

トポロジカル・イメージ・サイエンス

Topological Image Science

飯 田 武 揚*

Takeaki HIDA

従来までの画像工学は物の像をいかに複写するかにとりこんできた。ところがこれからの画像工学は情報処理の高次元化にともなって判断したり、記憶したり、学習したりする分野にまで拡張してゆく必要性が生じてきている。一般にこの分野の画像工学はパターン認識という概念でとりあつかわれはじめているが、このパターン認識は analytical におこなうものでなく、topological におこなわれるべきであると考え。そう考えると人間がおこなっている情報処理はまさしく topological なものであることがわかる。この解説においてはこれからの画像工学が位相数学 (topology) を使った Image Science であるべきことを述べ、新たにこのような分野を Topological Image Science として発展させることを提唱する。

1. はじめに

Reprography の分野はめざましい勢いで発展しつつある。この中心的役割は銀塩写真よりも非銀塩写真、いわゆる電子写真、ジアブ写真らがうけてきたといえよう。これらのプロセスは情報処理技術の一つの手段として今後も発展してゆくだろうが、この技術はよく考えてみると情報処理としては物を単に写しとるという低次元の技術である。これらの分野にたずさわる人々はこれらの技術の次にくるものは何であろうかといつも自問しているわけであるが、その解答は明確である。それは情報処理の一つのプロセスとして、さらに機能的向上を計る必要があるということである。例を上げるなら同じ複写でも写真電送は Image の長距離輸送が可能であるという点で物を単に複写するという技術にくらべて一次元機能を上げることになる。このような技術の実用化はテレビファクスや home facsimile などとして現在かなりわれわれの生活に入りこんできた。これらは新聞の家庭への電送、紙ビデオテープなどの多くの応用を開発してゆくものと思われる。

ところがこのような複写の技術も 10 年後にはさらに一つの飛躍が必要である。そこで直面する重大な問題は判断する画像工学をいかにするかということであろう。それは例えば手で書かれた漢字を即座に活字にするような漢字プリンターや郵便番号による手紙の選別などの技術である。このような機能は人間の頭脳は μsec の速さで処理しているが、現在のいかなる技術をもってきても手も足も出ないというのが現状であろう。これらはすべて情報処理技術の中の画像工学の問題である。

これらの判断したり、記憶 (記録とは意味がちがう) したりする画像工学は従来のような解析的立場に立ってはいは解決できない問題であり、一つの新しい立場に立っておしすすめてゆく必要がある。その新しい立場とは topology であると考え。topology (位相数学) は近年

になって発展してきた新しい数学であり、空間を連続写像を通して判断する基準を与える数学であるので、前述の判断する画像工学にはなくてはならない概念である。一般的に一つのパターンを判断することをパターン認識といっているようであるが、ここで提出する位相数学を使った画像工学はさらに一般的に Topological Image Science と呼ぶことにする。

この解説においてはそのような観点に立って画像工学を展開してみたいと思う。まず情報の定義、画像の定式化、その写像 (mapping) と画像処理の対応、さらに判断する画像処理をいかにするかについて述べてゆく。

2. 情報と集合と画像

現代の科学に重要な意味をもつ情報の概念は通信理論における Shannon¹⁾ らによって次式によって定義された。

$$I = - \sum_{i=1}^{i=n} P_i \log P_i$$

ここで P_i は i 番目の事象があらわれる確率で、 I は情報量 (bits) である。この情報量は Boltzmann の提出したエントロピーの概念とまったく逆の性質をもつことから負のエントロピー (Negentropy) とも呼ばれる。しかし情報という概念は通信理論だけにとどまらず、生物系など多くの事象に関与することがわかってきたので、Makishima²⁾ は次のように情報を定義化している。

“情報とは均一な back ground の中にわれわれの認識をとおして直接あるいは間接的に判別できる何らかの特徴”

この概念をもちいれば画像と情報の関係が明確になる。すなわち画像とは情報が物質的にか、エネルギー的にか 1 つのパターンをなすものであるということである。

画像のもつ他の一つの重要な概念は集合というものである。すなわち画像は一つの情報の集合であるといえる。画像の情報のメンバーは有限であるから、画像は一般に情報の有限集合で表現できる。物の像を考

* 東京大学生産技術研究所 第 4 部

えるとき、この集合という概念は重要になってくる。なぜなら、画像を判別する際に集合の概念を使えば、一つの画像にある共通の性質によって集合に属さないものと属しているものとを区別できるからで、画像処理工程一つの情報集合の写像と考えると集合の意味はさらに明確になる。

3. 画像の定式化

最初に時間変化しない2次元画像を考えると、画像は1) 画像面積の大きさ、2) 画像表面の各点の明暗、3) 画像表面の各点の色の3つの変数が決まれば決定される。そこでこの2次元画像を定式化するにはまず図1のような均一な back ground をもつ一つの平面を考え、

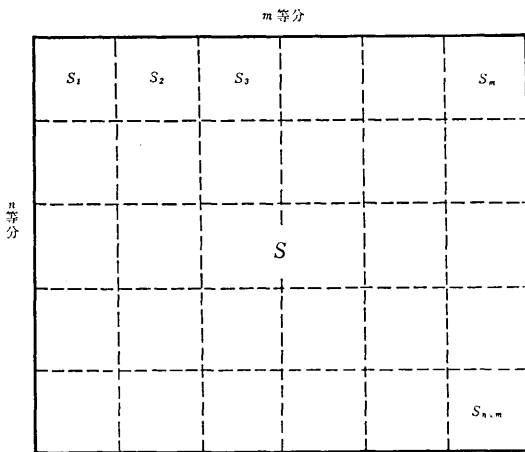


図1 画像の区分

その全面積を S であらわす。この平面の横軸を m 等分し、たて軸を n 等分して分けられた小面積を S_i とし、左上方から S_1, S_2, S_3, \dots と番号をつけると、この面積の集合 S は

$$S = \{S_1, S_2, S_3, S_4, \dots, S_{n \times m}\}$$

と表現できる。

大きさ S 上の画像を I^s とすれば次式が

$$I^s = D(d_1, d_2, d_3, d_4, \dots, d_{n \times m})$$

画像の明暗を示す式になる。ただし、 d_i は S_i 上の明るさを意味し、 D はその集合全体を意味する。

さらにこれに色を決める変数を導入するには次のように考える。 $(S_1, S_2, S_3, \dots, S_{nm})$ の小面積の中に明暗の因子と色の因子から任意にそのメンバーを選んで並べるのであるから、この概念は直積の考え方で同等である。すなわち色彩のある画像は明暗の集合と色合いの集合の直積になっていると考えることができる。差色、和色を含めたすべての色は Red, Yellow, Green, Blue で表現できるので、色の集合の独立のメンバーは4個と考えるとよい。

そこで時間変化しない2次元画像の定式化は次のようになる。

$$I^s = D(d_1, d_2, d_3, \dots, d_{nm}) \times C(R, Y, G, B)$$

さらに時間変化している画像は次のようになる。

$$I^s = D(d_1, d_2, d_3, \dots, d_{nm}) \times C(R, Y, G, B) \times T(t)$$

この式が色、明暗、画像の広さ、時間を含む2次元画像の式になり、これらのメンバーが決定されれば画像は完全に与えられることになる。

3次元画像への拡張は2次元画像 I^s へ z 方向への明暗の集合 $D_3\{d_i\}$ を直積すればよい。

$$I^v = I^s \times D_3\{d_i\}$$

一般の n 次元画像は1次元の明暗の集合 $D_i\{d_i\}$ を考え、

$$I^n = D_1 \times D_2 \times D_3 \times \dots \times D_n \times C \times T \\ = D^n \times C \times T$$

とすればよい。

4. 写像と画像プロセス

写像 (mapping) とは一つの集合を他の集合へ写す方法のことをいう。この概念は画像工程の転写と同等である。つまりある景色をカメラで撮るとことは視野の各点の集合を光によりフィルムの上の点の集合に写像していることになる。これを数学的にいえば、集合 X と Y とがあるとき X より Y への写像 f とは X の各メンバー x に Y のメンバー y を対応させる規則であり

$$f : X \rightarrow Y$$

と書かれる。カメラの例をとれば X とは視野の情報集合であり、 Y とはフィルム中の銀粒子の集合のことである。われわれが見ている視界というものは個々の分子に光が当たってそこに生じた情報の集合であるから一見連続のように見えるが、実は不連続の有限集合である。これを写像して作る画像も個々の分子によって作られているから、画像も不連続の有限集合である。しからば一般の画像処理は不連続写像かというそうではない。それは画像処理における写像は個々の点を他の空間にそれぞれ対応させており、一つの点の neighbor を忠実に他の空間の neighbor に写しているので、写像は有限集合 X から、有限集合 Y への連続写像と考えてよい。このような集合の写像は厳密には分離連続 (separately continuous) としてあつかうべきであろう。

しかしこの場合に X の点はすべて Y へ連続的に写像されるのではない。少なくともわれわれが画像といているのは人間の目を基準にして作られるものであるから、人間の目の分解能力、判断基準、時間的変化率には一定の許容範囲がある。この許容範囲を許容空間と呼び、人間はこの範囲内では判別できない不確定さをもっている。しかし画像 Y のメンバーは少なくとも視界 X のいずれかに属しており、 $Y \subset X$ という関係があるため画像処理は集合 X より Y への包含写像と考えられ、 X の濃度は Y より大きいにもかかわらず許容空間内での各点はそれぞれ写像されていると考える。よって画像処理にあら

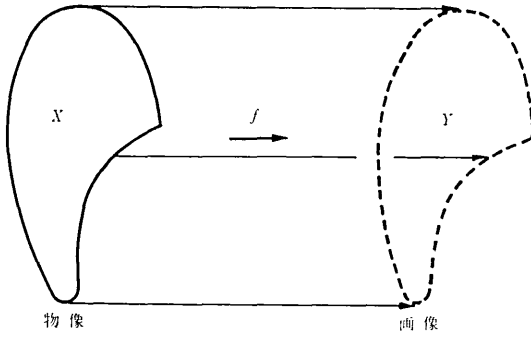


図2 物像の集合(X)と画像(Y)

われる写像操作はあらゆる意味でこの包含写像であるといえることができる。

写像関係 $f : X \rightarrow Y$ において集合 X は写像 f の領域 (domain) と呼ばれ、これは人間が見ている視野の範囲に相当し、集合 Y は f の値域 (range) とよばれ画像化される範囲に相当する。

銀塩写真のように視野の光像の逆の明暗をもつ、いわゆるネガ画像を作る写像に逆像の関係 $f' : X \rightarrow Y'$ で示され f' のことを逆像写像という。 X から Y への逆像の写像をするときに、 X の中のある点の明暗はまったく変化しないで Y へ写像される場合すなわち X と Y とでまったく等しい明暗の度合いがある。このような点を写像における不動点 (fixed point) と考えることができる。銀

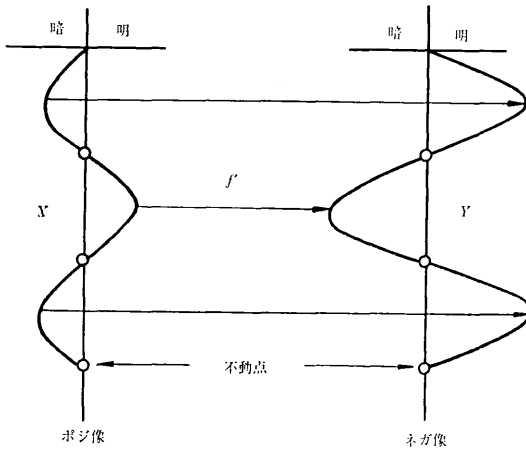


図3 銀塩写真のときの写像

塩写真のように逆像の関係を写像する画像処理は感光材料からの規制としてやむおえないプロセスのように考えられているが、実は似ている画像の判断のときに逆像の関係や不動点の考え方は重要な意味をもってくるものとする。

ところで一つの視野を画像化し、それをオリジナル・コピーとして他の画像を再現させる電子写真や写真電送の場合の写像はどのように考えるべきだろうか。これは写像のプロセスがまず視野を画像化する f とそれを土台にして複写をおこなう操作 g の2つのプロセスを考え、

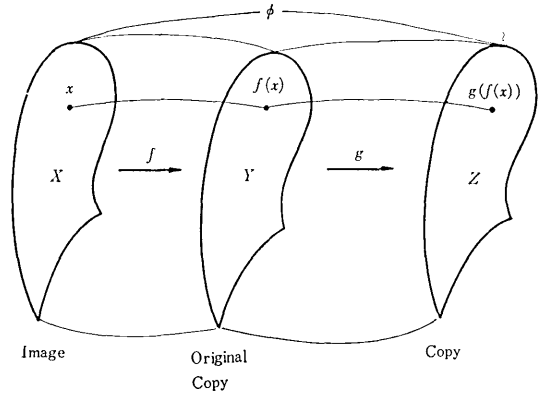


図4 写真電送のときの写像

それには結合関係があるとする。これらには

$$f(x) \in Y$$

$$g(f(x)) \in Z$$

の関係があり

$$g(f(x)) = \phi(x) \text{ とおくと}$$

$$\phi = g \circ f : X \rightarrow Z$$

となり、この ϕ は g と f の結合操作であるといえることができる。

画像処理においては像の拡大・縮小・増幅 (強度の拡大) などをおこなう場合が多いが、このプロセスは次のように考える。 A_1 を A の部分集合と考え、 $A_1 \subset A$ とする。写像 $f : A \rightarrow B$ に対して、 $g : A_1 \rightarrow B$ のような写

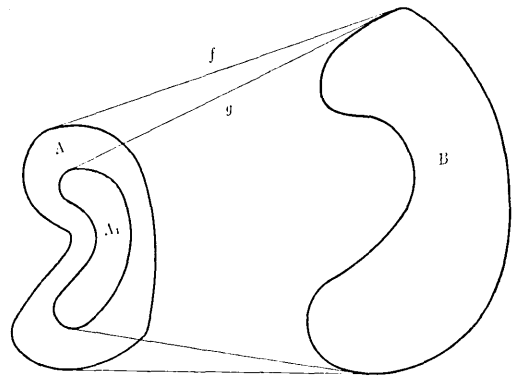


図5 画像の拡大と縮小

像を考えると g は A の部分集合 A_1 を B に写像しているのであるから、 f を g の拡大、 g を f の縮小といい、

$$g = f|_{A_1}$$

とあらわす。

このように画像操作のプロセスはいかなる場合もある情報集合から他の情報集合への写像としてあらわすことができる。

5. Topological Image Science によるパターン認識

これまで述べてきたことから画像工学はすべて情報集

合の画像操作に帰せられることがわかった。それではこれらの性質を使ってパターン認識のような画像を判断するような方法をいかにしてゆくべきか、これが一番の問題になる。例えば漢字の中の次のようなものを考えてみ



る。はじめの文字から字体はだんだんと変形してゆく。右から 2 番目までなら誰でも容易に解読できるが、一番右側の文字はかなり変形しており、このへんが変形の極大範囲でこれ以上変形すると他の情報を意味するようになってしまう。

上に示した 5 つの文字はいずれも“電”という文字であり人間には同一の情報として認識される。何の不思議もないようであるが、これを現代のエレクトロニクスでアナリティカルにとりあつかってみると大変なのである。なぜなら上述の文字を光電管で走査してその中から {0, 1} の信号をアナログ化して取り出したとすれば、まず最左の字と次の字とは異なっている信号であると記録する。すなわち最左の字の画像の式

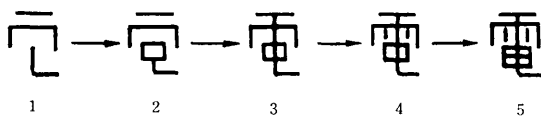
$$I^* = D(d_1, d_2, d_3, \dots, d_{nm})$$

に対して、次の字の画像の式

$$I^* = D(e_1, e_2, e_3, \dots, e_{nm})$$

において $d_i = e_i$ ならば同じ字であり、 $d_i \neq e_i$ なら違っていると判断する。この文字を同一の“電”だと認識させるのはコンピューターにあらかじめこれらのパターンを記録させるか、学習させねばならない。文字の変形は無数に考えられるので、まずこの方法によるパターン認識は不可能に近いことになる。

トポロジカルな観点に立つなら上述の困難はよほど軽減される。まずトポロジーでは変形そのものは問題にならず、そこに存在している“特徴”を指数化することによってそれを判断してゆく。電の字を例に上げれば、雨冠の中の点がなくとも、二つの平行線の下の線が 2 回屈曲（微分位相的にいえば、線の勾配、すなわち微分の不連続点が 2 個あるということになる）しており、田という文字を貫いて下がる右へ曲る線によって特徴づけられる（下図 1）、次の特徴としては中央に口を入れ（下図 2）、さらに雨冠のところに点を入れることによって（下図 4）ますます電の字らしくなってくる。



おそらく 3 番目の特徴まで抽出すればほぼこの字が当用漢字（1850 個）の何であるかを判定してくれるものと考えられる。トポロジーの場合変形は問題にならないのだから少々直線が曲線になっても、それを特徴づける指数は

一定に保たれる。Topological Image Science はそのような特徴を写像的観点から指数化（記号化でもよい）することから出発し、その指数あるいは記号を基準にしてパターンの認識をおこなう方法であるので、従来の解析的な画像工學にくらべてはるかに判断する画像工學に適しているものであることがわかる。

6. Topological Image Science における許容空間

人間のパターン認識にしても、画像認識の方法にしても、そこに一定の限界があり、それ以下の限界は問題にしないでいい場合がある。例えば視野の中心では黒の 2 点は角度が 1 分以内のときは 1 点に見え、視野の端の方では 2 点が 1 度以内にあるときは同じ点に見える。カラーの点の判別はこれより鈍いことがわかっている。さらに前述の“電”という文字がどこまで変形できるかという例もパターン認識の限界があるという一つの例となる。このような限界の範囲を一般に許容空間とよび、判断する画像工學の重要な概念になる。なぜならあらゆる情報集合は一つの情報を他へ写像する際に少なくともその限界を満足させて写像すれば十分であるという機能上の要求を問題にする場合が多いからである。

それでは画像の許容空間をトポロジー的に表現するにはどうしたらよいかということ述べる。まず画像の定式化のところ述べてから出発する。一定の面積 S をもつ色も時間も含まない一番単純な画像を考える。まず、たて軸の分割と横軸の分割が $m=1, n=1$ の場合に画面が白であるか黒であるかの 2 つの場合を考えてみよう。

表に示したようにそのときは画面にあらわれる絵の数は白か黒の $2^1=2$ しかない。この考えをすすめて白と黒の間を連続的に変化する明暗を d_1 とする。この d_1 は 2 次元の関数空間のベクトルと考える。このベクトルは表に示したように $(1/2, 0)$ のところに原点（不動点）を対応させると、連続的に変化する画面の明暗はこのベクトルの先端の点で表現されることになる。

人間はわずかな明暗の変化は判断できないので人間の判断が不確実になる明暗の度合いの範囲を $\eta_1 \sim \eta_1'$ とすると、この画面の場合許容空間は $\eta_1 \leq d_1 \leq \eta_1'$ となり、直線の微小な長さの空間に対応する。

以上のことを一般化してゆくと画面の分割が $m \times n$ 個の画素にわけられるときは一つの絵は $n \times m$ 次元の関数空間の中の 1 点に対応し、その許容空間は

$$\eta_1 \eta_2 \eta_3 \dots \eta_{n \times m} \leq d_1 d_2 d_3 \dots d_{n \times m} \leq \eta_1' \eta_2' \eta_3' \dots \eta_{n \times m}'$$

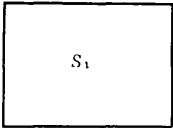
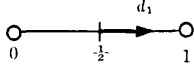
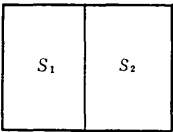
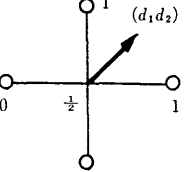
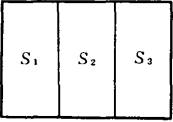
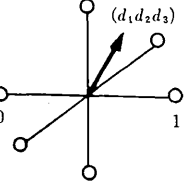
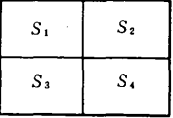
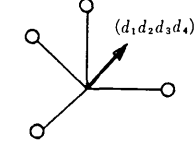
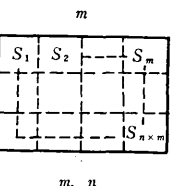
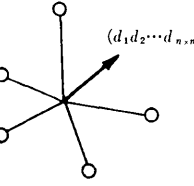
となるが、これをより簡単な表現にすると

$$\eta_i \alpha^{nm} \leq d_i \alpha^{nm} \leq \eta_i' \alpha^{nm}$$

となる。

画面の分割は画像の性質に依存している。非常に混み

表 1 画像の表現と許容空間

画面の分割	画面の数 (白, 黒)	関数空間	関数空間の次元	許容空間
 $m=1, n=1$	$2^1=2$		1次元	$\eta_1 \leq d_1 \leq \eta_1'$
 $m=2, n=1$	$2^2=4$		2次元	$\eta_1 \eta_2 \leq d_1 d_2 \leq \eta_1' \eta_2'$
 $m=3, n=1$	$2^3=8$		3次元	$\eta_1 \eta_2 \eta_3 \leq d_1 d_2 d_3 \leq \eta_1' \eta_2' \eta_3'$
 $m=2, n=2$	$2^4=16$		4次元	$\eta_1 \eta_2 \eta_3 \eta_4 \leq d_1 d_2 d_3 d_4 \leq \eta_1' \eta_2' \eta_3' \eta_4'$
 m, n	$2^{n \times m}$		$n \times m$ 次元	$\eta_1^{n \times m} \leq d_1^{n \times m} \leq \eta_1'^{n \times m}$

入った絵のときは画面の分割はかなり小さくする必要があるが、混み入っていない絵はかなり分割を荒くとることができる。要するに情報の密度に依存している。漢字はアルファベットより情報の密度が高く、高い密度の複写をしないと解読できないのもその例である。

7. 許容空間におけるパターン認識

5で述べたようにパターン認識は特徴抽出法が一番手っ取り早く、それを一つの記号に変換して処理する方法をとるべきことを述べた。6では一つの画像は画面分割の次元空間の1点としての代表 (representative) として

示されることを述べ、情報の許容 (allowance) が許される範囲を許容空間と定義した。それでは同じ情報を持ちながらパターンがわずかに変化しているときに、それを同じ物であると判断しなくてはならない場合はどうすべきかを考えてみよう。これは前述の“電”を変形させても同じ情報として判断するような場合である。

この時にはまず3次元的に構成された回路網を考えよう。この回路は何等かのビット数をもつ情報が入ったら、必ずこの3次元の空間の1点にゆきつくように作られているとする。そこで“電”という文字がいかようにも変形されて、変形されるごとにその特徴を抽出して記号化

されてくるとしよう。特徴は文字の変形によって異なって抽出されるが、その異なり方の範囲があまり大きくなく“電”という文字であるかぎり、一定範囲の空間内に納められるとする。そのような一定の範囲を許容空間にとって、その中へ入ったものは最終的には“電”であると判断するような回路網を構成すれば、文字の変形による同じ情報を判断することは可能になると考えられる。このような方式をトポロジーによる許容空間におけるパターン認識と呼ぶことにしよう。このような考え方はトポロジーの中では位相同型 (topological congruent) としてとりあつかわれている概念である。

む す び

従来までの画像工学が解析的な方法だけで画像処理を行ってきたが、漢字プリンターや高次の画像処理機能をもつ画像工学が必要になっている。このような分野を一般にパターン認識といっているようであるが、このパターン認識にはトポロジーの概念を導入すると多くの困難が除かれる。そのようなトポロジカルな観点に立つ画像工学を Topological Image Science と呼び、新しいアイデアのもとでのパターン認識工学を行なうべきことを述べた。それらの対応づけでわかったことをまとめる。

- 1) 画像とは情報集合が物質的にかエネルギー的にか1つのパターンをなすものである。
- 2) 画像は画面の大きさ、各点の明暗、色がきまれば決定され、それらの集合の直積で定式化される。
- 3) 画像処理のプロセスはことごとくトポロジーでいう写像の概念で表現され、対応づけられる。
- 4) 画像には必ずトポロジーでいう許容空間が存在し、それをいかに処理するかによって画像工程が複雑にも簡単にもなりえる。
- 5) 銀塩写真のようなネガ、ポジの写像の中には両方の

像にまったく明暗の等しいところがある。この点は写像の際に動かないので不動点と考えられる。

- 6) トポロジカルにパターン認識をするには画像の特徴だけを問題にし、それを一定の指数(あるいは記号)に変換して写像する。写像の際には許容空間内での許容をどのように取るかによって判断の基準化をおこなうので、従来の解析的な画像処理にくらべてはるかにその困難を軽減する。
- 7) 文字の変形などのパターン認識の困難さはトポロジカルな観点に立った許容空間の概念によってなんとか解決の糸口を持ちそうである。

以上この論文で述べたことについてまとめてみた。最後に現在の人間がやっている情報処理とエレクトロニクスでやれることのギャップの深さを考えてみると、そこにはかなり深い断層があるように思う。人間は学習によるパターンを土台に記憶をしている。画像工学においては学習どころか単に物を記録だけしているということである。記憶と記録は同じようにみても、その機能の間には簡単にとび越えられない断層があると思う。このようなギャップを一つ一つ解決していくことが、これからの画像工学の問題であろうかと思う。このような一つの大系の中でトポロジーが一つの役割を担ってくれるにちがいない。そうして画像処理が高次元の機能を持ちえたならば、絵→文字→思考といった一連の考える画像工学が誕生するにちがいない。筆者はそのようなことが実現するのはそんなに遠い将来でないと考えている。

(1970年3月25日受理)

文 献

- 1) C.E. Shannon, J. Instit. Elect. Eng. **93**, 429 (1946)
- 2) 牧島, Engineers (日科技連) No. 248, April (1969)
- 3) 河田敬義, 位相数学 (共立数学講座)
 亀谷俊司, 集合と位相 (朝倉書店)
 野口宏, トポロジーの世界 (ダイヤモンド社)
- 4) 野崎, 飯田, 第7回生研講習会テキスト, p. 107