

寄 書

視覚のパターン認識における一考察*

三重野博司**

Abstract

The aim of this paper is to give a theoretical expression to the process of visual pattern recognition. A man seems to perform his visual scanning, taking care of his whole visual field so as to compare the visual information with his inner "templete" pattern for its identification. Moreover, it is highly probable that he may switch one template to another with certain probability consciously or unconsciously. It might be well said that the dynamic process of this sort is very similar to the one that Gestalt psychology has proposed.

1. はじめに

視覚のパターン認識のプロセスについて数理的表示を試みたものである。この考察の特徴は知識を持ち、あるパターンをすでに学習している人間が、全体を考慮しつつパターン認識を行なっていると思われることに着目し、これを数学的に記述しようとした点である。その点ゲシュタルト心理学的パターン認識である。また人がさまざまな関心の持ちかたをする点から、いわば唯一のパターン原型ともいふべきものを意識しつづけるのではなく、ある確率で複数個のパターン原型を交互に期待しながら、視線の走査を行なっていくような人間のパターン認識過程に近いもの、いわば人間らしいものにした点である。

以下パターン認識の動的過程としてとらえたアルゴリズム図にしたがって記述する。(図1参照。)図中にある関心関数等の用語は本文中で定義して使用したものである。

本論文のようなパターン認識の動的把握は、動的画像入出力であるグラフィックディスプレイのソフトウェア開発の基礎ともなる。むしろ現在のそれは知的作業をするユーザーの立場から作られていない点で、このようなものが必要と思われる。

2. 関心関数

2次元ユークリッド空間に存在する入力図形の要素を x とする。1つの眼でとらえる点の位置は網膜上の2次元の点として刺激されるので、2次元の離散的な順序集合の要素として把握される。

x を意識できるのは、色の違い、明暗の違い、輝きの違いである。色はヤングの3原色説では、2次元面上に連続量として位置づけられる。また色相と称して正の整数として1次元で表現することも実用上考えられている。明るさ、輝きは1次元の連続量である。そこでこれら3者あわせると4次元となる。

ゲシュタルト心理学ではつねに全体を考えて判断する。基本図形内の未知の各 x は想定される基本図形から推定されつぎの条件があてはまる考えである。a. 基本図形と同じ色明るさ、b. 可動図形ではまとまって動く。上記のように x の特性値が4次元の実数軸からうる特性値空間において定義される場合について考える。つまり入力図方が色彩的属性を有する場合である。(色相彩度明度を色素番号で示せば3次元の整数軸。)

この時この特性値 $\mu(x)$ の集合を $\mu(x) (X \ni x)$ とすると、その部分集合の特定なものに人の視線が動くと考えられる。 μ を x から特性値の写像とするとその逆写像 μ^{-1} を関心をもちやすい特定な定義域について考えたとき $\{x\}$ を意識する。例えば赤い点(色番号

* A Study on Visual Pattern Recognition, by Hiroshi Mieno (Science University of Tokyo)

** 東京理科大学

号色相1)に関心のある人がこれを R^2 での定義域として見るとそれによって色どられている場所 $\{x\}$ が意識できるわけである。この特定な定義域を示す関数を特性値 $\mu(x)$ の関心関数 $f(\mu(x))$ と呼ぶことにする。これは色明るさ輝きそれぞれに、 $f_1(\mu_1(x))$, $f_2(\mu_2(x))$, $f_3(\mu_3(x))$ が考えられる。しかも、それは人共通の関数形と個人特有のものを含みあるいはその時、状況に応じて形が変わるであろう。特にここでは、パターン認識の過程によってこの関数の形が変わることを強調する。

3. 視野の時系列

視ることのできる点 x の集合を改めて X とする。その部分集合 $u_t (u_t \subset X)$ が t 時点での視野とする。 t が1から τ までの間にみることができ視野 U は

$$U = \{u_1, u_2, u_3, \dots, u_t, \dots, u_\tau\}$$

で示される順序集合でありかつ X の部分集合の族である。ただし $X \supset u_t (t=1, 2, \dots, \tau)$ 。

U がある形をなす図形 (x の集合であるから $\{x\}_t$ としておく) を含んでいれば、その図形があると意識できるはずである。人間のパターン認識はそれが大体含まれればその図形と判断して止めてしまう。それがおぼろげなときはそれまで見た u_1, u_2, \dots, u_t から判断して u_{t+1} を選ぶ。つまり視線をそちらにむける。このとき眼球運動はかりに連続的でも脳筋に送られる視覚情報は離散的であるから u_t と u_{t+1} にそれぞれ含まれる点は連続とは限らない。(他方人が2点を見る場合、ある値以上の角度でないと異なる2点と判断できないから数学的に連続であるか否かは本来認識できない。)

しかしながら $\{x\}_t \subset U$ になるように次の次点の u_{t+1} を選び、多くの場合 $u_t \cap u_{t+1} \neq \emptyset$ のように選んでゆく。といっても一番始めの段階では $\{x\}_t$ を想定できないから関心のある色の点の集積している u_t から始めよう。その意味で濃度 $|u_1| |u_2|$ は大きくばくぜんと全体をいくつか割った程度でとらえ $|u_1 \cap u_2|$ は小さくして早く全体を大づかみでとらえるであろう。これが初期の段階である。(図1ブロック①に相等。)

これによって発見された、関心のある色などを多く含む u_t を詳細に観察することになる。(図1ブロック②に相等。) また同時に関心ある色をもつ点を内容とする集合 u_t の他に境界点を多く含む u_t を詳細に観察する。(赤とそれ以外の色の境をたどって形を把握しようとする。)

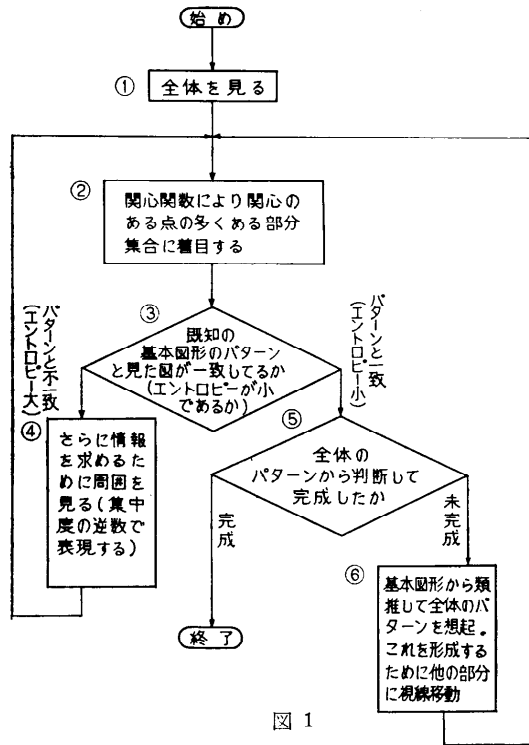


図 1

ここで改めて人の頭脳にはいつている知識としてのパターンを定義つけてみよう。

パターンとは特定な特性値 $\mu_f(x)$ を持つ点 x_f の集合 $\{x_f\}$ で、その集合のしかたつまり点の相互関係さらに点の集合間の相互関係 γ_f が特定の形式のものをいう。つまり $\{\lambda_f\}$ は形を定めるものであり $\{x_f\}$ はその位置を示すことになる。したがってパターン f の表現は

$$f = (\{x_f\}, \{\lambda_f\}, \{\mu_f(x)\})$$

で示される。 $\mu_f(x)$ は色などを示すが1つの形を整えた物でも2色以上からなるものの輝きの異なる物があるはずで、その物の一部から f が想定されると他の部分はこんな色こんな輝きであるはずだということが予想され、前述の関心関数の形が変わることになる。

γ_f の簡単な例を示すと U の中で点と点の相対的位置が一定方向である点の集合を求めると f は直線となる。ただし太さのある場合はすべての γ_f がその条件を満たさない。この場合は統計的に判断したり、境界点のみについて前記の条件を考えねばならない。

また、パターンは線分、弧などの集りが三角、四角などの図形をなし、さらにそれらはより複雑な図形を構成する。その階層に存在するパターンをそれぞれ

f_1, f_2, f_3 とすると

$$f_1 = (\{x_{f_1}\}, \{r_{f_1}\}, \{\mu_{f_1}(x)\}),$$

$$f_2 = (\{f_1\}, \{r_{f_1}\}),$$

$$f_3 = (\{f_2\}, \{r_{f_1}\})$$

で同様に示すことができる。

4. パターン認識のためのエントロピー

ある点に関心のある点か否かの確率をそれぞれ $p, (1-p)$ とする。そのときその点のあいまいさは

$$H(x) = p \log \frac{1}{p} + (1-p) \log \frac{1}{(1-p)}$$

前述の u_i はそれらの点の集まりとみて、 u_i が含む点のエントロピー $H(x)$ の点に関する総和が大なる u_i を求めて視線をむける。つまり関心のない赤以外の点からのみなっている部分には、 u_i を選ばない。赤があるのが確実な所も確認する必要がないかぎり選ばない。はじめは不確実な所こそ視線を向けて情報をとる必要がある。また境界を探すにはエントロピー大なる所に目をむければよい。そして境界点の多く存在する所は u_i の各点のエントロピーの総和も大である。境界点が多く発見できることは形状把握を早めることである。当然のことながら視線は情報量大な u_i を追って動き、ある情報量に U がなったとき止まることを示す。

エントロピーの概念を導入するには X と色などの属性 $\mu(x)$ によりボレル集合体などを定めそこにおける確率 p を定め確率測度空間を明確にしてから行わねばならないが詳細は割愛する。

つぎに、もう一つのエントロピーを考えよう。それは、基本図形 f が今見ようとしている画の中にあるか否かのエントロピーである。視覚でとらえた情報が U の広がるにつれて大となり、これにともなってこのエントロピーは少くなる。この判定が図1ブロックである。

5. 近傍依存

これまで用いていた画の中の点 x を改めて視覚でとらえて同一の点としか判別できない画の中の点の集合ということに定め、2次元 (i, j) で示されるセルのことをいうことにする。したがって x_{ij} は i, j なる位置にある点の集合である。 i, j はセルにつけられた整数である。

いま x_{ij} の $\mu(x_{ij})$ は未知でもその近傍

$$N = \{x_{i+\alpha, j+\beta} | \alpha\beta \text{ は } 0 \text{ を含まぬ有限の}$$

1 から始まる整数}

とすでに記憶ずみのパターン f の集合 $\{f\}$ から推定することができる。 N はもとより U の中に含まれたときの話である。 x_{ij} を含む u_i に視線を向けて $\mu(x_{ij})$ をみたとき推定したものと合わなかったときは想定した f を変えて試行錯誤するか、さらに別の u_{i+1} に視線を向けなければならない。 u_i から情報量のあまり得られないときは近傍のわくを広げて $\alpha\beta$ の数を増大するために、時おり視線をランダムに大きくふらせて周囲の情報をとる必要がある。事実人間の眼はこのように動くという。これは示された図によって広く見ないとわからない図とそうでないものがある。これを定量的に示す尺度として、視線の集中度 η なるものを考えてみよう。つまり η の小のものは周囲をみなければわからないということになる。 η はつぎの式であらわす。

$$\eta = \frac{1}{\left(\sum_{i=1}^r (\rho(u_i, u_{i+1}))^2\right)^{1/2}}$$

$\rho(u_i, u_{i+1})$ は両視野の中心間の距離。

なお近傍から推定する過程はセル・オートマトンの表現をかりることができよう。図1ブロック④はこれらのことを示す。

6. 想起したパターンの吸引力

パターン f は $\{x_f\} \{r_f\} \{\mu_f(x)\}$ で特徴づけられるかそれらで構成された空間の中の点としてあるパターンを考えられる。 $\{f\}$ はその空間の中の点の集合である。

このとき、それらの点の近傍はその点に対しエントロピーの小さい領域と考えられる。逆に2つのパターン f_1, f_2 は学習済みであれば、それらは判別をつけられるだけの距離にあるはずである。この空間は $\{f\}$ を逐次学習するのにしたがって軸の伸張をさせて判別しうるように2つのパターンを遠くへ引離すようになっている。

近傍に落ちこんだ図形はその点が示す f と同型と判断される。近傍の範囲は人の性格等でまちまちである。

また人が常に関心をもっている f は近傍の範囲が広くその分だけ引きこむ力が強い。その吸引力 a は

$$a = \left(\sum_f \left(\frac{w_f}{H_f^n}\right)\right)^{1/n}$$

で示してみた。 w_f は関心の強さを示したものとし、

w_f が実数で示しうるなら、 a も実数で示しうる。 H_f はパターン f に対する図形のもつエントロピーである。 n は心理実験によって定めねばならない数である。

以上のような空間にパターン f が固定され、画から得た U の情報で推定した点が $\{f\}$ の間をさまよいついて動くことになる。そして吸引力のあるいわゆるアトラクターに引かれて動き吸収されてしまう過程をふむ。これは力の場として位相力学の考えが導入できる。 f のもつ力でどういう軌道を描くかを考え、周期軌道・特異点の有無をしらべれば、初期情報からどんな判断過程をふむかが明確となる。

なお視線のランダム性(上記)とか $\{f\}$ を想起する不確定性もあって、初期情報が等しくても同一の結果となるとは限らない。

パターンはこのように階層の下からつみあげて判別されるだけでなく(点から線、線から図形へ、の逆)上から判断して試行錯誤されるものである。

以上のような事柄から図1ブロック⑤⑥が必要である。

7. あとがき

この種の研究で動的にあつかったものは少いと愚考するので、つぎに一般システムをあらわすフローチャートの開発と応用したものを許されれば発表し、さらに今後心理実験などで本論文の基礎的な事柄を実証しながら研究をすすめたいと思う。本学会レフリーの方に、ご親切なご指導をいただいたことを深く感謝いたします。

参 考 文 献

- 1) Bongard: Pattern Recognition, Spartan, 1970.
- 2) E. A. Barbashin: Introduction to the theory of Stability, Wolters-Noordhoff, 1970.
- 3) 飯島泰蔵: 図形空間の構成法に関する基礎理論, 電子通信学会誌 71/8, Vol. 54-C, No. 8.
- 4) 鹿取広人: 図形認知の発生条件, 日本心理学会, 昭43年.

(昭和46年12月7日 受付)

(昭和47年3月22日 再受付)