

JAVA言語による計算機シミュレーションで生じた風景画像の 理解場面での多段階連想形認識過程の異常現象

鈴木 昇一 川俣 博司 大槻 善樹

Abnormal Phenomena about Understanding Scene Images by Using Multi-Stage Associative Recognition under its Computer Simulation Carried Out by Program Language JAVA

Shoichi Suzuki Hiroshi Kawamata Zenju Otuki

あらかし

画像中の任意の1つの画素に視点を合わせ、視野の範囲に存在する形状情報を手がかりにその画素を認識を行う認識システムRECOGNITRONを考えよう。その認識結果、その画素に1つのカテゴリラベルが対応することになる。このようにして得られた帰属する各カテゴリラベルを各画素に貼り付けていき、1枚の画像を構成する全ての画素について適切なカテゴリラベルを付与することができれば、1枚の画像から、画素単位に画像から知識を抽出することができる。同一カテゴリラベルを持つ画素の集まりは1つの事物を表すことになる。このようにして、1枚の画像内にどのような事物が存在するかに関する知識が抽出される。同時に画像を意味のある有限個の領域に分割する領域分割 (region segmentation) の機能を実現することができる。

パターン情報処理に関する数理(SS理論)を適用し、この領域分割の機能を備え、風景画像から画素単位に各事物の抽出を行う画像理解システムIUS(Image-Understanding System)が既に構築されている [5] ~ [7]。IUSは画素数だけのRECOGNITRONからなっている。その際必要となる3基本要素(axiom 1~3を各々満たすモデル構成作用素T, 類似度関数SM, 大分類関数BSC)の構成が既に説明され [5], その計算機シミュレーションの結果が既に検討されている [6]。JAVA言語で実装化されたこの画像理解システムIUSの動作が簡単に説明され、IUSを稼動する方法も解説されている [7]。

Tには外界の事物に対応するパターンモデルをIUSの内部に形成する働きがあり、SMには外界の事物のパターンモデルが記憶内の各カテゴリの代表パターンのモデルとどの程度似ているかを測る働きがあり、BSCには外界の事物のパターンモデルが帰属するカテゴリの候補を複数個出力する働きがある。従来の画像を理解する方法と決定的に相違する点は、入力画像の意味はこの画像から3構成成分T, SM, BSCを使って想起される画像モデルであるとしていることである。

視野の大きさを決定する2助変数 p, q を導入し、事物 φ の存在する領域は、視野の中心としての視点を2次元直交座標値 $x = \langle x_1, x_2 \rangle$ とするような画素数が $(2p+1) \cdot (2q+1)$ であるような連結領域としよう。処理の対象とする問題のパターン φ は、

$$\varphi = \{ \varphi_k \equiv \varphi(x_1 + p, x_2 + q) \mid k = 0, \pm 1, \pm 2, \dots, \pm p, \ell = 0, \pm 1, \pm 2, \dots, \pm q \}$$

と表わされ、 $\varphi_k \equiv \varphi(x_1 + p, x_2 + q)$ は φ の第 $\langle k, \ell \rangle$ 画像成分である。領域分割をあからさまに行わない認識システム RECOGNITRON $\langle x_1, x_2 \rangle$ がJAVA言語で実現されている。IUSの1成分 RECOGNITRON $\langle x_1, x_2 \rangle$ は、視野の中心としての視点を2次元直交座標値 $x = \langle x_1, x_2 \rangle$ に持つ認識システムである。1枚の画像内の全画素数だけの認識システム RECOGNITRONの集まりとして解釈される本画像理解システム

$$\text{IUS} \equiv \{ \text{RECOGNITRON} \langle x_1, x_2 \rangle \mid x_1 = -50 \sim +50, x_2 = -40 \sim +40 \}$$

は、直交座標 $x = \langle x_1, x_2 \rangle$ が付置された2次元平面に投影された物体の濃淡値から得られる灰色形状情報のみを手がかりに理解の動作を行い、各カテゴリの代表パターンからの変形がある程度少ない事物であればかなり高い認識率を得られることがシミュレーションにおいて確認されている。

本論文は、文献 [5] の理論をほぼ忠実に、JAVA言語でシミュレーションした結果とその考察が報告された2文献 [6], [7] の続編である。文献 [6] では、風景画像内に空、木、車、家、道路、電柱(6個の事物カテゴリ)が存在するかどうかを RECOGNITRONが理解できるかどうかを検証された。更に、IUSを操作する方法、並びに、その動作の簡単な説明については文献 [7] で与えられている。

RECOGNITRONによる**SS多段階想起認識法**による認識結果において、

認識可能(該当するカテゴリが唯1個あること)・**認識不定**(該当するカテゴリが複数個あること)・**認識不能**(該当するカテゴリが1個もないこと)

という3結果の内、認識不定、認識不能を得るという異常現象が生じたが [6] , これは次のように説明される:

採用した複数の類似度関数SMがいずれもミックスチュア条件(付録C)を満たすことが保証されていないためである(文献 [4] の定理6.6を参照)。 □

SS多段階想起認識過程の振動(SS多段階想起認識の非収束)という**異常現象**も、実際生じたが [6] , これは次のように説明される:

採用した複数の類似度関数SMがいずれも直交条件(付録D)を満たすことが保証されていないため、カテゴリ帰属知識(パターンが入力された認識システム RECOGNITRONの、第t段階の作業状態) $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー $E(\varphi^t, \lambda^t)$ が認識段階番号tの非増加関数であるとは限らないからである(文献 [4] の4定理8.1~8.4を参照)。 □

以上の“SS多段階想起認識過程における2**異常現象**”が、本論文では、少し検討され報告される。このような異常現象が生じる原因の1つは、T, SM, BSCが適切に選定されていないことであり、いま1つの原因はBSC内の助変数に関する学習不足、或いは学習過剰であり、画像理解に必要な知識がIUSに適応的に反映されていないことである。要するに、IUSが画像内容を理解できるよう**発見的学習**(heuristic learning)がなされていない場合、異常な認識現象が発生する。

本IUSで提案されている画像処理方式の特徴のひとつに、処理画像が例えば、風景画像に限定されないという利点がある。画像内容に関する従来の理解法とは異なり、学習に用いる訓練パターンと、あらかじめ用意するカテゴリラベルを変更するだけで、顔画像の理解や、指紋解析、文字認識などに適用することも可能であるという一般性を挙げることができる。

キーワード

パターン認識の数学的理論(SS理論) モデル構成作用素 類似度関数 大分類関数
カテゴリ選択関数 構造受精変換 カテゴリ帰属知識 不動点連想形認識
画素単位の認識 風景画像からの知識の抽出 セグメンテーション
ファジィ推論ニューラルネット 最急降下学習法 類似度関数のミックスチュア条件
類似度関数の直交条件

Abstract

Consider a recognition system RECONITRON which recognizes a pixel in question by as clues using a configuration existing in a sight after RECOGNITRON fixed a viewpoint on the pixel of an image in question. We notice that an object may be formed by gathering all the pixels which are in possession of the same result of recognition. Such a pixelwise-recognition has a faculty of a region-segmentation.

An image-understanding system IUS has already been implemented by applying a theory of recognizing patterns suggested by S.Suzuki to an understanding of scenery images, which has two abilities of a pixelwise-recognition of the scenery input image and the above-mentioned region-segmentation[5]~[7]. **IUS is made up of as many RECOGNITRONs as the number of pixels.** Three model-construction operators Ts, Three similarity-measure functions SMs, a rough classifier BSC which must satisfy axiom 1, 2, 3 respectively are needed to build RECOGNITRON, and its explanation has already been given[5]. We then proceeded to discuss its result of simulation[6]. Its Specification about IUS and how IUS implemented with JAVA operates are provided. [7].

T makes a corresponding internal model of an object in the external world. SM measures how much a pattern-model of the object is similar to each of the memorized prototypical pattern-models of supposed categories. BSC outputs several candidates for categories to which the pixel in question may belong.

It seems obvious to us that a distinction from traditional understanding methods of images is that a meaning of the input image is expressed as a model recalled by using three fundamental components T, SM and BSC.

Introducing two parameters p and q which determine a size of sight, let us suppose that a region in which an object exists is a connected area whose viewpoint as a center of the sight is a two-dimensional coordinate $x = \langle x_1, x_2 \rangle$ and whose number of pixels is $(2p+1) \cdot (2q+1)$. An input pattern φ in question to be recognized is expressed as

$$\varphi = \{ \varphi_{kl} \equiv \varphi(x_1 + p, x_2 + q) \mid k = 0, \pm 1, \pm 2, \dots, \pm p, \ell = 0, \pm 1, \pm 2, \dots, \pm q \}$$

, where $\varphi_{kl} \equiv \varphi(x_1 + p, x_2 + q)$ is the $\langle k, \ell \rangle$ -th component of φ . A recognition system has been realized by use of JAVA language, which is not going explicitly to segment an image. A component RECOGNITRON $\langle x_1, x_2 \rangle$ of IUS has the two-dimensional coordinate $x = \langle x_1, x_2 \rangle$ as viewpoint. The set of RECOGNITRONs which are as many as the total number of pixels in an image

$$\text{IUS} \equiv \{ \text{RECOGNITRON} \langle x_1, x_2 \rangle \mid x_1 = -50 \sim +50, x_2 = -40 \sim +40 \}$$

is a system whose purpose is to understand a content of the image based upon gray-scaled configurations of the object projected on a two-dimensional plane having orthogonal coordinates $x = \langle x_1, x_2 \rangle$. A high rate

of recognition has been achieved in its computer simulation if the object was not very different to one of the memorized prototypical objects.

This paper follows up two literatures [6], [7] which reported the results of the JAVA-simulation and its analysis whose theory is described in literature [5].

Whether or not RECOGNITRON can understand that a sky ,trees ,cars,houses,roads and utility poles (six categories of objects) exist in the given image was reported in [6]. Moreover a method of operating IUS and a specification of IUS are shortly explained in [7].

The recognition-results of SS-multi-stage associative recognition are classified into three cases:

- (i) Capability of recognition (In the case of that the number of obtained categories is one).
- (ii) Indeterminate recognition (In the case of that the number of obtained categories is greater than two).
- (iii) Being not capable of recognition (In the case of that the number of obtained categories is zero).

□

Abnormal phenomena (ii) and (iii) are observed in simulation [6]. The reason why (ii) and (iii) occurs is that the SM employed here may not satisfy the condition of mixture (see theorem 6.6 [4]).

We also observed that the process of SS-multi-stage associative recognition was oscillated, i.e., the process was not converged [6]. Notice that the oscillation is abnormal. The reason why the oscillation occurs is as follows:

The SM employed here may not satisfy the condition of orthogonality. A potential energy $E(\varphi^t, \lambda^t)$ of categorical knowledge $\langle \varphi^t, \lambda^t \rangle$ (the t-th state of operation of RECOGNITRON which received the input pattern φ) was not necessarily decreased as the stage-number t increases (see four theorems 8.1 ~ 8.4[4])

□

The above-mentioned two abnormal phenomena are partly discussed and examined in this paper. Such abnormal phenomena occur if T, SM and BSC is not adequately selected.. Abnormal phenomena also occur if parameters of BSC are overlearned or lesslearned. The cause of abnormal phenomena is that the knowledges needed to understand the image has been not reflected to IUS. In actual fact, abnormal phenomena occur if heuristic learning which IUS necessitates to understand the image is not perfect.

Applying the theory proposed by S.Suzuki(SS theory [1]~ [4]) to the field of image-processing has many advantages. That images to be processed are not limited to scenery images is characteristic of the method of image-processing which has been adopted in our IUS.The method is applicable to the facial images, finger-print images and character images provided that training patterns, categories and its prototypical patterns for which we must make preparations only are exchanged for appropriate another even if T, SM and BSC are fixed.

Key words : a mathematical theory of recognizing patterns(SS-theory) model-construction operator(T) similarity-measure function(SM) rough classifier(BSC) category-selection function(CSF) structure-fertilization transformation categorical-membership knowledge associative recognition of searching a fixed-point pixelwise recognition scenery image extraction of knowledges from scenery images segmentation fuzzy- inference neural net gradient descent learning mixture-condition of SM orthogonality-condition of SM

1. はじめに

本論文はSS理論 [1]～[4] を適用して得られた研究を報告している3文献 [5], [6], [7] の続編であり, 用いられる諸記号, 諸設定はそのまま使用されている.

1.1 認識システムRECOGNITRONの行う多段階連想形不動点探索認識の万能性

S,Suzukiは, 30年間以上の年月を費やし, パターン認識の数学的理論(SS理論)を構築している [3], [4]. SS理論を適用すれば, 任意にパターン認識の働きが与えられた場合, 認識システムRECOGNITRONを構成して, 与えられたこのパターン認識の働きを多段階作業状態で再現可能である. このRECOGNITRON構成のためには, axiom 1～3を各々満たすモデル構成作用素T, 類似度関数SM,大分類関数BSCの3構成要素を適切に選定しなければならない. 得られたこの種の再現性は, 認識システムRECOGNITRONは万能性を備えている [3] ことを示している.

RECOGNITRONは, パターンモデル $T\varphi$ を見るとあたかも原パターン φ であるかのように錯覚する(パターンモデル $T\varphi$ と原パターン φ 間の同一知覚原理). このように, Tには外界のモデルを形成する働きがある. 同様に, SMにはモデル同士の似ている程度を測る働きがあり, BSCには外界の事物が帰属するカテゴリの候補を複数個出力する働きがある. 3構成成分T, SM, BSCさえ与えられれば稼動可能なRECOGNITRONが処理の対象とする問題のパターン φ に対し行う作業は, 或る構造受精変換の不動点を多段階にわたり連想的に探索し, 或るカテゴリ \mathcal{C}_j (第j番目のカテゴリ)の代表パターン ω_j のモデル $T\omega_j$ を出力する多段階連想形の不動点探索認識過程(SS多段階想起認識過程)である.

この多段階連想形不動点探索認識(SS多段階想起認識)方式では, 処理画像が例えば, 風景画像に限定されない構成法となっていることをその処理方式の特徴の1つとして, 指摘できる. 言い換えれば, BSCなどの学習に用いる訓練パターン(training pattern)と, あらかじめ用意するカテゴリラベルを変更するだけで, 顔画像の理解や, 指紋解析, 文字認識などに適用することも可能であるという一般構成性を挙げる事ができる.

1.2 SS理論が解決した5事項

4公理のaxiom1～4(SS公理系)を前提としているSS理論は次の5事項(1)～(5)を解決している:

(1)パターンというものを再帰的に定義可能な再帰領域方程式(reflective domain equation)を提案し, この方程式を解き, 処理の対象とする問題のパターン φ の集合 Φ を明示的に表示できたこと [3]

(2)認識システムRECOGNITRONが処理の対象とする問題のパターン $\varphi \in \Phi$ に対し持つカテゴリ帰属知識を各カテゴリについての知識へと直交直和分解(SS分解)したこと [3]

(3)或る連想形認識方程式(SS方程式)を提案し, 認識システムRECOGNITRONの多段階認識の働きはこの連想形認識方程式をSS分解に関し解く過程であり, このSS方程式の解は認識可能(入力パターン $\varphi \in \Phi$ の帰属する候補カテゴリが唯1個決まること), 認識不定(入力パターン $\varphi \in \Phi$ の帰属する候補カテゴリが複数個存在すること), 認識不能(入力パターン $\varphi \in \Phi$ の帰属する候補カテゴリが1つも存在しないこと)に3分類されること [4].

(4)直交性SMの概念, 並びにカテゴリ帰属知識のポテンシャル・エネルギー(SSポテンシャル)を提唱し, 多段階連想形認識過程の進展に伴なうSSポテンシャル減少性を証明したこと [4].

(5)mixture性SMの概念を提唱し, mixture性SMを採用していれば多段階連想形認識過程は認識

可能な状態に収束することを証明したこと [4]. □

また、各人の主観的感性が反映されるように、axiom1~3を各々満たすT, SM, BSCを選定でき、知能のある通常人の認識主観特性をsimulateし、analysis by synthesis的手法で浮き彫りにできる[10].

1.3 2つの情報システムMEMOTRON, FUZZITRON

4公理axiom1~4(各々, T, SM, BSC, 並びにカテゴリ選択関数CSFの満たすべき公理;CSFはSM, BSCで構成される)を前提としたSS理論を適用して、可分な一般抽象ヒルベルト空間 \mathfrak{H} 上で動作する2つの情報システム

- 出現順序を保ちパターン系列を記憶し、その任意の部分系列が入力されるとその後続するパターン系列を再生可能なニューラルネット(neural net)としての連想形記憶器

MEMOTRON [8]

- 問題解決器(problem-solver)としてのマルチメディア・プロダクションシステム

FUZZITRON [9]

も構築されている。

4公理axiom1~4は、各々, T, SM, BSC, 並びにカテゴリ選択関数CSFの満たすべき公理であり、最後のCSFは文献 [3] の付録Eの定理E1で示されているように、SM, BSCで構成される。

1.4 情報システムとしての画像理解システムIUS研究 [5], [6], [7]の続編としての本研究

パターン認識システムRECOGNITRONも、MEMOTRON, FUZZITRONと同じく、可分な一般抽象ヒルベルト空間 \mathfrak{H} 上で動作する情報システムであり、本画像理解システムIUS (Image-Understanding System)は画素単位で定義されたRECOGNITRONの、すべての座標値(画素)にわたる集合体である。

本論文は、axiom1~3を各々満たす特別なモデル構成作用素T, 特別な類似度関数SM, 特別な大分類関数BSCが採用された上述のRECOGNITRONをシミュレーションした結果、観察された異常な挙動に関し、報告するものである。本論文は、このRECOGNITRONによる風景画像の、画素単位の理解理論 [5] をJAVA言語でほぼ忠実に、シミュレーションした結果が簡単に報告された文献 [6] と、JAVA言語で実装化され、このシミュレーション結果を得た画像理解システムIUSの動作概要と、その稼動方法が解説されている文献 [7] との両文献の続編である。

本論文では、このシミュレーションで生じた**異常な認識現象**が報告される。

2つのこの異常現象

- (a) SSポテンシャルが減少しなくて、多段階連想形認識過程が終了しなかった場合があったこと、
- (b) たとえ終了したとしても、認識不定、認識不能なる多段階連想形認識結果が得られたことは採用された類似度関数SMが各々、

(a') 直交性を満たしているとは限らないこと

(b') mixture条件を満たしているとは限らないこと

から、SS理論で予想されていたこと(文献[4]の2定理8.3, 6.6)である。

計算機シミュレーションで採用されたSS多段階想起認識法は次のように簡単に説明される。より詳細な説明は付録Bにある。

処理の対象とする問題の入力パターン φ について、まずそのパターンモデル $T\varphi$ を求め、この $T\varphi$ について多段階にわたってパターンモデルの変換を行い(帰納推理変換)、構造受精変換の不動

点(あるカテゴリの代表パターンのモデル)を探索する形(不動点探索)で想起し、この不動点モデルの帰属するカテゴリを φ の帰属するカテゴリとするを認識する方法(想起認識)、つまり、不動点探索型・多段階パターンモデル帰納推理変換・想起認識法(SS多段階想起認識法)では、simplex探索法に似た手法で原パターン φ を認識しようとする。今、少し詳しく説明すれば、次のようになる:

(c) 1より小さい非負実数 $s_j(T\varphi)$ は、パターン φ について、得られたパターンモデル $T\varphi$ が第 $j \in \gamma$ 番目のカテゴリ \mathcal{C}_j の代表パターン ω_j のモデル $T\omega_j$ と似ている程度であり、

(d) J はすべてのカテゴリ番号の集合である

としよう。処理の対象とする問題のパターン φ について、そのパターンモデル $T\varphi$ を先ず、求めた後、初期単体の部分集合

$$\left\{ \sum_{j \in \gamma} s_j(T\varphi) \cdot T\omega_j \mid \gamma \subseteq J \right\} \quad (1.1)$$

$$, \text{where } [\forall j \in \gamma, 0 \leq s_j(T\varphi) \leq 1] \quad (1.2)$$

$$\wedge [0 \leq \sum_{j \in \gamma} s_j(T\varphi) \leq 1] \quad (1.3)$$

から探索を開始し、パターンモデルの各変換段階において帰納推理で選択された構造受精変換

$$TA(\mu)T, \text{ここに、} \mu \subseteq J \quad (1.4)$$

の操作を行い、候補カテゴリ知識(パターン ψ がカテゴリ集合 $\mathcal{C}_j, j \in \lambda$ 内の1つのカテゴリに帰属する可能性があるという認識システム RECOGNITRON が持っている知識)

$$\langle \psi, \lambda \rangle \quad (1.5)$$

を順次精練し改善することにより、カテゴリ知識の不動点の最適解(SSポテンシャルを最小にする不動点)を求め、原パターン φ を認識しようとする。□

事物の存在する領域を $(2p+1) \cdot (2q+1)$ 個の画素からなる画像成分と仮定し、領域抽出(セグメンテーション)が自動的になされる**画素単位の認識方式**を内蔵している認識システム RECOGNITRON の集合体としての画像理解システム IUS についてその性能を確かめる計算機シミュレーションが JAVA 言語で実行され、その結果が既に検討されている [6]。更に、IUS を操作する方法、並びに、その動作の簡単な説明については文献 [7] で与えられている。モデル構成要素 T 、類似度関数 SM 、大分類関数 BSC 、カテゴリ選択関数 CSF はそれぞれ、axiom 1~4 を満たすように構成されているので、2文献 [3]、[4] で解析されている RECOGNITRON の基本的な認識性能は保証されている。axiom 4 を満たすカテゴリ選択関数 CSF は SM と BSC と用いて構成され(文献 [3] の付録 E の定理 E1)、処理の対象としているパターンが帰属するであろう複数個の候補カテゴリを抽出する働きがあり、カテゴリ帰属知識の変換に使われる構造受精変換の性能を分析するときに必要な写像である。

画像中の(任意の1つの)画素に視点をあわせ、視野の範囲(視点を中心とする矩形の領域)に存在する形状情報を手がかりに視点にあるこの1画素の認識を行い、その結果得られた帰属するカテゴリラベルをその画素に貼り付けていく。貼り付けるこの方法で1枚の画像を構成する全ての画素について適切なカテゴリラベルを付与することができれば、1枚の画像から、画素単位の画像からカテゴリ知識を抽出することができる。同時に、同一カテゴリを持つ画素を集めて1つの領域として採用すると、各画素にカテゴリラベルを貼り付けるこの方法で、意味のある有限個の領域に画像を分割する領域分割 (region segmentation) の機能を実現することができる。パターン情報処理に関する数理(SS理論)を適用し、この領域分割の機能を備え、各画素にカテゴリラベルを貼り付

ける方法で、風景画像内のすべての事物の抽出を行う画像理解システムIUSが構築されている [5], [6], [7]. そこでは、事物の存在する領域を大きさ $(2p+1) \cdot (2q+1)$ (画素の個数)の画像成分と仮定し、領域抽出をあからさまに行わない認識方式が採用されている. 各カテゴリの代表パターンからの変形がある程度少ない画像であればかなり高い認識率を得られることがシミュレーション [6] において確認されている.

2文献 [6], [7] で採用されている上述の画像処理手法では、平面上の直交座標系 $\langle x_1, x_2 \rangle$ が付与された2次元画像面を見て、さまざまな外界の事物の知覚的認識を行う人間の視覚認識機能がJAVA言語により作成したプログラムによって実現されている. つまり、学習の結果得られた記憶(経験)に基づいて最もありそうなことを想起し、多段階の最終段階において想起されたパターンモデルが外界の事物のモデルであるとする「想起型認識 (associative recognition)」の働き(外界の事物の帰属するカテゴリと、その外界の事物から想起される形状との双方を出力する働き)を基盤として、風景画像を解釈・理解するシステムRECOGNITRONが構築され、その計算機シミュレーションが実施されている. **従来の画像を理解する方法と決定的に相違する点は、入力画像の意味はこの画像から想起される画像のモデルであるとしていることである.**

パターン情報処理に関する数理(SS理論)を適用し得られた文献 [5] で説かれている風景画の理解理論を基本的に具体化しなされた文献 [6] でのシミュレーションにおいては、上述(文献 [6] の7.の認識方法)による認識結果に、

認識可能(該当するカテゴリが唯1個あること)・認識不定(該当するカテゴリが複数個あること)・認識不能(該当するカテゴリが1個もないこと)という3結果が生じたが、これは、採用した類似度関数SMがいずれもミックスチュア条件を満たすことが保証されていないためである(文献 [4] の定理6.6を参照).

また、文献 [6] でのシミュレーションでは、上述(文献 [6] の7.の認識方法による7.7で説明されている)認識過程においては、認識過程が終了しないという振動が、実際生じたが、これは次のように説明される.採用した類似度関数SMがいずれも直交条件を満たすことが保証されていないため、7.6.3のカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー $E(\varphi^t, \lambda^t)$ が認識段階番号 t の非増加関数であるとは限らないからである(文献 [4] の3定理8.1, 8.3, 8.5を参照).

また、式(8.26)の多段階連想形認識過程に随伴する9章で説明されている4種の量

(イ)認識情報量 $RGetpy(\varphi^t)$

(ロ)認識情報量 $RAIN(\varphi^{t+1}, \varphi^t)$

(ハ)カテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー (energy) $E(\varphi^t, \lambda^t)$

(ニ)2つのカテゴリ帰属知識間の距離 $SMdis(\langle \varphi^s, \lambda^s \rangle, \langle \varphi^{s+1}, \lambda^{s+1} \rangle)$

をも、画像理解システムIUSは求めるが、振動する多段階連想形認識過程についてはこの4種の量の変遷が報告される. この4種の量は任意の画素の各認識段階について詳細に検討する際に有用なものである.

1.5 4つの付録A, B, C, D

尚、以下で引用されているaxiom 1~4については、付録Aで説明されている. また、本計算機シミュレーションで採用されたSS多段階想起認識法は付録Bで簡単に説明されている. 更に、axiom 2を満たす類似度関数SMに関するSMミックス条件は付録Cで説明されている. 最後に、類似度関数SMに関するSM-直交条件は付録Dで説明されている.

2. デジタルカメラ画像から右手整数値直交座標系表示による入力画像 φ への変換と、画像間の内積、画像のノルム

本章では、まず、デジタルカメラ画像が、右手整数値2次元直交座標系 $x = \langle x_1, x_2 \rangle$ で表示され、かつ x でのその濃淡振幅 $\varphi(x) = \varphi(x_1, x_2)$ が -128 から 127 までの値をとる入力画像(パターン) φ へ変換される手順が説明される。その後、パターンの諸性質やパターン同士の関係を数値的に計量するために必要とされる画像間の内積、画像のノルムも定義される。

2.1 デジタルカメラ画像から右手整数値直交座標系表示による入力画像 φ への変換

デジタルカメラ画像から入力画像 φ へ変換する手順は文献 [6] の2.1節で説明されており、この操作の結果、デジタルカメラ画像は、最終的に -128 (白) \sim $+127$ (黒)の値をとる濃淡値の集合として、2次元配列(単色モード256階調モノクロ画像) $\text{phi}[a_1][a_2]$ に変換される。その後、右手整数値直交座標系 $x = \langle x_1, x_2 \rangle$ を導入し、式(2.7)のように、デジタルカメラ画像は $\varphi(x) = \varphi(x_1, x_2)$ として表現される。

面積が 100×80 であるような矩形領域を考え、 x_1 軸(横軸)、 x_2 軸(縦軸)の値が

$$x_1 = 0, \pm 1, \dots, \pm 50, x_2 = 0, \pm 1, \dots, \pm 40 \quad (2.1)$$

であるような整数の値をとる右手直交座標系 $x = \langle x_1, x_2 \rangle$ を設ける。処理対象とする一枚の画像 φ は2の8乗(=256)階調の灰色画像であり、

$$\begin{aligned} \varphi(x) = \varphi(x_1, x_2) \in \{0, \pm 1, \pm 2, \dots, \pm 127, -128\}, \\ \text{ここに, } x = \langle x_1, x_2 \rangle \end{aligned} \quad (2.2)$$

と表わされる。画像の横幅width, 画像の縦幅heightを

$$\text{width} = 101, \text{height} = 81 \quad (2.3)$$

として、 x_1 座標知の最大値 x_{1_max} , x_2 座標知の最大値 x_{2_max} を

$$x_{1_max} = (\text{width} - 1) / 2 = 50, x_{2_max} = (\text{height} - 1) / 2 = 40 \quad (2.4)$$

と求め、パターンを記憶する2次元配列 $\text{phi}[a_1][a_2]$ の2つの指標 a_1, a_2 を、

$$a_1 = x_{2_max} - x_2 = 0 \sim 80 \quad (2.5)$$

$$a_2 = x_1 + x_{1_max} = 0 \sim 100 \quad (2.6)$$

と設定すると、この $\text{phi}[a_1][a_2]$ と式(2.2)の $\varphi(x) = \varphi(x_1, x_2)$ とは、

$$\varphi(x) = \varphi(x_1, x_2) = \text{phi}[a_1][a_2] \quad (2.7)$$

という関係にある。ここに、

$$x_1 = -x_{1_max} \sim x_{1_max} \quad (x_1 = -50 \sim 50) \quad (2.8)$$

$$x_2 = -x_{2_max} \sim x_{2_max} \quad (x_2 = -40 \sim 40) \quad (2.9)$$

である。

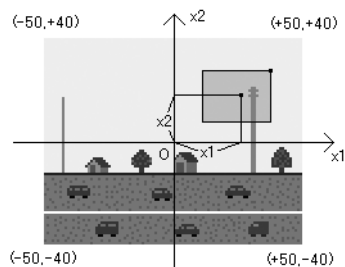


図2.1 採用した右手系直交座標系 $x = \langle x_1, x_2 \rangle$
Fig.2.1 an adopted right-handed orthogonal coordinate $x = \langle x_1, x_2 \rangle$

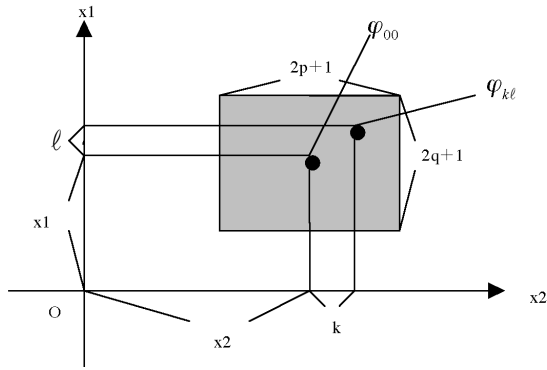


図2.2 パターン $\varphi(x_1, x_2)$ の第 (k, l) 番目成分 $\varphi_{k\ell}$
 Fig.2.2 The (k, l) th component $\varphi_{k\ell}$ of pattern $\varphi(x_1, x_2)$
 at coordinate $x = \langle x_1 + k, x_2 + \ell \rangle$

上の図2.1, 図2.2に, 採用した右手整数値直交座標系 $x = \langle x_1, x_2 \rangle$ が示されている.

以後 視野の中心(視点)の座標値 $x = \langle x_1, x_2 \rangle$ を固定し, 座標値 $x = \langle x_1 + k, x_2 + \ell \rangle$ でのパターン成分

$$\varphi_{k\ell} \equiv \varphi(x_1 + k, x_2 + \ell) \quad (2.10)$$

において k, ℓ を動かして得られるパターン

$$\varphi = \{ \varphi_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \} \quad (2.11)$$

を認識処理するように, 画像理解システム IUS が視点を座標値 $x = \langle x_1, x_2 \rangle$ に移したときのシステム RECOGNITRON $\langle x_1, x_2 \rangle$ が構成される. RECOGNITRON $\langle x_1, x_2 \rangle$ の視野の大きさを決める2整数値 p, q として, 本シミュレーションでは,

$$p=5, q=4 \quad (2.12)$$

が採用される. 図2.2には, 視点の座標値 $x = \langle x_1, x_2 \rangle$ にある画素成分 φ_{00} と, 式(2.11)のパターン φ 内の画素成分(第 (k, ℓ) 番目の成分) $\varphi_{k\ell}$ が示されている.

視野の中心である視点 $x = \langle x_1, x_2 \rangle$ をすべての座標値にわたり変えると, 視野の大きさを $(2p+1) \cdot (2q+1)$ とする認識システム RECOGNITRON $\langle x_1, x_2 \rangle$ の集合体

$$\{ \text{RECOGNITRON} \langle x_1, x_2 \rangle \mid x_1=0, \pm 1, \dots, \pm 50, x_2=0, \pm 1, \dots, \pm 40 \} \quad (2.13)$$

が得られるが, この集合体が本研究での画像理解システムである. RECOGNITRON $\langle x_1, x_2 \rangle$ は視点を座標値 $x = \langle x_1, x_2 \rangle$ に持つ多段階連想形不動点探索認識システムである.

2.2 画像間の内積 (φ, η) , 画像のノルム $\|\varphi\|$

本節では, パターンの諸性質やパターン同士の関係を数値的に計量可能にする類似度関数 SM を構成するために使われる画像間の内積(inner product), 画像のノルム(norm)が定義される.

本認識システム RECOGNITRON が注目している座標値 x_1, x_2 は固定されている状況下で, パターン φ と η の内積 (φ, η) は,

$$(\varphi, \eta) = \sum_{k=-p}^{+p} \sum_{l=-q}^{+q} \varphi_{k\ell} \cdot \eta_{k\ell} \quad (2.14)$$

と定義され, 同様に, パターン φ のノルム $\|\varphi\|$ は,

$$\|\varphi\| = \sqrt{(\varphi, \varphi)}$$

$$= \sqrt{\sum_{k=-p}^{+p} \sum_{l=-p}^{+q} (\varphi_{kl})^2} \quad (2.15)$$

と定義される．視野の中心座標値 $x = \langle x_1, x_2 \rangle$ にある画素がいかなるカテゴリを表す事物の成分であるかを決定する目的を持つシステム RECOGNITRON $\langle x_1, x_2 \rangle$ の各構成成分 (以下での, T, SM, BSC, CSF のこと) がこの内積 (φ, η) , ノルム $\|\varphi\|$ を用いて, 組み立てられる．

3. axiom 1 を満たす採用したモデル構成作用素 T

axiom 1 を満たす処理の対象とするパターン φ の集合 Φ とモデル構成作用素

$$T: \Phi \rightarrow \Phi \quad (3.1)$$

との対 $[\Phi, T]$ を導入する．T には, 外界の事物 φ のモデル $T\varphi$ を形成する働きがある．式 (3.1) の写像 T が axiom 1 の (i), (ii), (iii) の 3 後半, 並びに, (iv) を満たすように構成されたとき, 集合は式 (A1.14) のごとく決定される (付録 A の定理 A2.1 の (ロ))．

構築された画像理解システム IUS では, 文献 [3], [4] での axiom 1 を満たす 3.1, 3.2, 3.3 の 3 節で説明されている 3 種類のモデル構成作用素 T を内蔵している．以下で説明されるパターン φ のモデル $T\varphi$ は, $(2p+1) \cdot (2q+1)$ 個の振幅の集まりであり,

$$T\varphi = \{(T\varphi)_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q\} \quad (3.2)$$

と表示される．

パターン φ の意味 (表象) は φ から多段階にわたり想起される或るカテゴリ \mathcal{C}_j の代表パターン $\omega [j]$ のモデル $T\omega [j]$ であるとしているから, いかなるモデル構成作用素 T を採用するかは, 認識システム RECOGNITRON の性能に直接, 影響することに注意しておかなければならない．

3.1 振幅の最大値によるパターン規格化作用素 S を使ったモデル構成作用素 T

各 $(S\varphi)_{kl}$ については

$$(S\varphi)_{k\ell} \equiv \begin{cases} 0 & \dots \max_{r,s} |\varphi_{rs}| = 0 \text{ の場合} \\ \varphi_{kl} / \max_{r,s} |\varphi_{rs}| & \dots \max_{r,s} |\varphi_{rs}| > 0 \text{ の場合} \end{cases} \\ k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm 1 \\ r=0, \pm 1, \pm 2, \dots, \pm p, s=0, \pm 1, \pm 2, \dots, \pm 1 \quad (3.3)$$

と選び, $T\varphi$ を

$$T\varphi = S\varphi \text{ for any} \quad (3.4)$$

と定義する．写像 S は振幅の最大値 $\max_{r,s} |\varphi_{rs}|$ によってパターン φ を規格化するパターン変換作用素である．

3.2 振幅が 3 値のモデル構成作用素 T

各 $(T\varphi)_{k\ell}$ については

$$(T\varphi)_{k\ell} \equiv \begin{cases} 1 & \dots h^+_{k\ell} < (S\varphi)_{k\ell} \leq 1 \text{ のとき} \\ 0 & \dots h^-_{k\ell} \leq (S\varphi)_{k\ell} \leq h^+_{k\ell} \text{ のとき} \\ -1 & \dots -1 \leq (S\varphi)_{k\ell} < h^-_{k\ell} \text{ のとき} \end{cases}$$

$$k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \quad (3.5)$$

と定義され、閾値 h^+, h^- は、

$$h^+ \equiv |h^+_{k\ell}| \quad k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \quad (3.6)$$

$$h^- \equiv |h^-_{k\ell}| \quad k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \quad (3.7)$$

であり、 $h^-_{k\ell}, h^+_{k\ell}$ は不等式

$$-1 < h^-_{k\ell} \leq 0 \leq h^+_{k\ell} < +1 \quad (3.8)$$

を満たすものである。

本シミュレーションでは、

$$h^+_{k\ell} = 0.5 \quad h^-_{k\ell} = -0.5 \quad (3.9)$$

と与えている。 $(S\varphi)_{k\ell}$ については式(3.3)で定義されている。

3.3 振幅が5値のモデル構成作用素 T

各 $(T\varphi)_{k\ell}$ については

$$(T\varphi)_{k\ell} \equiv$$

$$\begin{cases} 1 \cdots 2/3 < (S\varphi)_{k\ell} \leq 1 \\ 0.5 \cdots 1/3 < (S\varphi)_{k\ell} \leq 2/3 \\ 0 \cdots -1/3 \leq (S\varphi)_{k\ell} \leq 1/3 \\ -0.5 \cdots -2/3 \leq (S\varphi)_{k\ell} < -1/3 \\ -1 \cdots -1 \leq (S\varphi)_{k\ell} < -2/3 \end{cases}$$

$$k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \quad (3.10)$$

と求める。 $(S\varphi)_{k\ell}$ については式(3.3)で定義されている。

3.4 パターンモデル $T\varphi$ を画像として表示する処理

画像理解システム IUS は式(2.11)の原パターン(処理の対象とする問題の入力パターン) φ に対応し、そのパターンモデル $T\varphi$ をシステム内部の表現として確保する機能を備えている。本システム IUS は、パターンモデル $T\varphi$ を見るとあたかも原パターン φ であるかのように錯覚する(パターンモデル $T\varphi$ と原パターン φ 間の同一知覚原理)。画像として表示する際の濃度値は、 $T\varphi$ のとる各値に各々 127 を掛け整数化したものであり、 $-127 \sim +127$ の濃度値を持つ画像として出力される。

4. axiom 2 を満たす採用した類似度関数 SM

本章では、座標値 $x = \langle x_1, x_2 \rangle$ に視点を持つパターン認識システム RECOGNITRON $\langle x_1, x_2 \rangle$ が内蔵している axiom 2 を満たす類似度関数 (similarity-measure function)

$$SM : \Phi \times \Omega \rightarrow \{s \mid 0 \leq s \leq 1\} \quad (4.1)$$

の5構成例が説明される。式(4.1)の SM は付録 A の A3 で説明されており、そこでの axiom 2 を満たさなければならないが、パターン集合 Φ と代表パターン集合 Ω との直積集合 $\Phi \times \Omega$ から 1 より大きくない非負実数値集合 $\{s \mid 0 \leq s \leq 1\}$ への写像である。ここに、

$$\Omega = \{\omega_j \mid j \in J\} \quad (4.2)$$

は、第 $j \in J$ 番目のカテゴリ \mathcal{C}_j の代表パターン $\omega [j]$ の集まり(代表パターン集合)であり、また、

カテゴリの総数を m として,

$$J = \{1, 2, \dots, m\} \quad (4.3)$$

はカテゴリ番号の全集合である。以下では、 ω_j を $\omega_j[j]$ と表すことがある。 $SM(\varphi, \omega_j)$ (= $SM(\varphi, \omega[j])$)はモデル $T\varphi$ が第 $j \in J$ 番目のカテゴリ \mathcal{C}_j の代表パターン ω_j (= $\omega[j]$)のモデル $T\omega[j]$ に似ている程度を反映した値である。 $P(\mathcal{C}_j)$ は \mathcal{C}_j の生起確率であり、カテゴリ総数を m として、等確率に

$$P(\mathcal{C}_j) = 1/m \quad (4.4)$$

と与えている。

本章では、2文献 [3], [4] での axiom 2を満たすの5種類の写像 SM が各節で説明される。

4.1 相互情報量を使ったSM

式(A3.1)の全カテゴリ集合 \mathcal{C} 内の第 $j \in J$ 番目のカテゴリ \mathcal{C}_j の持つ諸性質を典型的に代表する代表パターンを $\omega[j] = \{\omega[j]_k \mid k=0, \pm 1, \pm 2, \dots, \pm p, 1=0, \pm 1, \pm 2, \dots, \pm q\}$ と導入する。2つのパターンモデル $T\varphi, T\omega[j]$ について、 $T\varphi$ 内に $T\omega[j]$ が含まれる程度を情報量として計量するには、相互情報量 (mutual information) を

$$MI(T\varphi, T\omega[j]) \equiv -(1/2) \cdot \log_e [1 - |nip(T\varphi, T\omega[j])|^2] \quad (4.5)$$

と定義すればよい。式(4.5)に登場している $nip(T\varphi, T\omega[j])$ は $T\varphi$ と $T\omega[j]$ との間の規格化内積 (normalized inner product) といわれるものであり、

$$nip(T\varphi, T\omega[j]) \equiv \begin{cases} 0 \cdots \|T\varphi\| \cdot \|T\omega[j]\| = 0 \text{ の場合} \\ (T\varphi, T\omega[j]) / \|T\varphi\| \cdot \|T\omega[j]\| \cdots \|T\varphi\| \cdot \|T\omega[j]\| > 0 \text{ の場合} \end{cases} \quad (4.6)$$

と定義される。そして、この相互情報量 $MI(T\varphi, T\omega[j])$ を用いて、 SM が

$$SM(\varphi, \omega[j]) \equiv \begin{cases} P(\mathcal{C}_j) \cdots \sum_{j \in J} MI(T\varphi, T\omega[j]) = 0 \text{ のとき} \\ MI(T\varphi, T\omega[j]) / \sum_{i \in J} MI(T\varphi, T\omega[i]) \cdots \sum_{j \in J} MI(T\varphi, T\omega[j]) > 0 \text{ のとき} \end{cases} \quad (4.7)$$

と構成される。

4.2 各カテゴリに帰属する複数個のパターンを使ったSM①

本節では、第 $j \in J$ 番目のカテゴリ \mathcal{C}_j に帰属するパターン $\psi[j; q]$ の集合

$$\Psi_j \equiv \Psi[j] = \{\psi[j; 1], \psi[j; 2], \dots, \psi[j; r]\} \quad (4.8)$$

が導入される。各パターン $\psi[j; q]$ は、視点 $x = \langle x_1, x_2 \rangle$ が固定された状況下で、

$$\psi[j; q] = \{\psi[j; q]_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q\} \quad (4.9)$$

$$\psi[j; q]_{k\ell} \equiv \psi[j; q](x_1+k, x_2+\ell) \quad (4.10)$$

である。次に、 $i \neq j$ であれば、

$$\|T\omega[i] - T\omega[j]\| > 0 \quad (4.11)$$

であるように、各 $\omega[j]$ は選定されなければならない。更に、2つの相異なるモデル集合

$$T\Psi[j] \equiv \{T\psi[j; q] \mid q=1 \sim r\}, T\Psi[i] \quad (j \neq i)$$

も共通の要素を持たないように各 $\psi[j; q]$ を選定されていなければならない。

カテゴリ番号 $j \in J$ を助変数とする関数 $g_j(\varphi)$ を

$$g_j(\varphi) = \min \{ \|T\varphi - T\psi\| \mid \psi \in \Psi[j] \} \quad (4.12)$$

と定義する。 $\text{Min} \{f(z) \mid \text{condition } c\}$ はcondition c を満たす $f(z)$ の最小値の意である。ここで、

$$\text{ndis}(T\varphi, T\psi) = \|T\varphi - T\psi\| \quad (4.13)$$

と定義される $\text{ndis}(T\varphi, T\psi)$ はノルム距離と呼ばれ、パターン $T\varphi$ と $T\psi$ 間の相違の程度を測定する1つの方法である。

次に、関数 $f_j(\varphi)$ を

$$f_j(\varphi) = \min_{i \in J - \{j\}} g_i(\varphi) \quad (4.14)$$

と定義する。 $\min_{\text{cond}} f(z)$ は条件 cond を満たす $f(z)$ の最小値の意である。

その後、 $\text{SM}(\varphi, \omega[j])$ を

$$\text{SM}(\varphi, \omega[j]) \equiv \begin{cases} f_j(\varphi) / \sum_{i \in J} f_i(\varphi) & \cdots \quad \sum_{i \in J} f_i(\varphi) > 0 \quad \text{のとき} \\ P(\mathcal{C}_j) & \cdots \quad \sum_{i \in J} f_i(\varphi) > 0 \quad \text{のとき} \end{cases} \quad (4.15)$$

と定義する。

このとき、すべての $j \in J$ について

$$\text{SM}(\psi[j; q], \omega[j]) = 1 \quad (4.16)$$

がすべての $q (= 1, 2, \dots, r)$ につき成り立つ。また、

$$\text{SM}(\psi[i; q], \omega[j]) = 0 \quad (i \neq j) \quad (4.17)$$

がすべての q , すべての i, j につき成り立っている。

4.3 量子化SM

本節では、類似度 $\text{SM}(\varphi, \omega[j])$ の値を2つの閾値 $s_0(j)$, $s_1(j)$ を使い、 $s_0(j)$ より大きくない値は0に切り下げし、また、 $s_1(j)$ より小さくない値は1に切り上げするという量子化方法を述べる。Axiom 2 を満たす類似度関数 SM から得られた式(4.23)の量子化類似度関数 SM' は axiom 2 を満たすことを保存している。この量子化により、パターンの変形を吸収し、多段階認識過程の収束を速めることができる。

4.3.1 SMを量子化する方法

本項では、 $\text{SM}(\varphi, \omega[j])$ の値を2つの閾値 $s_0(j)$, $s_1(j)$ を導入し、量子化する方法(量子化方法)を述べる。

第 $j \in J$ 番目のカテゴリに帰属しているパターン

$$\varphi[j; t] = \{\varphi[j; t]_k \mid k = 0, \pm 1, \pm 2, \dots, \pm p, l = 0, \pm 1, \pm 2, \dots, \pm q\} \quad (4.18)$$

の集合

$$\Phi_j = \{\varphi[j; t] \mid t = 1 \sim N_j\} \quad (4.19)$$

を導入し、閾値 $s_0(j)$, $s_1(j)$ は、2つの不等式

$$s_1(j) \leq \min_{\varphi \in \Phi_j} \text{SM}(\varphi, \omega[j]) \quad (4.20)$$

$$\max_{k \in J - \{j\}} \max_{\varphi \in \Phi_k} \text{SM}(\varphi, \omega[j]) \leq s_0(j) \quad (4.21)$$

を満たすように決められる。その後、 $\text{SM}(\varphi, \omega[j])$ を

$$s(\varphi, j) \equiv \begin{cases} 1 & \cdots \quad s_1(j) \leq \text{SM}(\varphi, \omega[j]) \leq 1 \quad \text{のとき} \\ [\text{SM}(\varphi, \omega[j]) - s_0(j)] / [s_1(j) - s_0(j)] & \cdots \quad s_0(j) < \text{SM}(\varphi, \omega[j]) < s_1(j) \quad \text{のとき} \\ 0 & \cdots \quad 0 \leq \text{SM}(\varphi, \omega[j]) \leq s_0(j) \quad \text{のとき} \end{cases} \quad (4.22)$$

へと変換し、

$$\begin{aligned}
& \text{SM}'(\varphi, \omega[j]) \equiv \\
& \begin{cases} s(\varphi, j) / \sum_{k \in J} s(\varphi, k) \cdots \sum_{k \in J} s(\varphi, k) > 0 \text{ のとき} \\ p(\mathcal{C}_j) \cdots \sum_{k \in J} s(\varphi, k) = 0 \text{ のとき} \end{cases} \quad (4.23)
\end{aligned}$$

を求める．この $\text{SM}'(\varphi, \omega[j])$ の値が $\text{SM}(\varphi, \omega[j])$ を量子化した値である．

尚， $s_0(j) < s_1(j)$ が成り立っていないならばならないことに注意する．閾値 $s_0(j)$ ， $s_1(j)$ の選び方については，文献 [6] で説明されている．

4.4 各カテゴリに帰属する複数個のパターンを使ったSM②

本節4.4と2節4.5, 4.6は，節4. 2と同じく各カテゴリに帰属する複数のパターンを使ってSMを求める式(4.15)による方法である．従って，節4.2の，式(4.12)の関数 $g_j(\varphi)$ の与え方を変えるだけで，式(4.8)の Ψ_j や式(4.14)の関数 $f_j(\varphi)$ などはまったく同じであるため，以下では， $g_j(\varphi)$ の選び方のみ示す．

関数 $g_j(\varphi)$ は，

$$\begin{aligned}
& g_j(\varphi) = \\
& \begin{cases} \min_{\varphi \in \Psi_j} -2^{-1} \cdot \log_e |\text{nip}(\text{T}\varphi, \text{T}\psi)|^2 \\ \cdots \text{すべての } \psi \in \Psi_j \text{ について} \\ \|\text{T}\varphi\| \cdot \|\text{T}\psi\| > 0 \text{ かつ } |\text{nip}(\text{T}\varphi, \text{T}\psi)|^2 > \epsilon > 0 \text{ のとき} \\ -2^{-1} \cdot \log_e \epsilon \\ \cdots \text{ある1つの } \psi \in \Psi_j \text{ について} \\ \|\text{T}\varphi\| \cdot \|\text{T}\psi\| = 0 \text{ かつ } |\text{nip}(\text{T}\varphi, \text{T}\psi)|^2 \leq \epsilon \text{ のとき} \end{cases} \quad (4.24)
\end{aligned}$$

と，定義される．定義式(4.24)に登場している閾値 ϵ は，不等式

$$0 < \epsilon < \min_{i \in J} \min_{i' \in J - |j|} \{ |\text{nip}(\text{T}\psi, \text{T}\psi')|^2 \mid \psi \in \Psi[i], \psi' \in \Psi[j] \} \quad (4.25)$$

を満たすように，

$$\epsilon = (1 / |J|) \cdot \min_{i \in J} \min_{i' \in J - |j|} \{ |\text{nip}(\text{T}\psi, \text{T}\psi')|^2 \mid \psi \in \Psi[i], \psi' \in \Psi[j] \} \quad (4.26)$$

と与える．

4.5 各カテゴリに帰属する複数個のパターンを使ったSM③

前節4.4と同じく，本節では， $g_j(\varphi)$ を

$$g_j(\varphi) = \min \{ 1 - |\text{nip}(\text{T}\varphi, \text{T}\psi)|^2 \mid \psi \in \Psi[j] \} \quad (4.27)$$

と与える．ここで，

$$\text{idis}(\text{T}\varphi, \text{T}\psi) = 1 - |\text{nip}(\text{T}\varphi, \text{T}\psi)|^2 \quad (4.28)$$

と定義される $\text{idis}(\text{T}\varphi, \text{T}\psi)$ は内積相違度と呼ばれ，パターン $\text{T}\varphi$ と $\text{T}\psi$ 間の距離を計測する1つの方法である．

4.6 各カテゴリに帰属する複数個のパターンを使ったSM④

本節では，各カテゴリについて $a_j(j \in J)$ を，

$$a_j = 3^{-1} \cdot \min \min \{ \|\text{T}\psi - \text{T}\psi'\|^2 \mid \psi \in \Psi[i], \psi' \in \Psi[j], \Psi[i] \neq \Psi[j] \} > 0 \quad (4.29)$$

と求めた後，前節4.5と同じく $g_j(\varphi)$ を

$$g_j(\varphi) = \min \{1 - \exp[-a_j^{-1} \cdot \|T\varphi - T\psi\|^2 \mid \psi \in \Psi[j]\} \quad (4.30)$$

と選ぶ。

5. 各カテゴリの代表パターンの学習による決定

第 $j \in J$ 番目の代表パターン $\omega[j]$ を式(5.7)の訓練パターンの集合

$$\varphi[j;0], \varphi[j;1], \dots, \varphi[j;t] \quad (5.1)$$

に適応させる形式で決定する手法(学習による決定法; 文献 [3] での付録 I での学習アルゴリズム)が文献 [5], [6] で説明されている. この学習アルゴリズムにつき, 採用している終了条件もその文献で説明されている.

尚, 画像理解システム IUS は任意の訓練パターンから各代表パターンを視察で選ぶことも可能である.

6. axiom 3を満たす採用した大分類関数BSC

付録Aの章A4で解説されている大分類関数(rough classifier; binary-state classifier)

$$\text{BSC} : \Phi \times J \rightarrow \{0, 1\} \quad \text{ここに, } \Phi \text{ はパターン } \varphi \text{ の集合, } J \text{ はカテゴリ番号 } j \text{ の集合} \quad (6.1)$$

はそこでの axiom 3 を満たすように構成されなければならないが, このBSCには, 外界の事物 φ が帰属するかもしれないカテゴリの候補を複数個出力する働きがある.

axiom 3 を満たす文献 [3], [4] での大分類関数BSCは, 知識抽出ネットワーク (knowledge-extraction network) である. 画像理解システム IUS の各成分 RECOGNITRON の持っている知識が処理の対象としている画像内に存在するかどうかを検索するための知識抽出ネットワークとして, BSCは働き, ファジィ推論規則 (fuzzy inference-rule) を用いて近似的推論 (approximate reasoning) するニューラルネット (fuzzy-inference neural net) として, 以下のごとく構成される.

6.1 BSCの構造形式

2値関数BSCは $\Phi \times J$ から集合 $\{0, 1\}$ への写像

$$\text{BSC} : \Phi \times J \rightarrow \{0, 1\} \quad (6.2)$$

であり, 次のように定義される.

$$\text{BSC}(\varphi, j) \equiv \text{psn}(y(\varphi, j)) \quad (6.3)$$

ここに, 1実数値変数 u の2値関数psn(positive sign function)は

$$\text{psn}(u) = \begin{cases} 1 & \dots u \geq 0 \text{ の場合} \\ 0 & \dots u < 0 \text{ の場合} \end{cases} \quad (6.4)$$

と定義されており, 現実出力 $y(\varphi, j)$ は

$$y(\varphi, j) \equiv W\text{mean}(\varphi, j) - b[j] \quad (6.5)$$

と定義される. ここに, $W\text{mean}(\varphi, j)$ は,

$$W\text{mean}(\varphi, j) \equiv \sum_{k=-p}^{+p} \sum_{\ell=-p}^{+q} W[j]_{k\ell} \cdot p_{k\ell}(\varphi, j) \quad (6.6)$$

$$p_{k\ell}(\varphi, j) \equiv q_{k\ell}(\varphi, j) / \sum_{i \in J} q_{k\ell}(\varphi, i) \quad (6.7)$$

$$q_{k\ell}(\varphi, j) \equiv \exp[-\{(T\varphi)_{k\ell} - (T\omega[j]_{k\ell})\}^2 / (\sigma[j]_{k\ell})^2] \quad (6.8)$$

と、定義されている。

$y^\wedge(j)$ をカテゴリに関する理想出力(desired output)とすると、理想的な fuzzy 推論規則

$$\text{if } (T\varphi) \text{ is } T\omega[j] \text{ then } y^\wedge(j)_{k\ell} \text{ is } W[j]_{k\ell} - b[j] \quad (6.9)$$

を表すように、式(6.2)の大分類関数 $BSC(\varphi, j)$ の構造が設定されている。

上述の neuro-fuzzy system としての構造形式を備えている BSC 内の 3 助変数 $W[j]_{k\ell}$ (重み), $b[j]$ (閾値), $\sigma[j]_{k\ell}$ (標準偏差の $1/\sqrt{2}$ であるが、簡単に標準偏差と称していることに注意) を次節 6.2 の最急降下学習によって求める。

尚、本 RECOGNITRON では、各視点(注目している各画素) $x = \langle x_1, x_2 \rangle$ についても共通に学習しており、すべての視点(画素)についても共通の学習結果を使って認識を行っている。

6.2 BSC 内の $W[j]_{k\ell}$, $b[j]$, $\sigma[j]_{k\ell}$ の最急降下法による学習

現実出力 $y(\varphi[j_i, t], j_i)$ の、理想出力 $y^\wedge(j_i)$ からの違い $[y(\varphi[j_i, t], j_i) - y^\wedge(j_i)]$ の自乗の $1/2$ を最小とするように、6 式(6.3)~(6.8)で定義される BSC 内の重み $W[j]_{k\ell}$, 閾値 $b[j]$, 標準偏差 $\sigma[j]_{k\ell}$ を学習する方法が文献 [6] で説明されている。そのために、あらかじめ訓練パターン (training pattern) の系列

$$\varphi[j_i, t] = \{\varphi[j_i, t]_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, p, \ell=0, \pm 1, \pm 2, \dots, q\} \\ t=0, 1, \dots \quad (6.10)$$

を用意した状況下で、最急降下法(method of steepest descent)によって学習する方法が採用されている。ここに、は時刻 t に入力される訓練パターンが帰属するカテゴリ番号である。

大分類関数 BSC 内の助変数 $W[j]_{k\ell}$, $b[j]$, $\sigma[j]_{k\ell}$ はお互いに影響を与え合い、相互の関係の中で適切な値が定まってくるように、学習される。

すべてのカテゴリについて $W[j]_{k\ell}$, $b[j]$, $\sigma[j]_{k\ell}$ のそれぞれの最急降下法に基づく学習が終了するならば、改めて、 $b[j]$ の最適化を行い、学習処理全体が終了する。

6.3 非逐次学習

本研究では前節 6.2 で述べた逐次学習 (sequential learning) の他に、一括学習 (batch learning) を一部取り入れた非逐次学習 (non-sequential learning) も取り入れている。以下に、その説明をしておこう。

$W[j]_{k\ell}$, $b[j]$, $\sigma[j]_{k\ell}$ の値を初期化したら、時刻 t を固定し、カテゴリ番号の若い番号から順に $W[j_{i+1}, t+1]_{k\ell}$ の値を求めていく。すべてのカテゴリについて $W[j_{i+1}, t+1]_{k\ell}$ を求めたら、 $W[j_{i+1}, t+1]_{k\ell}$ の値を次の時刻までの一時的な $W[j]_{k\ell}$ として採用し時刻 t を 1 つ進める。次に、前段階での $W[j_{i+1}, t+1]_{k\ell}$ の値を、今回の $W[j_i, t]_{k\ell}$ として、今回の $W[j_{i+1}, t+1]_{k\ell}$ を求め、この処理を繰り返す。もし、終了条件を満たしたら $W[j_i, t]_{k\ell}$ を $W[j]_{k\ell}$ として採用し、以後そのカテゴリについては時刻 t が進んでも学習を行わない。全てのカテゴリが終了条件を満たしたら、 $W[j]$ の学習は終了し、次は $b[j]$ を $W[j]_{k\ell}$ と同様の方式で学習を行う。

$b[j]$ の学習が終了したら、 $\sigma[j]_{k\ell}$ を $W[j]_{k\ell}$ や $b[j]$ と同様の手順で学習を行う。

$\sigma[j]_{k\ell}$ の学習が終了したら、 $b[j]$ の最適化を行い、学習処理全体を終了する。

7. axiom 4を満たすカテゴリ選択関数CSFの構成

多段階連想形不動点探索認識の働きは節8.5で説明されるが、この働きを実現するのにも使われるカテゴリ帰属知識の変換に必要な節8.4の構造受精変換 $TA(\mu)T$ においては、2文献 [3], [4]でのaxiom 4を満たすカテゴリ選択関数 (category-selection function)

$$CSF : \Phi \times 2^J \rightarrow 2^J \quad (7.1)$$

が用いられる。付録Aの章A5で解説されているCSFは付録Aの定理A4.1で得られているものが採用されている。定理A4.1によれば、カテゴリ選択関数CSFはaxiom 2, axiom 4を各々満たす類似度関数SMと大分類関数BSCが与えられれば、自ずから決まってくることに注意しておく。

パターン φ がカテゴリ集合

$$\mathcal{C}_j, j \in \lambda \in 2^J$$

の1つのカテゴリに帰属しているかも知れない時、その内、意味あるカテゴリ (significant category)の番号のリストが $CSF(\varphi, \lambda) \in 2^J$ である。関数 $CSF(\varphi, \lambda)$ は、任意の $\varphi \in \Phi, \lambda \in 2^J$ について

① $\varphi = 0$ または $\lambda = \phi$ (空集合) のとき

$$CSF(\varphi, \lambda) \equiv \phi \text{ (空)} \quad (7.2)$$

② $\varphi \neq 0$ かつ $\lambda \neq \phi$ のとき

②-1 $\sum_{k \in \lambda} BSC(\varphi, k) = 0$ ならば

$$CSF(\varphi, \lambda) \equiv \{k \in \lambda \mid SM(\varphi, \omega_k) > 0\} \quad (7.3)$$

②-2 $\sum_{k \in \lambda} BSC(\varphi, k) > 0$ ならば

$$CSF(\varphi, \lambda) \equiv \{k \in \lambda \mid SM(\varphi, \omega_k) > 0 \text{ かつ } BSC(\varphi, k) = 1\} \quad (7.4)$$

と定義される。

8. カテゴリ帰属知識を使用したRECOGNITRONによる認識・理解の方法 (不動点連想形多段階認識法;SS認識法)

本章では、先ず、RECOGNITRONがパターンに対し持つカテゴリ帰属知識が説明され(節8.1)、次に、構造受精作用素 $A(\mu)$ が構成され(節8.2)、また、構造受精変換 $TA(\mu)T$ を定義する(節8.4)ために必要なカテゴリ帰属知識間に2元関係 $=_{\Delta}$ (節8.3)を導入する。多段階的に連想しながら構造受精変換の不動点を探索することによって認識するスキーマとして、多段階連想形不動点探索認識(SS多段階想起認識)の働きが説明される(節8.5)。この種の不動点探索を行う2つの方法として、線形探索節(節8.6)、修正線形探索(節8.7)が説明される。

8.1 RECOGNITRONがパターンに対し持つカテゴリ帰属知識

処理の対象とする問題のパターン φ に関し、RECOGNITRONが事前に全く知識を持っていないくて無知(ignorance)の場合、RECOGNITRONは、

パターン φ はカテゴリ集合

$$\mathcal{C}(J) \equiv \{\mathcal{C}_j \mid j \in J\} \quad (8.1)$$

の内の1つに帰属している

というカテゴリ帰属知識を持っているといえよう。一般に、処理の対象とする問題のパターン $\varphi \in$

Φ に関し、RECOGNITRONが、パターン φ は候補カテゴリの部分集合

$$\mathcal{C}(\gamma) \equiv \{ \mathcal{C}_j \mid j \in \gamma \subseteq J \} \quad (8.2)$$

の内の1つに帰属しているというカテゴリ帰属知識 (categorical membership knowledge) を持っている事態を、

$$\langle \varphi, \gamma \rangle \quad (8.3)$$

と表す。 $\gamma \subseteq J$ はRECOGNITRONによる探索で発見された候補カテゴリの番号のなすリストである。すべての $\langle \varphi, \gamma \rangle$ の集まりをカテゴリ帰属知識空間 (categorical membership knowledge space) といい、

$$\langle \Phi, 2^J \rangle \quad (8.4)$$

と表す。 2^J は

$$2^J \equiv \{ \gamma \mid \gamma \subseteq J \} \quad (8.5)$$

と定義され、カテゴリ番号の全集合 J のすべての部分集合のなす集合である。

多段階連想形不動点探索認識 (SS多段階想起認識) の働きで、

ある1つのカテゴリ番号 $j \in J$ を見つけ、 $\langle T\varphi, J \rangle$ を $\langle T\omega_j, [j] \rangle$ へと変換し、

入力されたパターン φ を第 j 番目のカテゴリ \mathcal{C}_j に帰属し、

$$\varphi \text{ は } T\omega_j \text{ の形で再生される} \quad (8.6)$$

と想起形認識することが、RECONITRONには課せられる。

8.2 構造受精作用素 $A(\mu)$ の構成

式 (8.2) のカテゴリ部分集合 $\mathcal{C}(\gamma)$ の内の1つに帰属するかも知れない様に、パターン $\varphi \in \Phi$ を変換することを考えよう。カテゴリ番号リスト $\mu \in 2^J$ を助変数に持つこのような変換目的に役立つ構造受精作用素

$$A(\mu) : \Phi \rightarrow \Phi \quad (\mu \in 2^J) \quad (8.7)$$

は、任意の $\varphi \in \Phi$ について、次の①、②で定義される (文献 [3] の節6.5を参照) :

① $\varphi = 0$ または $\mu = \phi$ (空集合) のとき

$$A(\mu) \varphi \equiv 0. \quad (8.8)$$

② $\varphi \neq 0$ かつ $\mu \neq \phi$ (空集合) のとき

②-1 $\sum_{k \in \mu} \text{BSC}(j, k) = 0$ ならば

$$A(\mu) \varphi \equiv \sum_{k \in \mu} \text{SM}(\varphi, \omega_k) \cdot T\omega_k. \quad (8.9)$$

②-2 $\sum_{k \in \mu} \text{BSC}(j, k) > 0$ ならば

$$A(\mu) \varphi \equiv \sum_{k \in \mu} \text{SM}(\varphi, \omega_k) \cdot \text{BSC}(\varphi, k) \cdot T\omega_k. \quad (8.10)$$

□

パターン φ が作用素 $A(\mu)$ で変換された結果 $A(\mu)\varphi$ は、その1次結合係数が1より大きくない非負実数 $\text{SM}(\varphi, \omega_k)$ 、或いは、 $\text{SM}(\varphi, \omega_k) \cdot \text{BSC}(\varphi, k)$ であるような $T\omega_k$ 、 $k \in \mathcal{C}(\mu)$ の非負1次結合であることに注意する。この事実は、付録Bの3式 (B.1), (B.2), (B.3) に反映されている。

8.3 カテゴリ帰属知識間の2元関係 $=_{\Delta}$

2つのカテゴリ帰属知識 $\langle \varphi, \gamma \rangle$, $\langle \eta, \mu \rangle$ が2元関係 $=_{\Delta}$ にあることを、

$$\langle \varphi, \gamma \rangle =_{\Delta} \langle \eta, \mu \rangle < \Phi, 2^J \rangle \quad (8.11)$$

と表す。ここに、

$$\langle \varphi, \gamma \rangle =_{\Delta} \langle \eta, \mu \rangle \Leftrightarrow [[\varphi = \eta] \wedge [\gamma = \mu]] \quad (8.12)$$

と定義される．パターン間の相等関係 $\varphi = \eta$ と候補カテゴリリスト間の相等関係 $\gamma = \mu$ とを表している2元関係 $=_{\Delta}$ は、カテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ の、 $\text{TA}(\mu)\text{T}$ による次節の変換結果式 (8.15) に登場している．

8.4 構造受精変換 $\text{TA}(\mu)\text{T}$ の定義

式 (8.7) の写像 $A(\mu)$ の両側にモデル構成作用素 T を配置して得られる写像

$$\text{TA}(\mu)\text{T} : \Phi \rightarrow \Phi \quad (8.13)$$

は、その定義域 Φ 、値域 Φ の双方を拡張して、

$$\text{TA}(\mu)\text{T} : \langle \Phi, 2^J \rangle \rightarrow \langle \Phi, 2^J \rangle \quad (8.14)$$

と考え直すことができる．それには、 $\text{TA}(\mu)\text{T}\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle$ を、

$$\text{TA}(\mu)\text{T}\langle \varphi, \gamma \rangle =_{\Delta} \langle \text{TA}(\mu \cap \gamma)\text{T}\varphi, \text{CSF}(\varphi, \mu \cap \gamma) \rangle \in \langle \Phi, 2^J \rangle \quad (8.15)$$

と定義すればよい．

カテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ の、 $\text{TA}(\mu)\text{T}$ による変換結果 $\text{TA}(\mu)\text{T}\langle \varphi, \gamma \rangle$ においては、 $\text{CSF}(\varphi, \mu \cap \gamma)$ は、 φ の帰属するかも知れない候補カテゴリの集合

$$\mathcal{C}_j, j \in \gamma \quad (8.16)$$

を

$$\mathcal{C}_j, \mu \cap \gamma \quad (8.17)$$

へと限定し、更にその意味ある候補カテゴリ集合

$$\mathcal{C}_j, j \in \text{CSF}(\varphi, \mu \cap \gamma) \quad (8.18)$$

へと限定している．この限定された候補カテゴリの集合の1つに帰属するかも知れないパターンモデルが、 φ のモデル $\text{T}\varphi$ から得られている $\text{TA}(\mu \cap \gamma)\text{T}\varphi$ である．

8.5 多段階連想形不動点探索認識 (SS多段階想起認識) の働き

本節では、カテゴリ帰属知識を連想変換して行ってこの変換の不動点を求める形式を備えた多段階認識の働きと、その認識結果とについて説明される．

8.5.1 RECOGNITRONによる認識

RECOGNITRONの注目している入力画像の画素 $x = \langle x_1, x_2 \rangle$ を中心とする $(2p+1) \cdot (2q+1)$ 個の画素の集合上の入力パターン

$$\varphi = \{ \varphi_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \} \quad (8.19)$$

に対し、多段階連想形不動点探索認識 (SS多段階想起認識) の働きで得られる第 t ($=0, 1, 2, \dots$) 認識段階における出力パターンを

$$\varphi^t = \{ \varphi^t_{k\ell} \mid k=0, \pm 1, \pm 2, \dots, \pm p, \ell=0, \pm 1, \pm 2, \dots, \pm q \} \quad (8.20)$$

と表現する．

パターン (1枚の画像内の事物) φ のモデル $\text{T}\varphi$ が、或るカテゴリ \mathcal{C}_j の代表パターン ω_j のモデル $\text{T}\omega_j$ からあまり変形していなくて、かつ、任意の $i \in J - \{j\}$ についてのカテゴリ \mathcal{C}_j の代表パターン ω_j のモデル $\text{T}\omega_j$ に一致していない場合の、ある特別な状況を考えよう．RECOGNITRONは、この $\text{T}\varphi$ を多段階連想形不動点探索認識 (multi-stage associative recognition based on searching for a fixed-point) の働きで、或るカテゴリ、 \mathcal{C}_j の代表パターン ω_j のモデル $\text{T}\omega_j$ へと連想形変換し、

$$\text{パターン } \varphi \text{ を 第 } j \in J \text{ 番目のカテゴリ } \mathcal{C}_j \text{ に帰属する} \quad (8.21)$$

と、認識できる場合がある(SS認識法において認識可能の場合)．情報が欠如している状況下でとりあえず矛盾のない結論を導出するというデフォルト推論(default reasoning)でなされるこの連想形変換は、8.1節で導入された

「RECOGNITRONがパターン $\varphi \in \Phi$ に対し持つカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ 」
に対し式(8.15)の構造受精変換 $TA(\mu)T$ を用いてなされる．

具体的には、次のように説明される．多段階認識の第0段階においてRECOGNITRONがパターン φ に対し持つカテゴリ帰属知識 $\langle \varphi^0, \lambda^0 \rangle$ を

$$\langle \varphi^0, \lambda^0 \rangle =_{\Delta} \langle T\varphi, J \rangle \in \langle \Phi, 2^J \rangle \quad (8.22)$$

と設定し、次第 t 認識段階における想起パターン φ^t と、そのパターン φ^t の帰属する可能性のあるカテゴリ番号リスト λ^t との対(pair)によって表現される“第 $t-1$ 段階でのカテゴリ帰属知識” $\langle \varphi^t, \lambda^t \rangle$ を、第 $t-1$ 段階での知識 $\langle \varphi^{t-1}, \lambda^{t-1} \rangle$ から

$$\langle \varphi^t, \lambda^t \rangle =_{\Delta} TA(\mu^{t-1})T\langle \varphi^{t-1}, \lambda^{t-1} \rangle, t=0, 1, 2, \dots \quad (8.23)$$

ここに、

$$\varphi^t = TA(\mu^{t-1} \cap \lambda^{t-1})T\varphi^{t-1} \quad (8.24)$$

$$\lambda^t = CSF(\varphi^{t-1}, \mu^{t-1} \cap \lambda^{t-1}) \quad (8.25)$$

と求め、つまり、段階的に

$$\langle \varphi^0, \lambda^0 \rangle \rightarrow \langle \varphi^1, \lambda^1 \rangle \rightarrow \dots \rightarrow \langle \varphi^{t-1}, \lambda^{t-1} \rangle \rightarrow \langle \varphi^t, \lambda^t \rangle \rightarrow \dots \quad (8.26)$$

と連想していき、このカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ が不動点となる認識段階番号(最終段階番号；final stage number) t が発見されたとき、すなわち、カテゴリ帰属知識の不動点方程式(fixed-point equation)

$$\langle \varphi^{t+1}, \lambda^{t+1} \rangle =_{\Delta} \langle \varphi^t, \lambda^t \rangle \quad (8.27)$$

、つまり、2不動点方程式

$$\varphi^{t+1} = \varphi^t \quad \text{かつ} \quad \lambda^{t+1} = \lambda^t \quad (8.28)$$

が成立するとき、カテゴリ帰属知識を用いた不動点連想形多段階認識過程が終了する．カテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ が、 $\langle \varphi^0, \lambda^0 \rangle$ から式(8.21)の構造受精変換 $TA(\mu)T\lambda^t$ 不動点として連想されたカテゴリ帰属知識 λ^t 不動点である．終了するとき、

入力パターン φ は不動点パターン φ^t として再生され(パターン連想の働き)、
カテゴリ番号の集合 λ^t のいずれか1つの番号 $j \in J$ を持つカテゴリ \mathcal{C}_j に
帰属する(パターン認識の働き) (8.29)

と、処理される．これが多段階連想形認識の結果である．

尚、カテゴリ帰属知識の不動点方程式(8.27)が成立する認識段階番号 t が発見され得ない場合があることに注意しておく．

8.5.2 認識結果における認識可能・認識不定・認識不能

カテゴリ帰属知識の不動点方程式(8.27)が成立する認識段階番号 t が発見される場合を考えよう．RECOGNITRONによるSS多段階想起認識法による認識結果において、画像内の個々の事物 $\varphi \in \Phi$ についてRECOGNITRONの持っている知識と学習能力とに応じ、3認識結果

$$\left\{ \begin{array}{l} \text{認識可能(該当するカテゴリが唯1個あること；} |\lambda^t| = 1) \end{array} \right. \quad (8.30)$$

$$\left\{ \begin{array}{l} \text{認識不定(該当するカテゴリが複数個あること；} |\lambda^t| \geq 2) \end{array} \right. \quad (8.31)$$

$$\left\{ \begin{array}{l} \text{認識不能(該当するカテゴリが1個もないこと；} |\lambda^t| = 0) \end{array} \right. \quad (8.32)$$

が得られ、何が何でも、認識可能、つまり、該当するカテゴリを唯1つ求めるという解釈結果を出

力することはない。ここに、 $|B|$ は集合Bにふくまれる要素の個数の意である。

任意であるが1つの画素について2式(8.31)～(8.32)の3解釈結果が得られるという異常現象がシミュレーションについてにおいて観測された事実を検討するのが本論文の目的である。特に、該当するカテゴリが1個もないこと(認識不能)はRECOGNITRONの持っている知識と学習能力が不足しているか、過剰であることが原因となって生じると考えられる。

8.6 本画像理解システムIUSにおいて採用した2つの探索法

式(8.25)におけるカテゴリ番号リスト μ^{t-1} は一般的には、帰納推理(inductive reasoning)の働きで探索の結果、定まってくるものである。本画像理解システムIUSでは、探索法として簡単に、線形探索、修正線形探索の2つを採用した。本節では、文献[3]の章G12の(イ)、(ロ)で説明されているこの2つの探索法が簡単に説明される。

8.6.1 線形探索

式(8.23)のカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ における、カテゴリ番号リスト μ^{t-1} を

$$\mu^{t-1} = \lambda^{t-1} \quad (8.33)$$

と選んだ場合を線形探索と呼ぶ。登場しているカテゴリ選択関数CSFについては、7章で説明されている。本シミュレーションでは、式(8.22)から、

$$\lambda^0 = J \equiv \{1, 2, \dots, m\}, \quad m=6 \quad (8.34)$$

と選んでいる。

8.6.2 修正線形探索

式(8.23)のカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ における、カテゴリ番号リスト μ^{t-1} を

$$\mu^{t-1} = \lambda^{t-1} - \{i\} \quad (8.35)$$

と選んだ場合を修正線形探索と呼ぶ。ここに、カテゴリ番号 i は

$$\textcircled{1} \quad |CSF(\varphi^{t-1}, \mu^{t-1} \cap \lambda^{t-1})| > 1 \quad \text{のとき}$$

$$i = \arg \min_k SM(\varphi^{t-1}, \omega_k) \quad (8.36)$$

$$\text{ここに、} k \in CSF(\varphi^{t-1}, \mu^{t-1} \cap \lambda^{t-1}) \quad (8.37)$$

$$\textcircled{2} \quad |CSF(\varphi^{t-1}, \mu^{t-1} \cap \lambda^{t-1})| \leq 1 \quad \text{のとき}$$

$$\{i\} = \phi \quad (8.38)$$

である。

9. 任意の画素の各認識段階について詳細に検討する際に求められる値としての、認識情報量 RGetpy(φ^t)、認識情報量 RAIN(φ^{t+1}, φ^t)、カテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー(energy) $E(\varphi^t, \lambda^t)$ 、2つのカテゴリ帰属知識間の距離 SMdis($\langle \varphi^s, \lambda^s \rangle, \langle \varphi^{s+1}, \lambda^{s+1} \rangle$)

本章では、8.5、8.6の両章での多段階連想形不動点探索認識(SS多段階想起認識)の働きが良好に発現しているかどうかを示す諸指標について、説明される。この諸指標は“式(8.26)の認識過程に付随する諸量”である。任意の視点(画素)についての不動点探索認識がどのような段階を経て行われているかを詳細に検討するのに効果的な諸指標を9.1～9.4の4節に掲げよう。

9.1 認識情報量 RGetpy(φ^t)

認識情報量 (recognition entropy) $R\text{Getpy}(\varphi^t)$ は

$$R\text{Getpy}(\varphi^t) = \sum_{j \in J} \text{SM}(\varphi^t, \omega_j) \cdot \log_e [\text{SM}(\varphi^t, \omega_j) / P(\mathcal{C}_j)] \quad (9.1)$$

と定義される。但し、

$$\text{SM}(\varphi^t, \omega[j]) = 0 \text{ のときは, } \text{SM}(\varphi^t, \omega_j) \cdot \log_e [\text{SM}(\varphi^t, \omega_j) / P(\mathcal{C}_j)] = 0 \quad (9.2)$$

として計算する。

認識が正常に進んでいるときは通常, 増加性

$$R\text{Getpy}(\varphi^t) \leq R\text{Getpy}(\varphi^{t+1}) \quad (9.3)$$

の成立が期待される。認識過程の進展に伴い(認識段階番号 t が増大するに従い), 入力画像の画素 $x = \langle x_1, x_2 \rangle$ を中心とする $(2p+1) \cdot (2q+1)$ 個の画素の集合上の, 式(8.19)の入力パターン φ が帰属すかもしれない候補カテゴリが適切に絞られていけば(探索という帰納推理の働きが正常に進んでいるときは) $R\text{Getpy}$ の値は増加していく。

9.2 認識情報量 $\text{RAIN}(\varphi^{t+1}, \varphi^t)$

φ^t を基準にしての φ^{t+1} の認識情報量 $\text{RAIN}(\varphi^{t+1}, \varphi^t)$ は

$$\text{RAIN}(\varphi^{t+1}, \varphi^t) = \sum_{j \in J} \text{SM}(\varphi^{t+1}, \omega_j) \cdot \log_e [\text{SM}(\varphi^{t+1}, \omega_j) / \text{SM}(\varphi^t, \omega_j)] \quad (9.4)$$

と定義される。但し、

$$\begin{aligned} \text{SM}(\varphi^{t+1}, \omega_j) = 0 \quad \text{または} \quad \text{SM}(\varphi^t, \omega_j) = 0 \quad \text{のとき} \\ \text{SM}(\varphi^{t+1}, \omega_j) \cdot \log_e [\text{SM}(\varphi^{t+1}, \omega_j) / \text{SM}(\varphi^t, \omega_j)] = 0 \end{aligned} \quad (9.5)$$

として計算する。

認識が正常に進んでいるときは通常, 減少性

$$\text{RAIN}(\varphi^{t+1}, \varphi^t) \geq \text{RAIN}(\varphi^{t+2}, \varphi^{t+1}) \quad (9.6)$$

の成立が期待される。

9.3 カテゴリ帰属知識のポテンシャルエネルギー $E(\varphi^t, \lambda^t)$

カテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー (energy) $E(\varphi^t, \lambda^t)$ は次の様に定義される:

$$\textcircled{1} \quad \varphi^t = 0 \quad \text{あるいは} \quad \lambda^t = \phi \text{ (空集合) のとき} \\ E(\varphi^t, \lambda^t) = 0. \quad (9.7)$$

$$\textcircled{2} \quad \varphi^t \neq 0 \quad \text{かつ} \quad \lambda^t \neq \phi \text{ (空集合) のとき} \\ E(\varphi^t, \lambda^t) = |\lambda^t| - \sum_{j \in \lambda^t} \text{SM}(\varphi^t, \omega_j). \quad (9.8)$$

□

認識が正常に進んでいるときは通常, 減少性

$$E(\varphi^t, \lambda^t) \geq E(\varphi^{t+1}, \lambda^{t+1}) \quad (9.9)$$

の成立が期待される。

9.4 カテゴリ帰属知識間の距離 $\text{SMdis}(\langle \varphi^s, \lambda^s \rangle, \langle \varphi^{s+1}, \lambda^{s+1} \rangle)$

2つのカテゴリ帰属知識間の距離 SMdis は

$$\text{SMdis}(\langle \varphi^s, \lambda^s \rangle, \langle \varphi^{s+1}, \lambda^{s+1} \rangle) \equiv -\log_e \sum_j [\text{SM}(\varphi^s, \omega_j) \cdot \text{SM}(\varphi^{s+1}, \omega_j)]^{1/2} \quad (9.10)$$

と定義される。ここに, \sum_j はカテゴリ番号

$$j \in \text{CSF}(\varphi^s, \lambda^s) \cap \text{CSF}(\varphi^{s+1}, \lambda^{s+1}) \quad (9.12)$$

にわたる総和である。

認識が正常に進んでいるときは通常、減少性

$$\text{SMdis}(\langle \varphi^s, \lambda^s \rangle, \langle \varphi^{s+1}, \lambda^{s+1} \rangle) \geq \text{SMdis}(\langle \varphi^{s+1}, \lambda^{s+1} \rangle, \langle \varphi^{s+2}, \lambda^{s+2} \rangle) \quad (9.13)$$

の成立が期待される。

10. 認識過程における振動発生の可能性

式(8.26)の認識過程において、連想的に不動点探索を行っているとき、

$$\sum_j \text{SM}(\varphi^t, \omega_j) = 0 \quad (10.1)$$

が各認識段階 t で成立していれば、つまり、選ばれている式(4.1)の axiom 2 を満たす類似度関数 SM が直交性であれば(付録Dを参照)、式(8.27)の不動点方程式は必ず成立し、連想形認識は有限の認識段階番号で終了する。ここに、式(10.1)においては、カテゴリ番号 j は、カテゴリ番号の部分集合

$$J - \text{CSF}(\varphi^{t-1}, \lambda^{t-1}) \quad (10.2)$$

についての和である。しかし上述の式(10.1)が認識の各段階で成立していなければ(axiom 2 を満たす類似度関数 SM が付録Dでの直交条件をみたしていなければ)、式(8.26)の認識過程が振動し続け終了しない可能性がある。また、7.7で説明されている認識過程の振動は、実際生じたが、これは次のように説明される。

採用した類似度関数 SM がいずれも直交条件を満たすことが保証されていないため、7.6.3のカテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー $E(\varphi^t, \lambda^t)$ が認識段階番号 t の非増加関数であるとは限らないからである(文献[4]の3定理8.1, 8.3, 8.5を参照)。

本シミュレーションにおいては、式(10.1)は必ずしも成立しておらず、稀に振動を発生し認識が終了しない場合が観測された。以下、多段階認識の異常な結果である認識不能、認識不能と、多段階認識過程の非収束(振動)とが生じたシミュレーション結果が報告される。

11. 認識過程における認識不定、認識不能、振動発生のシミュレーション例

本シミュレーションでは、各認識段階のカテゴリ番号リスト ($t=0, 1, 2, \dots$) を

$$\mu^t = J \equiv \{1, 2, \dots, m\}, m=6 \quad (11.1)$$

としている。

このように、不動点方程式(8.27)が成立する認識段階番号 t が見つからない場合が多段階認識過程の振動現象である。認識段階番号 t が見つかったとき、入力画像の、処理の対象としている画素 $x = \langle x_1, x_2 \rangle$ は、

$$\text{カテゴリ集合 } \mathcal{C}_j, j \in \lambda^t = \text{CSF}(\varphi^{t-1}, \mu \cap \lambda^{t-1}) \text{ の内の一つに帰属する} \quad (11.2)$$

という認識結果が得られるという。

$$(i) \text{ 認識段階番号 } t \text{ が見つからない場合 (振動の場合)} \quad (11.3)$$

見つかったときの λ^t が

$$(ii) \lambda^t = \phi \text{ (空集合)} \dots \text{認識不能} \quad (11.4)$$

$$(iii) \lambda^t = \dots \text{認識不定} \quad (11.5)$$

のいずれかの場合を以下で確認する。

11.1 シミュレーションの進め方

本研究で構築された画像理解システムIUSには様々なモデル構成作用素Tや、様々な類似度関数SM、様々な学習方法、様々な探索法などが用意されており、システムに与える各種パラメータや、訓練パターンや代表パターンの選び方、その大きさなど、多数の選択肢が存在する。これらはマウスをクリックすることによってその1つが選択され、実行されるようになっている。例えば、dispSelectorを1の値に設定すると、学習過程、認識過程を逐一、Dos画面に表示しながら、画像理解過程を遂行できる。また、dispSelectorを2の値に設定すると、最終理解結果のみを表示することになる。

採用された画像は、CASIOのデジタルカメラQV-10A、とSONYのモバイルコンピュータVAIO PCGC1-Rに内蔵されているCCDカメラで撮影した風景画像を、横×縦のサイズを101×81に縮小し、輝度値を取出して256段階単色画像に変換したものを使っている。濃度値の数値表現は-128（白）から+127（黒）となっている。

シミュレーションは、膨大な選択肢の組み合わせの中から、1つ1つの選択がどのような影響を及ぼすかを、予測と実験を繰り返しその傾向を読み取りながら進められた。どの選択も一長一短で、常に最良の結果を出すということはほとんどなく、対象とする画像の性質によっても理解の性能が変動する。ある程度、認識率の向上を犠牲にしながらも、学習したパターンに近いものはなるべく正確に認識することができて、なおかつできるだけ誤認識を少なくできると思われるものを採用し、シミュレーションが繰り返された。

表10.1にカテゴリの数、カテゴリの集合、視野の大きさ、画像の大きさが示されている。

表10.1 カテゴリの数、カテゴリの集合、視野の大きさ、画像の大きさ
Tab.10.1 The number of categories, the set of categories, the size of sight and the size of the input images

カテゴリの数	$m=6$
カテゴリの集合	$\mathcal{C}_1 = \text{空}, \mathcal{C}_2 = \text{木}, \mathcal{C}_3 = \text{車}, \mathcal{C}_4 = \text{家},$ $\mathcal{C}_5 = \text{道路}, \mathcal{C}_6 = \text{電柱}$
視野の大きさ	$p=5, q=4$ (11×9=99画素)
画像の大きさ	101×81 (=8181画素)

11.2 使用された6台のパソコン

画像理解システムIUSを稼働させると [7] , 画像解釈の詳細な結果が出力されるが、この結果のヘッダー部分にある「結果##-##-###」は、例えば「結果1-2-3-004」なら、1月2日に鈴木昇一研究室の番号3の計算機で、画像004.bmpを処理した、という意味である。計算機の番号については、表10.2のようになっている。

表10.2 計算機番号と対応する計算機

Tab.10.2 The computer number and the corresponding computer

1 : Gateway Performance 800 (Pentium) 800MHz)
2 : 東芝 BREZZA PV1010JA (Pentium 200MHz)
3 : akia Torando 8280/128 (Pentium) 800MHz)
4 : 富士通 FMV - DESKPEWOR model 936X (Celeron 366MHz)
5 : SONY VAIO PCG-C1 (MMX Pentium 266MHz)
6 : Gateway GATWAY SELECT (AMD Athlon 1.2GHz)

11.3 画像理解システムIUSを構成するJAVAプログラム

本研究でJava言語によって作成された画像理解システム「Image-Understanding System」の各JAVAプログラムを下に掲載する。

- (1) ImageUnderstanding.java … 画像理解システムの主要な処理を行うクラス。 帰属するJFrame1クラスのインスタンス(cif)から与えられた命令を実行する。
- (2) JFrame1.java … プログラム起動時に表示されるフレーム。 将来的に複数の画像理解システムを同時に作動させることを想定したプログラム設計上の構造では、 従属するImageUnderstandingクラスのインスタンス(IU)に人間が命令を与えるための窓口としての役割を持つ。 プロジェクト「ImageUnderstandingSystem(.vep)」におけるメインプログラムでもある。
- (3) JFrame2.java … 画像表示用フレーム。 保持している画像オブジェクト (currentDisp) を表示し、 任意の画素の認識や、 訓練パターンの切り出し、 マウスポインタの指す画素の座標や帰属するカテゴリを表示する機能を持つ。 保持している画像オブジェクトに拡大や保存の命令を与えることもできる。
- (4) JFrame3.java … テキスト表示・編集用フレーム。 実行結果の表示や、 文字の書き込み、 テキストの保存や読み込みを行う。
- (5) JFrame4.java … 画像理解システムの設定を変更する機能を持つフレーム。 正確には、 従属するImageUnderstandingクラスのインスタンス(IU)の機能設定を変更するためのクラス。
- (6) BmpDisplay.java … 画像クラス。 画像本体の濃度値データやサイズ、 解釈結果のラベルや拡大比率、 などを保持している。 画像理解システムが処理対象として実際操作する濃度値の2次元配列型データは、 このクラスのインスタンスが持つ画像本体の1次元配列の濃度値データを数値計算用加工し、 2次元配列にコピーしたものを処理している。 また、 拡大された画像については拡大画像自体の1次元データから拡大比率を元に対応する座標を割り出し、 本来の大きさの2次元配列を復元し、 処理対象としている。 このため、 拡大画像そのものの2次元配列は存在しない。
- (7) BitMapHandle.java … Microsoft Bit Map ファイル形式の24ビット無圧縮カラー画像データを読み込み、 バイト配列から画像情報を生成するためのクラス。
- (8) JaboutDialog.java … システムの所属・名称を表示するためのクラス。
- (9) T.java … モデル構成作用素Tの追加・拡張を行うクラス。 これを使用する場合は、 「ImageUnderstanding.java」 のメソッドT(double ϕ)の2行上の//T Tn=new T(this);文の最初の//をはずさなければならない。

11. 4 シミュレーションでの各種パラメータの設定方法

処理画像とその理解結果の下に書かれている各種の「Selector」は、学習・認識・解釈を行うときに与えた諸条件を示すものである。

各種「Selector」の意味は、次のようになっている。

- (1) Tselector：1(3値のT)，2(5値のT)，3(Sを採用するT)，4(移動)，回転，伸縮を修正するT)，5(カテゴリごとに修正するT)。
- (2) SMSelector：1(MIを使ったSM)，2(訓練パターンを考慮したSM①)，3(訓練パターンを考慮しないSM①)，4(訓練パターンを考慮したSM②)，5(訓練パターンを考慮したSM③)，6(訓練パターンを考慮したSM④)。
- (3) QUANTIZATION：true(SMを量子化する)，false(SMを量子化しない)。
- (4) omegaSelector：1(学習による $\omega[j]$ の決定)，2(視察による $\omega[j]$ の決定)。
- (5) bSelector：1(代表パターンのみを考慮した $b[j]$ の決定法)，2(訓練パターンを考慮した $b[j]$ の決定法)，3(全ての訓練パターンについてBSCを1にする $b[j]$ の決定法)，4(学習による $b[j]$ の決定法)。
- (6) linerSearchSelector：1(線形探索)，2(修正線形探索)。
- (7) learningSelector：1(逐次学習)，2(非逐次学習)。
- (8) dispSelector = 1(学習または認識処理の途中の計算結果の，DOS画面への逐一表示)，2(非表示)。









12. 認識過程における認識不定，認識不能，振動発生(異常な認識結果)のシミュレーション例

本章では，異常な認識が生じたシミュレーション結果のみを報告し，簡単な考察と検討を加える。

12.1 画像理解結果の表示方法

1枚の画像内に存在するカテゴリは \mathcal{C}_1 =空， \mathcal{C}_2 =木， \mathcal{C}_3 =車， \mathcal{C}_4 =家， \mathcal{C}_5 =道路， \mathcal{C}_6 =電柱の6種類である。それぞれのカテゴリ \mathcal{C}_j の番号 j と，結果表示のための文字および色との対応関係は表12.1の通りである。

表12. 1 カテゴリ番号，カテゴリと採用したカテゴリカラー
Tab.12.1 The category numbers, categories and the corresponding adopted category-colors

カテゴリ番号	カテゴリ	カテゴリカラー
1	空	
2	木	
3	車	
4	家	
5	道路	
6	電柱	
内部表現 -1	認識不定	
内部表現 0	認識不能	

12.2 画像理解システムの稼動条件

画像理解システムを稼動するにあたって設けた条件は次の通りである。

濃度値の数値表現：-128（白）から +127（黒）

訓練パターン読込フォルダー：5x4風景訓練パターン

視野の大きさを決める2パラメータ：p=5, q=4

カテゴリ総数：m=6

カテゴリごとの訓練パターンの数：r=12

代表的な訓練パターンの番号：o=5

カテゴリ：空 木 車 家 道路 電柱

カテゴリ番号：1 2 3 4 5 6

TSelector=2(5値のT)

SMSSelector=1 (MIを使ったSM) (QUANTIZATION=false;量子化しない)

omegaSelector=1(学習による各 $\omega[j]$ の決定)

bSelector=1(代表パターンのみを考慮したb[j]の決定法)

linerSearchSelector=1(線形探索)

learningSelector=1(逐次学習)

dispSelector=1

12.3 各学習の初期設定

学習の初期設定は次の通りである。各種諸記号の意味は、これまでの3文献 [5], [7], [7] にある。

$$\delta 2 = 1.0E - 10$$

$$\delta 3 = 1.0E - 10$$

$$\delta 4 = 1.0E - 10$$

$$\epsilon 2 = 0.1 * (1.0 / (1.0 + q)) \quad q = (\text{int}) (t/r)$$

$$\epsilon 3 = 1.0 / (1.0 + q) \quad q = (\text{int}) (t/r)$$

$$\epsilon 4 = 1.0 / (1.0 + q) \quad q = (\text{int}) (t/r)$$

$$\wedge y(j) = 0.625$$

$$W[j;0]_{kl} = 0.1$$

$$b[j;0] = 0.0$$

$$\sigma [1;0]_{kl} = 1.0$$

$$\sigma [2;0]_{kl} = 1.0$$

$$\sigma [3;0]_{kl} = 1.0$$

$$\sigma [4;0]_{kl} = 1.0$$

$$\sigma [5;0]_{kl} = 1.0$$

$$\sigma [6;0]_{kl} = 1.0$$

学習開始時刻：2001年3月14日21時22分31秒

学習終了時刻：2001年3月15日0時46分15秒

ω , W , b , σ についての学習結果の記載は省略される。

訓練パターン $\varphi[j, t]$ の集合が図12.1に示されている。



図12.1 訓練パターン $\varphi[j, t]$ の集合 $\varphi[j, t], j=1 \sim 6, t=1 \sim 12$
Fig.12.1 The set $\varphi[j, t], j=1 \sim 6, t=1 \sim 12$ of training patterns

また、使われた代表パターン $\omega[j]$ の集合が図12.2に示されている。

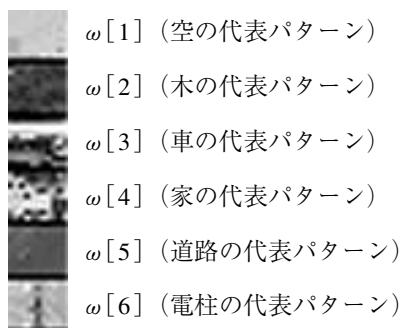


図12.2 代表パターン $\omega[j], j=1, 2, 3, 4, 5, 6$
Fig.12.2 The set $\omega[j], j=1, 2, 3, 4, 5, 6$ of prototypical patterns

12.4 処理結果03-14-6-002

「結果+03-14-6-002」なら、3月14日に鈴木昇一研究室の番号6の計算機で、画像002.bmpを処理した、という意味である。

12.4.1 認識不能、認識不定が得られた1例

図12.3に、理解の対象とした問題の風景画像002.bmpが示されている。



図12.3 理解の対象とした問題の風景画像002.bmp
Fig.12.3 The scenery image 002.bmp to be understood in question

節12.2の稼動条件での節12.3の初期条件下での、学習結果を用い、IUSを稼動させ、得られた理解結果が表12.1のカテゴリと採用したカテゴリカラーを用い、図12.4に示されている。

黒く塗りつぶされた画素があることは、それらの画素が認識不能であることを示している。全画素の内、約半数の画素が認識不能であることを示している。

白く塗りつぶされた画素があることは、それらの画素が認識不定であることを示している。

このように、**認識不定**、**認識不能を得るという異常現象**が生じたが、これは次のように説明される：

採用した複数の類似度関数SMがいずれも、付録Cのミックステュア条件を満たすことが保証されていないためである（文献[4]の定理6.6を参照）。



図12.4 風景画像002.bmpを理解した結果
Fig.12.4 The result of understanding the scenery image 002.bmp

12.5 座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素についての認識

図12.4の理解結果には、座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素についての多段階認識が収束しなくて、振動することが示されている。つまり、節8.5で説明されている認識過程の振動は、実際生じたが、これは次のように説明される。

採用した類似度関数SMがいずれも付録Dの直交条件を満たすことが保証されていないため、カテゴリ帰属知識 $\langle \varphi^t, \lambda^t \rangle$ のポテンシャルエネルギー $E(\varphi^t, \lambda^t)$ が認識段階番号 t の非増加関数であるとは限らないからである（文献[4]の4定理8.1～8.4を参照）。

振動という異常性を少し、詳細に検討しよう。

表12.2に、処理画像002.bmpでの、座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ の画素についての多段階認識過程の諸量が示されている。

表12.2 処理画像002.bmpでの、座標値 $x=\langle x1, x2 \rangle = \langle -25, 4 \rangle$ の画素についての多段階認識過程の諸量
 Tab12.2 The six quantities attendant on the multi-stage process of recognizing the pixel at the coordinate $x = \langle x1, x2 \rangle = \langle -25, 4 \rangle$ in the scenery image 002.bmp to be understood

認識段階番号 t	$RGetpy(\varphi^t)$	$RAIN(\varphi^{t+1}, \varphi^t)$	$E(\varphi^t, \lambda^t)$	$SMdis(\langle \varphi^t, \lambda^t \rangle, \langle \varphi^{t+1}, \lambda^{t+1} \rangle)$	λ^t	$\max_i SM(\varphi^t, \omega[i]), = SM(\varphi^t, \omega[j_t])$
0	0.31040	0.06183	5.00000		$\begin{bmatrix} 1 & 2 \\ 3 \\ 4 & 5 \\ 6 \end{bmatrix}$	$SM(\varphi^0, \omega[1]) = 0.39156(j_0 = 1)$
1	0.31722	6.00222	5.00000	0.01694	$\begin{bmatrix} 1 & 2 \\ 3 \\ 4 & 5 \\ 6 \end{bmatrix}$	$SM(\varphi^1, \omega[6]) = 0.332577(j_1 = 6)$
2	0.32257	0.03250	4.00114	0.00285	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^2, \omega[6]) = 0.34033(j_2 = 6)$
3	0.34331	0.03153	4.00164	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^3, \omega[6]) = 0.40236(j_3 = 6)$
4	0.32257	0.03250	4.00114	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^4, \omega[6]) = 0.34033(j_4 = 6)$
5	0.34331	0.03153	4.00164	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^5, \omega[6]) = 0.40236(j_5 = 6)$
6	0.32257	0.03250	4.00114	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^6, \omega[6]) = 0.34033(j_6 = 6)$
7	0.34331	0.03153	4.00164	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^7, \omega[6]) = 0.40236(j_7 = 6)$
8	0.32257	0.03250	4.00114	0.00800	$\begin{bmatrix} 1 & 2 \\ 3 \\ 5 & 6 \end{bmatrix}$	$SM(\varphi^8, \omega[6]) = 0.34033(j_8 = 6)$

表12.2を用いて、認識情報量 $RGetpy(\varphi^t)$ の移り変わりを棒状グラフにしたのが、図12.5である。

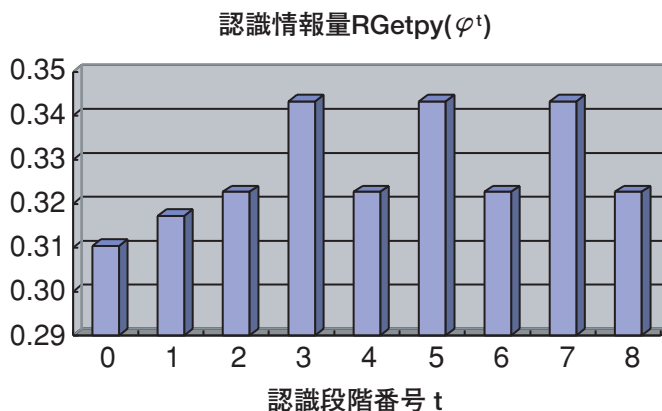


図12.5 認識情報量 $RGetpy(\varphi^t)$ の移り変わり
Fig.12.5 The transition of recognition informations $RGetpy(\varphi^t)$

表12.2を用いて、ポテンシャル E の移り変わりを棒状グラフにしたのが、図12.6である。

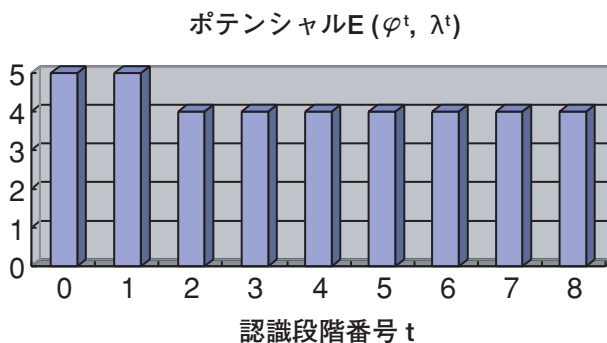


図12.6 ポテンシャル $E(\varphi^t, \lambda^t)$ の移り変わり
Fig.12.6 The transition of potentials $E(\varphi^t, \lambda^t)$

図12.7に、座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素についての認識過程 $\langle \varphi^t, \lambda^t \rangle$, $t=0, 1, 2, 3, 4, 6, 7, 8$ での、第 t 認識段階のパターンモデル φ^t が画像として表現されている。

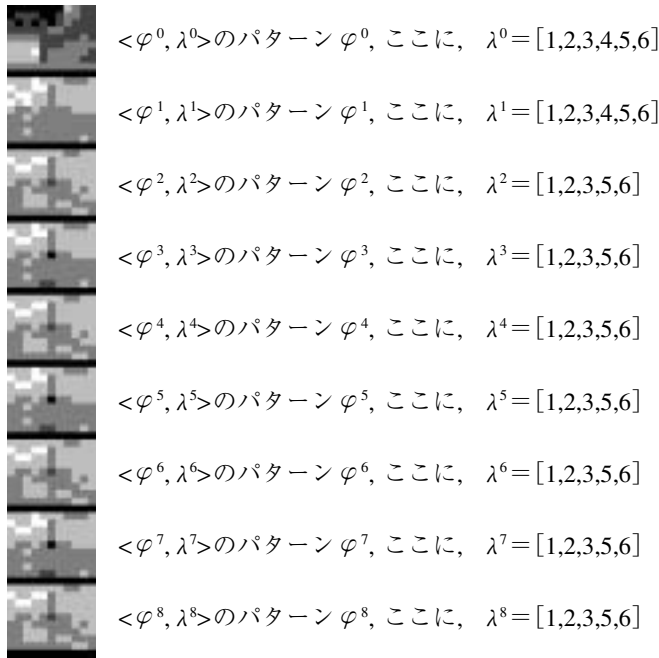


図12.7 座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素についての多段階認識過程 $\langle \varphi^t, \lambda^t \rangle$, $t=0, 1, 2, 3, 4, 6, 7, 8$

Fig.12.7 The multi-stage process $\langle \varphi^t, \lambda^t \rangle$, $t=0, 1, 2, 3, 4, 6, 7, 8$ of recognizing the pixel at the coordinate $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$

表12. 2の多段階認識結果については,

(0) $BSC(\varphi^t, j) = 0$, $j = 1 \sim 6$, $t = 0 \sim 8$ が成立しており, 大分類関数 BSC は有効に機能していないことが判明している.

表12.2について得られる考察と検討は, (1)~(6)の通りである.

(1) 先ず, $t=1$ から, $t=2$ への遷移においては, 第4番目のカテゴリ(家) \mathcal{C}_4 が除かれていることが読み取れる. また, $t=0$ においては, 処理の対象とする問題の, 座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素は, 第1番目のカテゴリ \mathcal{C}_1 (空) に帰属しているが, $t=1$ 以降では, 第6番目のカテゴリ(電柱) \mathcal{C}_6 に帰属している.

(2) $t=3$ から $t=4$ への遷移, $t=5$ から $t=6$ への遷移, $t=7$ から $t=8$ への遷移においては, $RGetpy(\varphi^t)$ は減少している.

(3) $t=3$ から $t=4$ への遷移, $t=5$ から $t=6$ への遷移, $t=7$ から $t=8$ への遷移においては, $RAIN(\varphi^{t+1}, \varphi^t)$ は増加している.

(4) $t=3$ から $t=4$ への遷移, $t=5$ から $t=6$ への遷移, $t=7$ から $t=8$ への遷移, $t=7$ から $t=8$ への遷移においては, $E(\varphi^t, \lambda^t)$ は増加している.

(4) $t=2$ から $t=3$ への遷移においては, $SNdis(\langle \varphi^t, \lambda^t \rangle, \langle \varphi^{t+1}, \lambda^{t+1} \rangle)$ は増加している.

以上は, 多段階認識過程の振動性から生じている.

(5) 「 φ^2, φ^3 」, 「 φ^4, φ^5 」, 「 φ^6, φ^7 」は, 多段階認識過程が振動していることを示していることが

読み取れる(カテゴリ数の10倍認識段階が進んだならば、認識の動作を中止している)。

$$\begin{aligned}(6) \text{ SM}(\varphi^0, \omega[4]) &= 0.00758, \text{ SM}(\varphi^1, \omega[4]) = 0.00000, \text{ SM}(\varphi^2, \omega[4]) = 0.00114 \\ \text{ SM}(\varphi^3, \omega[4]) &= 0.00164, \text{ SM}(\varphi^4, \omega[4]) = 0.00114, \text{ SM}(\varphi^5, \omega[4]) = 0.00164 \\ \text{ SM}(\varphi^6, \omega[4]) &= 0.00114, \text{ SM}(\varphi^7, \omega[4]) = 0.00164, \text{ SM}(\varphi^8, \omega[4]) = 0.00114\end{aligned}$$

からわかるように、明らかに、 φ^1 から φ^2 への過程は付録Dの直交性を保存していない。

□

以上、風景画像0002.bmpでの座標値 $x = \langle x_1, x_2 \rangle = \langle -25, 4 \rangle$ にある画素について、Gateway GATEWAY SELECT (AMD Athlon 1.2GHz)による認識過程における結果 03-14-6-002が説明された。

9. おわりに

S.Suzukiは、万能性認識システムRECOGNITRONに関する数理 [3], [4] を展開済みである。実は、節8.5の不動点探索の、RECOGNITRONによる認識過程はS.Suzukiが考案したある連想形認識方程式を解く過程である。既に、事物の存在する領域を大きさ $(2p+1) \cdot (2q+1)$ の画像成分と仮定し、領域抽出(セグメンテーション)が自動的になされる画素単位の認識方式を内蔵している認識システムRECOGNITRONについてその性能を確かめる計算機シミュレーションをJAVA言語で実行し、その結果が検討されている [6]。モデル構成要素T, 類似度関数SM, 大分類関数BSC, カテゴリ選択関数CSFはそれぞれ、SS公理系のaxiom 1~4を満たすように構成されているので、2文献 [3], [4] で解析されているRECOGNITRONの基本的な認識性能は保証されている。

このように、不動点方程式(8.27)が成立する認識段階番号 t が見つからなくて、振動する場合を確認した。この振動現象は、採用している類似度関数SMが付録Dの直交条件を満たさないことから生じることが理論的に明らかにされている。SMが付録Dの直交条件を満たさないことから、付録Bの多段階認識過程が振動し不動点方程式(8.27)が成立する認識段階番号 t が存在しないという振動現象も計算機シミュレーションで確認された。

また、不動点方程式(8.27)が成立する認識段階番号 t が見つかったとき、入力画像の、処理の対象としている画素 $x = \langle x_1, x_2 \rangle$ は、

カテゴリ集合 $\mathcal{C}_j, j \in \lambda^t = \text{CSF}(\varphi^{t-1}, \mu \cap \lambda^{t-1})$ の内の一つに帰属するという認識結果が得られるが、このとき、カテゴリ番号のリスト λ^t が0, 1, 2個の要素からなっている場合、各々認識不能、認識可能、認識不定になり、2個以上の場合、認識不定になるが、axiom 2を満たす式(4.1)の類似度関数SMが付録Cのミックスチュア条件を満たさないことから、認識不定、認識不能という異常な認識結果が生じることが、計算機シミュレーションで確認された。

本研究はSS理論を風景画像の理解に適用した計算機シミュレーション研究のまだ初期の段階のものであり、認識率の改善よりもむしろ画像解釈における限界点を模索し、問題点を露呈させることに、本研究は主力を注いだ。本研究では認識対象を自動的に縮小・拡大しマッチングをとる正規化処理や、2次元の投影像からもとの3次元形状を脳内で復元し新しい視点からの像を作り上げる感覚運動的な知能処理も取り入れていないし、遠近・高低・大小・明暗・色・カテゴリ同士の関係(雲は空にある等)などに関し人間の持つ常識も考慮させていない。それらの操作を扱えるようにするためには、モデル構成作用素、類似度関数、大分類関数等を改良することと、計算

機の処理速度や記憶容量などの向上とが必要とされる。とりあえず、これらのパターン変形を吸収・除去する形で処理し、RECOGNITRONに該当する代表パターンのモデルを想起させるかが今後の課題であろう。

無造作に撮影された自然画像を正しく理解させるためには、まずある程度規格化された画像内事物を理解させることができるようになってはじめてその可能性が見えてくる。本シミュレーションでは、ある程度規格化された風景画像の解釈がRECOGNITRONに可能かどうかの検討に重点を置いていることに注意しておく。今後の研究では、一層適切な適応的な設定を考えなければならない。

今回のシミュレーション結果を踏まえ、IUSの理解性能に関し改良を重ねなければならない。

文 献

- [1] 鈴木昇一：“認識工学”，柏書房，Feb.1975
- [2] 鈴木昇一：“ニューラルネットの新数理”，近代文芸社，Sept.1996
- [3] 鈴木昇一：“パターン認識の数理的一般解決”，近代文芸社，Jun.1997
- [4] 鈴木昇一：“認識知能情報論の新展開”，近代文芸社，Aug.1998
- [5] 鈴木昇一：“風景画像から知識を抽出し、解釈するシステムの、ファジィ推論ニューラルネットによる構成”，情報研究(文教大学・情報学部)，no.23, pp.183-265, 2000
- [6] 鈴木昇一，川俣博司，大槻善樹：“風景画の理解に関するJAVA言語によるRECOGNITRONの計算機シミュレーション”，情報研究(文教大学・情報学部)，no.27, pp.73-109, Mar.2002
- [7] 鈴木昇一，川俣博司，大槻善樹：“JAVA言語で実装化された画像理解システムIUSの動作概要と、その稼動方法”，情報研究(文教大学・情報学部)，no.28, pp.143-165, Dec.2002
- [8] 鈴木昇一：“連想形記憶器MEMOTRONと日本語母音系列の再生に関する計算機シミュレーション”，情報研究(文教大学・情報学部)，no.7, pp.14-29, Dec.1986
- [9] 鈴木昇一：“プロダクション・システムとしてのファジィ・マルチメディア・コンピュータと、空間多重パターンファジィ推論系”，情報研究(文教大学・情報学部)，no.24, pp.105-183, Dec.2000
- [10] 鈴木昇一：“各個人の感性を反映した認識システムRECOGNITRON”，情報研究(文教大学・情報学部)，no.24, pp.185-257, Dec.2000
- [11] 鈴木昇一：“SS大分類関数BSCの適応的構成への，計算論的学習理論の適用”，文教大学情報学部「情報研究」，no.25, pp.187-238, Mar.2001
- [12] 鈴木昇一：“量子力学の諸原理，多段階量子認識系と，心理状態を取り入れた想起に基づく部分空間認識法”，文教大学情報学部「情報研究」，no.25, pp.239-284, Mar.2001
- [13] 鈴木昇一：“Support Vector Machineを利用した大分類関数の構成”，情報研究(文教大学・情報学部)「情報研究」，no.26, pp.1-62, Dec.2001
- [14] 鈴木昇一：“2カテゴリ分類困難度の情報理論”，情報研究(文教大学・情報学部)，「情報研究」no.26, pp.63-160, Dec.2001
- [15] 鈴木昇一：“一般化類似度関数を用いた”導出原理による第1階述語推論”，文教大学情報学部「情報研究」，no.26, pp.27-71, Mar.2002

- [16] 鈴木昇一：“遺伝的アルゴリズムにおける適合度比例選択戦略を採用した進化方程式の、パターン多段階変換に基づく認識への応用”，文教大学情報学部「情報研究」，no.28, pp.37-67, Dec.2002
- [17] 鈴木昇一：“近傍を利用した音素認識のためのモデル構成作用素T，類似度関数SM，大分類関数BSCの諸構成と，SS不動点探索型多段階認識”，文教大学情報学部「情報研究」，no.28, pp.69-141, Dec.2002

**付録A. axiom1～4(SS公理系)を各々，満たさなければならないパターン集合Φ，
モデル構成作用素 Tの対【Φ，T】，類似度関数SM，大分類関数BSC，
カテゴリ選択関数CSF**

本付録では，処理の対象となる問題のパターンφの集合Φ，モデル構成作用素T，類似度関数SM，カテゴリ選択関数CSFについて説明される．対【Φ，T】のみたされなければならないaxiom 1と，類似度関数SMの満たされなければならないaxiom 2も説明され，Φの表示が明らかにされ，Φが構成的集合であることが明らかにされる．更に，大分類関数BSCの満たされなければならないaxiom 3も説明される．カテゴリ選択関数CSFが満たされなければならないaxiom 4も説明され，CSFの構造がSM，BSCを用いて決定されることが明らかにされる．

A1. axiom 1とパターン集合Φ，モデル構成作用素T

一般に，処理の対象とする問題のパターンφの集合Φは或る可分な [A1] な(separable)一般抽象ヒルベルト空間 \mathfrak{H} の零元0を含む或る部分集合である．例えば， $\overline{\eta}$ を η の複素共役として，

$$M; q\text{-次元ユークリッド空間}\mathbb{R}^q\text{の可測部分集合} \quad (\text{A1.1})$$

$$dm(x): \text{正値ルベーク・スティルチェス式測度} \quad (\text{A1.2})$$

$$x = \langle x_1, x_2, \dots, x_q \rangle \in M (\subseteq \mathbb{R}^q) : \text{実数値 } q \text{ 変数直交座標系} \quad (\text{A1.3})$$

を導入し，その内積 (φ, η) ，ノルム $\|\varphi\|$ を，

$$(\varphi, \eta) = \int_M dm(x) \varphi(x) \cdot \overline{\eta}(x) \quad (\text{A1.4})$$

$$\|\varphi\| = \sqrt{(\varphi, \varphi)} \quad (\text{A1.5})$$

とする線形空間(ベクトル空間)としての可分なヒルベルト空間 $\mathfrak{H} = L_2(M; dm)$ の特別な場合として，

$$M = \mathbb{R}^2 \text{(2次元全平面)} \quad (\text{A1.6})$$

$$dm(x) = [x_1^2 + x_2^2]^{-1} \cdot dx_1 dx_2 \quad (\text{A1.7})$$

を選ぶことができる [B7]，[B9]．

このような，並びに，写像

$$T : \Phi \rightarrow \Phi \quad (\text{A1.8})$$

は次のaxiom 1を満たさなければならない．このとき，写像Tはモデル構成作用素(model-construction operator)と呼ばれ， $T\varphi \in \Phi$ は $\varphi \in \Phi$ の代りとなり得るという意味で，パターン $\varphi \in \Phi$ のモデル(model)，或いはパターンモデルと呼ばれる．

下記のaxiom 1からわかるように，パターンモデル $T\varphi$ の集合 $T \cdot \Phi$ は，原パターン φ の集合 Φ への埋込性

$$T \cdot \Phi \equiv \{T\varphi \mid \varphi \in \Phi\} \quad (\text{A1.9})$$

を満たし、 Φ は原点(=0)を始点とし、 Φ の任意の点を通る半直線を含むような集合、つまり、錐(cone)であらねばならない。下記の式(A1.14)による Φ の表示が正に Φ が錐であることを明らかにしている。

Axiom 1を満たすパターン集合 Φ は実は、構成的集合(constructible set)である。S.Suzukiは形式と意味が互いに非分離であるようなパターンというものが満たされなければならない帰納的定義(再帰的定義)から Φ の集合論的再帰領域方程式(axiom 1を満たす最小の Φ の表現式; set-theoretic reflective domain equation)を提案し、この方程式を解き、 Φ の、構造と構成方法を明らかにしている(文献 [B3] の2.4節)。その結果は次のとおりである:

パターンと判明している元の集合(基本領域;basic domain) (axiom 1の(i)の前半から、 $0 \in \Phi_B$ を導入して、集合論的再帰領域方程式

$$\Phi = \Phi_B \cup T \cdot \Phi \cup R^{++} \cdot \Phi \quad (\text{A1.10})$$

ここに、

$$T \cdot \Phi \equiv \{T\varphi \mid \varphi \in \Phi\} \quad (\text{A1.11})$$

$$R^{++} \text{は正実数全体の集合} \quad (\text{A1.12})$$

$$R^{++} \cdot \Phi \equiv \{r^{++} \cdot \varphi \mid r^{++} \in R^{++}\} \quad (\text{A1.13})$$

の解 Φ は

$$\Phi = R^{++} \cdot [\Phi_B \cup T \cdot \Phi_B] \quad (\text{A1.14})$$

と表示される(文献 [B3] の式(2.56)を参照)。 Φ の表示式(A1.14)から、明らかに、2つの等式

$$\begin{aligned} \text{(a) } T \cdot \Phi &= T \cdot \Phi_B \subseteq \Phi \\ &\because \text{ axiom 1の(ii), (iii)の2後半} \end{aligned} \quad (\text{A1.15})$$

$$\begin{aligned} \text{(b) } R^{++} \cdot \Phi &= \Phi (= R^{++} \cdot \Phi_B \cup R^{++} T \cdot \Phi_B) \\ &\because \text{ axiom 1の(ii)の後半} \end{aligned} \quad (\text{A1.16})$$

が成り立つ。

Axiom 1 (パターン集合 Φ とモデル構成作用素 T との対【 Φ, T 】の満たすべき公理)

(i) (零元0の Φ -包含性と、零元0の T -不動点性;fixed-point property of zero element under mapping T)
 $0 \in \Phi \wedge T0 = 0.$

(ii) (Φ の錐性、 T の正定数倍吸収性;cone property)

$$\forall \varphi \in \Phi, a \cdot \varphi \in \Phi \wedge T(a \cdot \varphi) = T\varphi$$

for any positive real number a .

(iii) (Φ の埋込性(embeddedness)と、 T のベキ等性(idempotency))

$$\forall \varphi \in \Phi, T\varphi \in \Phi \wedge T(T\varphi) = T\varphi.$$

(iv) (写像 T の非零写像性;non-zero mapping property of T)

$$\exists \varphi \in \Phi, T\varphi \neq 0. \quad \square$$

A2. 処理の対象となる問題のパターン φ の集合 Φ とモデル構成作用素 T との対【 Φ, T 】の基本構成と、パターンモデル $T\varphi$ とパターン φ との間の同一知覚原理

原パターン φ が如何なる意味を備えているか、つまり、 φ が如何なる類概念(category)を表しているかを決定する働きをもつのが、認識システム RECOGNITRON である。RECOGNITRON がモデル $T\varphi$ を見たり聞いたりしたならば($T\varphi$ を感性的に受け取るならば)、原パターン $\varphi \in \Phi$ と同じに見

えたり聞こえたりすること(原パターン φ と錯覚し、原パターン φ と同じように感性的に受容すること)だと、解釈可能な対 $[\varphi, T]$ について説明しよう。

パターンモデル $T\varphi$ を出力する式(A1.8)の写像 T に要求されるのは、次の4性質①～④である[B3], [B4], [B6] :

①(零元不動点性;axiom 1の(i))

$$\varphi=0 \in \Phi \text{ については, } T\varphi=0.$$

②(正定数倍不変性;axiom 1の(ii)の後半)

任意の正実定数 a に対し,

$$\forall \varphi \in \Phi, T(a \cdot \varphi) = T\varphi.$$

③(ベキ等性;axiom 1の(iii)の後半)

$$\forall \varphi \in \Phi, T(T\varphi) = T\varphi.$$

④(非零写像性;axiom 1の(iv))

$$\exists \varphi \in \Phi, T\varphi \neq 0. \quad \square$$

零元 $\varphi=0 \in \Phi$ は背景が無のパターンである。上述の①～④はA1章のaxiom 1の(i), (ii), (iii)の3後半, (iv)である。

Φ は処理の対象とする問題のパターン φ の集合 Φ であり、 $T\varphi \in \Phi$ は $\varphi \in \Phi$ に対応するパターンモデルであって、原パターン $\varphi \in \Phi$ と同じ空間 Φ に埋め込まれている。モデル $T\varphi$ は、 $T\varphi \in \Phi$ を見たり聞いたりしたならばあたかも原パターン $\varphi \in \Phi$ かのように見えたり聞こえたりするようなものである($T\varphi$ と φ との間の**同一知覚原理**)。この同一知覚原理を達成するために、S理論 [B1]～[B6] では、式(A1.8)の写像であるモデル構成作用素 T が導入され、対 $[\Phi, T]$ はA1.のaxiom 1を満たしていなければならないことになる。このとき、写像 T はモデル構成作用素と呼ばれ、 $T\varphi \in \Phi$ は $\varphi \in \Phi$ の代りとなり得るという意味で、パターン $\varphi \in \Phi$ のモデルと呼ばれる。

処理の対象とするパターン φ の集合 Φ は或る可分なヒルベルト空間 \mathfrak{H} の、零元 0 を含む或る部分集合であり、この Φ 、並びに、式(A1.8)の写像 T の対 $[\Phi, T]$ は上記の4性質①～④((ii), (iii)の2後半、並びに(i), (iv))を含む形で、A1.のaxiom 1をみたさなければならない。

次の定理A2.1は、axiom 1を満たす対 $[\Phi, T]$ を決定している。

[定理A2.1] (パターン集合 Φ とモデル構成作用素 T との対 $[\Phi, T]$ の構成定理)

パターンと判明している φ の集合(基本領域) $\Phi_B(\ni 0)$ と、すべての正実定数の集合 \mathbb{R}^{++} とを用意する。

式(A1.8)の写像 T がaxiom 1の(i), (ii), (iii)の3後半、並びに、(iv)を満たすとしよう。このとき、次の(イ)、(ロ)が成り立つ:

(イ)処理の対象とする問題のパターンの集合 Φ を、式(A1.14)の如く設定すれば、2式(A1.15), (A1.16)が成立し、axiom の(i), (ii), (iii)の3前半を Φ は満たし、結局、対 $[\Phi, T]$ はaxiom 1を満たす。

(ロ)逆に、 $(0 \in) \Phi_B$ を部分集合に持つ Φ がaxiom 1の(i), (ii), (iii)の3前半を満たすとすれば、

$$\Phi \supseteq \Phi_B \cup \mathbb{R}^{++} \cdot \Phi \cup T \cdot \Phi \quad (\text{A2.1})$$

が成立するが、ここで、特に、包含式(A2.1)において等号が成立するような最小の Φ を採用すれば、つまり、領域方程式(A1.10)の成立を仮定すれば、axiom 1を満たす対 $[\Phi, T]$ の Φ は式(A1.14)のように表され、2式(A1.15), (A1.16)も成立する。

(証明) (イ)は文献 [B4], 付録1の定理A1.1である。(ロ)は文献 [B3], pp.64-66(2.4節)で証明されている。 \square

A3. axiom 2と類似度関数SM

任意のパターン $\varphi \in \Phi$ が、記憶されている代表的なパターンからなる有限個の元からなる集合内の任意の代表パターンとどの程度似ているか、違っているかを計量する手段を設定することが、認識の働きを確保するために必要とされる。計量のための手段が類似度関数SMである。

“正常なパターン” (well-formed pattern) は、ある1つのカテゴリ \mathcal{C}_j (第 $j \in J$ 番目の類概念) のみに帰属しているものとし、このような \mathcal{C}_j の集まり (有限集合)

$$\underline{\mathcal{C}} = \{\mathcal{C}_j \mid j \in J\} \quad (\text{A3.1})$$

を想定する。 \mathcal{C}_j の備えている性質を典型的に持っている (第 $j \in J$ 番目の) 代表パターン (prototypical pattern) $\omega_j (\neq 0)$ を1つ選定する。 \mathcal{C}_j は、典型 (prototype) としての代表パターン ω_j を中心とした緩やかな (第 $j \in J$ 番目の) カテゴリであることを仮定したことに注意しておく。ここに、

$$\Omega \equiv \{\omega_j \mid j \in J\} \subset \Phi \quad (\text{A3.2})$$

が式 (A3.1) の全カテゴリ集合 $\underline{\mathcal{C}}$ に1対1に対応する代表パターンの集合である。式 (A3.2) の系 Ω は、

$$\text{複素定数 } a_j \text{ の組 } \{a_j \mid j \in J\} \text{ について } \sum_{j \in J} a_j \cdot \omega_j = 0 \Rightarrow \forall j \in J, a_j = 0 \quad (\text{A3.3})$$

が成立しているという意味で、1次独立 (linearly independent) でなければならない。 Ω を視察で決定できる場合があるが、訓練パターン系列から Ω を適応的に決定する方法については、文献 [B3] の付録 I で説明されている。

Axiom 1 を満たす式 (A1.1) のモデル構成作用素 T によって、式 (A3.2) の代表パターン集合 Ω が変換されて得られる系

$$T \cdot \Omega \equiv \{T\omega \mid \omega \in \Omega\} = \{T\omega_j \mid j \in J\} \quad (\text{A3.4})$$

も1次独立であると要請する。このとき、類似度関数 (similarity-measure function)

$$SM : \Phi \times \Omega \rightarrow \{s \mid 0 \leq s \leq 1\} \quad (\text{A3.5})$$

を導入し

$$\begin{aligned} SM(\varphi, \omega_j) = 1, 0 \text{ に従って、パターン } \varphi \in \Phi \text{ は各々、} \omega_j \text{ と確定的な} \\ \text{類似度関係、相違関係にあり、また、} 0 < SM(\varphi, \omega_j) < 1 \text{ の場合は、} \\ \text{あいまいな類似・相違関係にある} \end{aligned} \quad (\text{A3.6})$$

と、SMを解釈しよう。

関数SMは次のaxiom 2を満たすように構成されねばならない。Axiom 2の(i)では、クロネッカー (kronecker) の δ 記号

$$\delta_{ij} = 1 \text{ if } i=j, = 0 \text{ if } i \neq j \quad (\text{A3.7})$$

が導入されているが、特にaxiom 2の(i)なるこの正規直交性は、候補カテゴリの分離・抽出が効果的に行われ、

$$\text{候補カテゴリの鋭利な削減 (a sharp reduction)} \quad (\text{A3.8})$$

をもたらすために要請されている。

Axiom 2 (類似度関数SMの満たすべき公理)

(i) (正規直交性; orthonormality)

$$\forall i, \forall j, SM(\omega_i, \omega_j) = \delta_{ij}.$$

(ii) (規格化条件, 確率性, 正規性; probability condition, normalization)

$$\forall \varphi \in \Phi, \sum_{j \in J} SM(\varphi, \omega_j) = 1.$$

(iii) (写像Tの下での不変性; invariance under mapping T)

$$\forall \varphi \in \Phi, \forall j \in J, SM(T\varphi, \omega_j) = SM(\varphi, \omega_j). \quad \square$$

上述の axiom 2 の (i) ~ (iii) について簡単に説明しておこう。

SM の解釈式 (A3.6) の下で，(i) は，相異なるカテゴリの代表パターン同士は確定的な相違関係にあり，同一カテゴリの代表パターン同士は確定的な類似度関係にあることを要請している。(ii) は，任意のパターン φ について，すべてのカテゴリについての類似度の総和は 1 であることを要請している。つまり，パターン φ は少なくとも 1 つのカテゴリ \mathcal{C}_j に帰属していることを要請している。(iii) は，パターンモデル $T\varphi$ は原パターン φ と任意のカテゴリについて同一類似度を持つことを要請している。ということは，パターンモデル $T\varphi$ を見たり，聞いたりするならば，原パターン φ と同じように見えたり，聞こえたりすること (同一知覚原理；A2. を参照) を要請していることになる。

尚，第 $j \in J$ 番目のカテゴリ \mathcal{C}_j の生起確率である非負実数 $p(\mathcal{C}_j)$ を，2 条件

$$[\forall j \in J, 0 < p(\mathcal{C}_j) < 1] \wedge [\sum_{j \in J} p(\mathcal{C}_j) = 1] \quad (\text{A3.9})$$

を満たすものとして導入しておく。

A4. axiom 3 と大分類関数

本章では，ある 1 つのカテゴリに帰属するかどうかを決定する 2 カテゴリ分類器としての大分類関数 BSC は，axiom 3 を満たすように構成されなければならないことが説明される。

式 (A3.5) の類似度関数 SM が式 (A3.8) でいう “候補カテゴリの鋭利な削減” を持つためには，axiom 2，(i) の正規直交性を満たす必要があることが A3. で指摘されたが， $SM(\varphi, \omega_j)$ の代りに $SM(\varphi, \omega_j) \cdot BSC(\varphi, j)$ を用いれば，パターン φ が帰属するかも知れない候補カテゴリを益々，鋭利に削減できると期待される。

大分類関数 (rough classifier, binary-state classifier) と呼ばれる **2 値関数**

$$BSC : \Phi \times J \rightarrow \{0, 1\} \quad (\text{A4.1})$$

を，次の axiom 3 を満たすものとして導入し，解釈

パターン $\varphi \in \Phi$ の帰属する候補カテゴリの 1 つが第 $j \in J$ 番目の \mathcal{C}_j であるならば，

$$BSC(\varphi, j) = 1 \text{ であることが望ましい} \quad (\text{A4.2})$$

を採用しよう。この際，注意すべきは，

$BSC(\varphi, j) = 0$ であっても，パターン $\varphi \in \Phi$ の帰属する候補カテゴリの 1 つは，

$$\text{第 } j \in J \text{ 番目の } \mathcal{C}_j \text{ でないとは限らない} \quad (\text{A4.3})$$

としていることである。また，axiom 3 の (i) からわかるように，**カテゴリ間の相互排除性** (the mutual exclusion of the one category from the other categories)

$$\forall j \in J, \forall i \in J - \{j\}, BSC(\omega_i, j) = 0 \quad (\text{A4.4})$$

を公理として要請していない事実に注意しておこう。この事実を補うのが実は，式 (A3.5) の類似度関数 SM が満たさなければならないとしている axiom 2 の (i) (正規直交性) である。

Axiom 3 (大分類関数 BSC の満たすべき公理)

(i) (カテゴリ抽出能力；category separability)

$$\forall j \in J, BSC(\omega_j, j) = 1.$$

(ii) (写像 T の下での不変性；invariance under mapping T)

$$\forall \varphi \in \Phi, \forall j \in J, BSC(T\varphi, j) = BSC(\varphi, j). \quad \square$$

A5. axiom 4 と、カテゴリ選択関数 CSF の構造形式

認識システム RECOGNITRON がパターン に対し，

「パターン $\varphi \in \Phi$ が、式 (A3.1) の全カテゴリ集合 \mathcal{C}_j の部分集合

$$\underline{\mathcal{C}}(\varphi) \equiv \{\mathcal{C}_j \mid j \in \gamma\} \quad (\text{A5.1})$$

内の何れか1つのカテゴリ \mathcal{C}_j に帰属する可能性がある」 (A5.2)

という “パターン $\varphi \in \Phi$ のカテゴリ帰属知識 (categorical membership-knowledge)” を持っているとする。この知識を、

$$\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle \quad (\text{A5.3})$$

と表す。

$$\langle \Phi, 2^J \rangle \equiv \{\langle \varphi, \gamma \rangle \mid \varphi \in \Phi, \gamma \in 2^J\} \quad (\text{A5.4})$$

は、カテゴリ帰属知識空間 (categorical membership-knowledge space) と呼ばれ、すべてのパターン $\varphi \in \Phi$ と、すべてのカテゴリ番号のリスト $\gamma \in 2^J$ とのなす順序の付いた対リスト (an ordered pair list) $\langle \varphi, \gamma \rangle$ の集合である。ここに、集合 J のすべての部分集合のなす集合、つまり “ J のべき集合 (power set)” を 2^J で表している。

カテゴリ選択関数 (category-selection function) と呼ばれる写像

$$\text{CSF} : \Phi \times 2^J \rightarrow 2^J \quad (\text{A5.5})$$

は、包含関係 (inclusion relation)

$$\forall \varphi \in \Phi, \forall \gamma \in 2^J, \text{CSF}(\varphi, \gamma) \subseteq \gamma \subseteq J \in 2^J \quad (\text{A5.6})$$

を満たし、然も、次の axiom 4 を満たすものとし、設定されたとしよう。

Axiom 4 (カテゴリ選択関数 CSF の満たすべき公理)

(i) $\varphi = 0 \vee \gamma = \emptyset$ の場合

如何なるカテゴリ番号 $k \in \gamma$ も $\langle \varphi, \gamma \rangle$ の有効な候補カテゴリの番号ではない。

(ii) $\varphi \neq 0 \wedge \gamma \neq \emptyset$ であり、かつ $\text{SM}(\varphi, \omega_k) = 0 \wedge \text{BSC}(\varphi, k) = 0$ の場合

カテゴリ番号 $k \in \gamma$ は、 $\langle \varphi, \gamma \rangle$ の有効な候補カテゴリの番号ではない

(iii) $\varphi \neq 0 \wedge \gamma \neq \emptyset$ であり、かつ $\sum_{k \in \gamma} \text{BSC}(\varphi, k) = 0$ の場合

$\text{BSC}(\varphi, k) = 0$ であっても、 $\text{SM}(\varphi, \omega_k) > 0$ であるようなカテゴリ番号 $k \in \gamma$ は、 $\langle \varphi, \gamma \rangle$ の有効な候補カテゴリの番号である。

(iv) $\varphi \neq 0 \wedge \gamma \neq \emptyset$ であり、かつ $\sum_{k \in \gamma} \text{BSC}(\varphi, k) > 0$ の場合

(iv-1) $\text{BSC}(\varphi, k) = 0$ なるカテゴリ番号 $k \in \gamma$ は、 $\text{SM}(\varphi, \omega_k) > 0$ であっても、 $\langle \varphi, \gamma \rangle$ の有効な候補カテゴリの番号ではない。

(iv-2) $\text{SM}(\varphi, \omega_k) = 0$ であっても $\text{BSC}(\varphi, k) = 1$ なるカテゴリ番号 $k \in \gamma$ は、

$\langle \varphi, \gamma \rangle$ の有効な候補カテゴリの番号ではない。 □

次の定理 A4.1 では、式 (A5.5) の写像 CSF は、式 (A2.5) の類似度関数 SM、式 (A3.1) の大分類関数 BSC を使用する形式で、

その定義域が $\Phi \times 2^J$ であり、その値域が、パターン $\varphi \in \Phi$ のカテゴリ帰属知識 $\langle \varphi, \gamma \rangle$

$\in \langle \Phi, 2^J \rangle$ の “有効な” 候補カテゴリの番号リスト (a list of significant category-numbers)

の集合である (A5.7)

の如く、構成されている。

次の定理 A4.1 は、axiom 4 を満たすように、式 (A5.5) のカテゴリ選択関数 CSF の構造を決定したものである。

[定理 A4.1] (カテゴリ選択関数 CSF の構成定理)

次のように定義される式 (A5.5) の1つの写像 CSF は式 (A5.6) と上述の axiom 4 を満たす:

$$(i) \varphi = 0 \vee \gamma = \phi \text{ の場合} \\ \text{CSF}(\varphi, \gamma) = \phi \quad (\text{A5.8})$$

$$(ii) \varphi \neq 0 \wedge \gamma \neq \phi \text{ の場合} \\ \text{CSF}(\varphi, \gamma) = \begin{cases} \{k \in \gamma \mid \text{SM}(\varphi, \omega_k) > 0\} & \text{if } \sum_{k \in \gamma} \text{BSC}(\varphi, k) = 0 \\ \{k \in \gamma \mid \text{SM}(\varphi, \omega_k) > 0 \wedge \text{BSC}(\varphi, k) = 1\} & \text{if } \sum_{k \in \gamma} \text{BSC}(\varphi, k) > 0 \end{cases} \quad (\text{A5.9})$$

(証明) 文献 [B3] の定理E1である。 □

定理A4.1の写像CSFについて、次のように解釈できる:

処理の対象とするパターン $\varphi \in \Phi$ がカテゴリ $\mathcal{C}_j, j \in \gamma$ の何れか1つに帰属する可能性があると思定した場合、更に絞り込んで、その内のカテゴリ $\mathcal{C}_j, j \in \text{CSF}(\varphi, \gamma) \subseteq \gamma$ の何れか1つに帰属する可能性があると思定すると帰納推論 (inductive reasoning) できる機能を備え、その出力 $\text{CSF}(\varphi, \gamma)$ はパターン $\varphi \in \Phi$ の有効な候補カテゴリの番号のリストを与えている。

付録B. SS多段階想起認識法の詳細

SS多段階想起認識法は次のように述べられる。

【SS多段階想起認識法】

処理の対象とする問題のパターン φ について、そのパターンモデル $T\varphi$ を先ず、求める。ここに、axiom 1を満たす式(A1.8)のモデル構成作用素 T を導入している。その後、

①初期単体の部分集合

$$\left\{ \sum_{j \in \gamma} s_j(T\varphi) \cdot T\omega_j \mid \gamma \subseteq J \right\} \quad (\text{B.1})$$

ここに、 $s_j(T\varphi)$ は、 $T\varphi$ が第 $j \in \gamma$ 番目のカテゴリ \mathcal{C}_j の代表パターン ω_j のモデル $T\omega_j$ と似ている程度であり、

$$[\forall j \in \gamma, 0 \leq s_j(T) \leq 1] \quad (\text{B.2})$$

$$\wedge [0 \leq \sum_{j \in \gamma} s_j(T\varphi) \leq 1] \quad (\text{B.3})$$

の内部の探索を開始する。

②パターンモデルの各変換段階においてカテゴリ番号のリスト $\mu \subseteq J$ が帰納推理で選択されて確定する構造受精変換

$$TA(\mu)T, \text{ ここに, } \mu \subseteq J \quad (\text{B.4})$$

の操作を候補カテゴリ知識

$$\langle \psi, \lambda \rangle$$

(パターン ψ がカテゴリ集合 $\mathcal{C}_j, j \in \lambda$ 内の1つのカテゴリに帰属する

可能性があるという認識システム RECOGNITRON が持っている知識) (B.5)

に対し実行する。カテゴリ帰属知識空間 $\langle \Phi, 2^J \rangle$ 上の2元関係(等形式関係) $=_{\Delta}$ を

$$\langle \varphi, \gamma \rangle =_{\Delta} \langle \psi, \lambda \rangle \quad (\text{恒等的に等しい})$$

$$\Leftrightarrow \varphi = \psi \wedge \gamma = \lambda \quad (\text{B.6})$$

と導入し、少し詳しく説明すれば、候補カテゴリ知識を順次改善すること(想起)、つまり、

$$(\text{初期段階}) \langle \psi_0, \lambda_0 \rangle =_{\Delta} \langle T\varphi, J \rangle \quad (\text{B.7})$$

$$\begin{aligned} & (\text{帰納推理段階}) \quad \langle \psi_{s+1}, \lambda_{s+1} \rangle =_{\Delta} \text{TA}(\mu_2)T \cdot \langle \psi_s, \lambda_s \rangle, \\ & s=0, 1, 2, \dots, t \end{aligned} \tag{B.8}$$

と、各カテゴリ帰属知識 $\langle \psi_s, \lambda_s \rangle (s=0, 1, 2, \dots, t+1)$ を想起して行くことにより、最適解(SSポテンシャルを最小にする不動点解))として、構造受精変換Tの不動点方程式

$$(\text{終了段階}) \quad \text{TA}(\mu_1)T \cdot \langle \psi_t, \lambda_t \rangle =_{\Delta} \langle \psi_t, \lambda_t \rangle \tag{B.9}$$

を満たすという意味で不動点となるカテゴリ帰属知識

$$\langle \psi_t, \lambda_t \rangle \tag{B.10}$$

を求め、原パターン φ を認識しようとする。

③このようにカテゴリ帰属知識を多段階にわたって変換してゆく多段階認識が正常に終了した場合、

$$\text{SM}(\psi_t, \omega_j) = s_j(T\psi_t), j \in \lambda_t \tag{B.11}$$

であるような“axiom 2を満たす類似度関数SM”の1-0条件

$$[\exists j \in \lambda_t, \text{SM}(\psi_t, \omega_j) = 1] \tag{B.12}$$

$$\wedge [\forall k \in J - \{j\}, \text{SM}(\psi_t, \omega_k) = 0] \tag{B.13}$$

が成立する認識段階番号tが存在することがわかり、入力パターン φ が $T\omega_j$ として再生・想起される事態

$$\psi_t = T\omega_j \tag{B.14}$$

も成立する。このようにして、認識システムRECOGNITRONは、

$$\begin{aligned} & \text{RECOGNITRON can assign pattern } \varphi \text{ to the } j\text{th category } \mathfrak{C}_j \text{ with the maximum} \\ & \text{similarity 1 and he can associate } T\omega_j \text{ with } \varphi \end{aligned} \tag{B.15}$$

という“構造受精変換の不動点を求め、想起認識する動作”を遂行する。□

付録C. 類似度関数SMに関するSMミックスチュア条件

類似度関数SMに関するSM-チュアミックス条件は次のように述べられる。

【類似度関数に関するSM-ミックスチュア直交条件】

$$\forall k \in \mu, 0 < b_k < 1 \wedge 0 < \sum_{k \in \mu} b_k \leq 1 \tag{C.1}$$

を満たす実定数 b_k の組 $\{b_k \mid k \in \mu\}$ について、

$$\mu \supseteq \{\ell, m\} (\ell \neq m) \tag{C.2}$$

が成り立つような $\ell, m \in J$ が少なくとも存在する任意の $\mu \in 2^J$ に対し、ミックスチュア条件

$$\exists j \in \mu, \text{SM}(\sum_{k \in \mu} b_k \cdot T\omega_k, \omega_j) \neq b_j. \tag{C.3}$$

が成立しているような式(A3.5)の類似度関数SMはミックスチュア性類似度関数(ミックスチュア条件を満たす類似度関数)であるという。□

付録D. 類似度関数に関するSM-直交条件

類似度関数SMに関するSM-直交条件は次のように述べられる。

【類似度関数に関するSM-直交条件】

任意の $\mu (\neq \phi) \in 2^J$ についての, 実定数 a_i の組 $\{a_i \mid i \in \mu\}$ が, 正条件

$$\forall k \in \mu, a_k > 0 \tag{D.1}$$

を満たすとしよう. このとき, 直交条件

$$\forall j \in J - \mu, \text{SM}(\sum_{i \in \mu} a_i \cdot T\omega_i, \omega_j) = 0 \tag{D.2}$$

が成立しているような式 (A3.5) の類似度関数 SM は直交性類似度関数 (直交条件を満たす類似度関数) であるという.

(著者 鈴木昇一 (文教大学情報学部専任教員), 川俣博司 (富士ソフトABC株式会社), 大槻善樹 (文教大学情報学部非常勤講師), 論文題目 JAVA言語による計算機シミュレーションで生じた風景画像の理解場面での多段階連想形認識過程の異常現象, 文教大学情報学部情報研究no.29投稿論文, 投稿年月日 2003年2月15日(土))