

パターン認識・マルチメディア技術

—今後の技術と新たな産業創出—

9

村上篤道・石川 泰●三菱電機(株)

はじめに

パターン認識の研究開発は古い歴史を有し、文字認識、画像認識、音声認識などで実用化例も多くみられるようになった。本稿では、パターン認識技術を中心にマルチメディア技術のこれまでの発展を概観し、実用化例とその背景となるコンピューティング環境の変化の考察から、今後の方向性を論ずる。さらには、新たな産業創出の方向性を検討する。

パターン認識技術の歴史

■パターン認識マルチメディア技術とは

パターン認識は、画像、文字、音声などを人間が理解する過程への興味、その工学的利用への期待から、長い研究の歴史を有する技術分野である。70年代に本格化し、その困難性から現在でも情報処理の主要課題であり、将来への期待も依然として大きい。パターン認識は、一般に、観測されるパターンから、その利用上規定された意味(概念)を抽出・分類する技術である。したがって、パターン認識は、1) 扱うパターン、2) 扱う意味あるいは分類、3) 識別のための表現、特徴ベクトル抽出、データの構造化、4) 意味、分類カテゴリを表すモデル、5) モデルと入力パターンを照合・識別するための評価尺度と識別方式、を課題として論じる必要がある。

またマルチメディア技術は映像と音声など、複数のパターンを扱う技術であり、伝送・蓄積・活用など多様な応用を目的とする広い技術領域である。それぞれに個別の技術課題もあるが、効率的な表現方法、人間の知覚や理解のレベルでの評価など、パターン認識の技術課題が主要な課題をカバーする。したがって、パターン認識を論ずることで、マルチメディア技術も考察可能である。

そこで、本稿では、パターン認識を論ずることとし、次節ではパターン認識技術のこれまでの歴史を概観する。個別の技術の詳細は優れた解説^{1)~3)}に譲り、ここでは、今後の技術を論じるため、大局的な視点で、変化と動向の把握を試みることにする。

■パターン認識アルゴリズムの研究

パターン認識の初期、70年代から80年代の研究では、対象とするパターンは、枠内に書かれた文字、スイッチを押した後に発声された単語音声、対象物だけが撮影された映像など、認識を目的として加工、あるいは、作成されたパターンに限定することが多かった。扱う意味も、簡単な言語情報、あるいは製品の欠陥の有無など、あいまい性が少ない少数の意味カテゴリを対象とした。研究の主なテーマは、規定した意味と対応性の高い特徴量の抽出方法(特徴ベクトルの縮退)、距離尺度、また、ヒューリスティクスに基づく識別方法にあった。

90年代になると、大規模なデータが扱えるようになるとともに高速な演算処理が可能となった。これにより、パターン認識アルゴリズムに大きな変化が生じた。パターンに意味を付与した学習データを用い、数理モデルの識別性能を最大化する学習を行う方法、すなわち統計的な手法が基本的な方法となった。これにより、認識性能が向上するとともに、複雑なパターン、意味が扱えるようになった。対象物を含む画像や、連続音声などからの対象パターン抽出も、主要な研究課題のひとつとなった。

初期のヒューリスティックな方法では、特徴量はなるべく意味との対応がよいベクトル量に縮退することが重要であったが、統計的学習手法を用いる場合は、適切なモデルを仮定すれば、扱う特徴量はなるべく豊富である必要がある。意味の文脈依存性を解消するためのサイドインフォメーションを含めた豊富な情報量の活用が行われるようになった。

統計的な手法では、一般に、モデルの学習のためのデータが大量であればあるほど、よりよい結果が得られるため、学習データの構築方法、意味情報の付与方法の研

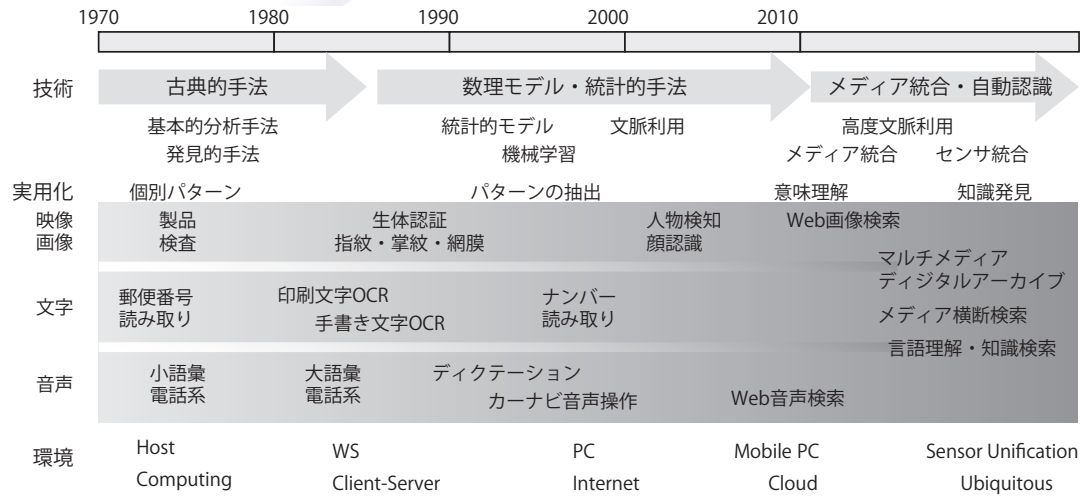


図-1 パターン認識の研究と実用化、その歴史と今後

究も精力的に行われるようになった(図-1)。

■実用化の動向

前述のように、初期のパターン認識では、入力のために作成あるいは事前処理されたパターンを対象とし、あいまい性のない少数の意味を扱うものであった。郵便番号の認識、製品や農作物の形状の検査システム、数字だけを対象とする電話系音声認識などが主な実用化成功例である。これらは、コンピュータがデータを扱う際の人手による入力の代替手段を目的とするものであった。産業上の成功の鍵は、技術的な制約の上で、省力化による利益が大きい対象を選択することにあった。

統計的な手法が用いられるようになると、人間のパターン処理の代替ではなく、より複雑で人間では扱いにくいパターンの利用や、自然界の映像や音から、対象パターン自体の抽出が可能となり、その実用化が進展した。前者の事例としては指紋、掌紋、網膜などのパターンを用いる個人認証があり、後者の事例としては、走行車輛のナンバー読み取り、デジタルカメラの顔検出、さらには、音声認識を用いた用件抽出によるコールセンタの回線振り分けシステムなどが挙げられる。

文脈の利用、さらに複雑な対象・意味を扱う実用化例も増えている。カーナビにおける音声認識は、対話の枠組みの中で、文脈情報を規定することで実用化したと考えられる。さらに、最近の音声による Web 情報検索は、Web 上に存在する莫大な言語資源を、統計言語モデルとして音声理解におけるサイドインフォメーションとして活用しているのみならず、目的とする検索自体が、Web 情報を音声の認識の意味的制約として利用したと解釈でき、今後のパターン認識技術を考える上で興味深い。

同様に、Web 上の画像検索は、画像というパターンから得られる特徴量ではなく、同一のページ内のテキス

ト、あるいは、画像に作成者が付与したアノテーション情報のみを用いて認識したと解釈できる。技術視点では、「パターン認識」の実用化成功事例というには抵抗を感じる読者も多いかもしれないが、今後の実用化を考察する上では重要な事例であろう。

パターン認識の今後の展開

■パターン認識とコンピューティングの進化

前章に述べたように、技術視点では、統計手法、大量データの活用という大きな発展があり、実用的な成功はそれに対応し

- 1) 明確な入力用データから明確な少数のカテゴリを識別する、人間のパターン認識の代替、省力化
- 2) 統計的な手法、高速演算で可能となった人間が扱うことが困難なパターンの認識、対象パターンの抽出
- 3) 文脈やインターネット上の大量な情報の活用と時代を追って大別することができる

これは、規定された目的のために設計されたデータ処理系、ネットワークコンピューティング、さらにはクラウドコンピューティングに対応付けて考えることが可能である(図-2)。このことは、今後の真の意味のユビキタスコンピューティングに、次の時代のパターン認識の方法論、実用化成功の鍵があることを示唆する。

■パターン認識アルゴリズムの深化

パターン認識を人間の認識過程の実現と考えると、現状においても、技術課題は山積している。統計的数理モデルとその学習手法については、継続して、モデリング、学習、探索手法をさらに深化させる必要がある。利用可

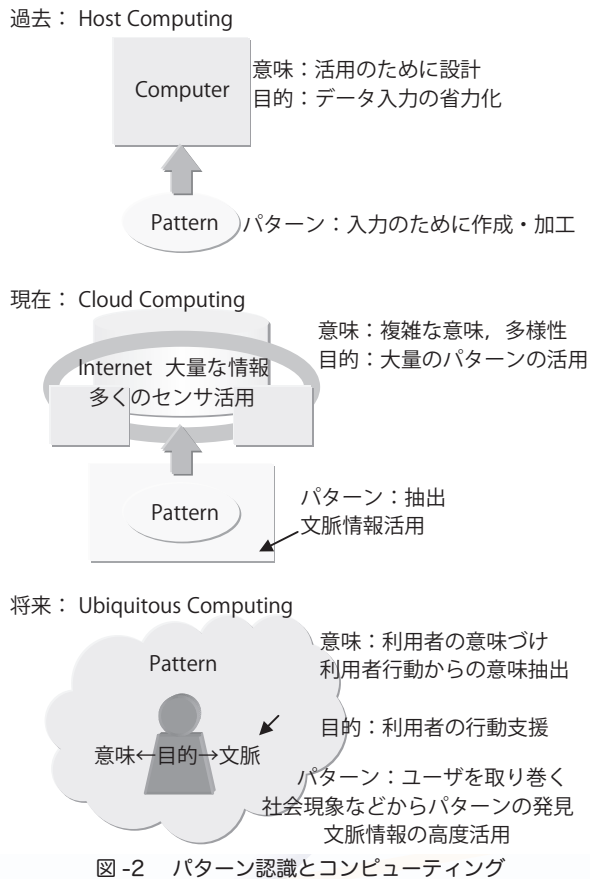


図-2 パターン認識とコンピューティング

能な情報量は今後さらに大量になり、センサ情報などを
含め、パターンの文脈として活用できる情報量もますます
増大する。有効なモデリング、効率的な探索、適応化、
適応学習などに進歩を期待したい。

また、統計手法で問題となる、学習データの取得方法、
より根源的な問題である意味の表現についての研究の重
要度はさらに増すだろう。意味を言語情報として与える
上では、そのあいまい性、文脈依存性をどのように解消
するかが重要である。オントロジーなどの意味知識の表
現と活用、Web上の大量の言語資源から言語知識を構
築する方法など、自然言語研究との連携のもと、意味モ
デルの研究がパターン認識の深化をもたらすだろう。

人間の行動・動作認識、医療画像の認識、マルチメ
ディアコンテンツの関連付け、検索など、高い目標をか
かげ、技術的なチャレンジを地道に続ける必要がある⁴⁾。
従来の手法に加え、認知科学、脳科学の進展が、物理モ
デルや心理尺度などに新たな方法論を提供する可能性も
あり、学際的取り組みが必要である。

■実用化・産業創出の新たな展開

一方、産業上の成功を考える上では、再度、1)扱うパ
ターン、2)扱う意味あるいは分類、を今後の真のユビキ
タスコンピューティング環境で考察する必要がある。人

間ができるパターン認識の実現は、ロボットなどで重要
ではあるが、情報産業としての価値は見出しにくい。

今後のコンピューティングで扱うことができる、社会
現象、経済活動、自然現象など、複雑で多様な意味が内
在するデータを対象とし、パターン認識技術でその意味
を抽出・発見できたとき、大きな産業的価値を創造でき
る。ユーザの生活から得られるパターンの中から意味、
行動目的や価値基準を抽出、その意味に適合する情報を
莫大なネットワーク情報から抽出、支援できて初めてユ
ビキタスコンピューティングは、価値を持つだろう。パ
ターン認識こそが新たなコンピューティングのキー技術
と言っても過言ではないだろう。

おわりに

パターン認識の基盤技術を進歩させるには、解明され
ない人間のパターン認識などを対象に地道なチャレンジ
が必要である。学際領域であり、研究者が相互に意見交
換ができるコミュニティとして学会の意義も、産学連携
の重要性も高まるだろう。一方で、その技術を、社会現
象などの複雑なデータから意味を抽出する技術として捕
らえたとき、新たな実用化展開と産業創出の可能性が生
じる。情報を、事前にその意味や価値を規定できるもの
とすると、Webの情報量の爆発を危惧しなくてはなら
ないが、多くの意味と価値が内在するパターン認識の
対象として考えたときは、さらに多くの情報が蓄積利用
できることに可能性と期待を感じるのである。

参考文献

- 1) 日浦，他：パターン認識・メディア理解 15 年の進歩，信学会誌，Vol.92, No.8, pp.647-655 (2008)。
- 2) 柳井：一般物体認識の現状と今後，情処論文誌，Vol. 48, No. (SIG_16 (CVIM_19)), pp.1-24 (2007)。
- 3) 古井：音声認識の動向 [I]：話ことば認識，信学会誌，Vol.89, No.8, pp.746-751 (2006)。
- 4) 馬場口，他：今なぜグランドチャレンジか，信学会誌，Vol.92, No.8, pp. 640-642 (2008)。

(平成 21 年 12 月 26 日受付)

村上篤道 (正会員)

Murakami.Tokumichi@eb.MitsubishiElectric.co.jp

1971 年東北大学・通信工学科卒。同年、三菱電機(株)入社。情報、通信、マルチメディア技術の研究・開発に従事。ITU、ISO/MPEG 等の国際標準策定に貢献。現在、開発本部役員技監、博士(情報科学)。IEEE 論文賞、米国 R&D100 賞、映像情報メディア学会功績賞、電子情報通信学会業績賞等を受賞。IEEE Fellow、映像情報メディア学会フェロー、本会フェロー & 副会長 & 評議員、日本工学アカデミー会員。

石川 泰 (正会員)

Ishikawa.Yasushi@eb.MitsubishiElectric.co.jp

1982 年東京工業大学・物理情報工学専攻修了。同年三菱電機(株)入社。音声処理、HMI の研究開発に従事。工学博士。