ニューラルネットワークの学習とエネルギー極小化による動的領域抽出

Active-Region Extraction by Learning and Energy Minimization Using Neural Networks

岩田 治幸^{†*}, 正会員 長橋 宏[†]

Haruyuki Iwata^{†*} and Hiroshi Nagahashi[†]

Abstract In the process of extracting nonuniform target regions such as faces, it is necessary to identify their features. This paper proposes an active-region extraction method based on pattern recognition using a layered neural network. The neural network is used to extract features of previously given sample patterns. During the extraction, the neural network outputs the suitability of the extraction based on the features. Experimental results show that it is possible to extract some occluded parts associated with visible parts of target regions. It is also possible to extract face regions using this method.

キーワード：ニューラルネットワーク，パターン認識，動的領域抽出，誤差逆伝播アルゴリズム，変形度

1. ま え が き

画像の領域分割処理において，物体領域と背景領域との分割を行う場合に，その物体に関する知識をあらかじめ持つことで，良好な結果が得られる場合が多い．動画像を対象とする場合には，オプティカルフローやフレーム間差分画像を用いることで，物体領域を自動的に抽出できる場合がある¹⁾．しかし，静止画像を対象とする場合に，オプティカルフローやフレーム間差分画像を用いることはできない．従って，静止画像において，複数の均一な領域で構成される物体領域を，一まとまりにして抽出するためには，物体領域内のそれぞれの領域が持つ特徴や，領域間の関係などの情報を効率的に利用する必要がある．また，それらの特徴は，自動的に決定できることが望ましい．従来手法としては，画像を複数の均一な領域に分割した後に，一部の領域を統合することで，物体領域を構成する手法が考えられる．この場合のように，処理を分割と統合の2段階に分けてしまうと，ばやけや途切れのために分割の段階で得られない輪郭線は，処理の最後まで得ることができない．

筆者らは，Hopfield 型ニューラルネットワーク²⁾を用いた動的領域分割手法を既に提案した³⁾．文献 3) の手法は，Hopfield 型ニューラルネットワークのエネルギー関数値極小化原理を用いた，正則化に基づく手法である．この手法は，動的モデルを用いる手法⁴⁾⁵⁾と比較して，高精度な領域抽出が可能である．しかし，物体領域の特徴として色の均一性を仮定していたことから，複数の均一な領域で構成される物体領域を一まとまりにして抽出するためには，シードを複数与えて抽出した領域を，更に統合する必要がある．

そこで，本論文では，サンプルとしてあらかじめ与える物体領域の特徴を，階層型ニューラルネットワークの学習によって自動的に獲得した後に，獲得した特徴との適合性に基づき，動的に領域を抽出する手法を提案する．一般的に物体領域の特徴は，変形に伴い次第に失われていくと考えられる．そこで，変形の度合い（変形度）を，特徴に関する適合性を調べるために用いる．本手法における階層型ニューラルネットワークは，この適合性を出力するように学習する．ネットワーク構造および入力信号については 4 章で，学習データおよび学習過程については 5 章で述べる．

階層型ニューラルネットワークが獲得した特徴に基づく領域抽出は，階層型ニューラルネットワークが記憶したパターンの連想⁶⁾に相当する．したがって，連想するために何らかのキーが必要である．本論文では，従来の動的モデルで用いられてきたエッジ強度等に基づく画像エネルギーの最小化をキーとする．文献 3) で定義した目的関数において，階層型ニューラルネットワークの出力値を用いるこ

1997 年 11 月 20 日，電子情報通信学会技術研究報告，パターン認識・メディア理解研究会で発表．

1997 年 12 月 24 日受付，1998 年 3 月 25 日再受付，1998 年 5 月 7 日採録
† 東京工業大学 工学部画像情報工学研究施設

(〒 226-8503 横浜市緑区長津田町 4259, TEL 045-924-5478)

* 現在，陸上自衛隊 富士学校総合研究開発部

† Imaging Science and Engineering Laboratory, Faculty of Engineering, Tokyo Institute of Technology

(4259, Nagatsuta, Midori-ku, Yokohama 226-8503, Japan)

* Research & Development Division, Fuji School, Japan Ground Self-Defense Force

とで、変形度の認識と動的領域分割処理を統合する。そして、画像エネルギーの最小化と、階層型ニューラルネットワークが獲得した特徴との適合、すなわち、変形度の最小化を同時に満たすような抽出を、互いにキーとすることで求める。

実験では、階層型ニューラルネットワークが獲得した特徴との適合性を重視すれば、オクルージョンによる不可視部分についても、可視部分からの連想によって抽出できることを示す。また、不可視部分を抽出する場合としない場合との切り替えが、パラメータの変更によって可能であることを示す。不可視部分を抽出できる能力は、従来の分割・統合の2段階処理では困難であった、ぼやけや途切れのある輪郭線検出に対して効果的である。更に、本手法を用いた顔領域抽出実験の結果について述べる。

2. 動的領域抽出手法の概要

本論文における動的領域抽出手法は、最初に、抽出する物体領域内の1画素をシード Q とする。次に Q を含む任意形状領域 A とその他の領域 B とに画像を分け、 A に属する画素にラベル1を、 B に属する画素にラベル0を割り当てる。 A あるいは B 内で、他方の領域に隣接する画素に対して、Hopfield 型ニューラルネットワークの1つのユニットをそれぞれ割り当て、ユニットの状態更新結果に従ってラベルを変更していくことで領域を変形させる。最後に、収束状態におけるラベルの状態を領域抽出結果とする。

本手法では、初期領域内に物体領域が2つ以上存在していても、1つの物体領域、すなわちシードを含む領域を分離抽出することができる。本論文における Hopfield 型ニューラルネットワークの目的関数は、境界線に沿った領域特徴量の連続性、領域の隣接する画素におけるエッジ強度、そして、領域の変形度に基づき定義される。ただし、本論文における近傍領域は、注目画素を中心とする 5×5 画素の領域とする。

処理する画像が、図1の P に示すような $X \times Y$ 画素の濃淡画像 $P(x, y) (x \in \{0, 1, \dots, X-1\}, y \in \{0, 1, \dots, Y-1\})$ である場合に、本論文における動的領域抽出アルゴリズムは、次の(1)から(4)に示す処理ステップからなる。

図1 AとBの隣接画素とユニット
Adjacent pixels of A and B, and units.

- (1) 抽出しようとする領域にある1画素 Q を与える。
- (2) 領域 A あるいは B 内で、他方の領域に隣接する画素に対して、図1に示すように、0または1を出力する Hopfield 型ニューラルネットワークの1つのユニットをそれぞれ割り当てる。図1の $U_{R,j} (R \in \{A, B\}, j \in \{0, 1, \dots, N_R-1\}, N_R: \text{ユニット総数})$ は、領域 R を選んだ場合に割り当てるユニットを表している。
- (3) ランダムに選んだ第 j ユニットの状態更新の結果、その出力値 $U_{R,j}$ が0の場合に、ラベルを0ならば1に、あるいは1ならば0に変更する。
- (4) 3章で述べる Hopfield 型ニューラルネットワークのエネルギー関数値が収束するまで、処理ステップ(2)と(3)を繰り返す。

3. 領域変形度認識と動的領域抽出の統合

本手法は、領域 A の特徴がサンプル物体領域の特徴にできるだけ適合するように、領域 A を動的に変形させる。ここで、サンプル物体領域を変形させることを考える。サンプル物体領域を変形させた領域を入力信号とした場合に、変形度に応じて小さな値を出力するような階層型ニューラルネットワークを構成できれば、その出力値を大きくするように入力信号、すなわち領域 A を変形させることで、最終的にサンプル物体領域の特徴に最も適合する領域 A が得られる。ただし、できるだけ多くの変形過程を学習用データとして呈示すると共に、汎化能力を有する学習アルゴリズムを用いる必要がある。本手法では、学習用データを自動的に生成し、誤差逆伝播アルゴリズム⁷⁾を用いて学習を行う。

一方、Hopfield 型ニューラルネットワークにおいては、 n 回目の更新における第 j ユニットの出力値 $U_j(n)$ を(1)式に従って決定すれば、(2)式に示すエネルギー関数 E_{NN} の値は決して増加しない。

$$U_j(n) = \begin{cases} 1 & \text{if } \sum_{k=0}^{N-1} w_{jk} U_k(n-1) - h_j > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$E_{NN} = -\frac{1}{2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} w_{jk} U_j U_k + \sum_{j=0}^{N-1} h_j U_j \quad (2)$$

w_{jk} : 第 k ユニットから第 j ユニットへの結合荷重。

h_j : 第 j ユニットの閾値。

本論文では、第 j ユニットを割り当てる画素 $P_{unit}(x_j, y_j)$ の近傍領域で、 $P_{unit}(x_j, y_j)$ と同じラベルを割り当てられている画素を $P_{same}(x_j + u, y_j + v), (u, v \in \{-2, -1, 0, 1, 2\})$ 、異なるラベルを割り当てられている画素を $P_{diff}(x_j + u, y_j + v)$ 、 $P_{same}(x_j + u, y_j + v)$ の平均階調値を $\overline{P_{same}}(x_j, y_j)$ 、 $P_{diff}(x_j + u, y_j + v)$ の平均階調値を $\overline{P_{diff}}(x_j, y_j)$ とする。

階層型ニューラルネットワークを用いた領域変形度の認

識と、Hopfield 型ニューラルネットワークを用いた動的領域抽出手法を統合する。複数の領域を一まとまりにして抽出する場合には、異なる領域が隣接する部分で、領域特徴量が輪郭線に沿って不連続になる場合がある。そこで、統合に際して、文献 3) で、最適な領域抽出において各領域が満たすべきであると仮定した、仮定条件 1「 $\overline{P_{same}}$ は境界線に沿って連続 (continuous) である。」を、「 $\overline{P_{same}}$ は境界線に沿って、ある一定区間において連続 (continuous) である。」に変更する。更に、文献 3) の仮定条件 3「領域 A において、領域 B に隣接する画素の近傍領域とシード近傍領域の平均階調値が適合 (fit) している。」を、「領域 A の変形度が小さい。」に変更する。すなわち、階層型ニューラルネットワークがサンプル物体領域から獲得した特徴に、領域 A の特徴が適合していることを仮定する。したがって、階層型ニューラルネットワークの出力値が大きくなる、すなわち、1.0 に近づくように領域 A を動的に変形させることになる。仮定条件 2「 P_{unit} におけるエッジ (edge) 強度は、その他の画素よりも大きい。」については、輪郭に沿って異なる領域が隣接する部分で、多少の不整合を生じる場合がある。実験では、仮定条件 3 を重視することで、ある程度これに対処できることを確認したが、最適な仮定条件 2 については、今後更に検討が必要である。

$U_{R,j} = 0, (R \in \{A, B\})$ であるとき、 $P_{unit}(x_j, y_j)$ が領域 R に属さない、すなわち、ユニットを割り当てる画素として存在しないことを、 $U_{R,j} = 1$ であるとき、存在することを仮定すれば、(3) 式から (6) 式で定義する E_{cont} , E_{edge} , そして E_{fit} の値が小さくなるほど、仮定条件 1 から仮定条件 3 にそれぞれ近付いたといえる。

注目画素 $P_{unit}(x_j, y_j)$ から、境界線に沿ってある一定距離離れた 2 つの $\overline{P_{same}}(x_k, y_k)$ の、少なくとも一方が $\overline{P_{same}}(x_j, y_j)$ と等しい場合に、値が 0 になるような E_{cont} を (3) 式に示すように定義する。本論文では、予備実験の結果に基づき、この距離を 5 画素に設定する。

$$E_{cont} = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(\delta_{j+5,k} \cdot |\overline{P_{same}}(x_j, y_j) - \overline{P_{same}}(x_k, y_k)|, \delta_{j-5,k} \cdot |\overline{P_{same}}(x_j, y_j) - \overline{P_{same}}(x_k, y_k)|) U_j U_k \quad (3)$$

$$f(a, b) = \begin{cases} a & \text{if } a \leq b \\ b & \text{otherwise} \end{cases} \quad (4)$$

$\delta_{j,k}$: クロネッカーのデルタ。

$|\overline{P_{same}} - \overline{P_{diff}}|$ を、 P_{unit} におけるエッジ強度とする。第 j ユニットの割り当てた画素のラベルを変更したときのエッジ強度の変化 $\Delta|\overline{P_{same}} - \overline{P_{diff}}|$ を用いて、 E_{edge} を (5) 式に示すように定義する。

$$E_{edge}$$

$$= \sum_{j=0}^{N-1} \Delta|\overline{P_{same}}(x_j, y_j) - \overline{P_{diff}}(x_j, y_j)| U_j \quad (5)$$

n 回目の更新における抽出パターンを入力した場合の階層型ニューラルネットワークの出力値を O_n , 第 j ユニットの割り当てた画素のラベルを変更した後の抽出パターンを入力した場合の階層型ニューラルネットワークの出力値を $O_n(j)$ とする。そして、 E_{fit} を (6) 式に示すように定義する。

$$E_{fit} = \sum_{j=0}^{N-1} (O_n(j) - O_n) U_j \quad (6)$$

Hopfield 型ニューラルネットワークの目的関数 E_{obj} を、(7) 式に示すように定義する。

$$E_{obj} = \alpha E_{cont} + \beta E_{edge} + \gamma E_{fit} \quad (7)$$

α, β, γ : 正の実数。

E_{obj} の最小値を与えるラベルの状態を、最適な領域抽出結果であると仮定する。 E_{obj} が E_{NN} と恒等的に等しくなるように、係数比較によって w_{jk} および h_j を決定して Hopfield 型ニューラルネットワークを構成すれば、ランダムに選んだユニットの状態を (1) 式に従って更新していくことによって、 E_{obj} が極小値になる U_j を求めることができる。 w_{jk} を (8) 式に、 h_j を (9) 式に示す。

$$w_{jk} = -2\alpha f(\delta_{j+5,k} \cdot |\overline{P_{same}}(x_j, y_j) - \overline{P_{same}}(x_k, y_k)|, \delta_{j-5,k} \cdot |\overline{P_{same}}(x_j, y_j) - \overline{P_{same}}(x_k, y_k)|) \quad (8)$$

$$h_j = \beta \Delta|\overline{P_{same}}(x_j, y_j) - \overline{P_{diff}}(x_j, y_j)| + \gamma \{O_n(j) - O_n\} \quad (9)$$

4. 階層型ニューラルネットワークの構造と入力信号

本論文における階層型ニューラルネットワークは、図 2 に示すように、入力層、隠れ層、出力層の 3 つの層からなる。

変形の途中にある領域 A だけを用いて特徴を獲得すると、例えば、図 3 に示すような場合に、太い線で表す領域 A が物体領域の内部であることを認識できない。そこで、本手法では、図 2 に示すように、領域 A と B , それぞれについて入力ユニットを用意して、領域 A と B の変形によって失われる特徴を共に獲得する。

本手法では、物体領域サイズの影響を受けないようにするために、画像 P 全体ではなく、領域 A を含む P の一部を階層型ニューラルネットワークへの入力パターンにする。しかし、図 4 の I で示すような矩形領域を入力パターンにすると、 A の輪郭と I の辺とが一致する部分では、 A が物体領域の内部であることを認識することが困難になる。そこで、本手法では、 A に外接する矩形領域よりも、大きめの領域を入力パターンとして用いる。本論文で

は、図2に示すように、縦横10画素分大きな $W \times H$ 画素($W = x_{max} - x_{min} + 10, H = y_{max} - y_{min} + 10$)の領域を入力パターンにする。そして、 A を含む $W \times H$ 画素の領域を $L \times L$ 個のブロックに分割する。本論文では、図2に示すように $L = 9$ に設定した。図2に示す $d_{x,0}, d_{x,1}, d_{x,2}, \dots, d_{x,8}$, および $d_{y,0}, d_{y,1}, d_{y,2}, \dots, d_{y,8}$ は、次式に従って求める。ただし、 $(W\%9), (H\%9)$ は、 W, H をそれぞれ9で割った余りを表す。

$$d_{x,0} = d_{x,8} = W/9 + (W\%9)/2 \quad (10)$$

$$d_{x,i} = W/9, (i = 1, 2, 3, \dots, 7) \quad (11)$$

$$d_{y,0} = d_{y,8} = H/9 + (H\%9)/2 \quad (12)$$

$$d_{y,i} = H/9, (i = 1, 2, 3, \dots, 7) \quad (13)$$

そして、各ブロック内の画素の平均階調値を $[0.0, 1.0]$ に正規化した値を階層型ニューラルネットワークの入力ユニットへ、図2の矢印で示すように入力する。したがって、階層型ニューラルネットワークの入力層は、 $L \times L \times 2$ 個の入力ユニットからなる。本論文では、162個である。隠れ

図2 階層型ニューラルネットワークの構造と入力信号
Structure and input signals of the layered neural network.

図3 問題点1
Problem 1.

図4 問題点2
Problem 2.

ユニット数を多くすれば、学習用入力データに対しては正しい結果を出力するようになる。しかし、隠れユニット数が多過ぎると、一般的に汎化能力が低下する。隠れユニット数の決定法は、幾つか提案されているが⁹⁾、本論文では試行錯誤的に隠れユニット数を3に決定して、自動的決定は今後の課題とする。出力ユニットは1つにする。

ユニットの入出力関数は、(14)式に示すロジスティック関数 $f(u)$ とする。ただし、 u はユニットへの入力を表す。

$$f(u) = 1/(1 + e^{-u}) \quad (14)$$

5. 学習用データ自動生成と学習過程

本手法では、学習用データを自動的に生成する。変形一回につき、領域の隣接する1画素のラベルを変更する。サンプル画像における領域の隣接する画素の数を N 、変形回数を T とすれば、変形パターンは N^T 個存在する。すべての変形パターンを一括して学習できることが望ましいが、膨大な数の変形パターンを、学習用データとしてあらかじめ用意することは、記憶容量の限界のために困難である。そこで、変形パターンの生成および学習を1ステップとして繰り返す追加学習⁹⁾を行う。追加学習においては、文献8)で述べられているように、前回の例題を学習して得られた結合荷重および閾値を初期値として、新たな例題を完全に学習すると、新たな例題だけを学習した場合と同じになってしまう。そこで、本手法における追加学習では、学習が完全に行われることを仮定しない。できるだけ多くの変形パターンを用いることが望ましいが、本論文では2つの変形パターンを一括修正によって学習して、新たに生成した変形パターンを追加的に学習する方法を採用する。学習過程は次の(1)から(6)に示す処理ステップからなる。

- (1) 4章で述べた階層型ニューラルネットワークの結合荷重および閾値をランダムな値に設定して保存する。
- (2) 領域の隣接する画素をランダムに選択して、割り当てられたラベルを変更する。これを T 回繰り返して領域を変形させる。本論文における T は、物体領域と同じ面積の正方形領域を、物体領域と中心が一致するように重ねた場合に、重ならない部分の画素数とする。言い換えれば、物体領域を変形させて、この正方形領域にするための最小ラベル変更回数である。 T を9で割った回数毎、すなわちラベル変更

回数 $T/9 \times t$, ($t = 0, 1, 2, \dots, 9$) における変形領域を学習用データの入力部とする。ただし、背景領域の階調値には制約がない。すなわち背景領域の階調値が如何に変化しても物体領域を抽出できる必要がある。そこで背景領域における階調値は、0 から 255 のランダムな値にする。変形度を $0.1 \times t$, 出力部すなわち教師信号を $1.0 - 0.1 \times t$ とする。したがって、変形前の領域に対する教師信号は 1.0, 変形回数 T の領域に対する教師信号は 0.1 である。結果的に、変形領域とその変形度の組合せを自動的に生成する。これを学習用データセット I とする。

- (3) 処理ステップ (2) と同様にして、学習用データセット II を生成する。
- (4) 階層型ニューラルネットワークの結合荷重および閾値を読み込み、誤差逆伝播アルゴリズムを用いて学習を行う。学習用データセット I と学習用データセット II を提示した後に、結合荷重および閾値を一括修正する。本手法における追加学習では、学習が完全に行われることを仮定しない。そこで、本論文では、修正回数を 100 回に設定する。結合荷重の修正量 Δw は (15) 式および (16) 式に従って求める。ただし、閾値については、常に -1.0 を出力するユニットを入力層および隠れ層に、それぞれ 1 つずつ追加して、それとの結合荷重として修正する。

$$\begin{aligned} \Delta w_{nm}(t) &= -\eta \sum_{I, II} \{ (O_n - D) \cdot (1 - O_n) \cdot O_n \cdot O_m \} \\ &\quad + \theta \cdot \Delta w_{nm}(t-1) \end{aligned} \quad (15)$$

$$\begin{aligned} \Delta w_{ml}(t) &= -\eta \sum_{I, II} \{ (O_n - D) \cdot (1 - O_n) \cdot O_n \cdot w_{nm}(t) \cdot (1 - O_m) \cdot O_m \cdot O_l \} \\ &\quad + \theta \cdot \Delta w_{ml}(t-1) \end{aligned} \quad (16)$$

w_{nm} : 第 m 隠れユニット ($m \in 0, 1, 2, 3$) と出力ユニット間の結合荷重。

η : 学習係数。

D : 教師信号。

O_m : 第 m 隠れユニットの出力。

θ : 慣性項係数。

w_{ml} : 第 l 入力ユニット ($l \in 0, 1, 2, \dots, 162$) と第 m 隠れユニット間の結合荷重。

O_l : 第 l 入力ユニットの出力。

- (5) 階層型ニューラルネットワークの結合荷重および閾値を保存する。
- (6) 処理ステップ (2) から (5) を、すべてのサンプル画像について順に行い、追加例題を最初に提示した時の、階層型ニューラルネットワークの平均 2 乗誤差があらかじめ与える値 ϵ よりも小さくなるまで繰

り返す。

ただし、2 回目以降の処理ステップ (4) においては、新たに生成する学習用データセットを追加例題として、追加学習を行う。したがって、処理ステップ (2) および (3) で生成した学習用データセットを、追加例題として最初に提示した時の平均 2 乗誤差は、その時点における汎化能力を反映していると考えられる。

本論文では、予備実験の結果に基づき、学習係数 η の初期値を 1.0 にして、200 個の追加例題を学習する毎に、(17) 式に従って変更する。

$$\eta = \eta / 10.0 \quad (17)$$

また、慣性項係数 θ は、0.01 にする。

本論文では、図 5 に示す 4 つのサンプル画像を用意した。図 5(b), (c), (d) は、(a) に対する簡単なオクルージョンを想定して、それぞれ (a) の右部、上下部、左部を削除することで生成した。乱数の初期値を変更することで膨大な量の学習用データを生成できる。図 5(a) について生成した学習用データセットの例を図 6 に示す。

学習が進むにつれて、変形度が大きいほど小さな値を出力するような階層型ニューラルネットワークの結合荷重および閾値が得られる。すなわち、階層型ニューラルネットワークは、変形によって失われる、サンプル物体領域の特徴を獲得する。

6. 実験結果

図 5 に示したサンプル画像について学習を行った後に、得られた階層型ニューラルネットワークを用いて、図 7(a) に示す 70×70 画素 256 階調の濃淡画像において領域を抽出した。図 7(a) は、図 5(a) を家であると仮定して、その左後方にビル、右前方に樹木を配置したシーンを表すパターン画像である。Hopfield 型ニューラルネットワークの目的関数におけるパラメータは、予備実験の結果に基づき、 $\alpha = 0.1, \beta = 1.0, \gamma = 100.0$ に設定した。図 1 に示したように、+ で Q の位置を、正方形で囲んだ領域であらかじめ与えた領域 A の初期状態を表している。ユニットを割り当てた画素は白で表示している。画像エネルギー最小化と、サンプル画像から獲得した特徴に対する適合を同時に満たす抽出として、図 7(b) に示す結果が得られる。更に、サンプル画像の特徴に対する適合性をより重視して、パラメータ $\gamma = 3000.0$ にすると、図 7(c) に示す結果が得られ、木で隠れている部分を推定できることがわかる。したがって、階層型ニューラルネットワークの出力値を重視すれば、オクルージョンによる不可視部分を、可視部分からの連想によって抽出することが可能であると考えられる。図 7(c) は、分割・統合といった従来手法では、決して得られない結果である。図 5 に示したようなサンプル画像について学習すれば、階層型ニューラルネットワークは、サンプル画像の特徴として、上部が暗く下部が明るいといったような

図 5 サンプル画像
Sample images.

図 6 学習用データセットの一例
An example of training data set.

特徴を獲得すると考えられる。

以上の実験をふまえ、本論文では、顔領域を一まとまりにして抽出した結果について述べる。まず、図 5 のサンプル画像を学習した階層型ニューラルネットワークを用いて、図 8(a) に示す人物の顔領域を抽出した。図 8(a) の 100×100 画素の画像は、カラー標準画像“ミュージシャン”の R 成分である。特に頭髮部と背景の階調値がほぼ等しいために、分割・統合といった従来手法では抽出が困難な例であるが、図 8(b) に示すように良好に抽出することができる。更に、図 8(b) の顔領域を新たに学習して、図 9(a) および図 10(a) に示す 70×70 画素の画像において抽出した結果を、それぞれ、図 9(b) および図 10(b) に示す。背景が複雑な例であるが、顔領域の特徴を用いること

図 7 抽出結果
Extraction results.

図 8 顔領域抽出結果（標準画像）
An extraction result of a face region. (Standard image)

図 9 顔領域抽出結果（低画質画像 1）
An extraction result of a face region. (Low quality image 1)

で、良好に抽出できることがわかる。また、図 10(b) において、顔領域の左下部がうまく抽出できていない原因としては、背景に強いエッジが存在することと、図 8(b) における顔領域の特徴、特に首部分の特徴との適合が考えられる。したがって、画像エネルギーの最小化と、図 8(b) をサンプルとして獲得した特徴との適合を共に満たす抽出結果であるといえる。そこで、図 8(b) と図 9(b) の顔領域をサンプルとして学習した後に、図 10(a) の顔領域を再び抽出すると、図 10(c) に示す結果が得られる。したがって、さまざまな顔領域を抽出するためには、幾つかのパターンをあらかじめ学習しておく必要があると考えられる。

パラメータ γ 決定の指針について簡単に述べる。抽出しようとする物体領域が、サンプル画像との適合および画像エネルギーの最小化を同時に満たしている場合には、 E_{cont} , E_{edge} , そして、 E_{fit} の最小値を与える解は一致するため、 γ の影響は少ない。しかし、 E_{cont} , E_{edge} の最小解と E_{fit} の最小解が一致しないような場合には、 γ の値に応じて異なる抽出結果が得られる。例えば、図 7 の場合、 γ を大き

(a) Low quality image 2 (b) An extraction result 1

(c) An extraction result 2

図 10 顔領域抽出結果 (低画質画像 2)
Extraction results of a face region. (Low quality image 2)

くすることで、図 7(c) の結果が得られたが、この抽出結果は、画像エネルギーの最小値を与えるものではない。したがって、学習するサンプル画像と、オクルージョン等による物体領域の変化を考慮して、 γ を決定する必要がある。

7. む す び

学習によってあらかじめ獲得したサンプル物体領域の特徴を用いて、複数の均一な領域で構成される物体領域を、一まとまりにして抽出できる新しい手法を提案した。本論文の実験では、顔領域を抽出したが、より多くの学習用データを学習すれば、さまざまな物体領域を抽出することが可能であると考えられる。更に、ニューラルネットワークや遺伝的アルゴリズムによる物体抽出手法¹⁰⁾¹¹⁾を用いて、物体領域を矩形領域として抽出することで、物体領域の位置および大きさを推定できる。これらの手法を利用することで推定した位置・大きさと、学習用データ自動生成における変形回数 T に基づき、領域 A の初期状態を与えれば、自動抽出システムを構築することも可能である。

今後は、本手法をカラー画像に適用すると共に、学習と抽出精度の関係について考察する予定である。また、高速化と大域的解探索のための方策として、文献 3) と同様にピラミッド画像を用いることについても検討する予定である。

【文 献】

- 1) 荒木 昭一, 横矢 直和, 岩佐 英彦, 竹村 治雄: “複数物体の抽出を目的とした交差判定により分界する動的輪郭モデル,” 信学論 (D-II), **J79-D-II**, 10, pp.1704-1711 (Oct. 1996)
- 2) J. J. Hopfield and D. W. Tank: “Neural computation of decisions in optimization problems,” Biol. Cybern., **52**, pp.141-152 (1985)
- 3) 岩田 治幸, 長橋 宏: “ニューラルネットワークを用いたカラー画像における動的領域分割,” 信学論 (D-II), **J80-D-II**, 11, pp.2995-3003 (Nov. 1997)

- 4) M. Kass, A. Witkin and D. Terzopoulos: “Snakes: Active Contour Models,” IJCV, **1**, 3, pp.321-331 (1988)
- 5) 坂上 勝彦, 山本 和彦: “動的な網のモデル Active Net とその領域抽出への応用,” テレビ誌, **45**, 10, pp.1155-1163 (Oct. 1991)
- 6) T. Kohonen: “Self-organization and associative memory,” Second Edition, Springer-Verlag, Berlin (1987)
- 7) D.E. Rumelhart, G.E. Hinton, and R.J. Williams: “Learning representations by back-propagating errors,” Nature, **323**, 9, pp.533-536 (1986)
- 8) 池田 和司: “追加学習の漸近論的解析,” 信学論 (D-II), **J80-D-II**, 7, pp.1913-1918 (July 1997)
- 9) N. Murata, S. Yoshizawa, and S. Amari: “Network Information Criterion,” IEEE Trans., Neural Networks, **5**, 6, pp.865-872 (Nov. 1994)
- 10) 安居院 猛, 古久保 良隆, 長橋 宏, 長尾 智晴: “2 段階神経回路網を用いた白黒階調静止画像からの顔領域の抽出,” 画電学誌, **23**, 3, pp.167-175 (June 1994)
- 11) 榊永 慎哉, 長尾 智晴: “静止階調画像を対象とした人物の顔領域の自動抽出,” 印刷誌, **33**, 6, pp.377-382 (June 1996)

いわた はるゆき
岩田 治幸 1990 年, 防衛大学校応用物理学教室卒業。1998 年, 東京工業大学大学院博士課程修了。現在, 陸上自衛隊富士学校総合研究開発部勤務。画像処理, ニューラルネットワークなどの研究に従事。工学博士。

ながはし ひろし
長橋 宏 1975 年, 東京工業大学工学部電気工学科卒業。1980 年, 同大学大学院博士課程修了。1981 年, 山形大学工学部電子工学科助手。1987 年, 同助教授。1990 年, 東京工業大学工学部情報工学研究施設助教授。1998 年, 同教授。現在, 画像計測による 3 次元形状のモデリング, 画像処理エキスパートシステム, 自然言語処理, ニューラルネットワークなどの研究に従事。工学博士。正会員。