

階層型ニューラルネットワークによる 地域気象の内部表現

前多 修二

(受理：1995年11月30日)

On a possible internal representation of areal climate by artificial
multilayered neural network

Shuji MAEDA

摘 要

今年4月上旬から11月上旬までの間の166日の正午の気象を対象として、気象衛星の雲画像が含む地域気象の情報を、多層ニューロネットワークによって有効に引き出す枠組みを試みた。構造が最も簡単な3層ネットワークを用い、北海道中央の矩形部分の赤外線強度分布を入力層とし、出力層において同時刻の矩形地域内3都市の天気表現されるよう、ネットワークに学習を行わせた上で、雲画像の情報が中間層においてどのように変換されているかを調べた。その結果、同じ出力値(3都市の天気の組み合わせ)を導く雲画像が、中間層において複数のクラスターを形成していることがわかった。各クラスター内では、互いに似た情報の表現になっている。もしこれが気象の何らかの分類に対応しているとする、ニューラルネットワークがその内部的表現において、カテゴリーを自律的に形成する事例と考えられる。

キーワード： 地域気象、衛星雲画像、3層ニューラルネットワーク、内部表現、クラスタリング

緒 言

気象衛星の雲画像は気象現象の監視や天気予報等に広く活用されているが、そこでは主に、別の観測データや気象に関する既知の概念と雲画像との対応関係を通して大気の様子が把握されている。一方、近年の異常気象は、大気運動の本質的な非線形性との結びつきを含み、従来の知識体系ではとらえ難い要因の存在を示唆している。そこで、衛星で観測される大気に関する

大量な情報を従来の知識にとらわれない形で解析する枠組みが、相補的な解析手順として望まれる。

ニューロコンピューティングは人工知能の一つの柱をなし、近年多方面に応用が試みられている。その利点は、言葉を経由することなしに、専門家がおこなっている判断の仕組みを計算機の中に作り上げることにある。即ち、事例に基づき、学習というメカニズムを通すことによって、ブラックボックス的に判断論理を作り上げる。ニューロシステムを気象解析に応用した

¹帯広畜産大学 教養課程 (〒080 帯広市稲田町)

²Liberal Arts, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido, 080, Japan

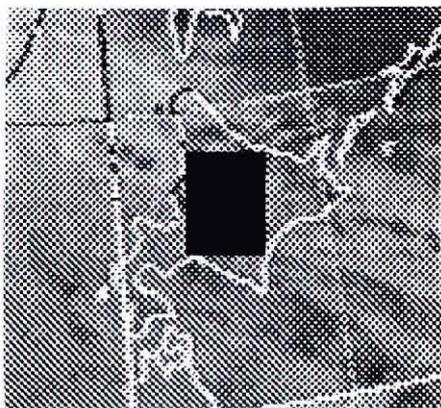
ときに、その論理が具体的な気象の状況を内部的にどのように表現し、どのように把握しているかが見えれば、大きな力を持つと考えられる。その意味で、ニューラルネットワーク独自のカテゴリー分けが自律的に形成されるか否かは重要なポイントである。

本研究では、北海道中央部の山岳地帯を含む矩形地域の雲画像を、多層ニューラルネットワークの出力層において同時刻の地域内の天気と強制的に結び付けたときに、雲画像のデータが中間層においてどのように変換されているかを調べた。雲画像に含まれる大量の地域気象情報が自律的に分類されるか否かの手がかりを得るために、中間層における表現ベクトル間の相似性においてクラスタリングが現れるか否かに焦点を当てて調査を行った。

方 法

3層のニューラルネットワークモデルでは、ネットワークに入力された情報は入力層から先ず中間層へ変換され、その後中間層から出力層へ変換される。入力層のデータとして雲画像を採用し、出力層をその地域の複数箇所の天気に対応させた。

図1 赤外線雲画像取得地域



雲画像として、図1に示す矩形部分を覆う縦27×横21=567地点の正午の赤外線強度を採用した。この地域は北海道の中央の山岳地帯（大雪山系、夕張山地、及び日高山脈の北半分）を含む。出力層を、同時刻の、この山岳地帯周辺の3都市（旭川、岩見沢、帯広）の天気に対応させるよう設定した。ただし、解析の対象事例としては

1. 快晴、
2. 晴、
3. 曇、
4. 雨

のいずれかである場合を採用した。

赤外線強度の尺度は0～63の64段階がとられている。一般に、数値がより大きいと、雲がより多く存在すること、あるいは、より高い位置に存在することを意味する。式の都合上、これらの値の、中央値32からのずれを入力層の各ユニット (x_i , $i = 1 \dots 567$) に入る値とした。出力層は、12個のユニット (z_k , $k = 1 \dots 12$) からなる。例えば、旭川が晴、岩見沢が曇、帯広が快晴であることは

0	1	0	0
0	0	1	0
1	0	0	0

という順の12個の数値で表現される。

本研究の趣旨から、中間層の次元数として入力層の次元数と出力層の次元数の中間にとり、 $10 \times 10 = 100$ とした (y_j , $j = 1 \dots 100$)。従って、入力層→中間層→出力層、という順に情報圧縮がすすむことになる。

入力層の情報は中間層、出力層において

$$y_j = f \left(\sum_{i=1}^{567} x_i w_{ij} \right)$$

$$z_k = f \left(\sum_{j=1}^{100} y_j v_{jk} \right)$$

という形に逐次変換されていく。ただし、関数 f はシグモイド関数

$$f(x) = 1 / (1 + \exp(-x))$$

で、階段関数を有界な疑似線形変換部分を持つように拡張したものである。 w_{ij} は入力層と中間層のユニット間を結ぶ荷重、 v_{jk} は中間層と出力層のユニット間を結ぶ荷重である。中間層、出力層のユニットには0から1までの間の値が入る。荷重の初期値は一様乱数によって設定した。

各雲画像に対して「正しい」出力値 p_k を出力しなければならないという条件がネットワークに対して課せられる。「正しい」出力値 p_k とネットワークが実際

に出力する値との誤差を

$$r = \sum_k (z_k - p_k)^2$$

で定義する。出力層のユニット値全体を一つのベクトルとみなすと、 r は、ニューラルネットワークが雲画像の表現として出力するベクトルと「正しい」ベクトルとの距離の2乗である。この誤差はバックプロパゲーション¹⁾というメカニズムにより、即ち

$$w_{ij} \rightarrow w_{ij} + \Delta w_{ij}$$

$$v_{jk} \rightarrow v_{jk} + \Delta v_{jk}$$

$$\Delta w_{ij} = -\varepsilon \partial r / \partial w_{ij}$$

$$\Delta v_{jk} = -\varepsilon \partial r / \partial v_{jk}$$

に従って荷重 w_{ij} 、 v_{jk} を逐次変化させていくことにより、減少させることができる。ただし、 ε は適当に小さい正の値を選ぶ。 ε の値を必ずしもすべての荷重に対して同じものにとる必要はないが、本研究では簡単のため共通にとることとした。このメカニズムは、層から次の層への情報の変換の仕方を、最後の出力層において「正しい」答を出すように修正するので、ニューラルネットワークの学習と呼ばれる。

降雪の少ない1995年の4月から10月までの期間(213日)における3都市それぞれの正午の各天気の日数は表1のようであった。

表1 4～10月の天気統計

天気	旭川	岩見沢	帯広
快晴	18	17	21
晴	52	55	66
曇	107	108	83
霧雨	0	1	0
雨	31	30	41
雷	3	0	0
雪	2	2	2

ニューラルネットワークの画像が取得でき、また天気上記の条件を満たした166日間の正午のデータを採用した。

これら166日の3都市の天気を正解とし、これにできる限り近い出力を導くように約10万回の学習を行わせた。

表2に3都市の天気の組み合わせの頻度の多いものを示す。最も多かったのは3都市すべてが曇りという場合である。

表2 天気の組み合わせ統計

旭川	岩見沢	帯広	日数
曇	曇	曇	40
曇	曇	雨	11
晴	晴	晴	11
曇	晴	晴	10
雨	雨	雨	9
曇	曇	晴	7
晴	曇	晴	6
晴	晴	曇	6

一般には、3地点の天気と同じという条件下において、様々な気象の状況がありうる。雲画像をとった領域内の他の地点の天気の違いとか、時間的に前後する日々の天気との組み合わせとか、温度の違いとか、季節の違いとか等、様々な相違点の影響が考えられるが、このようなカテゴリー分けが、ニューラルネットワークによって自律的になされるか否かは興味ある問題である。もしなならぬカテゴリー分けが形成されると、ニューラルネットワークの論理に基づく指標が形

表3 採用雲画像の取得月日

N	月日	N	月日
1	4/01	21	7/16
2	4/06	22	7/29
3	4/16	23	8/02
4	4/18	24	8/06
5	4/22	25	8/12
6	5/03	26	8/14
7	5/20	27	8/16
8	5/24	28	8/24
9	5/29	29	9/01
10	6/02	30	9/02
11	6/08	31	9/12
12	6/16	32	9/27
13	6/19	33	9/29
14	6/21	34	10/03
15	6/22	35	10/04
16	6/24	36	10/05
17	6/28	37	10/15
18	7/01	38	10/16
19	7/04	39	10/24
20	7/09	40	10/29

成されることになる。このような問題意識から、表2で最も多い組み合わせである「3都市すべてが曇り」の場合について、中間層における情報表現の解析を行った。採用した月日は表3のとおりである。

中間層のユニット値の全体を一つの(次元数100の)ベクトル(表現ベクトル)とみなし、40日分の雲画像の内部表現である40個のベクトルの相互の距離の2乗 R_{NM} を計算した。 R_{NM} は第N日の雲画像の中間層表現と第M日の雲画像の中間層表現の相似性の尺度となっている。従って、お互いの中でこの値が小さいベクトルのグループがあれば、それらのグループはなんらかのカテゴリーを表現していることを意味する。

結 果

図2に、約10万回のバックプロパゲーションによる学習によって出力層の誤差 r がどのように減少したかを示す。縦軸は r 、横軸の一つの目盛りは1万回を意味する。この図では、荷重を乱数によって設定した後まず10回の学習を行い(図の左端)、その後の学習の状況を示してある。7万回まで、 $\epsilon = 0.0000001$ を採用した。 r が減少する速度が次第に減少する一般的傾向が現れている。この後75000回まで $\epsilon = 0.00001$ を採用し、その後は $\epsilon = 0.0001$ を採用した。85000回から9万回まで $\epsilon = 0.001$ を採用した場合も示してある。この場合は、 r が減少せず、大きな値に跳ね上がっている。このように、学習の“歩幅”を大きくし過ぎると、谷間への道を見失ってしまうことがあるので、 ϵ のとりかたには一定の注意を要する。

上記のような形でな115000回の学習を行った結果、ニューラルネットワークは166日のすべてに対して、矩形の雲画像を3都市の天気へ、表4に示された程度の正確さで、結び付ける能力をもった。表4は典型的な出力例(6日分)で、各行上段に正解が示されている。

R_{NM} の実際の値は0から95までの間の数であり、かつ、この2つの数値の間を平均的に埋めるような分布をしている。

表5-1から表5-5に、 R_{NM} が10より小さい場合を○で、5より小さい場合を◎で、すべて示した。第何日目の雲画像に対するベクトルかということは縦横の座標の数で示してある。まとめると、 $R_{NM} \leq 10$ によってベクトル間が相似であることを定義した場合、40個のベクトルのうち7個を除く33個が全体として5つのグループを形成している様子が現れている。

このうち、表5-1にあげた7個のベクトルは非常にコンパクトなクラスタリングを示している。表ではこのうち、 $N=2$ と $N=39$ の間の相似性だけが小さいように見えるが、 $R_{2,39} = 11.5$ であってほとんど他の部分と同等と見てよい。このグループのベクトルと例えば表5-2のグループ内の $N=20$ との相関をみると、 $R_{2,20} = 81.3$ 、 $R_{5,20} = 85.1$ 、 $R_{7,20} = 85.0$ 、 $R_{18,20} = 85.9$ 、 $R_{22,20} = 86.3$ 、 $R_{37,20} = 81.4$ 、 $R_{39,20} = 80.9$ となっている。また、例えば $N=5$ のベクトルからみた場合、表5-1の他のベクトル群との R 値の平均は3.7、表5-2のベクトル群とは72.2、表5-3のベクトル群とは80.6、表5-1以外の159

図2 学習による誤差の減少
横軸1目盛りが1万回、縦軸の下端は0

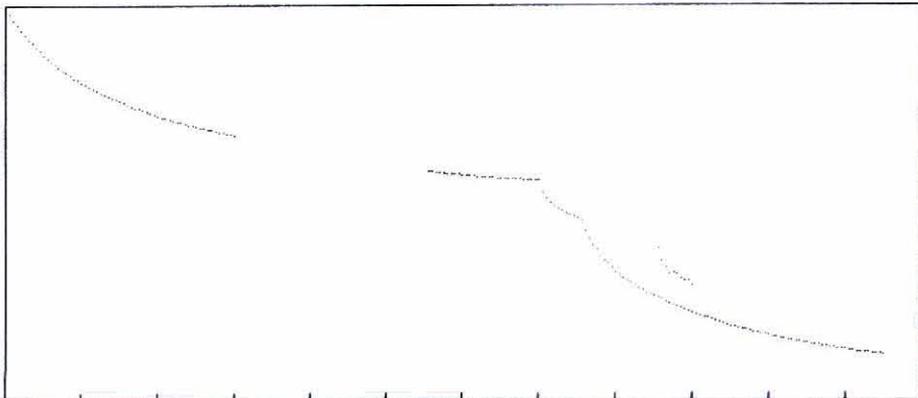


表4
学習進度
上から旭川, 岩見沢, 帯広
各上段が正解
列は左から快晴, 晴, 曇, 雨に対応

4/06			
0.000	0.000	1.000	0.000
0.000	0.022	0.825	0.137
0.000	0.000	1.000	0.000
0.006	0.000	0.948	0.042
0.000	0.000	1.000	0.000
0.001	0.003	0.938	0.054
4/07			
0.000	0.000	0.000	1.000
0.000	0.038	0.154	0.846
0.000	0.000	1.000	0.000
0.029	0.034	0.762	0.201
0.000	0.000	1.000	0.000
0.003	0.001	0.883	0.061
4/08			
0.000	0.000	1.000	0.000
0.004	0.014	0.953	0.047
0.000	0.000	1.000	0.000
0.025	0.080	0.997	0.100
1.000	0.000	0.000	0.000
0.857	0.105	0.136	0.022
11/01			
0.000	0.000	0.000	1.000
0.000	0.001	0.032	0.818
0.000	0.000	0.000	1.000
0.007	0.005	0.073	0.917
0.000	0.000	1.000	0.000
0.020	0.040	0.828	0.079
11/05			
0.000	0.000	1.000	0.000
0.035	0.105	0.440	0.398
0.000	1.000	0.000	0.000
0.032	0.701	0.104	0.172
0.000	1.000	0.000	0.000
0.176	0.576	0.099	0.011
11/06			
1.000	0.000	0.000	0.000
0.666	0.469	0.036	0.024
0.000	1.000	0.000	0.000
0.021	0.935	0.009	0.000
1.000	0.000	0.000	0.000
0.699	0.464	0.084	0.001

表5
中間層表現ベクトル間の相似性
○は強い相似性, ◎はとくに強い相似性
表5-1

	2	5	7	18	22	37	39
2		◎	○	○	○	○	
5	◎		◎	◎	◎	◎	○
7	○	◎		◎	◎	◎	○
18	○	◎	◎		◎	◎	◎
22	○	◎	◎	◎		○	○
37	○	◎	◎	◎	○		◎
39		○	○	◎	○	◎	

表5-4

	13	15	1	36	32
13		○			○
15	○		○	○	
1		○		◎	○
36		○	◎		◎
32	○		○	◎	

表5-5

	6	34
6		○
34	○	

表 5 - 2

	4	8	20	21	38	9	30	11	14	28	27
4		○	◎	○							
8	○		◎						○		
20	◎	◎		○							
21	○		○		◎	◎					
38				◎		◎	○			○	
9				◎	◎						
30					○			○		○	
11							○				
14		○									
28					○		○				◎
27										◎	

表 5 - 3

	10	24	26	35	17	23	29	31
10		◎	○	○	○	○		
24	◎		○	○				
26	○	○		○			○	○
35	○	○	○					
17						○		
23					○		○	
29			○			○		○
31			○				○	

のベクトルとは57.5であった。これらのことから、7つのベクトルが非常に近似したベクトルどおしのグループを形成していることがわかる。

一方、表5-2のグループの11個のベクトルは、相似性の定義をより厳しくした場合、さらに3つ程度のサブグループに分かれる様子が見える。表5-3から表5-5のグループは比較的弱い結びつきのグループといえる。表6-1に5つのグループ間の距離の2乗のを示した。これは、グループ間でのベクトルどおしの距離の2乗の平均をとったものである。表6-2は、平方根をとって“距離”の次元にしたものである。(距離の絶対値は中間層のユニット数に依存するので意味をもち、相対値のみ意味をもち)。これを見ると5つのグループそれぞれが比較的明確なまとまりをなしていることがわかる。

表6-1 中間層クラスター間の距離の2乗

	1	2	3	4	5
1	5.2	27.1	71.6	78.8	41.8
2	27.1	7.7	46.5	80.8	69.5
3	71.6	46.5	13.7	37.5	74.4
4	78.8	80.8	37.5	14.3	41.7
5	41.8	69.5	74.4	41.7	8.3

表6-2 中間層クラスター間の距離

	1	2	3	4	5
1	2.3	5.2	8.5	8.9	6.5
2	5.2	2.8	6.8	9.0	8.3
3	8.5	6.8	3.7	6.1	8.6
4	8.9	9.0	6.1	3.8	6.5
5	6.5	8.3	8.6	6.5	2.9

これら5つのグループがどのような内部表現になっているかをパターンの的に示したのが図3-1から図3-5である。これは表5-1から表5-5のそれぞれのグループ内でベクトルを平均したものを表示したもので、ユニット値が0.7以上のユニットを黒丸で、0.3以下のユニットを白丸で表している。中間のユニットは白紙の形にしている。図3-6は166すべてのベクトルで平均をとったものである。この場合は、ユニット値のほとんどが0.5に近い値をもつ。このことは、全体的には中間層のユニットには様々な値が入ってい

ることを意味する。この事実を考慮すると、上記5つのグループが比較的明確なパターンを持つといえる。第1のグループがとくに明確なパターンになっていることが注目される。

図3 中間層クラスターのパターン表示
図3-1

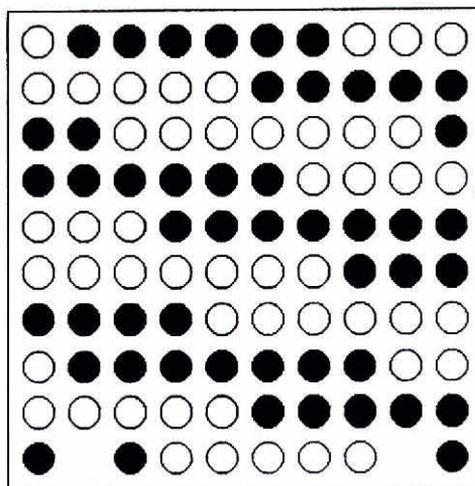


図3-2

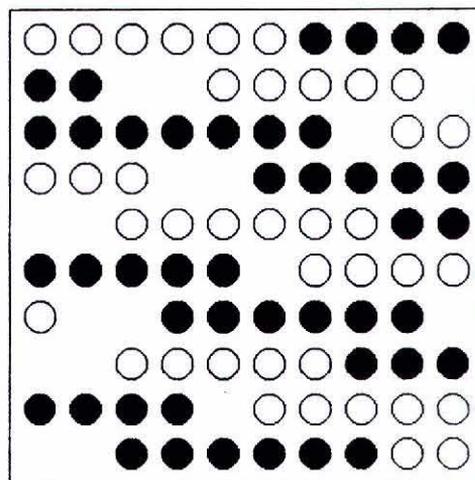


図 3-3

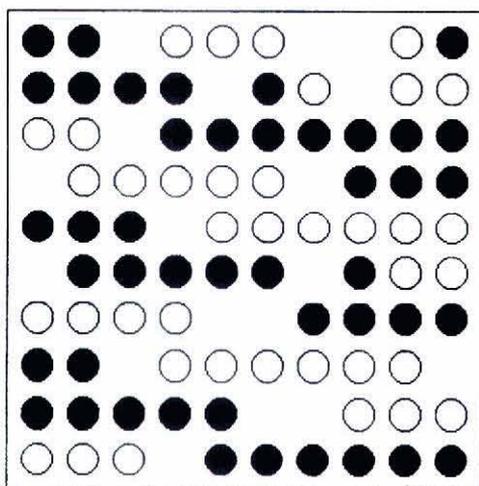


図 3-5

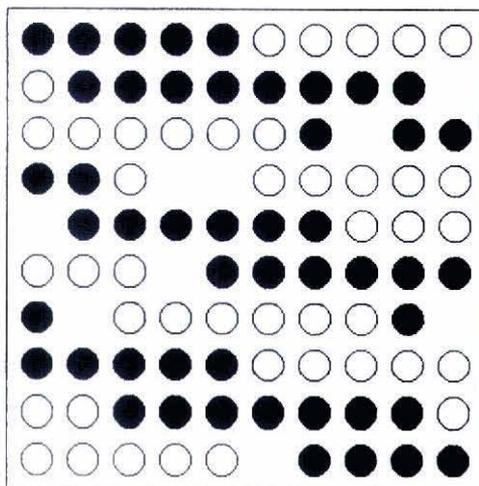


図 3-4

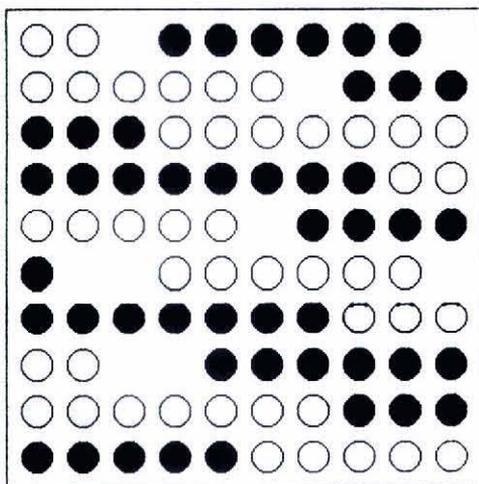
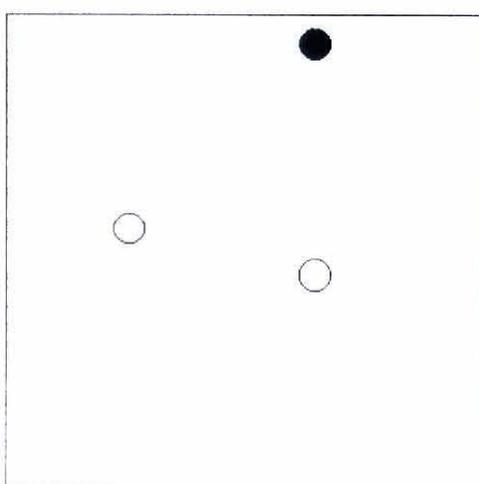


図 3-6



考 察

気象現象はマクロな物質系の極めて自由度の大きい複雑な現象であり、気象現象に含まれる周期性やその他の規則性が人間の知識体系と少し離れたところでとらえられることは望ましいことである。また、気象に関して取得できるデータの量は膨大なものになってきており、いかに少量の有効な情報に変換するかということが重要な側面となってきている。入力層の次元に対して中間層の次元を少なくすることで不要な情報を切り捨てる手法は、画像圧縮の技術などに応用されて

いる。²⁾

本研究の対象とした北海道中央部の雲画像は、非常に多くの情報を含むはずであり、これに比較して、出力層において設定した3箇所の天気は自由度が小さい情報である。従って、中間層の次元を入力層と出力層の中間の大きさにとった場合、出力層の自由度からもれた情報が中間層に現れるのは自然なことである。一般に、中間層における新しい情報の自由度の現れが明確なクラスタリングを伴うものである場合、期待されることは、そのクラスタリングが、出力層の設定による境界条件の下で現れた気象現象のなんらかの規則性

に対応するものではないかということである。

もし同じクラスターに属する雲画像がおしが、入力層においても(上に定義した意味での)相似性をもつ、という関係が存在するならば、中間層におけるクラスタリングの意味あいはいさいと見なされるだろう。しかし、入力層から中間層への変換は非線形なものであるから、このような関係は非常に高い相似性の場合でなければ保証されない。また、中間層におけるユニットの出力値は、変換がシグモイド関数を経由することにより、学習が進行するにつれて0または1の2元的なものに収束するという特徴をもつ。実際、今回の事例では、中間層のクラスタリングに単純に対応する入力層のクラスタリングは存在せず、むしろ両層における相似性の関係はまったく異なったものであるように見える。

本研究の事例で中間層における表現ベクトルが、とくに表5-1に現れたような形で、いくつかのかたまりに分かれたことは、多層ニューラルネットワークの中間層が内部表現の役割を担い得ることを示唆していると思われるが、グループに属する雲画像を詳しく解析せずに結論的なことはいえない。また、このような中間層における表現の事例を蓄積することが重要と思われる。ただ、このような計算は大量の計算時間を要するので、事例の蓄積を遂行する上でコンピュータの高速化が重要な要素であると思われる。

謝 辞

辻修氏、藤島宏和氏からひまわりの雲画像の赤外線強度の数値を引き出すプログラムを提供して頂いた。ここに感謝の意を表します。

引用文献

- 1) D. E. Rumelhart, G. E. Hinton and R. J. Williams (1986) : Parallel Distributed Processing. MIT Press.
- 2) G. W. Cottrell, P. Munro, D. Zipser (1989) : Image compression by back propagation. Advances in Cognitive Science, Vol.3. Norwood, NJ.

SUMMARY

We have analysed the information for areal climate contained in the in-frared cloud image

of the geostational meteorological satellite, in the framework of a 3-layered artificial neural network. After some 100 thousand iterations of learning, there have appeared in the intermediate layer a classification for 40 images belonging to the same weather (cloudy) at three different points which are located inside the image area. This fact can be viewed as an indication of possible role of internal representation played by the intermediate layers of the artificial neural network.

Key word : areal climate, satellite cloud image, 3-layered artificial neural network, internal representation, clustering

Res. Bull. Obihiro Univ., **19** (1996) : 203~211