

# デジタル図形の相似及びアフィン不変認識に関する研究

相良, 哲生

<https://doi.org/10.11501/3181887>

---

出版情報 : 九州芸術工科大学, 2000, 博士 (工学), 課程博士  
バージョン :  
権利関係 :

# 第1章

## 序論

2000年7月、我が国において九州・沖縄サミット（G8, 主要国首脳会議）が開催された。その主要なテーマの一つとなったのは、IT(Information Technology, 情報技術)である。計算機内の仮想的な空間において商取引をはじめとする様々な行為を行っていかうとするものである。計算機は、これまでの研究・開発中心の利用から、より一層身近な分野にまで浸透している。計算機を様々な分野で利用することは有用であり、もはや当たり前のことである。この計算機を利用した技術の一つとしてデジタル画像処理があり、本論文では、計算機を用いたデジタル画像処理における重要な分野であるデジタル図形の不変認識に関する研究について述べる。本章では、まず本研究の位置づけや目的および本研究を行うに至った背景について述べ、次いで本論文の全体の概要および構成を示す。

### 1.1 研究の目的とその背景

半導体技術のめざましい進歩に伴い、計算機の能力は飛躍的に向上している。この計算機の能力向上に関して、よく知られている話として「ムーアの法則」がある。これは米国 Intel 社の創設者の一人 Gordon Moore 博士が 1965 年に唱えたものであり、「半導体の性能と集積は、18 ヶ月から 24 ヶ月ごとに 2 倍になる」と言うものである[1]。この法則によれば、半

導体技術の集大成ともいえる計算機の処理能力は、1年半から2年という比較的短い期間に、指数関数的に増加することになる。実際に計算機の能力は、ほぼ「ムーアの法則」の通りに向上してきている。図 1.1 は半導体の集積度と処理能力の推移を表すグラフである。縦軸左側では一つのMPU (Micro Processor Unit, マイクロプロセッサ) に使用されているトランジスタの数で半導体の集積度を示し、縦軸右側では1秒間に処理できる命令数 (MIPS 値) で処理能力を表している。

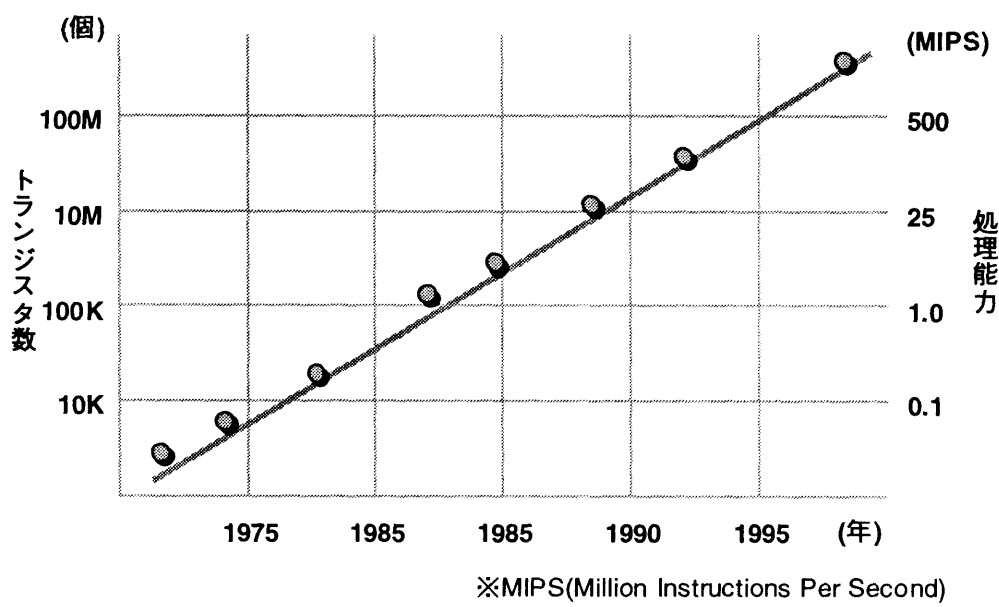


図 1.1 半導体の集積度と処理能力の推移

計算機の能力の向上にともない、その応用分野も広がった。計算機で取り扱うデータは、従来、テキストベースの情報が中心であったが、近年、マルチメディアデータの一つとして画像情報を取り扱うことが盛んになってきた。計算機で画像情報を扱うとき、問題となるのはテキストベースの情報が必要とする情報に比べ、情報量がはるかに大きいことである。たとえば、文字情報として漢字を扱う場合、一般的には、1文字あたり16ビット程度の情報量ですむ(表 1.1)。これを画像情報として取り扱おうとするときには、1文字が16ピクセル(画素)四方の大きさの場合で256ビット

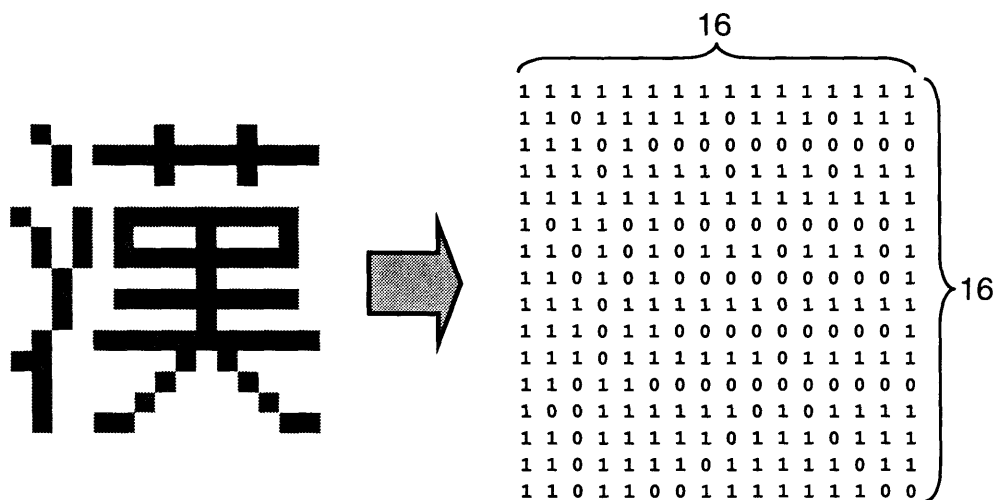
の情報量が必要であり (図 1.2), 文字情報として扱う場合の 16 倍となるため, 多くの文字を扱う場合は取り扱う情報の量は膨大なものとなる。

表 1.1 漢字を文字情報として表現した場合の情報量と値の例

文字コード体系[2]~[9]	情報量 (ビット)	文字“漢”の コード値
シフト JIS コード(MS 漢字コード)	16	8ABF
日本語 EUC コード	16	B4C1
JIS(区点)コード	14	20 区 33 点
JIS(JIPS)コード	16	3441
Unicode	16 以上	U6F22

※ EUC(Extended Unix Code, 拡張 UNIX コード)

JIPS(Japanese Information Processing System, 日本語情報処理)



黒画素=0, 白画素=1 で表すと,  $16 \times 16 = 256$  (ビット)

図 1.2 漢字一文字を画像情報とした場合の情報量

しかしながら、前述したような半導体の処理能力の向上により、計算機は時が経つにつれ、

- 同じコストならより高性能
- 同じ性能ならより低コスト

なものとなってきている。このことにより画像情報を取り扱うのに十分な能力を持った計算機が身近なものとなった。

画像を計算機で取り扱う分野は次の二つに大きく分けて考えることが出来る。一つは、計算機の内部の処理において生成したデジタル画像を外部へ出力するものであり、もう一つは、外部の画像をデジタル化して計算機へ取り込み、何らかの処理を行うものである。前者は主にコンピュータグラフィックスと呼ばれる。テレビ放送でのテロップ画像の表示や、芸術作品の表現方法等において幅広く利用されている。一方、後者は画像処理と呼ばれ、次のような分野がある。

### 1. 画像の変換

目的に応じて雑音、歪みが混入した画像の画質を改善し「見やすい」画像に変換することである。この分野では、幾何学変換、フィルタリング、階調の補正等の処理が用いられる。

### 2. 画像の符号化・圧縮

画像情報は膨大なデータ量となるが、これは通信や、メモリ・データベース等への記憶の際に障害になる。このデータの内、必要な情報を効率的に表現しなおす（符号化する）ことで画像の持つデータ量を削減することである。

### 3. 画像認識

画像の中から対象物を抽出して、その面積や個数、明るさなどを計測したり、対象物の不変的な特徴を抽出したりして、それらの処理から得られた情報をもとに、それが「何であるか」を認識・理解することである。

画像の変換や、符号化・圧縮の分野では十分な研究がなされ、たとえ

ば身近なところでは、人工衛星から撮影された画像に対して各種変換や解析を行った気象情報の提供等が実用化されている。GIF(Graphics Interchange Format)形式やJPEG(Joint Photographic Experts Group)形式など、画像圧縮技術[10]を用いたフォーマットはインターネット上で公開するWeb ページに画像を掲載する際の標準的な画像フォーマットとして普及している。

画像認識は、製造現場における工業部品の検査や選別などを中心に発展し、現在もさらなる応用分野の拡大が進められている。たとえば、交通の分野では、ITS(Intelligent Transport Systems, 高度道路交通システム)という交通システムにおける情報化[11]-[14]が行われており、その一つの柱としてAHS(Advanced Cruise-Assist Highway System, 走行支援道路システム)の研究・開発が進められている[15],[16]。このシステムでは、道路情報の収集に画像認識の技術が重要な役割を担っている。画像認識は今後も社会における情報処理技術としてより一層重要度を増していくことが考えられ、この分野の研究は更に発展していくものと思われる。

計算機で画像を取り扱う場合、通常は画像を数値データ化して扱う。一般的なデジタル画像では、2次元のアナログ画像を、格子状に細かく区切った各領域、すなわち点の集合とした上で、各点の輝度等の属性を数値化したデータの集合として用いる。我々人間は、視点の変化や対象物自身の変位などにより、網膜上での位置や大きさが異なったものとなってしまった同一の画像を、同一の対象物として認識できる。一方、計算機は基本的に数値の演算・比較程度の処理しか行うことができないので、これらのデータの集合を単純に演算・比較しても、同一なものとして認識するのは困難である。このため、画像の位置や大きさなどの変化に対して不変な認識方法を計算機に対して指示する必要がある。このような対象物の画像の変化を、画像の変換と考えると、次のような種類のものがある。

- 対象物の位置の変化 … シフト変換
- 対象物の回転 … 回転変換
- 対象物の大きさの変化 … スケール変換
- 対象物のある一方向への引き伸ばしや、四角形をひし形にする

このような各種の変換に対して不変な特徴を抽出することで、画像は画像空間から特徴空間へ写像される。ここで抽出された特徴を、計算機で取り扱うことが容易な特徴ベクトルとした上で、ベクトルどうしの比較により各種の不変認識を行うことができる。このような不変認識の実現のために、対象物の画像に含まれる輪郭線などの線図形情報をもとにした特徴がよく用いられている。たとえば、線図形に対してフーリエ記述子のパワースペクトルを求める方法[19]~[21]や、モーメント不変量を用いる方法[25],[26]などが提案されている。さらに、CSS (Curvature Scale Space)を用いる方法[27]や、複素自己回帰モデルを用いる方法[28]、デジタル曲線を多角形で近似する方法[29],[30]なども提案されており、これらはシフト・回転・スケール変換に対して不変な認識を実現している。また、モーメント不変量を用いたアフィン変換に対して不変な認識方法[31]も提案されている。

デジタル画像における線図形は、複数の点が“つながった”ものとして取り扱われるため、点と点とのつながり、すなわちある点の次にどの点がつながり、さらにその次はどの点がつながっているかといった情報が通常必要となる。しかしながら、前述したような従来のデジタル図形の不変認識方法では、互いにある変換の関係にある図形同士を同一のものと認識するためには、図形を構成する点のつながりに関する先験的な情報が必要であり、この情報がない場合、曲線をパラメータ表記することが出来ない。しかしながら、このような点と点のつながりに関する情報を得ることは一般にはむずかしい。また、対象となるデジタル図形は、単連結なものに限られる。

本研究では、「自己距離関数」および「自己三角形関数」という、新たな概念を用いた、デジタル図形の不変認識方法を提案する[32]~[38]。ここで提案する方法では、点と点のつながりに関する情報を一切必要としない。つまり、前述のような従来の認識手法では困難であった非単連結なデジタル図形に対しても適用が可能である。

## 1.2 論文の概要と構成

本論文では、「自己距離関数」および「自己三角形関数」という、新たな概念を用いた、点と点のつながりに関する情報を一切必要としないデジタル図形の認識方法を提案する。

本論文は次に示す6つの章から構成されており、本研究における認識手法や本手法を用いた実験結果をまとめたものである。

第1章(本章)では、まず本研究を行うに至る背景と扱う問題について述べ、次いで本論文の概要について述べる。第2章では、まず基礎理論として、パターン認識における不変認識とその特徴について述べる。次いで、一般的によく用いられている輪郭線の記述法および、距離の定義やデジタル図形を取り扱う場合の諸問題について述べる。

第3章では、自己距離関数を用いてシフト・回転・スケール変換に対して不変な図形認識法を提案する。自己距離関数法は、デジタル図形を点の集合と定義した上で、デジタル図形を構成するすべての点の、2点の組み合わせをとり、その2点を両端とした直線の長さのヒストグラムをもとにしている。このため、点の順序やつながりに関する情報は一切捨象されるので、点のつながりに関する情報は不要である。したがって、従来困難であった非単連結曲線にも適用できるという特長がある。この自己距離関数をベクトル化した自己距離ベクトルを用いて認識実験を行い、良好な認識結果が得られることを述べる。

第4章では、自己三角形関数を用いて、アフィン不変な図形認識法を提案する。図形のアフィン不変認識はその重要性にも関わらず、これまで有効な方法が見出されていなかった。本方法は、前章の自己距離関数を用いた方法と同様に、非単連結な曲線に対しても適用可能である。自己三角形関数は、デジタル図形を構成するすべての点の、3点の組み合わせをとり、その3点を頂点とした三角形の面積のヒストグラムをもとにしている。2つの三角形の面積の比はアフィン変換の下で不変であるので、自己三角形関数はアフィン不変となることが容易に示される。自己三角形関数は、前章の自己距離関数における2点間の直線の距離を、3点から構成される三角形の面積に置



き換えたものである。この自己三角形関数をベクトル化した自己三角形ベクトルを用いた認識実験を行い、良好な認識結果が得られることを述べる。

第5章では、第3章で提案した自己距離関数を改良した中心距離関数について述べ、さらにその中心距離関数を用いた多値画像（図形）の相似不変認識法を提案する。中心距離関数は、図形の重心点と、図形を構成する各点との距離のヒストグラムをもとにしており、自己距離関数に比べ大幅に計算量を削減している。多値画像の認識方法においては、図形をあらわす点は、点の階調値を表現する第三番目の座標と考えた、3次元空間上の点として記述される。この中心距離関数をベクトル化した中心距離ベクトルを用いた実験により、本方法が、多値画像の認識にも有用であることを述べる。

最後に、第6章では、本研究において得られた結果のまとめと今後の課題について述べる。