

顔画像認識における基礎的顔領域選択手法の検討

S013024

川嶋賢二

応用情報学講座

田中研究室

1. 序論.....	3
2.1. 顔画像認識の複雑さ.....	4
2.2. 人間領域の検出手法.....	4
2.3. 顔領域の選択手法.....	5
3. 人間領域の選択.....	6
3.1. 人の領域の検出.....	6
3.2. 表色系.....	6
4. 実験・検証.....	9
4.1. 開発環境・実験環境.....	9
4.2. 人間領域の検出.....	9
4.2.1. 肌色検出実験.....	9
4.2.2. 肌色検出を用いた人間領域検出.....	12
4.3. 顔領域の選択.....	13
4.3.1. 肌色ピクセルの連続に注目した顔領域選択手法.....	13
4.3.2. 顔の縦横比に注目した顔領域選択手法.....	18
5. 結論.....	21
5.1. まとめ.....	21
5.2. 今後の課題.....	21
6. 参考文献.....	23

1. 序論

現在、顔画像認識はインタフェイスやセキュリティなどの分野において利用がされ始めている。例えば、インタフェイス分野では表情の認識などを用いた新しいインタフェイスの構築が考えられており、セキュリティ分野では顔の他にも虹彩・指紋・DNA などを利用したバイオメトリクス技術による認証を応用したシステムなどが注目を集めている。例えば顔画像認識システムでは米国の Identix 社の FaceIt は誤秒以内でスキャンを完了し、その誤認識率は 1% 以下でユーザの顔を認識することができる [1]。本研究は、このような顔画像認識システムの前処理に関する研究である。

従来の顔認識手法では、制限付きの環境下での顔画像に対する認識技術としてはすでに実用出来るほどになっており、現在では顔の向きや撮影場所などによる環境の変化に対応するための様々な手法が研究されている。顔認証システムにおける顔画像認識は静止画像や動画から顔をいかに発見するかであり、画像全体から背景を取り除き、顔領域を抽出することで、全体の無駄な情報を減らし、また顔の特徴となる情報を残すことである。また、単純な背景からの顔の切り出しであればコンピュータでも今となっては容易であるが、人間の顔は常に同じ表情をして同じ方位を向いてはいない上に、カメラなどで写した画像では写し方にも影響を受け、単純な背景ばかりではない。このような問題のため、人間が無意識に行っている背景と顔との区分も、コンピュータ上では容易でない。

本研究では顔画像認識における顔領域の選択部分に重点をおき、その手法の実験・検証を行った。背景から人の領域を抽出するための HSV 表色系と YCC 表色系を用いた肌色領域抽出に関する手法 [2][3]と、抽出された肌色領域からの顔領域の選択に関する手法を検討した。実験では、いくつかの環境（背景、顔の向き、表情など）におけるそれぞれの手法の有用性を検証する。

2. 顔画像認識

人間が普段特に意識することなく行っている顔の認識をコンピュータに行わせるという試みは、顔のパターン認識の問題として古くから研究が行われてきた。近年では、コンピュータの性能向上とともに画像処理技術も発達し、リアルタイムでの画像認識や合成が可能になったため、顔情報処理システムに関する多くの試みがなされてきた。しかし、現在でも顔の発見、個人識別や表情認識などの課題がある。

2.1. 顔画像認識の複雑さ

人間は特に苦労することなく瞬時に人の顔を認識することが可能であるが、コンピュータにこの処理をさせることは難しい問題となる。

顔画像認識において特に重要と考えられる処理が二つある。一つは顔の検出である。画面中に人間が存在するかないかを判別し、人間がいるならその顔領域を切り出し後の処理につかえるかを判断する。もう一つは、顔の識別のための特徴抽出である。顔領域抽出の後、その中から顔の特徴を抽出し、データベースに格納された顔の特徴データと整合する必要がある。

これらが固定の画像パターンで表されないため、顔の検出や認識が困難となっている。認識の対象となる顔が、いつ、どんな距離で、どのような照明条件や背景において現れるかわからない。さらには、顔の表情や向き、髪型やひげ、眼鏡などによって同一人物の顔であってもおびただしい数の多様性が生じる

以上のように、認識のたびに認識対象となる顔を同じの向きや表情にすることは通常困難であるため、これらの影響を受けにくい方法を採用することが顔画像認識システムには求められている。さらに、顔認証システムでは顔画像認識の結果を即座に利用する必要がある場面が多いため、実時間で認識ができることが要求される。

以上より、顔画像認識システムの要求条件は実時間認識、顔の向き・表情変化への対応、照明の変動への対応、経年変化への対応である。

2.2. 人間領域の検出手法

2.1.で挙げたように、人間はぼけた写真などからも人の顔を即座に見つけることが可能である。また、髪型や顔の向きや表情が異なっていたとしても同じ人物であることを識別でき、さらに怒っている、笑っている、悲しんでいるなどの表情については相手が変わったとしても同じ表情であると理解できる。これらの人間の認識能力を、いかにコンピュータ上で実現するかが顔画像認識の問題となる。

人間領域の検出において単調な背景に比べ複雑な背景の場合では、顔の輪郭と背景の区別が困難になる。画像内の人の領域と他の物体とを区別し切り出す必要がある。顔の発見は顔画像認識において処理対象領域を決定する上での前処理である。この前処理を省略・単純化するために単純な背景を用いるなど、カメラと対象となる顔の位置関係、光源など

を制限して行うことが多かったが、現在では環境の限定が少なくなるような様々な手法が研究されている。この中で、画像の色情報はカメラの位置によらず比較的扱いが容易であるため頻繁に使われている。色情報を扱う手法では画像中の人の肌の色に注目し、その領域を切り出して次の処理を行う。コンピュータに取り込まれるカラー画像は、ほとんどの場合 RGB (赤緑青) の 3 成分に分解されるが、肌の色を抽出に利用する場合にはこの 3 成分ではあまり有効ではない。なぜならば、RGB 値を利用する場合同じ色であっても色の明るさによって RGB の値が大幅に異なってしまうため、RGB 値のどの範囲を肌色領域とするかの決定が難しい。そのため、本研究では RGB から HSV (色相、色彩、明度) の 3 成分による色表現に変換し、人の肌の色の色相に注目しその領域を抽出する手法 [2] と HSV 表色系の代わりに YCC 表色系をもちいた肌色領域抽出手法の検証を行った。

2.3. 顔領域の選択手法

本研究で用いる肌色検出による人の検出は、背景から顔領域の候補になるものを画像から検出するが、肌色検出では顔だけでなくそれ以外の肌色領域を抽出してしまう。例えば手や喉なども抽出してしまうため、その中から顔の領域のみを選択する必要がある。また、顔の中でも目や鼻や口などは個人を識別するために重要な情報であるため、目鼻口を含む領域でなければならない。さらに目鼻口などを含む領域を抽出するだけでなく、この後の処理の為に可能な限り顔以外の無駄な領域を排除する必要がある。そのため、本研究では肌色領域の肌色ピクセルの連続に注目した顔領域選択手法と、肌色領域における顔の縦横比に注目した手法の実験・検討を行った。

3. 人間領域の選択

3.1. 人の領域の検出

顔画像認識をするためにはまず認識対象の画像から人の領域を抽出する必要がある。また、顔を含んでいない画像や顔領域が小さすぎるなどの認識に適さないと判断した画像は、対象の画像に人はいないということとして破棄する。ここでは大まかに背景と人物の区別を行うための、明るさや背景に影響されにくいにおける3つの表色系の検討を行う

3.2. 表色系

色を表現する方式のことを表色系 (Color Specification System) とよぶ。表色系の代表的なものに RGB 表色系があり、これは赤・緑・青を3原色と定め他の色をこの3色の混合で表現するものである。コンピュータ上においても、RGB 表色系は一般的に用いられているカラー画像の表現方法でもある。

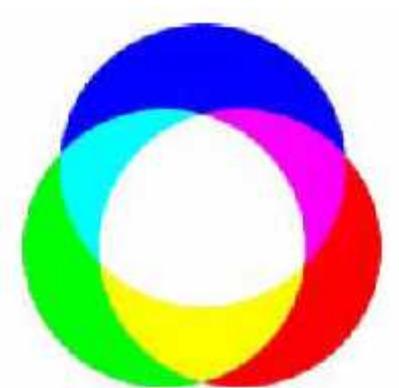
色の表現方法で代表的なものうち RGB 表色系と表現方法が異なるものとしては、HSV 表色系や YCC 表色系が挙げられる。RGB 表色系が基本となる色の混ぜ合わせ方で色を表現するのに対し、HSV 表色系は色の特性 (色相、彩度、明度) の組み合わせにより色を表現する。また、YCC 表色系では明るさと基本となる色の差の組み合わせによって色を表現する。RGB 表色系と HSV 表色系による主な色は図 3.2.1、図 3.2.2 のように表現できる。

この3つの表色系の特徴を表にすると以下の表 3.2.1 のようになる。

表 3.2.1

	色の表現法	各要素			利点
RGB	基本三色の混合	R 赤の強さ	G 緑の強さ	B 青の強さ	数値として計算しやすい
HSV	色合い、鮮やかさ、明るさ	H 色合い	S 鮮やかさ	V 明るさ	各要素が独立している
YCC	輝度、色み(色差)	Y 輝度	Cr 輝度と赤みの色差	Cb 輝度と青みの色差	処理の負荷を減らしやすい

また、実験で用いたプログラムでの、RGB 表色系から HSV 表色系への変換式は表 3.2.2 を[2]、RGB 表色系から YCC 表色系への返還式は表 3.2.3 のものを用いた[4]。HSV 値の範囲は $0 < H < 359, 0 < S < 100, 0 < V < 100$ 、YCC 表色系の値の範囲は $0 < Y < 255, -127 < Cb, Cr < 255$ とする。



	R	G	B
Red	255	0	0
Green	0	255	0
Blue	0	0	255
Yellow	255	255	0
Magenta	255	0	255
Cyan	0	255	255
White	255	255	255
Black	0	0	0

図 3.2.1 RGB 表色系の主な色表現



	度	R	G	B
Red	000	255	0	0
Green	120	0	255	0
Blue	240	0	0	255
Yellow	60	255	255	0
Magenta	300	255	0	255
Cyan	180	0	255	255

図 3.2.2 HSV 表色系の主な色表現

表 3.2.2 RGB HSV の変換式[2]

```

max = max((int)R[i], (int)G[i]);
max = max(R, G, B);
min = min(R, G, B);
delta = max - min;

// 色相の算出
if(R == max)
    Hue = 60*(G[i] - B[i])/delta;
else if(G == max)
    Hue = 60*(B[i] - R[i])/delta +120;
else if(B == max)
    Hue = 60*(R[i] - G[i])/delta +240;

// 色彩の算出
Sat = (delta /max) *100;

// 明るさの算出
Val = max /2.55;

```

表 3.3.3 RGB YCC の変換式[4]

```

// 輝度の算出
Y      = 0.29900*R +0.58700*G +0.11400*B;
// 赤との色差
Cr     = 0.50000*R -0.41869*G -0.08131*B;
// 青との色差
Cb     = -0.16874*R -0.33126*G +0.50000*B;

```

4. 実験・検証

本研究では、顔画像認識における人の領域の検出手法と顔領域選択手法の検討をおこなう。顔領域選択は以下の手順で行われ、各手順における手法の検討を行う。

1. 背景と人の領域候補となるものの区別

肌色検出を用いた人間領域検出手法

2. 顔領域の選択

肌色ピクセルの連続に注目した顔領域選択手法

顔の縦横比に注目した顔領域選択手法

4.1. 開発環境・実験環境

以下のソフトウェア、ハードウェア、開発環境を利用し実験を行った

・ハードウェア

CPU : Celeron 2.0GHz

Memory : 512MB

Camera : CCD 画像センサー 30万画素 VGA640x480 USB 接続

・ソフトウェア

開発環境

VisualC++.net

DirectX SDK 9.0

OS

Windows 2000 SP4

4.2. 人間領域の検出

肌色検出実験に用いた画像は、六人の人間の室内・室外で撮った明るさなどの環境の違う画像を使い、画像中の両頬の肌の1ピクセルずつの情報を取得し総数51種類のRGB値を求め、またHSV値・YCC値を計算によって求めた。[2][3]

4.2.1. 肌色検出実験

肌色検出におけるRGB表色系とHSV表色系、YCrCb表色系を用いた手法の検討を行う。コンピュータに人の領域の検出をさせるためには、まずプログラムに何が肌色であるかを判別させる必要がある。カメラで撮影した人物画像から人間の肌のRGB値を調べた結果、表2のようになった。HSV値とYCbCr値はRGB値を変換したものである。YVbCr値については、この表において小数点以下を切り捨てている。

表 4.2.1 肌の色の RGB 値・HSV 値の例

肌の色								
R	165	222	208	193	254	139	92	175
G	126	170	145	153	238	115	81	135
B	121	159	131	155	213	115	75	125
H	7	11	11	356	38	0	21	13
S	27	28	38	21	16	17	18	29
V	65	87	82	76	100	55	36	69
Y	137	184	162	165	239	122	83	145
Cb	-9	-14	-17	-5	-15	-4	-4	-11
Cr	19	26	32	19	10	12	5	20

使用した画像は天気、時刻、室内・野外など撮影が同じ環境にならないように撮影した。表より RGB 値と HSV 値を比較すると、RGB 表色系は値のばらつきが大きいことがよく分かる。HSV 表色系や YCrCb 表色系では色合いと明るさ（輝度）が独立しているため、画像の明るさに影響されることなく色合いを調べることができることが分かる。

表 4.2.1 で挙げた例のほかに 43 通りの顔の肌色情報（RGB 値）を用意し、計 51 個の肌色情報から RGB HSV 変換を行い HSV 値、RGB YCbCr 変換を行い YCbCr 値を算出した。そのうちで明るさを示す値（HSV の V 値、YCrCb の Y 値）以外は図 4.2.1 と図 4.2.2 のように分類された。

表 4.2.1 においてグラフの横軸は色相の範囲（度）で縦軸はその範囲内に分類された肌色の頻度（個数）を示し、図 4.2.2.においてはグラフの横軸は Cb または Cr の値の範囲で、縦軸はその頻度（個数）を示す。

この実験の結果（図 4.2.1、図 4.2.2）から、人の肌の色が HSV 表色系では H：0～30 度、YCbCr 表色系では Cb：-20～0、Cr：16～64 の値になることが分かる。ただし、表 3.2.1 に示した通り YCbCr 表色系の利点として負荷が減らせることが挙げられるが、これは Cb,Cr の値が極端な場合をのぞきある程度の範囲にまとまるため、データがほとんど欠けずに範囲を制限することができるからである。このことから Cr：0～64 という範囲は広く、また Cr ,Cb の片方だけで肌の色を表現できないと考えられるため、本研究では以後の実験は HSV 表色系による肌色検出手法を用いることとする。

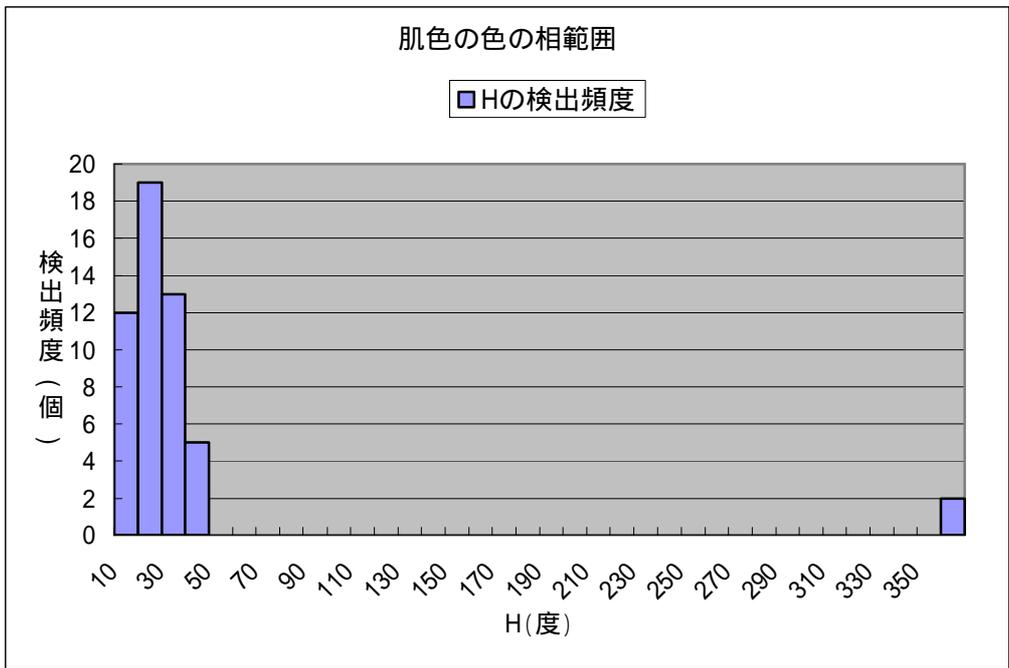


図 4.2.1 H 値別の肌色の検出頻度

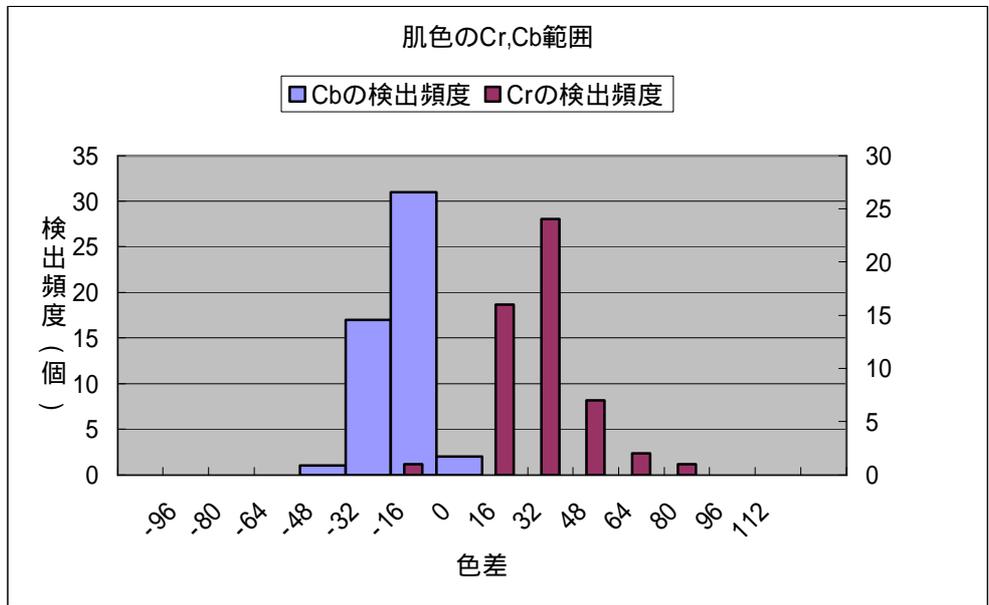


図 4.2.2 色差別の肌色の検出頻度

4.2.2. 肌色検出を用いた人間領域検出

肌色領域検出手法にはHSV表色系が最も有効と考えられるが他の方法とも比較するために、RGB・HSV・YCbCr表色系それぞれの方法について検証した。

HSV表色系では、Hの値0~30までを肌色と考え画像から肌色を抽出した。図4.2.3の左図が元の画像となり右図がHSV表色系による肌色領域検出手法の結果である。



図 4.2.3

YCbCr表色系では、Cbの値-48~0・Crの値16~80までを肌色とした。図4.2.4の左図が元の画像となり右図がYCbCr表色系による肌色領域検出手法の結果である。



図 4.2.4

図4.2.3と図4.2.4よりHSV表色系とRGB表色系はどちらもほぼ正確に顔領域を取得できていることが分かる。HSV表色系はHの値のみで表現できるため、この後の実験でHSV表色系をもちいた肌色検出を用いている。

4.3. 顔領域の選択

顔領域選択実験に用いた画像は、6人の顔画像を4枚ずつ撮影し、各画像から肌色領域を抽出したものを用いた。撮影時の制限としては、1つの画像に1人しか写っていないことと顔が傾いていないことの両方を満たした画像とした。

4.3.1. 肌色ピクセルの連続に注目した顔領域選択手法

X軸方向(またはY軸方向)に連続した肌色ピクセルの一行に注目し、あるピクセル数(しきい値とする)以上の肌色領域がある列からなくなるところまでを顔領域として考えその領域が最大となる部分を顔領域として選択する。具体的な手順は以下のようになり3.から6.をフローチャート図4.3.1に示す。またしきい値と規定値は以下のような値である。

1. X軸について平行な列について上から走査する
2. 画像全体を走査しそのうちで連続した肌色ピクセル列の長さが最大となる列を始点とする
この列の連続した肌色ピクセルの左端部分の座標を Left、右端部分の座標を Right、長さ(連続した肌色ピクセル数) Lとする
3. 始点から下に向かって走査していく
4. 規定値以上の連続した肌色領域がなくなるまで走査する
規定値以上の連続した肌色領域がなくなる直前の列の座標を Bottom とする
5. 始点から上に向かって走査していく
6. 規定値以上の連続した肌色領域がなくなるまで走査する
規定値以上の連続した肌色領域がなくなる直前の列の座標を Top とする
7. 以上より Left, Right, top, bottom でつくられる矩形領域を顔領域として選択する。

しきい値 : 注目している肌色領域が顔領域であるか判断する基準で以下の式で求める

規定値 : 実験によって経験的に決定する値 ($0 < \text{規定値} < 1$)

$$\text{しきい値} = X \text{軸に平行な行の肌色ピクセルの最大連続数} \times \text{規定値}$$

以上の実験を13の顔を含む画像に対し実験を行い、規定値を変えることで顔のどの部分を切り出すことが出来るかを調べ、顔領域選択として適した規定値を経験的に求める。規定値を0.05 ~ 0.95の範囲で変えて実験を行う。X軸方向に連続した肌色ピクセルの一行に注目する手法(手法1とする)とともに、Y軸方向に連続した肌色ピクセルの一行に注目した手法(手法2とする)の検討もおこなう。

しきい値 : 注目している肌色領域が顔領域であるか判断する基準
しきい値 = X 軸に平行な行の肌色ピクセルの最大連続数 × 規定値
規定値 : 実験によって経験的に決定するあたり (0 < 規定値 < 1)

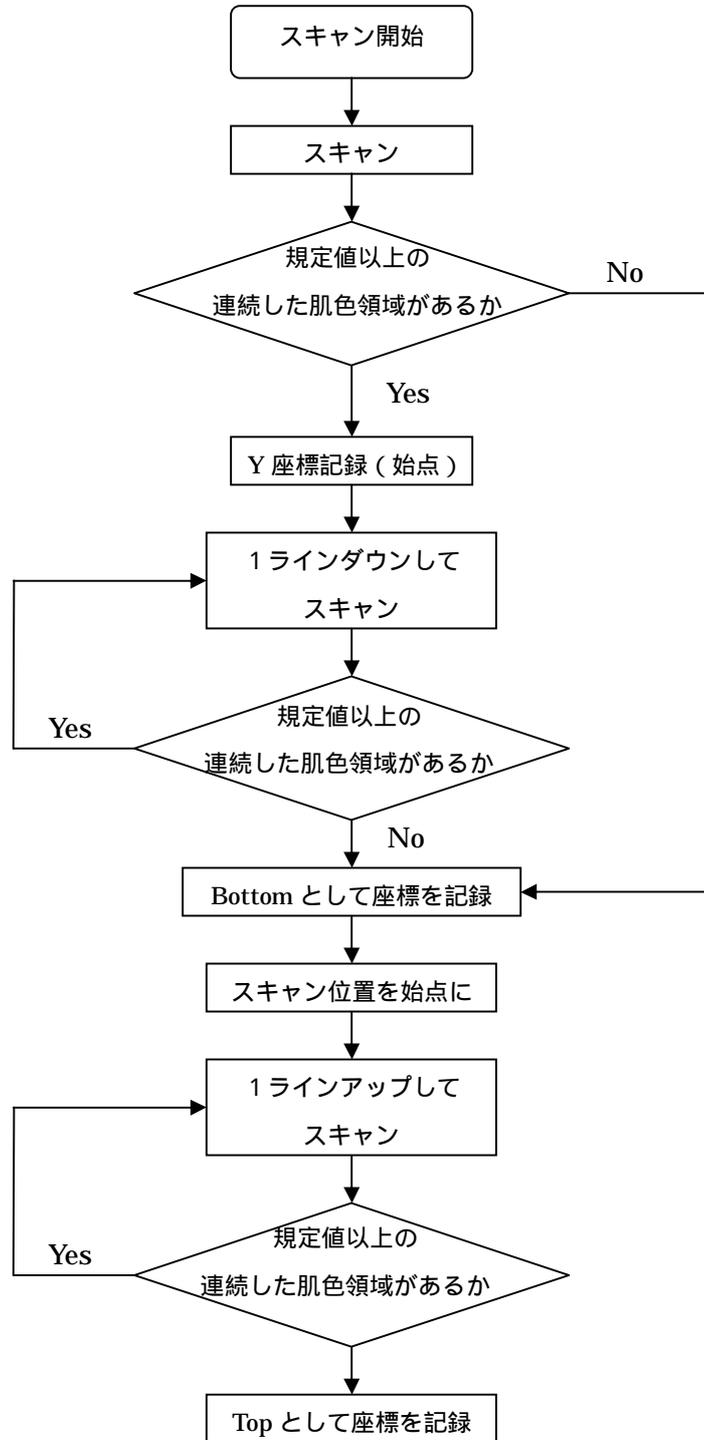


図 4.3.1 手法1フローチャート

実験の結果で、目鼻口や余分などそれぞれに分類された個数を表にしたものについて、X軸についての肌色領域ピクセルの連続に注目したものを表 4.3.1 に Y 軸について考えたものを表 4.3.2 に示す。表の項目は以下で示す範囲を顔領域として抽出したものとしている。それぞれ分類は人間の手で行い、範囲内であっても範囲外との境界ぎりぎりにある場合は欠けていることとした。また、それぞれをグラフにしたものを図 4.3.2 と図 4.3.3 に示す。グラフは各規定値における分類結果を 100%積み上げグラフで表したものである。

- 目鼻口 : 目と鼻と口を含んだ部分を抽出したもの、
- 余分 : 目鼻口以外に喉等の余分な部分を含んだもの
- 欠 : 目鼻口のうちどれか一つが欠けてしまったもの
- × : 目鼻口のうち二つ以上欠けたもの

実験の結果より、手法 1 では規定値 0.40 で最も良い結果が得られ、手法 2 では規定値によらずほとんど変わらない結果となっている。しかし、手法 1 では最も良い結果を出す規定値に設定しても全体の約 6 割しか目鼻口だけの領域を選択できず、余分な部分が含まれたものをあわせても 8 割程度となっているため精度の良いものでないと判断する。また手法 2 は手法 1 よりも全体的に結果が悪いことが分かる。

規定値：注目している肌色領域が顔領域であるか判断する基準で、この規定値以上の連続したピクセル数をもつ列を顔領域とする。

総標本数 24 画像

表 4.3.1 手法 1 の規定値の変化による検出割合の変化

規定値	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
目鼻口	0	0	0	0	6	9	13	14	15	14
余分	24	24	24	24	18	15	9	8	6	4
欠	0	0	0	0	0	0	2	2	3	6
×	0	0	0	0	0	0	0	0	0	0

規定値	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
目鼻口	12	8	4	1	0	0	0	0	0
余分	5	3	1	0	0	0	0	0	0
欠	7	11	15	4	2	1	0	0	0
×	0	0	4	19	22	23	24	24	24

表 4.3.2 手法 2 の規定値の変化による検出頻度の変化

規定値	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
目鼻口	4	4	4	4	4	4	4	4	4	4
余分	15	15	15	15	15	15	15	15	15	13
欠	5	5	5	5	5	5	5	5	5	7
×	0	0	0	0	0	0	0	0	0	0

規定値	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
目鼻口	2	2	1	0	0	0	0	0	0
余分	6	5	4	3	0	0	0	0	0
欠	15	14	12	6	3	2	1	0	0
×	1	3	7	15	21	22	23	24	24

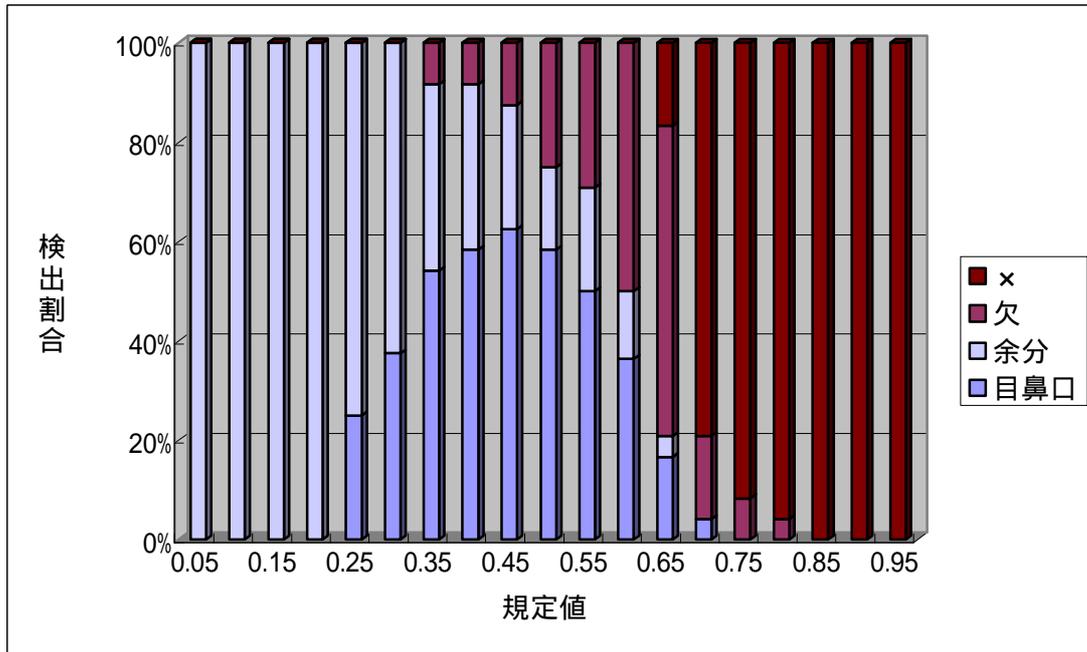


図 4.3.2 手法 1 の規定値の変化による検出割合の変化

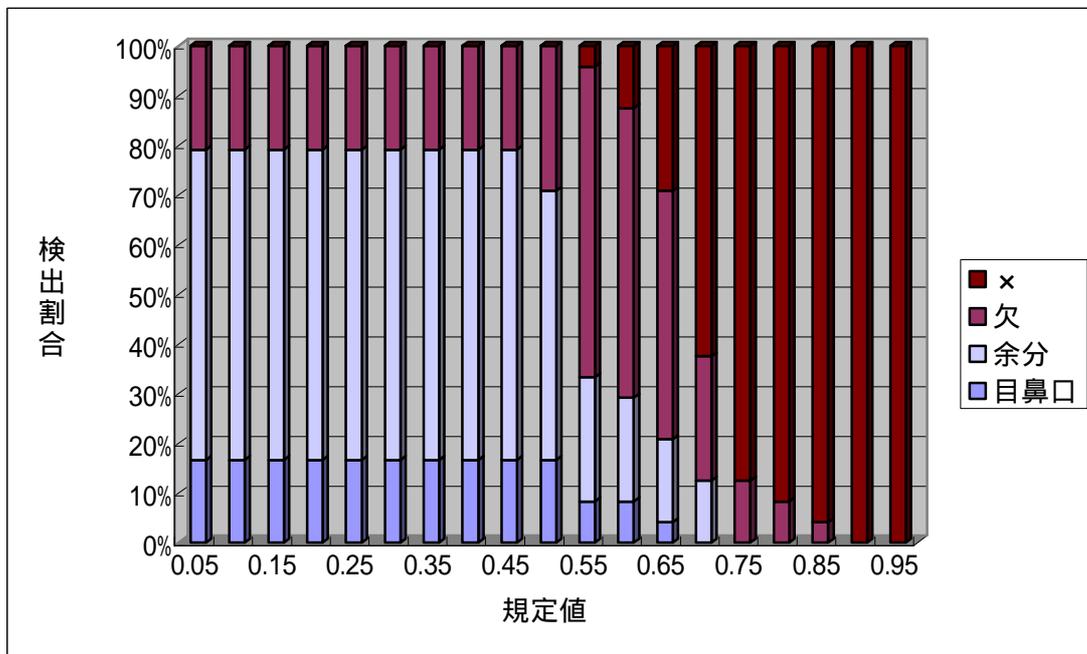


図 4.3.3 手法 2 の規定値の変化による検出割合の変化

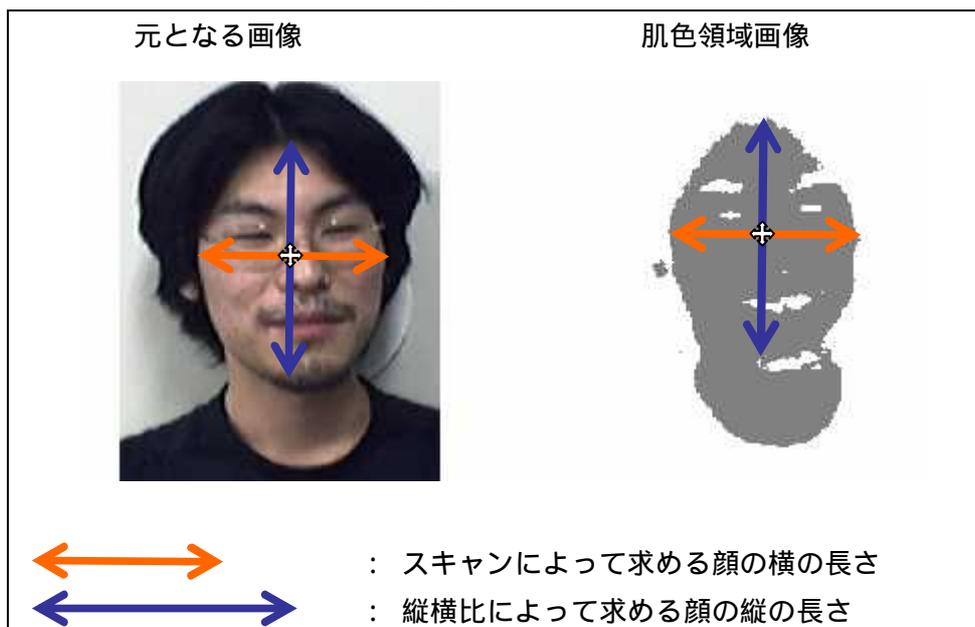
4.3.2. 顔の縦横比に注目した顔領域選択手法

顔の縦横比について注目し、X軸に平行な横一列ずつを考えその一列中にある肌色ピクセル数が最大となる個所を顔の横軸と考え、そこから縦軸の長さを考え矩形に囲まれた部分を顔領域として選択する。具体的な手順は以下のようになり、図 4.3.2 のように縦横比をつかって顔領域を選択する。

1. X軸について平行な列について上から走査する
2. 画像全体を走査しその一列中の肌色ピクセル数が最大となる列を顔の横軸とする
この列の左端にある肌色ピクセルのX座標を Left、右端部分のX座標を Right、長さ(連続した肌色ピクセル数) Lとする
3. 顔の縦軸の長さを横軸の長さから求め横軸の中心と縦軸の中心を合わせる
中心を合わせたあとの縦軸の上端のY座標を Top とし、下端のY座標を Bottom とする
4. 以上より求められた Left, Right, top, bottom によって囲まれた矩形領域を顔領域として選択する

以上の実験を 13 の顔を含む画像に対し実験を行い、縦横比を変えることで顔のどの部分を切り出すことが出来るかを調べ、顔領域選択として適した規定値を求める。縦横比は横：縦を 1:1 ~ 1.45:1 と変えて実験を行った。Left, Right, Top, Bottom で作られる矩形領域を顔領域として選択する手法(手法3とする)とともに、Left, Right, Top, Bottom で作られる囲まれた楕円形(円形)領域を顔領域とする手法(手法4とする)の検討を行う。

図 4.3.4



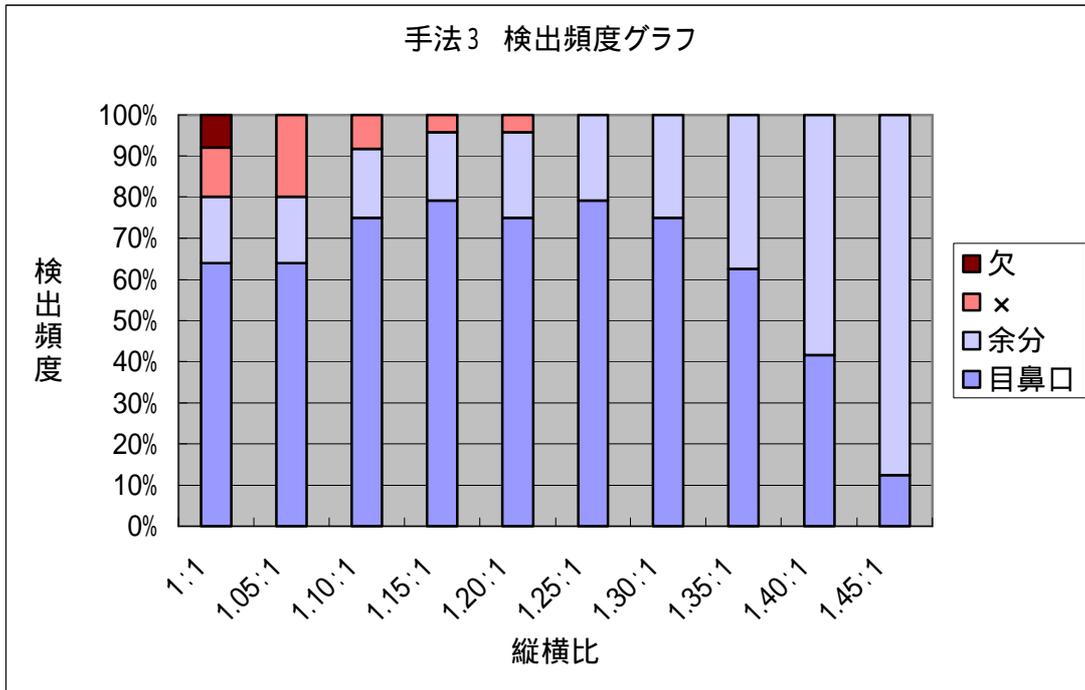


図 4.3.5 手法3 矩形による領域選択)の縦横比の変化による検出頻度の変化

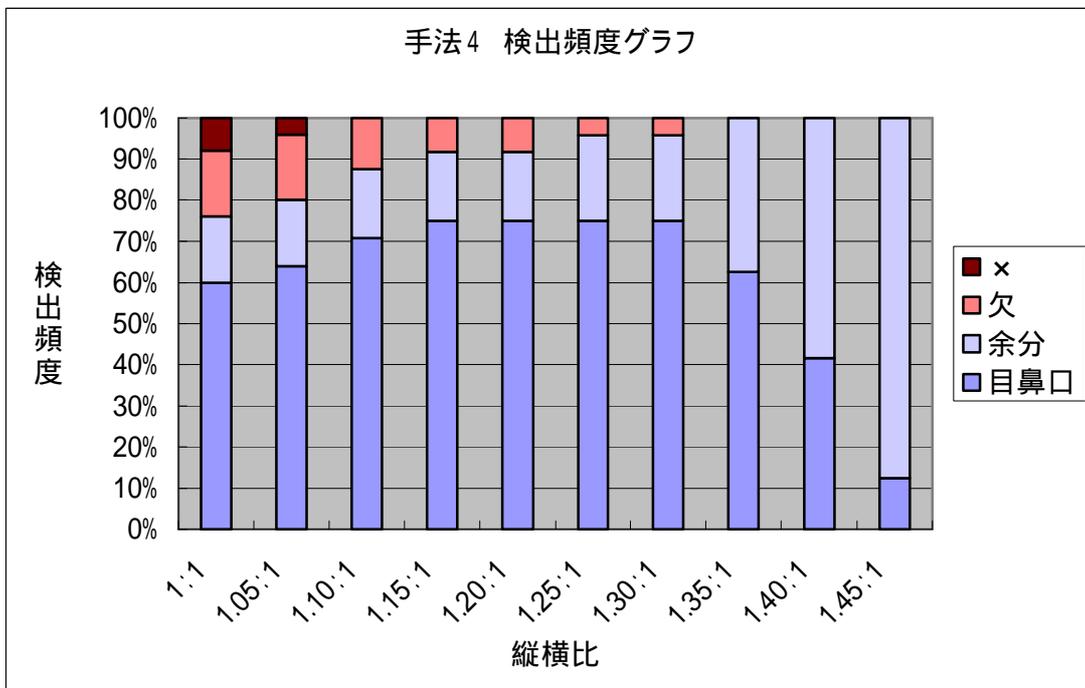


図 4.3.6 手法4 (楕円形による領域選択)の縦横比の変化による検出頻度の変化

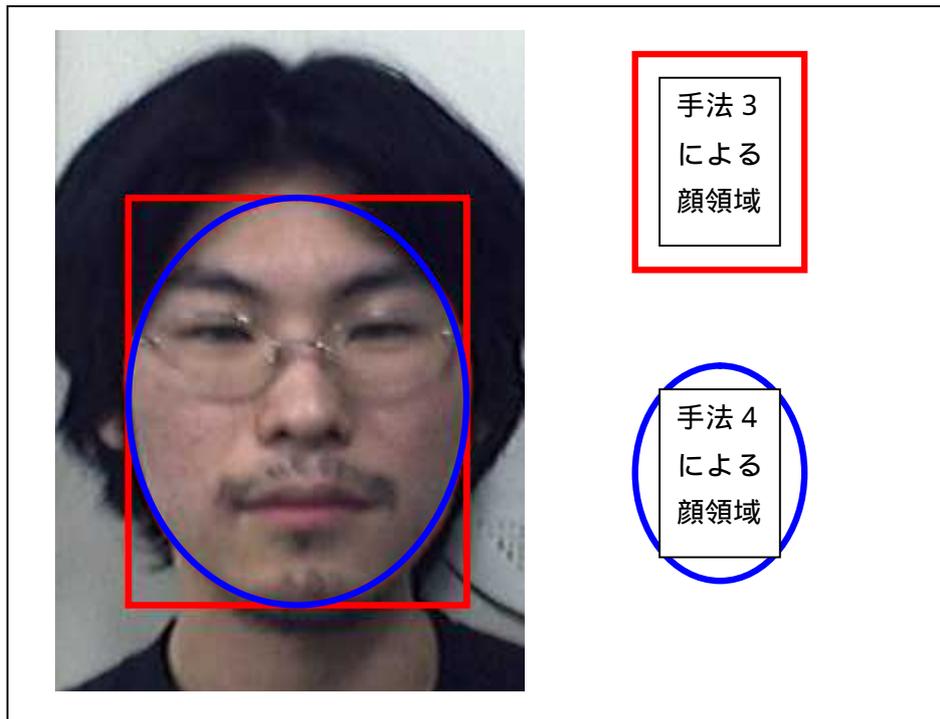


図 4.3.7 手法 3、手法 4 の選択領域の比較

5. 結論

5.1. まとめ

本研究では、認証システムやセキュリティーシステム、新しいインタフェイスなど幅広く応用できると期待されている顔画像認識技術における、顔領域の選択手法における手法についてそれぞれの有効性を検討したものである

実験においては、人間領域の検出で利用した肌色領域検出では、RGB 表色系や YCC 表色系に比べ HSV 表色系を用いることの有効性について実証できた。また、4 章において提案した手法をもちいた顔領域の選択では、確実とはいえないが顔領域を抽出できることを確認した。

5.2 今後の課題

顔領域選択手法において、顔の向きによっては正しく顔領域を取得できない場合があったためそれを解決する必要がある。また、本研究では他の手法との比較を行っていないため、他の研究との比較・検討の必要がある。

謝辞

本研究にあたり、田中章司郎教授には最後まで熱心なご指導をいただき、心より御礼申し上げます。

また、田中研究室の院生、学部生のみなさまには、本研究に関して数々の御協力と御助言をいただき、厚く御礼申し上げます。

なお、本論文、本研究で作成したプログラムおよびデータなど全ての著作権・知的財産権を、本研究者の指導教官である田中教授に譲渡いたします。

6. 参考文献

- [1] Identix : SoftWare Development Kits (BioEngine, Face It),
<http://www.identix.com/>
- [2] 貫目 洋一 : 顔認識の為のリアルタイム特徴抽出 ,
平成 13 年度島根大学総合理工学部数理情報システム学科修士卒業論文
- [3] Saori TAKEOKA, Masahiro OZAKI, Yoshinori ADACHI
Human Face Extraction for User Identification and Face Direction Recognition
名古屋女子大学 紀要 50(人・社) 145 ~ 151 2004
- [4] Studio-ID(ISIHARA WATARU) : 表色系 ,
<http://www.infonet.co.jp/apt/March/syllabus/bookshelf/COS.html>