

マシンビジョンの基礎知識

最終回

ヒューマンビジョンについて

FAビジョン(株)
丸地三郎

マシンビジョンの開発のポイントと難しさを記してきたが、それは、なぜ、そんなに難しいのだろうか？ どの程度難しいのだろうか？

これは、比較の基準となる人の目（ヒューマンビジョン）の性能の良さに依存しているように思える。

人の目は、どんな程度のもので、優秀なのか？ そうでないのか？ 一度、考えてみる価値がある。

マシンビジョンとヒューマンビジョン

1. ビジョンの対比

マシンビジョンが実際に使われている分野は多

い。それらの使用例では、精確・高速・連続・高信頼性、ミクロンの微小サイズ対応の面で、人の目に比べて優れている。しかし、これらのアプリケーションは、技術者が適切な画像を取れるようにセットし、個々の認識方法、計測方法を決め、プログラムした結果、可能になったもの。どちらかと言うと、個々のアプリケーションを開発した技術者の努力とプログラムが賞賛されるべき問題。一般的にマシンビジョンとは言い直すことには違和感がある。

2. ビジョンとCAD/CAMの対比

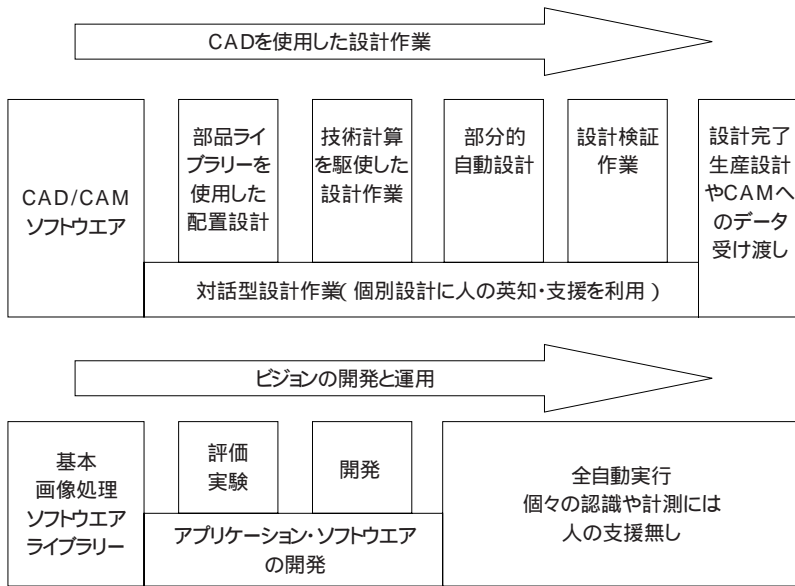
同じように形状をハンドリングするシステムに、CAD/CAMシステムがある。このシステムが、世の中に紹介された当初、ある人たちは、自動設計システムと訳して、CADで自動的に設計ができるものと夢想した。現実には、コンピュータがAssist＝援助する設計システムで、人が考え、設計することを、コンピュータが助けるのがCADであった。決して、人がこう作りたいと思ったことを自動的に設計してくれるシステムではない。

しかし、対話型の操作に加えて、計算した結果で形状を自動的に決定したり、半導体：LSIの設計では、ファンクション設計からゲートレベルに自動展開することが、可能になった。対話操作と自動設計を組み合わせたシステムになっている。しかし、設計の基本コンセプトからの展開は“人＝設計者”が行っている。

マシンビジョンも、CAD/CAMと同じように、コンピュータで、形状認識と計測を行うものであるが、CAD/CAMのようにスマートに支援できてい

第1表 マシンビジョンとヒューマンビジョンの処理機能の比較

	マシンビジョン	ヒューマンビジョン
形状認識	モデルと比較 精密な比較が可能	記憶・概念との比較と頭脳内での立体シュミレーション比較、複雑な認識が可能。
計測	得意 デジタル処理・数値化可能	不得手 アナログ比定・数値化に難
外観検査	プログラムされたもの可 プログラムに時間がかかるが、できれば、連続・高速処理可	言葉とサンプルで教示可 すぐに適応できるが、連続・継続処理に難あり。高速性も問題あり
文字認識	可	優秀
バーコード	可	不可
2Dコード	可	不可
道路・標識・障害物の認識	不可(部分的に可)	可
連続単純作業	得意 高速	不得意 中速度
立体認識	困難	可
多種・複合認識	困難	可
広い視野の認識	困難	可
総評	設定された条件の中では、プログラムさえ組めれば、人の目より正確・高速・連続で優位点あり できることに制約があるが、可能なものは優位性あり、ただし、変化に対応できない。	正確性・連続・高速に関しては得意でないが、適応力・判断力・柔軟に優れ、業務の修得や変更に対応力あり。 何でも良くでき、変化への対応力があるが、連続・高速性に難あり。



第1図

る状態にない。ビジョンの認識実行自体に、会話型操作というCAD/CAMで有効だった方法が使えない。そこで、人が、コンピュータをAssist=支援して、自動認識できるように、プログラムを作っている。プログラムの実行時には、人の支援はもう無い。

「CAD/CAMはコンピュータが、人を支援して設計する。マシンビジョンは、人が、コンピュータを支援して認識・計測させる。」

3. ビジョン“開発”に対話型の手法

マシンビジョンの、このプログラムの開発の作業に、会話型の操作が取り入れられてきた。認識などの命令を指定すると、即座に画面上に結果が表示され、コンピュータが正しく認識できたかどうか確認できる。順次命令を加えて行き、結果を画面で確認し、プログラムを作成する。そのプログラムをそのまま自動認識=マシンビジョンとする方法で開発が大変やり易くなった。人が確認し、支援した方法でそのままビジョンが認識することができるようになったので。

マシンビジョンでは、人の目にできることをコンピュータにやらせようと

している訳なので、“人の目”と同様のことを期待してしまう。そこで、どうしても、人の目との比較をされる。

ビジョンの基本機能と人の成り立ち

ビジョンの基本的な機能に関して、マシンビジョンと“人の目”を比較して行くとその差異が歴然としてくる。

神経細胞は、細胞体に樹状突起が多数出ており、長く伸びるニューロンにつながって、その先端はシナプスと呼ばれる信号交換の機能を持つ機関（I/Oピン）で外の神経細胞と接続されている。電気信号で高速に処理が行われている。

脳細胞1個は、平均10,000ものI/Oピンを持つ巨大なLSIで、そのLSIが脳には140億個あるという。

人の目は、脳を構成する一部分で、視細胞を含む網膜自体が10層になっておりその各層で処理を受け持っている。その処理結果が視神経から脳に伝わり、認識などの処理が行われる。

数年前（2001年冬）にある学会の講演で聞いた人間の形状認識の方法の話が印象に残っている。人が

第2図 マシンビジョンとヒューマンビジョンの基本性能比較

		マシンビジョン	ヒューマンビジョン
カメラと素子	構成するもの	カメラとコンピュータ	目と脳と身体全体
	受光素子 素子数とカラー	CCD(CMOS) 640×480画素=0.3M 1300×1000=1.3M カラー(同上)	視細胞(桿体・錐体)、網膜 濃淡用桿(杆)体=120M個 カラー用錐体の数=6M個 多階調・ダイナミックレンジ10の6乗 網膜で10層の処理
	1画素の性能 一次処理	256階調 なし(転送のみ)	
コンピュータ	処理系	コンピュータ	網膜・大脳・小脳
	処理セル I/Oピン数 / 個	LSI 16 ~ 256	神経細胞 数百 ~ 10数万ピ(シナプス) 平均1万ピ(シナプス)
	セル数	数百	大脳:140億個 小脳:1,400億個
認識の問題	メモリ	G(1,000M)バイト	上記(各シナプスも記憶能力あり)
	参照モデル モデル数	2次平面モデル 数十 ~ 数百	立体モデル 数万単位?
	立体認識 形状認識	複数カメラで形成 モデル比較 形状計測と比較 形状ごとのプログラミング	常時リアルタイム処理 記憶・概念との比較と頭脳内 シミュレーション比較
	空間・座標の認識 計測	形状ごとのプログラミング プログラムによる座標系 デジタル処理	成長と共に訓練・修得 空間認識・座標系知覚能力保有 アナログ処理

コップを認識するときは、手の運動を司る神経系まで反応しているという。握れるサイズか、運動神経まで使ってコップの認識をしているとのこと。

この話を聞いていて、チンパンジーは、どうしているのかなと思いをめぐらした。森の中で、木の枝を認識するときは、握れるか？ 安全な太さか？ という運動神経を駆使して、形状認識していると思うと、妙に納得した記憶がある。

さて、マシンビジョンとヒューマンビジョン“人の目”を比較してみる。受光部分に関して、1画素と視細胞1つを比べて良いのかも疑問だが、敢えて比べてみると、百万単位と億単位で2桁違う。

人の目の場合、分解能が高い部分が、視覚の中心部0.5度の領域に集中している。個々の視細胞の能力とさらに高い認識能力を持たせたい部分の密度は、人工物がまだ到達しえない能力と言える。

ほかの動物と比べて、人の目の優れている点は、立体認識とカラー認識といわれる。地上を走る草食動に比べ、猿・類人猿の目は、2つの目の視野の重なりが大きい。視野を重ねることで立体視の能力を高めている。生活基盤である高い樹林の中で、枝から枝へ移動するための能力は、葉の緑と枝の色を見極めるカラー識別と立体視能力に依存していたことに異論は生じない。人類は、500年前にアフリカ大陸で発生した地殻と気候変化に伴い、住み慣れた高い樹林から地上に降りることで生まれたという。高い樹林で培った視力・認識能力が、人の目をこれだけ高い能力にしているとのこと。

1980年代の考古学、DNA遺伝学などの成果が人の成り立ちを明瞭にしてくれたことに感謝する。

ヒューマンビジョンの仕組み

1. 視野をコントロールするシステム

視野を決めるには、足から首までの多数の関節と筋肉（ロボットでいう軸）を使って頭の位置と向きを動かし、視野をきめる。さらに、眼球を動かす筋肉を使い、適切な視野を確保している。これは、究極の検査ロボットの例として示したものの機能。外観検査ではワークの位置と照明と視野の角度を変えながら、キズ等の検査をすることがある。この機能が人には備わっていることになる。

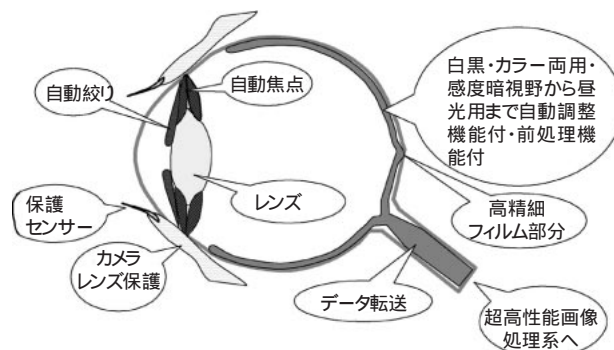
2. 眼球

- まぶた：レンズカバー
- 水晶体：自動焦点機能付レンズ

- 涙腺：自動レンズ洗浄装置
- 虹彩：自動絞り機能
- 視細胞：錐体＝高性能カラーフィルム、桿（杆）体＝白黒フィルム（明るさにより自動ゲイン切り替え機能付）、中心部分は特に高解像力
- 網膜（アマクリン細胞・双極細胞・水平細胞・神経節細胞など）ノイズ処理、エッジ強調、フィルター処理などを実施。

3. 脳細胞

視細胞から送られた情報の処理・認識が行われている。但し、その方法には、筆者に納得のゆく説明が見つかっていない。そこで、筆者の想像する方法を示す。そんな見方もある程度に、気軽に読んでください。



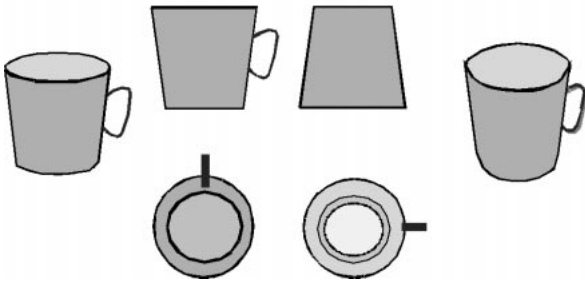
第3図 眼球の構造

人の目の認識の仕方

まず、例として、柄のついたカップの認識を考えてみよう。こんなカップの認識は誰でもできる簡単なことと思う。（実は、マシンビジョンには、まだ困難なことである。）



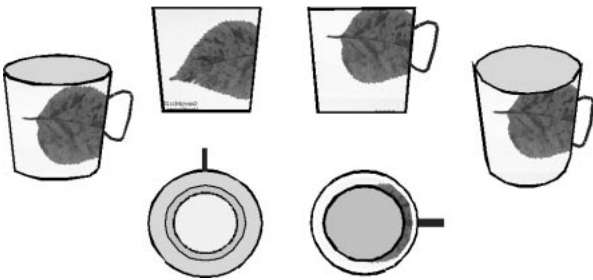
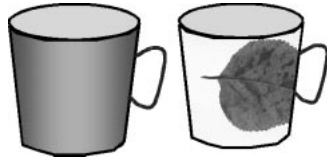
- さて、このカップは、斜め上方から見ている。
- 真横から見ると、台形である。伏せておいてあると下辺が広く。また、柄は見える向きと見えない向きがある。
- 真上から見ると、内側の面が見える丸になり、伏せておいてあると、外側の面が見える丸に見える。底の形が丸く見え、ふくらみがあると2重円に見える。
- 斜めから見る場合も、もう少し高めから見ると楕円が円に近くなる。



第3図 カップの形・位置向きの変化

現実の世界では3次元形状。立体の形状があたり前で、平面は極特殊な例になる。この立体形状を認識するには、人はどう認識しているのだろうか？

- まず、形を認識するにあたって、画面上（視野の中の）連続している部分とそうでない部分の境目、エッジ（端点、端面）を探す。そのエッジをつなぎ、ひと塊りの部分を仮に取り出す。
- 図柄が入るとエッジの量が増え、複雑になり過ぎるので、これ以上、論議しないが、図柄も考えるとエッジを先に抽出しているように想像される。
- 左右の目を使って立体認識を行い、各エッジの遠近を算出。



第4図 カップの形・位置向きの変化

- 近似する形状を発生させる可能性のある形状をデータベース上（形状モデルの記憶）で検索。その形状の位置、向き、サイズをシミュレーションし、該当する形状になるか比較し、適切な形状を選択し、カップと判断し、その位置、向き、サイズまで認識する。
- 近似する形状を探すには、状況判断も入ると思わ

れる。台所にカップがあることは、常識ですぐに認識するが、街の家屋の屋根の上に巨大なカップ状のものがあった場合は、それが、巨大な広告用のカップで、屋根の上にあること認識するには、時間を要する。意外なものは、形状データベースのサーチ時間もかかる。

- 形状に関しては、モジュール化して、記憶しているかも知れない。また、形状の記憶には、名称＝言語も役割を果たしていると思われる。ちなみに、雪（空から降る雪）の形状は、東京人には、数種類、多くて7～8種類の名前しかないが、エスキモーには降る雪は数十種類に言い分けられているとのこと。言語は形の認識にまで立ち入っていると想像している。
- コップの位置と向きとの認識には、人の別の感覚器の情報、平衡感覚も使われている、
- 色情報、図柄、掴めるか否かといった情報も使用しているはず。

マシンビジョンの認識の仕方

- 人の目の認識と同じようなことができるか、マシンビジョンで認識するとしたらどうなるか？検討してみる。
- モデルデータと比較。モデルデータは、2次元・固定サイズのため、比較するには、位置・向き・傾き・サイズなど全ての状態のモデルを（多数）予め登録する必要がある。
- 色や明るさの異なるカップでは、別のモデル群が必要。（多数×多数）
- モデル比較以外の方法としては、画面上でエッジ抽出する。そのエッジ形状から、形状認識を行うことになるが、立体視と平面画像のままの2通りを示す。
- 2つのカメラで捉えた画像から、立体視し、形状を認定するには、カメラ①で捉えた画像からエッジ抽出形状とカメラ②で捉えた画面から同じように抽出した形状を比較し、十分似ているか判定。エッジ部分の同一性と差異を調べる、同一ポイントと比定できた2つの位置を調べ、位置の差から、距離を3角測量の方法で算出し、遠近を全ポイントで測定。
- 立体または平面画像の測定ポイントを直線・円弧、楕円（二次曲線）に構成し直す。（不定点列を、直線、曲線に比定することも、やっかいな仕事：CADのように理想形状からの比定すること

は、方法があるが、カケや歪みのある画像の点列から比定することは、困難なこと)

- 比定された直線、円弧、曲線を持つ形状をデータベースから検索する。この検索手段はあるのだろうか？ または、極限られた数の中から、カップの形状を選択したとして、形状の比較・比定を行う。このためには、データベースの形状の位置・角度・向き・サイズを変化させ、エッジ形状から抽出した形状データと合致度を調べ、合致したときに形状認識できたことになる。
- この全体の処理を、リアルタイムにやっているのがヒューマンビジョンで、マシンビジョンはこれだけの処理をリアルタイムにはできそうもない。汎用性のある、アルゴリズムは、まだまだ展望も見えていない。

終わりに

先日、米国のフォーチュンにビジョンのトップメーカーの創設者が取り上げられた。その人の言を記す。「人の目にできることは、基本的には全てマシンビジョンにできる。」「但し、実現には1000年かかるだろう。それまで、私は生きられない。」

回りくどい言い方をしているが、マシンビジョンはヒューマンビジョンにかなわないとの意。

マシンビジョンの仕事（認識の仕方を機械に教える仕事）は、なくなる。新しい概念や発明が期待される分野である。

極めて優秀な人の目と比べ、「もうちょっとどうにかならぬのかね」などと言われながら、マシンビジョン屋さんは、仕事を続けることになりそうです。

ビジョンに興味を持って読んでいただけた方に感謝します。このチャレンジできる分野を、一緒に開拓して行けたら幸いです。

筆者紹介

丸地三郎

FAビジョン(株)

〒337-0043 埼玉県さいたま市見沼区中川68-1

TEL : 048-682-4192 FAX : 048-682-4191

E-mail : smaruchi@fa-vision.com

URL : <http://www.fa-vision.com>

日本工業出版(株)インターネットホームページのお知らせ

当社では、インターネットのホームページを運営しております。

<http://www.nikko-pb.co.jp/>

月刊技術誌に加え更に広く情報受発進を行い、明日の技術に貢献してまいりたいと存じますので、是非一度アクセスしていただきますようお願い申し上げます。

また、合わせてe-mailによる、当社刊行物へのご意見ご要望もお待ちしております。

e-mail : info@nikko-pb.co.jp (本社)

e-mail : info-n@nikko-pb.co.jp (日本橋事務所)

日本工業出版(株)

インターネット係

TEL : 03-3944-1181 FAX : 03-3944-6826