1. 総 論

正会員 富 永 英 義[†]

1. ま え が き

画像情報は、それを取り扱うメディアによって多様なものがあり、発展経過を見ると、それぞれ専門的分野において、その目的に最も適した専用的技術が採り入れられ駆使されてきた。

古典的なメディアとして、その経緯に即して分類すると次のようになる。

- (1) 手書き入力絵画、手紙、文書
- (2) 印刷
 - (a) 活字文書、新聞、雑誌、書籍
 - (b) 静止画印刷
 - ・図、表、網かけ写真印刷
 - ・カラー印刷、グラビア、オフセット
- (3) 白黒やカラーの写真
- (4) 映画
- (5) ファクシミリ(G I, G II, G III)

これらのメディアは、入力系と出力系が直結した物理現象を利用しており、画像として記録、保存することの目的に合った製作手段を見い出している。

ところで、画像入力という概念が生まれ、確立されたのは、コンピュータや通信系のシステムに入力する情報として取り扱うことを意図した画像を対象としているからである。

今日の新聞、雑誌をはじめとする印刷メディア、テレビやビデオ、ファクシミリなどの放送・通信メディアのいずれをとってもコンピュータ機能を使っており、入力され、蓄積、編集、伝送され、最終的な出力を得る。これらの画像情報は、基本的にデジタル化され、フレームメモリーやファイルメモリーに蓄積されることを前提として「画像入力」という概念が存在

する。

すなわち、画像入力という概念は、画像データベースなどをもつセンター機能と、情報を入出力するマルチメディアターミナルとを結ぶ通信網機能のそれぞれに対応したものである。

画像データベースシステムの全体像を見ると、コンピュータを援用して、画像および図形情報を入力あるいは作成し、蓄積し、必要に応じて検索抽出し、それを变形したり、処理を加えたり、遠隔地へ伝送するなどして利用するシステムである。

本特集では、画像入力の概念をそれぞれのメディアに応じた課題としてとらえるとともに、異なるメディアで共通し利用できる技術や異なるメディア間で情報を相互に変換したり、複合したりするために入力される情報を部分的には画像入力としてとらえている。

2. 画像データベースシステムによる入力機能

図1に画像データベースシステムの概念を示す。ここで入出力機器とは、画像情報を電気信号に変換する部分、あるいは逆に電気信号を可視化情報に変換する部分であり、静止画の場合は画像そのものを人間あるいは機械が何らかの手段によって創り出すものや、画像をいろいろな形に加工するものも含まれる。

入力は、従来カメラやスキャナが中心であったが、最近では感圧式等の描画タブレットや電子黒板、CRTとライトペン、キーボードを使ったもの等、多様化が進んでいる。超音波による地層探査データや衛星からのリモートセンシングによる雲観測データ等も広義には画像データといえるし、今後可視化した方が都合が良い情報はすべてシステムへ入力できることを前提とし、検討を進める必要がある。

また、これまではセンター側で考えられてきた種々の機能が、端末側にも備えられるようになりつつあり、技術課題も共通的なものが増えてこよう。イメージスキャナについては、撮像管や固体素子、ファクシ

[†] 早稲田大学 理工学部

"Image Input Technologies; 1. Introduction" by Hideyoshi Tominaga (Faculty of Science and Engineering, Waseda University, Tokyo)

図1 画像データベースシステムの概要

ミリ等で使われているイメージセンサは高精細化、高速化、カラー化等が進められており、今のところ寿命や安定度、コスト等に問題があるものが多いものの、ディスプレイ等を含めた必要な性能はほぼ実現できる状況にある。今後は固体化したカラーキャナの高性能化、低コスト化が課題である。描画装置はまだ紙とペンのような簡易さ、確実さや自由度を持ったものがなく、改善が望まれる他、リアルタイムで文字情報を認識したり漢字や外国語に変換する等のソフトウェアとの組合せ、あるいは既存の画像情報との組合せ等によって、総合的に機能を向上させる工夫が必要となる。

さらに広く考えると、パッケージメディア化されたメモリの移動も技術的には通信の部分を除けば他の端末とほとんど共通なものとなる。

また、かなり高度な判断能力、思考能力を持つコンピュータが端末に導入されたり、人間の意志を言葉(音声)や簡単な描画で入力できるようになれば、マン・マシンインタフェースは知能化されたものとして整理される。

3. 画像の認識合成技術とマルチメディア入力

画像通信端末に対する画像認識・合成技術の適用には2つの狙いがある。第一は、画像情報のディジタル化、キャラクタ化により画像処理を容易にすることであり、第二は、大幅な情報圧縮を図り、伝送経費や画像センターファイルの容量を削減することである。

画像認識のうち、活字認識はコンピュータ処理の分野で広く実用化されているが、現状では文字の種類やフォントサイズなどが限定されており、不特定者の手書き文字認識のためには、認識アルゴリズムの研究など、まだ多くの課題が残されている。

図形認識や合成技術は、CAD/CAMの分野で広く実用化されており、画像通信端末に充分適用できる安価な画像処理プロセッサや画像認識を可能とする端末の実現が手の内にある。写真などの自然画像の認識技術はインテリジェント符号化などの対象として活発な研究がなされている。

端末装置の持つ処理能力や機能は、VLSI技術の発展、プログラム開発の進歩により、認識や合成技術の大半が端末装置に適用できるようになった。

3.1 動画像

動画像における認識と処理の分野は大きく分けて2つになる。ひとつは、動画像を処理して画質改善、高能率伝送、ファイリングするもので、他のひとつは画像における動きを手がかりに物体認識、パターン認識等を行うものである。後者の分野は、静止画処理では限界のある物体認識の発展のブレイクスルーとなると考えられる。前者については、HDTV、放送テレビに代表される標準テレビおよびテレビ会議用の符号化の3つに大別できる。すなわち、

- (1) HDTV 符号化
- (2) 放送テレビに代表される標準テレビ方式の符号化
- (3) テレビ会議用符号化

である。

3.2 静止画

- (1) 静止画符号化方式

静止画は、動画像と異なりリアルタイム処理の制約が少ないので、より高度なコンピュータ処理が実施できる。そのため静止画特有の符号化アルゴリズムが考えられる。静止画符号化技術としてはカラー自然画を主体とする方式やファクシミリを中心とした2値画を主体とする方式がある。いずれにしても、イメージプロセッサやFCD(Facsimile Codec Processor)等の専用LSIの開発が進み、アルゴリズム開発テーマはほ

図 2 入力画像とその処理

図 3 静止画端末入出力装置の代表例

ほ出揃った感がある。

(2) 静止画像処理

パターン情報処理の応用分野は多岐にわたり、それぞれの適用分野に合うアルゴリズムが開発されつつある。今後は、工業応用分野や事務設計分野が大きな伸びを示すと予測される。また図 2 に示す画像パターン処理の過程に対応して、それぞれの技術課題が存在する。

静止画像の入出力装置の機能的分類をすれば次のようになる。これを複合的なシステムとして概念を示せば図 3 のようになる。実用的な将来のシステムは、これらの機能が統合されたいわゆるマルチメディアターミナルとなる。

(a) 文字入力系

(a.1) 光学的文字認識による入力：これは原理的にはファクシミリスキャナの入力を用いて文字切り出しを行い、文字認識を行う。

(a.2) オンライン手書き文字認識：これはディジタイザ等の Telewriting と同じ入力システムを用い描画の手順を認識して符号化する。

(b) イメージ入力系

イメージセンサにより白黒で 64 階調のイメージデータに変換する(例：指紋・顔写真など)。イメージスキャナとしては CCD が一般的で、密着型とレンズ系による縮小型とがある。特殊用途にはラインスポットスキャナやドラムスキャナなどがある。

(c) 図形入力系

(c.1) デジタイザ入力：これは座標入力装置によって各点の座標をマニュアル入力する。精度は高く、安定しているが、労働力集約型で時間コストがかかる。

(c.2) 自動入力：これは既存図面のイメージデータから有意点を自動抽出する。アルゴリズム原理と検出精度が課題である。

(d) 静止画入力系

(d.1) ビデオカメラ入力：これは動画像信号（30フレーム/秒）の1フレームをフリーズする。

(d.2) カラーイメージスキャナ：これは、200～400 dpiで、色分解方式には光源切替え方式とフィルタ方式とがある。

4. サンプルングと分解能の課題

対象とする画像の原理的な性質はアナログ情報である。これをデジタル処理するためには、サンプルングした値をA/D変換するプロセスが必要である。

サンプルングには指の動きなど1次元的なもの、描画像のように曲線上のもの、写真など2次元的な広がりのあるもの、さらに動画像のように2次元の密度を時間軸まで拡張した3次元的な広がりに対する密度を考えてサンプルング間隔を決定する必要があるものがある。

入力すべき情報から原情報にいかにかに忠実に再現するかを目安として、入力系と出力系の分解能の関係も重要である。しかし、一般に出力系の表示の分解能より入力系の分解能が高い必要があり、蓄積系、伝送系のシステムとして合理的なバランス条件を考慮した入力系の分解能が要求される。

2次元的な画像に対するサンプルングの問題は、HDTVにおける符号化方式やビデオ画像の高効率圧縮技術、さらに放送番組の制作過程におけるクロマキー等のデジタル化などのさまざまな技術において、基礎的な問題として扱われており、多くの研究報告がなされている。

しかし、時間軸まで考慮に入れた動画像の3次元的なサンプルングの問題は、必ずしも十分な研究がなされているわけではない。

すなわち、人間の視知覚構造を明らかにし、その構造を考慮にれてカメラから得られる映像情報をモデル化して、実際の3次元空間の動き情報・構造情報・奥行き情報を抽出するメカニズムを明らかにすることが研究課題である。

人間の視知覚構造は、ビデオカメラによって入力される映像情報のように、時間の変化に対して一定の空間解像度を持っていない。すなわち、人間の視知覚構造は注視する点に対して解像度は高く、動く対象物に対しては追従してそれを捕捉する構造で視野を得ている。つまり、認知のために物体をとらえる解像度が時間的にかつ部分的に変動している。このような人間の視知覚特性をモデル化し、画像情報を用いた物体認識に当てはめて考える研究が必要となる。

また、このとき形状情報の他に、2次元の画像情報からでは正確な検出が難しいとされている物体の動き情報を検出し、それを利用する形で奥行き情報（物体そのものの奥行きと物体間の前後関係）も同時に検出し、これらの情報を基にして3次元物体の構造を認知する。時間軸方向の画像変化に対する充分に考慮したサンプルング方法の確立が必要となる。

画像処理による3次元シーンの認識はコンピュータビジョンという名のもとに研究されている。画像の局部処理としては、エッジの抽出のための微分処理、平滑化のための積分処理などがある。3次元形状を抽出するものとして、カメラを1つしか用いないもの（単眼視）、カメラを2つ用いるものがある。これらのどちらも処理の対象としているのが、ある時点での画像（静止画）であり、時間変化をとまなう画像情報を充分に生かした方式となっていない。また、画面中の有意部分とそうでない部分とで、また、時間軸上で考えた場合にも有意部分とそうでない部分とを同一の解像度で扱っており、抽出される物体の解像度向上の課題がある。特に色情報の特徴を扱うには、まだ多くの研究課題がある。

現時点では、3次元構造認識においては、人間の視知覚特性に近い入力手法が確立しているとは言い難い。

5. 画像入力技術が直面する課題

—あとがきにかえて—

大量の既存画像データを効率的に入力するためには画像データ群と画像入力装置をどのように結びつけるかが重要な課題となる。また、入力装置の性能の向上、操作性の向上の他、蓄積に至る前の段階での処理手法について検討する必要がある。ここでは、それぞれの現時点で直面している課題を述べる。

5.1 入力装置の課題

対象画像、用途により要求される性能は異なる。まず、静止画像入力装置に対しては、次のような点で性能向上が必要となる。

入力速度：大量の画像データを効率的に入力するためには、可能な限り入力速度を速くすることが必要である。例えば、ドラムスキャナによる画像入力では、現状技術では A3 サイズ 1 枚で約 5 分の時間がかかるが、これは 1~2 桁以上高速にすることが望ましい。

解像度：一般文書に対しては、50 本/mm 程度以上、高忠実性・高精度の要求される X 線フィルム（医用）等に対しては、100 本/mm 程度以上の解像度が必要とされる（現状では、一般文書を対象としたファクシミリ入力装置で 16 本/mm、写真を対象としたドラムスキャナ、レーザビームスキャナで 40~60 本/mm である）。

また、装置として、高解像度テレビカメラに対しては、カラーで 5000 走査線以上、タブレットに対しては 50 ドット/mm 以上の高解像度が必要とされよう。

階 調：階調数として、濃淡・カラー・自然画に対しては、各画素当たり 8~9 ビット以上が必要である。また、A/D 変換における直線性の保持が重要である。

忠実度：原画像情報を忠実に入力する必要がある。なお、図面中の汚れや線のかすれ等に対しては、これらを原情報として忠実に入力するか、あるいは雑音として適宜の処理を行うか難しい問題となる。

動画像入力装置に対しても高性能化が必要であり、例えば VTR に対しては、録画状態の画質を保持して、現状の 30 フレーム/s あるいは 25 フレーム/s の再生速度の 2 倍以上の高速再生が要求されよう。また、映画用フィルムを入力する際には、現状技術のテレシネでは画質が大きくそこなわれており、これを高画質で電気信号変換する装置が必要となる。このための一案として高精細テレビ信号への変換装置が挙げられよう。

以上述べたように、入力装置の速度、解像度、階調、忠実度の点では従来技術の性能向上が必要となる。さらに効率的な画像入力を行うためには、入力に伴う手間ひまをいかに少なくするかが重要となり、操作上の自動化が必要となる。例えば、書類等に対しては、自動ページめくり、大きさの異なる文書に対しては、読み取り領域の自動調整が必要となる。

また、写真、フィルム等の原画像装填時に手間のかかるドラムスキャナ、VTR（1 インチ）では、写真、フィルム等の自動装填が必要とされる。

これら入力時の自動化に加えて、各種画像データに適した形で入力装置の階調、分光特性、照明むら等を自動的に補正・校正することが必要となろう。

5.2 デジタル処理の課題

（1）前処理の高速化

入力後のデータに対しては、データの切り分け、すなわち文字、写真、図面の自動領域分割が必要となる。さらに、原データの内容を忠実に表現することも必要であるが、劣化画像の品質向上を目的とした、雑音除去、復元、強調等の各種デジタル処理を高速に行うことが重要となる。

（2）符号化における課題（画像データとして取扱う場合）

大量のデータを蓄積する際には、高能率符号化技術によるデータ圧縮が必要となる。この場合、忠実性が重視されるデータに対しては、可逆符号化が、多少の画質劣化が許容されるデータには不可逆符号化が行われる。これら符号化技術に関しては、アルゴリズムの標準化および高速処理のための LSI 化が必要となる。

（3）画像データの記号情報/意味情報への変換

画像データが含む本質的な情報のみを取り出し、これを記号情報や意味情報などの抽象化された情報に変換することにより、画像データを簡潔に表現して蓄積のための記号容量を減らしたり、また、画像の内容を反映した検索手法を開発したりすることが可能となる。具体的には、パターン認識技術、画像からの特徴抽出・スケッチ化、画像の記述手法などが該当する。現在、開発が進められている技術としては、文字認識、図面の自動読取りなどがある。

5.3 画像の蓄積

蓄積装置の大容量化、高密度化、操作性の向上などのハード的面で改善が必要となる。また、ソフト的面として、データ管理方法、データ構造、データフォーマットなどは、入力・処理手法および検索などと密接な関連を持つものであり、画像データベースシステムの全体的な立場から、注意深い設計が要求される。個々の画像データに対しては、他の画像データと区別するために、識別子、付属情報、特徴抽出画像（例えばスケッチ画）などの情報を付与する必要がある。画像入力時に、これらの情報をどのようにデータベース中に入力するかも重要な課題である。

とみなが ひでよし
富永 英義 昭和 37 年、早稲田大学理工学部通信学科卒業。39 年、同大学院修士課程修了。同年、NTT 電気通信研究所に入所。46 年、早稲田大学助教授。51 年、同教授となり、現在に至る。53~54 年、英国エセックス大学客員教授。主に ISDN、テレマティクスサービス等の研究に従事。工学博士。正会員。