロナ社, 1999 .
- [16] 小国 力, Fortran 95, C & Java による新数値計算法 -数値計算とデータ分析-, 株式会社 サイエンス社, 1997 .
- [17] 森 友一郎ほか, RTLinux リアルタイム処理 プログラミングハンドブック 株式会社 秀和システム, 2000 .
- [18] 栗田多喜夫ほか, Log-Polar 画像の高次局所自己相関特徴を用いた大きさ・回転に不変な 2 次元図形及び顔画像の認識, 信学技報, Vol.PRMU96-212,1997.