

部分空間法と自己組織型ネットワークを用いた複数顔領域の切り出し*

奥 真一 橋詰 隆平 有木 康雄

龍谷大学 理工学部

〒 520-21 大津市瀬田大江町横谷 1-5

あらまし 本論文では、顔の大きさや存在場所に制約を課することなく、入力した人物の集合写真から髪型や眼鏡、ひげの有無に関係なく、複数の顔領域を精度良く切り出すことを目的として、部分空間法と自己組織型ネットワークを用いて顔領域を切り出す手法を提案している。部分空間法で顔領域を切り出す場合に、顔の存在場所と大きさに関係しないで精度良く切り出すためには、次のような工夫を行なっている。(1) 画像をピラミッド化して解像度を多重化しておく。これにより、大きな顔も小さな顔もただ一つの部分空間で切り出すことができる。(2) 切り出しの精度を上げるために顔データと顔でないデータの両方を学習させる。(3) 自己組織型ネットワークにより、目・鼻・口といった局所の特徴を認識することで顔領域を確定する。

和文キーワード： 部分空間法、自己組織型ネットワーク、ピラミッド、顔領域

Extraction of Facial Regions by Subspace Method
and Self-Organizing Network.

S. Oku, R. Hashizume and Y. Ariki

Faculty of Science and Technology, Ryukoku University
1-5, Yokotani Ooe-Cho Seta Otsu-Shi 520-21, Japan

Abstract The purpose of this study is to precisely extract human facial regions from input images without imposing any constraints on face size or the position. For that purpose, we used a subspace method and self-organizing network, and investigated the effectiveness through the experiments. The study points are summarized as follows; (1) An image pyramid is used to produce multi-resolution images in order to extract any size of facial regions. Due to this pyramid, one subspace for human faces is enough to search the facial region. (2) Face data and non-face data are used to produce the face subspace and non-face subspace in order to precisely extract the facial regions. (3) The facial regions extracted by the subspace method can be confirmed by self-organizing network which can recognize face parts such as eyes, nose and mouth.

英文 Key words: Subspace method, Self-organizing network, Pyramid, Facial region

1 はじめに

画像中からそこに映っている人の顔領域を切り出してくる処理は、個人識別を行なうセキュリティシステムにとって必要であるばかりでなく、入力画像の情景を理解する上でも必要とされている。どちらの場合においても、顔の存在する場所、顔の大きさについて制約を課することは好ましくないため、顔領域の切り出しを精度良く実行することは、極めて困難な処理と考えられている。

本研究では、顔の大きさや存在場所に全く制約を課することなく、入力した人物の集合写真から髪型や眼鏡、ひげの有無に関係なく、複数の顔領域を精度良く切り出すことを目的としている。

顔領域の切り出しにおいては、これまでに次のような幾つかの研究が行なわれている。

1. 画像中のエッジを抽出し、そのエッジの持つ形状の特徴や、お互いの位置関係を用いて、テンプレートマッチングにより、顔及び、顔器官の切り出しを行なう手法 [1][2][3][4][5][6]。

この手法では、画像中のノイズ、焦点のボケなどによって、正確にエッジが抽出できない場合が多く、顔領域の切り出しがエッジの抽出精度に依存してしまうといった問題、また、顔の表情の変化によってテンプレートマッチングができない場合なども出てくるといった問題がある。

2. 画像中の色彩情報を利用し、顔の持つ特有の色彩情報から顔の切り出しを行なう手法 [7]。

この手法は、白黒の濃淡画像には適用することができないため、顔に関する本質的な特徴抽出と切り出し処理にはなっていないと考えられる。

3. 距離画像によって与えられる3次元情報を利用する方法。

この手法は、3次元情報を得ることによって問題を簡単に行っているが、一枚の画像から人間は容易に顔を抽出できることを考えると、やはり、顔の切り出しに対して、本質的な処理とはなっていないと考えられる。

4. 濃淡情報を圧縮した表現において、顔領域の切り出しと識別を行なう手法。

例えば、KL展開による画像圧縮を用いて顔認識を行なう手法 [8][9] や、顔データをモザイク化し、顔領域を切り出して認識する手法 [10] がある。これらの手法は、一定の制約条件はあるものの、良好な実験結果が得られている。

これらの研究を踏まえて、本研究では、1枚の濃淡画像を対象として、部分空間法を用いて顔領域を切り出す手法を提案している。濃淡画像から、部分空間法

で顔領域を切り出す場合に、顔の存在場所と大きさに関係しない精度良く切り出すためには、次のような工夫が必要である。

1. 顔の大きさに依存しないで顔領域を切り出すために、画像をピラミッド化して解像度を多重化しておく。これにより、大きな顔も小さな顔も切り出すことができる。
2. 顔の存在場所を探すために、画像上で移動しながら部分空間法により、顔としての出力値が大きいところを顔領域として切り出す。
3. 切り出しの精度をあげるために、顔データと顔でないデータ（以後、非顔データと呼ぶ）の両方の部分空間をつくる。これにより、単に顔としての出力が大きい領域を切り出すのではなく、顔でないとする出力値と比較して顔領域を切り出すことにより精度が向上する。

人間の目の機能には中心視と周辺視がある。部分空間法による顔領域の検出は、顔の細部を認識しない検出法であり、周辺視に近いと考えられる。これに対して、目、鼻、口といった局所的な特徴を認識しつつ顔領域を精度良く検出する中心視の考え方も成立する。

本稿の後半では、部分空間法によって切り出した顔領域に対して目、鼻、口といった局所の特徴を認識しつつ顔領域を確定する手法について述べる。

2 部分空間法による顔領域の学習と判定

2.1 部分空間法

入力画像内のある領域が、顔であるか非顔であるかを判別する問題を考える。部分空間法では、まず始めに顔と非顔のカテゴリに対して学習データより正規直交基底 $V_i = \{v_{i1}, \dots, v_{ir}\}$ を求める。ここで、 r は部分空間の次元数である。顔データに対する正規直交基底は、学習に用いる顔データとの全距離が最小になる軸として求められる。実際には、顔データの相関行列を固有値分解したときの固有ベクトルとして求めている。同様に非顔のカテゴリに対しても正規直交基底を求める。

部分空間法により、入力画像のある領域が顔であるか否かを判定するには次のようにして行う。まず、入力画像のある領域を1次元データ x として表し、このデータ x と顔/非顔のカテゴリ ω_i の部分空間 V_i との距離を考える。入力データ x を部分空間 V_i に射影する射影行列を P_i とすると、距離は次式で示される。

$$\text{Dist}(V_i, x) = \|P_i x - x\| \quad (1)$$

射影ベクトル $P_i x$ は $\sum_{k=1}^r v_{ik}(v_{ik}^T x)$ と表されるので、射影行列 P_i は次のように定義される。

$$P_i = \sum_{k=1}^r v_{ik} v_{ik}^T \quad (2)$$

式(2)は次のようにも書くことができる。

$$P_i = V_i V_i^T \quad (3)$$

式(3)は、射影行列 P_i が観測空間における顔/非顔の学習データに対する正規直交基底より求められることを示している。いったん射影行列 P_i が求まると、入力データ x と顔/非顔の部分空間との距離は式(1)を用いて計算できる。今、顔の部分空間への射影行列を P_1 、非顔の部分空間への射影行列を P_2 とする。式(1)により、入力データ x と顔の部分空間との距離、入力データと非顔の部分空間との距離を求め、入力データに最も近い部分空間のカテゴリを判定結果とする。

実際には、部分空間との距離のかわりに、部分空間への射影ベクトルの大きさ $\|P_i x\|$ を計算している。入力データ x は、次式のように部分空間に射影した射影ベクトルの大きさが最大になるカテゴリとして分類される。

$$x^T P_1 x > x^T P_2 x \quad \text{then face} \quad (4)$$

2.2 学習と判定の処理過程

2.1で述べた部分空間法による顔判定のステップは次のようにまとめられる。

[Step1] 顔の学習データ $x_k (1 \leq k \leq N)$ から相関行列 S_i を求める。

$$S_i = \frac{1}{N} \sum_{k=1}^N x_k x_k^T \quad (5)$$

[Step2] 相関行列 S_i を固有値分解して固有ベクトルを求め、これを列ベクトルとする直交行列 V_i を求める。

$$S_i = V_i \Sigma V_i^T \quad (6)$$

[Step3] V_i より射影行列 P_i を求める。

$$P_i = V_i V_i^T = \sum_{k=1}^r v_{ik} v_{ik}^T \quad (7)$$

[Step4] 次の規則により判定を行う。

すべての $j \neq i$ にたいして

$$x^T P_i x > x^T P_j x = \|P_j x\|^2 = \sum_{k=1}^r (x^T v_{jk})^2 \quad (8)$$

ならば、 x をカテゴリ ω_i に属するものとして判定する。

2.3 学習データ

学生アルバムから顔画像を作成した。画像サイズは 72×90 画素であり、学習データとして図1(a)に示すように背景を極力含むことのないように顔画像を切りとった。また顔の大きさには個人差があるため、切りとった画像を全て 8×10 画素のサイズにアフィン縮小変換した。本研究では男 50 人、女 50 人の計 100 人分の顔画像データを作成した。

また非顔データとしては、処理対象画像から任意の大きさで背景を切り出し、顔画像と同じく 8×10 画素の大きさにアフィン縮小変換して、100 種の学習データを作成した。背景は、図1(b)に示すように、研究室内の一般的な背景である。例えば、白い壁、ワークステーションのディスプレイ、机などが含まれている。

また、体のうち顔以外の部分を切り出すことを防ぐため、顔の一部分、体の一部分、服なども背景画像として含めている。



(a) 顔データ

(b) 非顔データ

図1: 学習用データ

3 部分空間の決定と次元数

3.1 実験条件

部分空間としてどのような空間を設定すれば顔と非顔を精度よく判定できるかを実験により調べた。濃淡画像、モザイク画像、微分画像、微分モザイク画像を用いて実験を行なった。各画像の特徴を以下に示す。

1. 濃淡画像

256 階調の画像である。

2. モザイク画像

濃淡画像を平滑化した画像で、微細な特徴やノイズをとり除くことができる。(低域通過画像)

3. 微分画像

濃淡画像を微分してエッジを抽出した画像で、濃淡の濃さに依存せず特徴を表現できる。(高域通過画像)

4. 微分モザイク画像

濃淡画像を微分して平滑した画像で、濃淡の濃さに依存せず特徴を表現できるとともに、平滑化によって、微細な特徴やノイズをとり除くことができる。(帯域通過画像)

3.2 実験結果

4種類の画像に対して、顔と非顔の学習データ各々100枚を用いて部分空間を作成した。部分空間の次元数は顔と非顔では同数とし、新たに顔と非顔のデータ各々100枚を用意して判定を行なった。部分空間の次元数をパラメーターとして、判定結果を表1から表4に示す。表中、顔=顔と非顔=非顔は、顔と非顔を各々正しく判定した割合である。認識率はその平均を表している。

表 1: 濃淡画像 (%)

次元	顔 = 顔	非顔 = 非顔	認識率
2	100	98	99.0
4	100	98	99.0
6	100	98	99.0
8	100	99	99.5
10	100	100	100.0
20	100	100	100.0
30	100	99	99.0
40	100	99	99.0
50	99	99	99.0
60	100	96	96.0

表 2: 微分画像 (%)

次元	顔 = 顔	非顔 = 非顔	認識率
2	93	95	94.0
4	97	95	96.0
6	98	93	95.5
8	99	92	95.5
10	99	93	96.0
20	98	89	93.5
30	95	86	90.5
40	95	83	89.0
50	93	66	79.5
60	86	51	68.5

表 3: モザイク画像 (%)

次元	顔 = 顔	非顔 = 非顔	認識率
2	100	98	99.0
4	100	97	98.5
6	100	100	100.0
8	100	98	99.0
10	100	99	99.5
20	100	99	99.5
30	100	100	100.0
40	100	98	99.0
50	100	98	99.0
60	100	97	98.5

他のデータの場合と比べて微分画像の場合の認識率が低い。これは、微分した画像を8×10画素という小さい画像に縮小しているため、情報の欠落があったと考えられる。モザイク画像の場合、100%を示す次元数が6次元と30次元にばらつきを生じ、かつその前後の次元において非顔の認識率が低い。微分モザイク画像の場合には、20次元前後の次元数では認識率が低く信頼性に欠ける。

濃淡画像の場合、10または20次元を頂点とし、次元数が低くなくても高くなっても認識率は低下する。

安定性から考えて濃淡画像の10次元での判定結果が高い認識率を示し、最良の次元数であると考えられる。

表 4: 微分モザイク画像 (%)

次元	顔 = 顔	非顔 = 非顔	認識率
2	96	96	96.0
4	100	97	98.5
6	100	98	99.0
8	100	98	99.0
10	100	99	99.5
20	100	100	100.0
30	100	97	98.5
40	100	98	99.0
50	100	98	99.0
60	98	96	97.0

4 部分空間法による顔領域の切り出し

4.1 サーチ用ピラミッド

顔の大きさに依存しない顔領域を切り出すために、入力画像を順次縮小した多階層のピラミッド構造を用いる。指定した階層数を n とすると、入力画像に対して次の倍率

$$\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, \frac{n}{n} \quad (9)$$

でアフィン縮小変換し、図2に示すようなピラミッドを形成する。このピラミッド群に対して8×10画素の同一サイズで顔領域をサーチする。

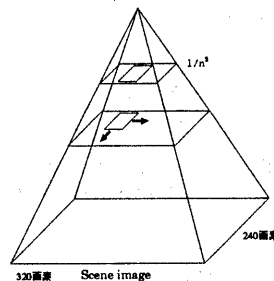


図 2: 多階層ピラミッド

4.2 複数階層の統合方法

ピラミッド構造の各階層で、8×10画素の領域を窓領域としてラスタ走査する。この窓領域内のデータを80次元ベクトル x とし、部分空間法の判定規則に従い、顔と非顔の部分空間へ射影する。この射影成分による分類に棄却類をもたせるため、射影成分を x のノルムで正規化する。こうして正規化された顔、非顔の部分空間への射影成分が小さい場合には、顔でも非顔でもない判定する。また、どの程度顔らしいかの信

信頼度を式(10)により算出する。 P_1 は顔の射影行列であり P_2 は非顔の射影行列である。

$$\text{顔らしさの信頼度} = \frac{x^T P_1 x}{x^T P_1 x + x^T P_2 x} \quad (10)$$

この信頼度を、ピラミッドの各階層、各窓位置で計算する。もし $x^T P_1 x > x^T P_2 x$ の条件を満たしていれば、原画像の対応する位置にこの信頼度を格納する。ピラミッドの全ての階層で信頼度の計算を終了すると、全階層で対応する画素上の信頼度を全て足し合わせて統合する。最後に、この値がある閾値より高い所を顔として判定し切り出す。

5 顔領域の切り出し実験と評価

3で述べた実験結果を基に、切り出し実験を行なった。実験では2.3で述べたように、顔データ100例と非顔データ100例の濃淡画像から、各々10次元の部分空間を作成した。切り出しの対象画像は、図3に示すような 320×240 画素の画像で、顔の大きさや人数の異なる画像50枚である。図3の入力画像に対して、信頼度を統合した結果を図4に、また切り出し結果を図5に示す。

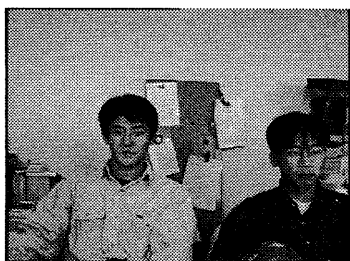


図 3: 顔領域の切り出し実験に用いた入力画像

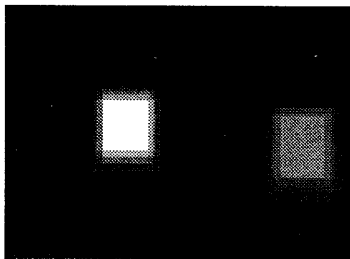


図 4: 各階層の信頼度を統合した画像

表5に切り出し精度の結果を示す。表5で検出率とは真に存在する顔領域のうち、正しく切り出せた割合で

ある。検出もれとは、切り出せなかった顔領域の割合である。また、誤検出とは、顔として切り出した領域のなかで、顔でない領域の割合である。閾値とは、図4に示すような統合画像から、顔領域と非顔領域を分けするためのものである。実験結果から、閾値は80が最良であることが分かる。閾値が高くなると検出もれが増加し、閾値が小さくなると誤検出が増える。

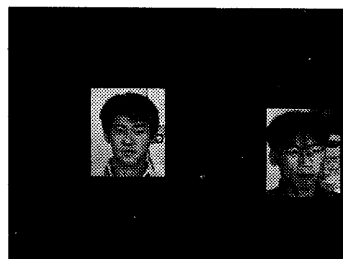


図 5: 顔領域切り出し結果

表 5: 部分空間法による切り出し結果 (%)

閾値	検出率	検出もれ	誤検出
60	80	20	26
80	80	20	24
100	78	22	24

6 特徴の認識に基づく顔の検出

6.1 中心視と周辺視

人間の視野には解像度の違いがあり、機能的な役割が異なっていることが知られている。すなわち視野の周辺では、対象の特徴を取り出す役割を持っており、中心部では対象の認識を行なうために、特徴を注視する役割を持っている[11]。

この目の機能的な役割の違いを考えると、顔全体を一つのパターンとして検出する部分空間法は顔の局所的な特徴を用いた処理に対応しており、人間の目でいう周辺視にあたると考えられる。これに対し中心視では、顔の局所の特徴である目や鼻、口などの特徴を用いて精度良く切り出しや認識を行なっているものと考えられる。

この節では、部分空間法により抽出した顔領域に対して、顔の局所の特徴を検出しながら顔領域を確定する手法について述べる。

6.2 自己組織型ネットワークによる顔領域の確定

顔領域を確定するには二つのステップを実行する。はじめのステップは、その顔領域から目、鼻、口とい

た特徴を検出する処理であり、次のステップは検出された特徴とその存在場所（位相情報）を用いて顔を確定する処理である。顔の目、鼻、口といった特徴を検出するために、競合学習を用いて検出ネットワークを構成する。こうしてできたネットワークを目、鼻、口に対応付けて、それぞれ目ネットワーク、鼻ネットワーク、口ネットワークと呼ぶ。

顔画像を入力として与えたとき図6に示すように目、鼻、口、それぞれのネットワークの出力が位相情報とともに現れる。この出力結果を、相互結合学習により位相情報を学習させたネットワークに入力し、顔領域を確定する。

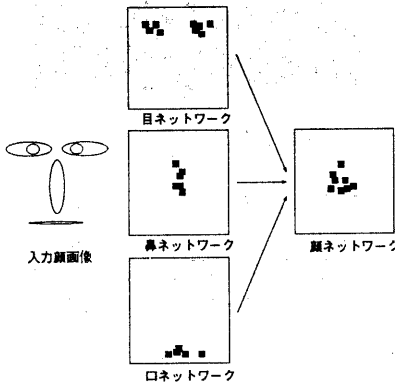


図6: 顔検出ネットワークの構成

6.3 競合学習による特徴抽出ネットワーク

目、鼻、口といった局所の特徴を学習させるために、競合学習ネットワークを用いる。ネットワークは図7のような構造をしている[12]。

j 番目の特徴抽出細胞 n_j の受容野を A_j とし、 i 番目の入力細胞 m_i との結合を w_{ji} とする。また、近傍の細胞との側抑制を u_{jk} 、競合領域を B_j とする。

特徴抽出細胞の出力値は次式によって求める。

$$y_j = \frac{V_j}{\sum_{k \in B_j} u_{jk} \cdot V_k} \quad (11)$$

$$V_j = \frac{1}{\sum_{i \in A_j} (w_{ji} - x_i)^2} \quad (12)$$

$$u_{jk} = \exp\left(-\frac{d_{jk}^2}{\sigma_u}\right) \quad (13)$$

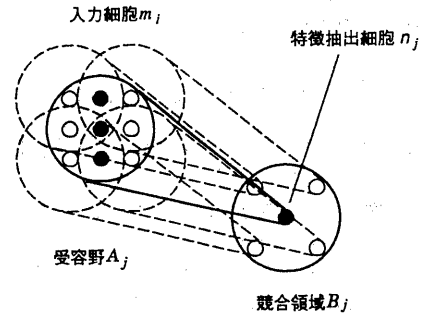


図7: 特徴抽出細胞

- x_i : 入力細胞 m_i の出力値
- y_j : 特徴抽出細胞 n_j の出力値
- V_j : 細胞 n_j のポテンシャル
- σ_u : 分散
- d_{jk} : 細胞 n_j と細胞 n_k の相対的な距離

ネットワークの側抑制の強さは、(13) 式のように距離の単調減少関数となっていて、これによって近くの細胞ほど強く抑制が効くようになっている。ポテンシャル V_j は結合値と受容野の入力値のユークリッド距離の総和の逆数であり、類似度を表している。つまり、入力値が結合値に近いほどポテンシャルは高くなり、一致した場合 ($w_{ji} = x_i$)、 V_j は無限大となる。

出力 y_j については、細胞 n_j の受容野の入力値と結合値が一致し、かつ細胞 n_j の競合領域内 B_j で、結合値が受容野の入力値と一致する細胞が細胞 n_j 以外に存在しない時のみ $y_j = 1$ となり、他の競合領域内の出力細胞を抑制する。逆に n_j が $V_j \neq \infty$ で、その他の細胞の内1つでも $V_j = \infty$ であれば $y_j = 0$ となって抑制される。

このネットワークの学習は、以下の式で表されるように最急降下法で求めることができる。すなわち、入力細胞 m_i から特徴抽出細胞 n_j への結合 w_{ji} は、エネルギーを小さくするように修正する。

$$\Delta w_{ji} \propto -\frac{\partial E}{\partial w_{ji}} \quad (14)$$

$$= -\frac{\partial}{\partial w_{ji}} \left(\sum_l \frac{1}{\sum_{k \in B_l} u_{lk} V_k} \right) \quad (15)$$

係数を c_w とすると、

$$\Delta w_{ji} = g_j \cdot (x_i - w_{ji}) \quad (16)$$

$$g_j = c_w \sum_{i \in B_j} \frac{u_{ij} \cdot V_j^2}{(\sum_{k \in B_i} u_{ik} V_k)^2} \quad (17)$$

となる。

(16) 式は、結合値 w_{ji} を係数 g_j で入力 x_i に近づけることを意味している。この係数 g_j は (17) 式に示すように、近傍の細胞との競合状態によって求められ、結合 w_{ji} に与える影響を全てたしあわせた値となっている。

このネットワークを用いて、目、鼻、口をそれぞれ別々のネットワークにより学習させる。このようにして学習し終えたネットワークをそれぞれ目ネットワーク、鼻ネットワーク、口ネットワークと呼ぶことにする。顔画像を入力として与えた時、それぞれのネットワークの出力値は、その存在場所（位相情報）とともに現れる。これらの出力を図 6 に示すような顔ネットワークの入力として用いることで、顔のあるなし、さらに顔が画像中のどの位置にあるのかを検出することができる。

目、鼻、口ネットワークでは局所的な特徴を特徴抽出細胞で表現しているため、複数の特徴抽出細胞が同時に発火 ($y_j = 1$) する。従って、特徴抽出細胞が複数個発火していれば、目、鼻、口と判断することが可能となる。

6.4 顔検出ネットワーク

顔検出ネットワークは図 8 に示すように、目、鼻、口ネットワークからの出力結果をもとに位相構造を考慮しつつ顔領域を検出するネットワークである。実際には相互結合型ネットワークを用いて各細胞間の結合荷重を学習し認識する。

相互結合ネットワークの構成を図 8 に示す。

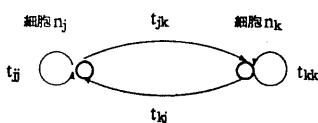


図 8: 相互結合ネットワークの構成

図中の丸が特徴抽出細胞を表しており、これが二次元平面上に並んでいる。これら特徴抽出細胞の出力値は以下の式に従って決定される。

$$y'_j = U\left(\sum_k t_{jk} \cdot y_k\right) \quad (18)$$

$$U(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (19)$$

ここで、 y_j は特徴抽出細胞 n_j の更新前の出力値であり、 y'_j は更新後の出力値である。ネットワークは、次のような Hebb の学習則に従って更新される。

$$\Delta t_{jk} = \begin{cases} c_i(s_j - t_{jk}) & (y_k > 0.0) \\ 0 & (y_k = 0.0) \end{cases} \quad (20)$$

$$s_j = \begin{cases} 1.0 & (y_j > 0.0) \\ -1.0 & (y_j = 0.0) \end{cases} \quad (21)$$

この顔検出ネットワークでは、目、鼻、口などの特徴抽出細胞の出力値が高い所で学習が進む。また、出力値の低い所では学習されず負の結合値となる。この顔検出ネットワークでは、入力の出現頻度に応じて学習されるため、目、鼻、口ネットワークの検出で生じた誤検出を解消することができる。

7 特徴の認識に基づく顔検出実験

7.1 目、鼻、口の検出実験

目、口についてはそれぞれ右側、左側のネットワークに分けて、目、鼻、口あわせて五つのネットワークを用いている。学習データは男性の目、鼻、口、3 人分をランダムに学習させる。また、検出実験で用いるテストデータは、100×120 画素の顔データで、濃淡値は 256 階調のモノクロ画像である。テストデータには、男性 60 人、女性 40 人の計 100 人が含まれている。



図 9: 目、鼻、口の検出結果

図 9 に目、鼻、口ネットワークによる検出結果の例を示し、表 6 に検出実験の結果を示す。また、このときの特徴抽出細胞の受容野の大きさは 5×5 画素である。図中、検出率とは目、鼻、口を正しく検出した割合で

表 6: 受容野の大きさが5×5 画素の場合の検出率 (%)

	検出率	検出もれ	誤検出
右目	88.0	12.0	47.0
左目	62.0	38.0	59.0
鼻	86.0	14.0	30.0
口の右側	41.0	59.0	80.0
口の左側	62.0	38.0	53.0
平均	67.8	32.2	53.8

あり、誤検出とは目、鼻、口でない所を検出した割合である。検出もれは画像中で、目、鼻、口を検出できなかった割合である。これらはテストデータの画像単位で計算される。例えば一枚のテスト画像で、検出及び誤検出の両方がある場合には、それぞれの数値がカウントされる。

特徴抽出細胞の受容野の大きさと検出率の関係を調べるために、特徴抽出細胞の受容野の大きさを、8×8画素と少し大きめのものを実験を行なった。検出結果を表 7に示す。

表 7: 受容野の大きさが7×7 の場合の検出率 (%)

	検出率	検出もれ	誤検出
右目	70.0	30.0	55.0
左目	84.0	16.0	54.0
鼻	90.0	10.0	18.0
口の右側	84.0	16.0	39.0
口の左側	71.0	29.0	45.0
平均	80.6	19.4	42.2

以上の結果より、特徴抽出細胞の受容野の大きさが5×5画素より、7×7画素の方が高い検出率を示すことが分かった。さらに、誤検出の割合も減っていることが分かる。

7.2 顔検出実験

表 8に顔検出ネットワークによる顔の検出結果を示す。テストデータとしては、顔データ 10 枚、非顔データ 10 枚をネットワークに与えた。表中の検出率は顔データを顔と判定した割合、検出もれは顔データを顔と判定しなかった割合、さらに誤検出は非顔データを顔データと判定した割合である。

表 8: 顔ネットワークによる顔の検出率 (%)

検出率	検出もれ	誤検出
90	10	10

8 おわりに

部分空間法と自己組織型ネットワークを用いて、顔データを学習し、シーン画像より顔の存在場所や大き

さに依存しないで顔領域を切り出し確定する方法を提案した。また、実験により濃淡画像が部分空間法に適しており、部分空間法の次元数は 10 次元が最良であることが分かった。

また、競合学習による複数のネットワークを用いて顔の検出実験を行ない、局所の特徴を検出しながら顔を検出することが可能であることを確かめた。今後はさらに学習データを増やし、顔の横向きや大きさに対応できるようにすることが課題である。

参考文献

- [1] 長谷川 浩史, 志水 英二: "パターンマッチングを用いた顔画像照合の一検討", 信学技報, IE92-40, (1992).
- [2] 長谷川 浩史, 那須 靖弘, 志水 英二: "多重解像度画像処理と snakes を用いた顔画像の特徴点抽出の方法", 信学技報, IE90-28, (1990).
- [3] 長谷川 浩史, 志水 英二: "Hough 変換を用いた顔画像照合の方法", 信学技報, PRU90-153, (1990).
- [4] 長谷川 浩史, 那須 靖弘, 志水 英二: "多重解像度画像処理 (全域的, 局所的) を用いた顔画像の特徴点抽出の方法", 信学技報, PRU89-26, (1990).
- [5] 宮沢 智司, 江島 俊朗: "顔画像におけるエッジ抽出の一手法", 信学技報, PRU87-115, (1987).
- [6] 矢敷 潤, 美濃 導彦, 池田 克夫: "並行に抽出された要素候補による投票を用いた線画の顔認識", 信学技報, PRU93-12, (1993).
- [7] 佐々木 努, 赤松 茂, 深町 映夫, 末永 康仁: "正面顔画像の自動識別法の検討", 信学技報, IE-91-50, (1991).
- [8] 赤松 茂, 佐々木 努, 深町 映夫, 末永 康仁: "KL 展開によるパタン記述法の顔画像識別への応用の評価", 信学技報, PRU90-152, (1990).
- [9] 赤松 茂, 佐々木 努, 深町 映夫, 末永 康仁: "濃淡画像マッチングによるロバストな正面顔の識別法", "フーリエスペクトルの KL 展開の応用", 信学論, J76-D-II No.7, pp.1363-1373, (1993).
- [10] 小林 信: "シーンの中の顔の探索と認識", 信学技報, PRU91-104, (1991).
- [11] 吉田 千秋, 豊田 雅信, 佐藤 幸男: "視野領域が機能分化した視覚系のモデル", 情報処理学会論文誌 Vol.33, No.8, pp.1032-1040(1992)
- [12] 久保 守, 岩田 彰, 鈴木 宣夫: "局所の特徴抽出細胞の自己形成とその相互結合回路", 信学技報 NC91 - 103, pp.34-42