

特集

画像認識技術の 実用化への取り組み

- 1 実社会での利用が広がる文字認識技術
- 2 半導体産業を支える画像応用検査・計測技術の現状と展望
- 3 安心安全を守るバイOMETRICS技術
- 4 安全性と利便性を両立する静脈認証技術
- 5 健康を守るマンモグラフィ診断支援
- 6 車の智能化のための画像認識技術の現状と今後
- 7 人を見る画像認識技術
- 8 デジタル機器向けオブジェクト認識技術の動向

0

編集にあたって

佐藤 敦¹ 大田 友一²

¹NEC ²筑波大学

本特集の目的

実世界の事象を把握するとともに有益な情報を抽出・分析し、処理結果を実世界にフィードバックする技術は、産業全体の生産性やエネルギー効率の向上、さらには安心安全かつ快適な社会の実現のために重要である。視覚情報は、人間が外部から受ける知覚の8割以上を占めると言われており、重要かつ豊富な情報を含んでいる。この視覚情報を工学的に取り扱う画像認識技術は、目視作業という単純作業から人間を解放するだけでなく、人間の判断を支援したり、人間が取り扱えないほどの膨大な量の画像を解析したりすることで、新たな価値を提供することができる。

画像認識技術は、画像の撮影条件の変化に対していかにロバスト(頑健)であるかがきわめて重要であり、実用化に向けては、理論だけでなく認識対象に応じたさまざまな工夫が求められる。そこで、本特集では、これまで実用化されているいくつかの画像認識技術について解説し、製品開発の裏に隠されたさまざまな工夫についてもスポットを当てることで、実用に耐え得る技術に仕上げるための知恵を共有することを目的とする。

画像認識技術とは

画像には、信号の空間的变化である静止画像と、これに時間的变化を加えた動画の2種類があるが、ここでは区別せずに画像と呼ぶことにする。画

像認識技術とは、広義としての画像処理技術の1つであり、実世界の事象を画像として捉え、そこに映る認識対象の属性を記号として返す技術である。これを有効に活用することで、実世界における効率化、安心安全、あるいは快適性を高めることが可能になる。

画像認識技術は、図-1に示すような処理過程から成り立っている。観測とは、実世界の認識対象を処理可能な電気信号やデジタル信号に変換する過程であり、通常のカメラや赤外線カメラ、イメージスキャナや奥行き情報も測定できる3次元スキャナなど、認識対象に応じた画像入力装置が用いられる。前処理とは、認識対象の位置や大きさを揃えたり、あるいはノイズを低減したりする過程である。特徴抽出とは、認識対象の画像から識別に有効と思われる特徴を取り出し、特徴ベクトルと呼ばれる多次元のベクトルで表現するとともに、必要最小限の次元数に低次元化する過程である。識別とは、特徴ベクトルがどのクラスに属するかを判定する過程であり、判定されたクラス名を出力する。ここでクラス名とは、認識対象が持つ属性を記号で表したもので、目

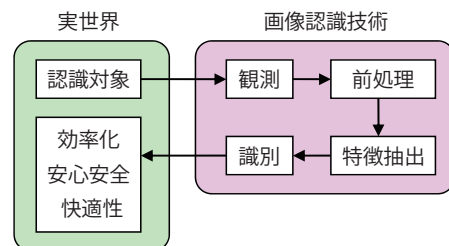


図-1 画像認識技術の処理過程

的に応じて人間が定める。

処理過程の特性

画像認識の処理過程は、後段になるほど抽象度が高くなり、数的手法との整合性が良くなる。たとえば、識別処理では、ベイズ決定理論に従って判定が行われる。最も簡単な例では、クラスごとに事前に特徴ベクトルを登録しておき、新たに入力される未知の特徴ベクトルとの類似度が最大となるクラスに判定する。近年は、機械学習の研究が進んだこともあり、大量のデータを事前に学習することで、精度の良い識別器を自動設計することが可能になってきた。80年代後半からのニューラルネットブームを経て、現在ではサポートベクトルマシンを始めとするマージン最大化に基づく手法や、カーネル関数を用いた非線形手法の有効性が示されている。

一方、処理過程の前段ほど対象依存性が高くなり、数的手法との整合性が悪くなる。認識対象ごとに、その分野の専門家が試行錯誤的にアルゴリズムを考案せざるを得ないのが実情である。認識対象に特化した処理になるため汎用性が低く、認識対象が変わるたびに多大な工数をかけてアルゴリズムを再開発することも少なくない。しかしながら、人間の持つ先験知識を導入することが可能で、認識対象の素性をよく観察し、対象固有の知見を得ることができたならば、それこそが技術の強みになるという一面もある。

画像認識の難しさと技術動向

画像には、認識対象そのものの変化のみならず、撮影条件が変化することで、その見えが大きく変わる性質がある。このような画像変動は、認識精度を下げる方向に働くので、実用にあたってはこの変動をいかに除去するかが重要な課題である。

画像変動の最も大きな要因の1つは、そもそも3次元である実世界を、画像として2次元平面に投影していることにある。つまり、認識対象の姿勢や

照明の変化によって、画像としての見えが大きく変化してしまう。この問題に対し、画像認識技術の初期段階では、実世界に存在する2次元のモノを認識対象とした。文字は紙という2次元平面に書かれたものであり、指紋は3次元物体である指を平面に押し付けることで2次元化したものである。顔については、カメラに正対してもらうことで2次元の正面顔画像と見なし、さらには照明条件を揃えるなどの強い制約を課すことによって3次元的な変動を除去し、実用化を進めてきた。現在では、認識対象に関する3次元モデルの構築や、大量の画像データを利用する方法、あるいは直接3次元情報を計測するなどの方法で、この制約を緩和する研究が進められている。

実空間を2次元の画像に投影しているもう1つの弊害として、遮蔽の問題がある。たとえば、画像からの人物検出では、腕などが本人の体に隠れて見えなくなってしまう場合がある。さらに、その人物の前に何か物体があれば、体そのものが遮蔽されて検出できない。このような情報の欠落に対してはいかんともしがたく、動画像であれば前後のフレームから情報を得て推定するか、あるいは遮蔽が生じないような位置から撮影するなどの工夫が求められる。

認識対象そのものの変化については、対象ごとに個別に対処する必要がある。手書き文字を例にとると、人によって書き癖が異なるため、同じ文字でもさまざまなバリエーションがあるのは周知の通りである。そのため、認識精度を上げるには、何らかの手段で形を揃える必要がある。手書き漢字は、縦と横のストロークで構成される文字が多いことに着目し、各ストロークの間隔を等間隔に揃える前処理を施すことで、認識精度が大きく改善されることが示されている。また、指紋の場合は、隆線の分岐点や端点を特徴点として捉えることで、指を押し付けたときの歪みなどの変形に頑健な照合を実現している。このように、認識対象に固有の変動は、認識対象を注意深く観察し、変動に頑健な前処理や特徴抽出を行うことが重要である。その一方で、近年のコンピュータ処理能力の向上に伴い、考え得るかぎりの大

量な特徴量の中から、識別に有効と思われる特徴を自動的に選択する研究も進められている。

画像認識のもう1つの問題として、認識対象の切り出しの問題がある。画像認識では、基本的に2つの画像を重ね合わせたときの類似性に基づいて判定が行われるので、位置や大きさをあらかじめ揃えておく必要がある。そのため、画像中から認識対象を正しく切り出す必要があるが、正しく切り出すためには、それが認識対象だとすでに知っている必要がある、「ニワトリと卵の問題」と同様の矛盾が生じる。そこで、たとえば文字認識では、紙面に枠を設定し、その枠内に1文字ずつ記入してもらったり、あるいは動画像の場合は、あらかじめ背景を覚えておき、背景との差分によって、前景にある認識対象を切り出したりする処理を行ってきた。近年は、コンピュータの処理能力の向上に伴い、総当たり探索が可能になってきている。すなわち、画像中のあらゆる場所において、さまざまな大きさで画像を切り出し、これを認識対象の候補と見なして処理する方法である。認識対象を回転やスケール変化に頑健な局所特徴量を持つ特徴点の集まりで記述し、登録画像の特徴点との対応を探索することで認識対象を検出する手法も提案されている。また、特徴点が含まれる頻度をヒストグラムで表現することで位置不変な特徴とし、映っているシーンを判定する研究も盛んに行われている。

このように、画像認識技術のそれぞれの処理過程で進展が見られるが、最近の特徴的な動向として、コンピュータ処理能力の向上以外に、インターネットの普及による画像データの蓄積とオープンソース化が挙げられる。これらの影響で、世界中の誰もが、共通の画像データベースに対して、複数のアルゴリズムを容易に試せるようになってきた。インターネット上の画像には、風景や人物などのスナップ写真が多いことから、最近是一般物体認識の研究も活発に行われている。一方で、課題の抽出や分析には、その目的に合った画像データベースを持つことがきわめて重要であり、独自に画像データベースを構築することが、技術の強みにもなり得る。

本特集の構成

本特集は、さまざまな認識対象について、実用に耐え得る技術にまで作り上げた8件の事例を紹介する記事で構成した。それぞれの事例において、認識性能を低下させる要因に対して、どのような工夫でロバスト性を高めているのか、技術的な共通点はあるのか、などの視点で読んでいただくと、興味深いのではないと思われる。

- (1) 人間が知識の蓄積や伝達のために発明した文字を機械に読ませる、文字認識技術を紹介する。文字は紙という2次元平面に書かれ、しかも人間が読みやすいようにデザインされた創作物であるため、古くから画像認識研究の対象とされてきた。認識率を上げるために、文字枠に1文字ずつ記入させるなど書き手に制約を課してきたが、現在は自由に書かれた文字列も認識できるようになってきている。物流や金融などへの応用に加え、看板や道路標識の認識といった情景中の文字の読み取りなど、文字認識技術に対するニーズは多様化している。
- (2) 目視作業に代わる技術として実用化されている、半導体LSIの画像検査技術を紹介する。微小欠陥に対する検出感度の向上、配線パターンの寸法に関する計測精度の向上という基本性能だけでなく、誤検出や誤計測を低減するロバスト性や、検査アルゴリズムに関するパラメータ設定の容易さも求められる。配線パターンの微細化に伴い、増大する画像データ量を高速に処理しなければならず、総合的な技術開発が必要とされる。
- (3) 生体情報を用いて本人を特定するバイオメトリクス認証技術の中で、私たちの安心安全を守るために重要な指紋認証と顔認証を紹介する。指紋は事件現場に残る生体情報として長年にわたり研究が進められており、近年はノイズの多い低品質な残留指紋の自動照合が可能になってきた。顔は3次元物体のため撮影条件の影響を受けやすいが、正面顔に限れば出入国管理などで実用化されるほど技術開発が進んでいる。日本の指紋・顔認証技

術は、世界でもトップレベルであることが、米国立標準技術研究所が主催するベンチマークで示されている。

- (4) 体内の情報を使うことで、安全かつ確実に本人であることを認証する静脈認証技術を紹介する。指や手のひらをセンサにかざすだけで認証が行える利便性と、非接触であることの受容性から、すでに銀行等で実用化されている。体内にある静脈を可視化するセンサの開発とその小型化により、静脈認証の市場開拓を日本が牽引している。
- (5) 近年実用化が進んだマンモグラフィの診断支援システムを中心に、画像認識を応用した診断支援技術について紹介する。医用画像特有の問題として、大量の臨床画像が集めにくいことや、医師でなければ良悪性が判断できないことなどが挙げられる。これらの問題をクリアし、微小な病変までの高精度な検出を実現することで、読影医の見落としや負担を軽減している。
- (6) 自動車の安全性や利便性を向上させる、車の知能化のための画像認識技術について紹介する。走行環境の画像認識では、精度向上だけでなく実時間で処理することも求められる。ロバスト性を高めるためには、電波レーダなどの複数のセンサを融合する必要もある。運転支援に必要な画像認識技術と課題について概観し、実例を交えながら課題解決のための方向性や工夫を述べる。
- (7) 人の動きや振る舞いを見る画像認識技術を、要素技術と応用技術の両面から紹介する。近年、屋内外に監視カメラが多く設置されるようになった

が、主に事後に現況を把握するための記録用として使われている。今後は、何か異常なことが起こったときに即時的に警報する、あるいは事前に警報を出すような機能も求められる。これを実現するための基礎技術と、公共分野や家電製品などでの実用化について、事例を交えて解説する。

- (8) 最後に、デジタルカメラや携帯電話など、デジタル機器への搭載が進んでいるオブジェクト認識技術について紹介する。本技術の搭載によって、これまでのフィルムカメラではできなかった、新たな価値を提供できるようになった。これを可能にした技術背景について解説し、オブジェクト認識技術の今後の展開と課題について述べる。

画像認識技術の今後への期待

認識は人間が有する重要な機能の1つであり、これを工学的に代替することは、人間を単純な作業から解放するだけでなく、人間には不可能な大量データの解析などによって、新たな価値を提供できる。本特集でも分かるように、認識対象に特化した技術開発を行うとともに、何らかの制約を加えることで、画像認識技術の実用化を進めてきた。制約を緩めるとまだ人間の能力には敵わないのが実情であるが、今後も継続して画像認識のロバスト性を高める研究が進められ、やがては人間を上回る認識能力を持つ機械が実現するであろう。このチャレンジングな課題に、ぜひ多くの方が挑戦していただきたい。

(平成22年10月8日)