

高機能 3次元視覚システム VVV

独立行政法人 産業技術総合研究所
知能システム研究部門 3次元視覚システム研究グループ長 兼
関西産学官連携センター 知的機能連携研究体長
富田文明
(画像処理、コンピュータビジョン、知能ロボット等の研究に従事)

ロボットの自律行動と眼

知能ロボットの研究、特に画像による視覚認識の研究に携わって20年以上になる。研究をはじめた当初は非常に高価なコンピュータを使って、今とは比べ物にならないレベルでの画像処理を行う程度のものであった。当時産業用のロボットが使われはじめ、さらにVALといった言語によってプログラミング制御が可能な第2世代といわれるロボットが登場しはじめていた。日本では産業分野でのロボットが急激に普及しロボット大国といわれるまでになっていく端緒であった。当時は産業界のみならず、ロボットが家庭やオフィスといった生活分野まであっという間に普及していくと考えられていた。

今日、21世紀になって、個人がアミューズメントの対象として購入して遊べる様々なロボットまで登場している。展示会レベルではあるが、昨年の愛地球博をはじめとして、様々なロボットが展示され好評を博している。あと少しで、我々が夢に描いたロボットの時代が到来するよう感じられる。

ところが実際に動かしている舞台裏を見ると、大半のロボットシステムは、人間がリモコンで制御していたり、あらかじめ決められた手順通りに動いていたたりするだけで、ロボット自身は何も考えていない。よくて、床に張られたマークを検出して動作を補正する程度である。このままでは、二足歩行ロボットに家事を手伝ってもらうどころか、犬型ロボットを番犬にすることすら難しい。

ロボットが人間の制御なしに自分で判断する「自律行動」の実現には多くの課題があることが知られている。その中でもロボットの眼一視覚機能の実現が重要と考えている。ロボットによっては、眼以外の様々なセンサが取り付けられていて、一定の柔軟性を持っているものもある。しかし人間が利用する大半(80%以上)の情報が視覚情報と言われていることからわかるように、ロボットが人間と同じように、眼で自分の置かれた環境を把握したり、対象物を見分けて、それがどこにあるか把握したりできなければ、人間の役に立つ自律行動の実現は不可能と考えてよい。

従来の画像処理システム

当研究グループへの見学者から「画像処理は、すでに工場では広く使われていて完成した技術ではないのですか?」と質問を受けることがある。産業界での利用を主なターゲットにした、画像処理やロボットに関連した多くの展示会でも、多くの産業用の画像処理システムが展示されている。例えば、半導体や回路基板の位置決め装置や検査装置として、産業界では以前から大活躍している。

しかし、注意深く見てみると成功例は少品種を大量生産する分野が圧倒的に多い。この原因は、こうして実用化

されたシステムが、カメラ1台(単眼視)で対象を2次元平面上での、明るさや色の集合として扱う「統計的解析」にその原理が偏っているためである。

「統計的解析」とは、例えば画像上のある部分が赤い色なら良品で、そこが緑色なら異常と判断する方法とか、明るさで特定の部分を切り取って、その中心を計算し、右に1mmずれているので位置を補正するといった方法を考えてもらうとわかりやすい。もちろん、実際には、この統計量を得るための画像処理にも様々なテクニックがあり、良否の判別や学習のためには、ニューラルネット、遺伝的アルゴリズム、サポートベクターマシン…と多くの手法が開発されていて、広義にはパターン認識として研究分野を形成している。

大量生産では、限られた対象物に対して、何ヶ月も同じ作業をする。そこで、このシステムを作るには、まず認識用のカメラを含めたラインを組み立ててみて、環境や対象を限定して、その中で考えられる限りのパターンの画像を収集する。これを画像処理し、必要な性能が得られるまで学習させて、これらを長期に渡って利用する。

ところが、近年は、日本の多くの工場に変種変量生産システムが圧倒的に多くなってきて、こうした画像処理システムのスキームが破綻してきている。少量生産のために学習できるほどの時間もデータも用意できない。多品種のために対象物の色や形が頻繁に変更するごとに、専用の画像処理システムやソフトウェアを準備できない。工場内での配置の変更や、隣接する環境の変化が大きく、変化のたびに性能が低下する。

さらに省力化の面では、以前から3次元立体物としての認識ができていない。また、統計的解析では不得意な、画像の変化が大きい、光沢のある部品や透明なパーツの組立てにはいまだに対応できていない。こうした理由から、むしろ従来よりも、画像処理システムが経済的にも技術的にも対応できない場合が増えている。

かつて企業で研究開発として活躍してきた企業の中堅幹部から、過去の画像処理システムの導入が如何に失敗であったかという苦言を聞かされることがある。新しい電機メーカーの工場を見学してみるとわかるが、画像処理装置を撤去するどころか自動化のシンボルであったはずの産業用ロボットまで、人間にとってかわられる皮肉な現象まで起きている。

2次元統計的解析から3次元構造的解析へ

一般の画像処理システムが工場で使えないのと全く同じ理由で、ロボットの自律行動に2次元画像処理は使えない。我々が共存する生活分野では、工場でのように環境や対象を制限し、可能性をすべて尽くすようなことは、さらに困難なことは明白だろう。

この問題を解決する手段の一つは、2次元の視覚をやめて3次元の視覚に切り替えることだ。人間はテレビや写真といった2次元の画像からその3次元の情報を得ることができる。しかし、これは人間に高度な脳の機構と膨大な先験知識によっているために現在の技術での実現は難しい。写真の中の餅が、果たして絵に書いた餅かどうか識別は不可能であることからわかるように、理論上に解けない問題つまり設定不良問題になっている。さらに、計測精度の面では、対象物までの正確な距離や大きさは、対象物についての事前知識がないと求めることができない。人間の眼と同じように2眼のステレオカメラを製作して、立体視をすれば解決する。環境が変化しても、3次元の形は照明条件や表面の模様や反射によらないので多くの問題が解決できる。言い換えれば、これまでの2次元画像処理が、わざわざ片目をつぶって見えにくくして、問題を難しくしてきただけでも言える。

もう一つの手段は、対象の構造を解析する手法である。画像の色や明るさといった統計量ではなく、対象物の連結性やその形状そのものを解析する方法への転換だ。例えばコップであれば、水の入る部分があって、横に取手がついていることを見つければよい。それは表面の模様や背景にはよらないし、わざわざ事例を用意して学習させる必要もない。こうすれば、コップを探したり、他のものと識別したりできる。工場での利用であれば、3次

元の形がわかる設計データを教えるだけで、対象環境から識別して認識するシステムが実現できる。

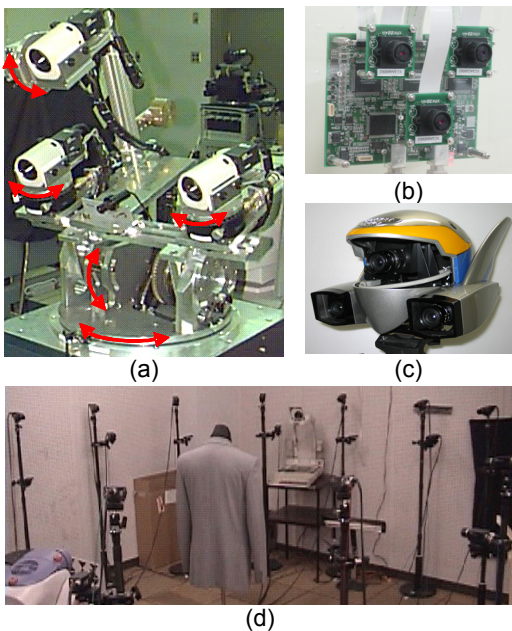
高機能 3 次元視覚システム

著者は、旧電子技術総合研究所時代から、「高機能3次元視覚システム VVV (Versatile Volumetric Vision)」を開発してきた。ステレオカメラの画像を処理する 3 次元画像処理であることと、構造的解析技術が中心になっている。

VVV の特徴は、(1) 多様な環境や目的に利用できる汎用システムであること。(2) 任意の形状の物体を対象にできること。(3) 3次元距離計測、3次元形状表現、3次元物体認識、3次元運動追跡の3次元画像処理として必要な全ての機能を持っていること。(4) 高精度に実行できることが挙げられる。これらは研究グループの研究者をはじめ、企業等からの共同研究者の協力で、一貫したシステムを形成し、機能と汎用性の点で他に類を見ないと自負している。

VVV の多種の視覚センサと利用対象

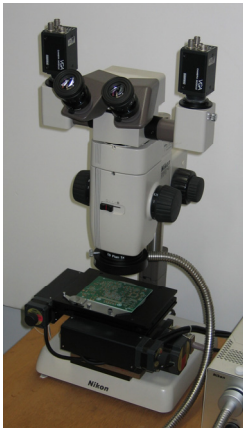
3次元視覚センサは、複数台(標準3台)のカメラを用いるステレオビジョンが基本になっている。用途に応じて、いろいろなタイプのシステムを開発している。



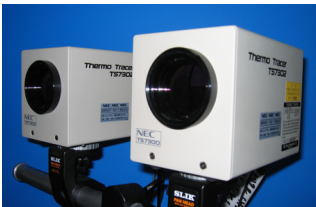
VVV を利用するステレオ視覚センサ

- (a) 注視用アクティブステレオカメラシステム
3台のカメラを独立に回転させることと、ズーム等のレンズを制御することが独立にでき、環境全体の観測と、対象物の精密に計測・認識を同時に可能にする。現在空中撮影用等の多くのバリエーションがある。
- (b) 小型3眼デジタルステレオカメラシステム
小型の PC 等で使用することを前提に、当グループおよび後述するベンチャーで独自に開発した3眼カメラシステム。
- (c) ヒューマノイド HRP-2 用ステレオカメラシステム
産総研で開発しているヒューマノイドロボットの頭部で、3台のステレオカメラが露出するようにシールドをはずした状態。環境認識から、対象物の探索・ハンドリングを行うために使用している。
- (d) 全方位マルチステレオカメラシステム
8組のステレオシステム(24台のカメラ)から構成されるマルチステレオカメラシステム。それぞれのステレオ

がクラスタによって同時に並列処理される。動きのある対象等を全周方向から同時に 3 次元計測することができる。なお、さらにドーム状に更正したシステムもある。



(a)



(b)

(a) ステレオ顕微鏡

微細な対象物であってもステレオ法による3次元認識ができる。

(b) 遠赤外線サーモステレオカメラ

ステレオ法は可視光以外に限らず、遠赤外線による温度像でも、紫外線、X線等の画像に変換できるものであれば原理的すべて VVV を利用することができる。

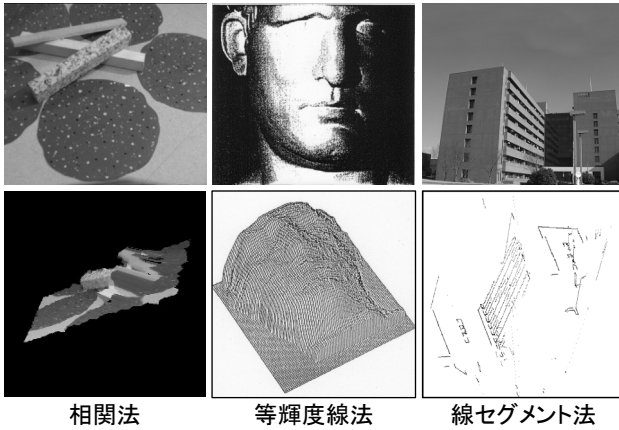
カメラキャリブレーションと高精度計測

ステレオ法による3次元で規則では、三角測量の原理を用いて距離を計測する。このためには、カメラレンズの性質による画像歪みや明るさ、カメラの位置・姿勢のパラメータ、カメラ内部の撮像面までの関係等を正確に求めて、校正(キャリブレーション)する必要がある。キャリブレーションの精度はステレオの 3 次元計測の精度ばかりではなく、認識や追跡の能力を大きく左右する。

VVV では、目的に応じた既知のパターンを利用した多種のキャリブレーション手法が利用できる。また、ステレオカメラを使用中に発生する物理的な誤差を、パターンを使わずに自己修正するセルフアジャストメント等の手法で高い精度を維持できるシステムを持っている。

3次元距離計測

ステレオ法は、各カメラ画像間で対象物の同じ位置を検出する必要がある。VVV では3種の方法が利用できる。表面に模様(テクスチャ)のある対象物に有効な相関法、陰影(シェイディング)に有効な等輝度線法、工業部品や建築物等の境界がはっきりしていて、その部分を高精度に求められる線セグメント法がある。また必要に応じて、これらを組み合わせて使うことができる。3次元距離計測は防犯等のロバストな侵入検知、自動車等での障害物の検出等に用いられている。



相関法

等輝度線法

線セグメント法

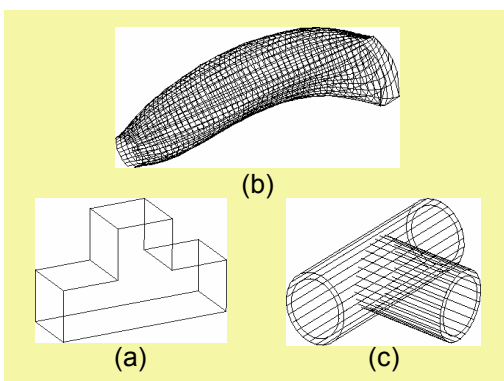
各距離計測方法と、その元画像(上段)および3次元距離結果(下段)

3次元形状計測

距離計測の段階では、対象とするものが何か?どのような構造を持っているのかはわからない。VVV では構造解析によって、それらを一般の3次元CADで広く使用されているデータフォーマットと同等の境界線表現 (Boundary-Representation) に変換される。

構造解析によって、境界線が明確な多面体は、境界線によって面に分割し、その境界線は頂点で辺に分割して表現される。明確な境界線を持たない自由曲線を持ったものは、面をネットで表現することができる。こうした物体はさらに鈹のような間接構造や対称性の情報を含めた複合体としても記述できる。

この機能は、計測結果をCADデータとして利用するモデリング(リバースエンジニアリング)にも使用されるが、次節の3次元画像認識のモデルデータとして使用することもできる。



VVV の多様な形状表現

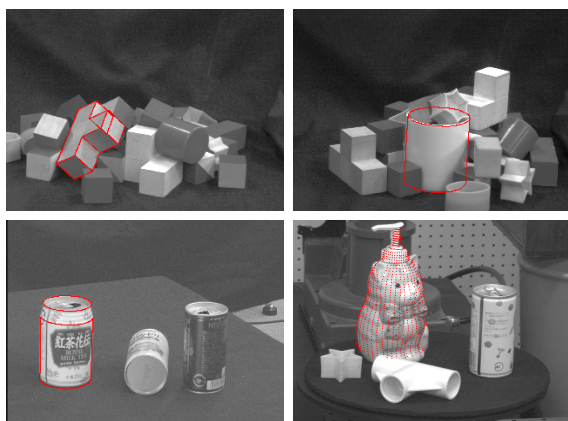
(a) 多面体、(b)自由曲面体、(c)複合体

3次元物体認識

VVV を最も特徴付ける機能が3次元物体認識である。あらかじめ3次元計測や、認識対象物のモデルデータを与えれば、その対象物を実空間で識別して、その3次元位置と姿勢を正確に求めることができる。

構造的な解析手法によっているので、基本的には事例ベースの学習等の過程を必要としない。認識対象物の

データは3次元形状計測でのデータから生成することもできるし、部品その他を製造した場合のCADデータから変換して作成することもできる。また、VVVでは認識データ作成に特化したCADシステムも持っている。構造解析による特徴としてその形状を認識の鍵としている。このために、原理的に背景がどのような状況であっても、照明が違っていても対応できる。(図左上、多面体のモデルデータを与えて認識させた場合。認識結果を赤線で重ね書きしている)。また、表面の光沢はもちろん、表面のテクスチャや色には依存しない。図(左下)のように、缶の形状データを与えれば、同じ形の缶であればメーカーや品種が違っていても認識する。また、認識自体が完全である必要はないので、一部が隠れていても認識できる(図、右上)。VVVのモデルデータとして記述できれば全て認識が可能で、多面体や自由曲面から構成される対象物も認識できる(図:右下)。

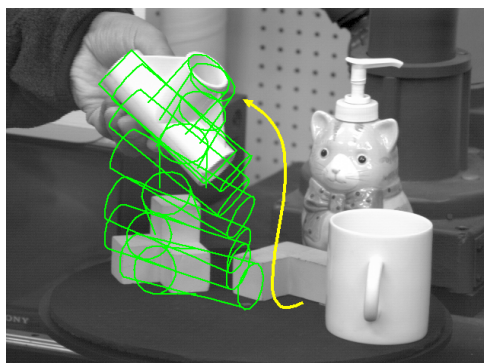


3次元物体認識例

3次元運動追跡

VVVはリアルタイムで、動いている対象物の動作を認識することができる。動いている認識対象物に対して、高速に画像との照合処理を繰り返すことによって実現している。VVVが対象物を一度認識できれば、その対象物の構造がわかっているのので、対象物のどの部分がステレオを構成する各画像中のどの位置にどのように見えるのかが予想できる。予想される近傍のみを探索することで、直前のビデオフレームからの運動を正確に求めることができる。

図のように対象物を人間が操作している様子を、リアルタイムで観測することが可能になり、ロボットへの教示システムとしての夢であった、動作を見せて学習させるシステム(Teaching-by-Showing)も実現できる。



3次元追跡結果(緑色が計測した3次元位置・姿勢)

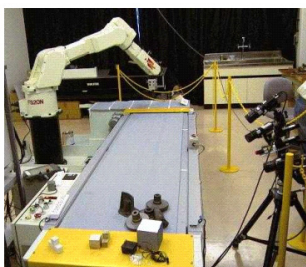
VVV 応用システム

ヒューマノイドロボットプロジェクト(HRP-2)は、例えば 2005 年に開催された愛地球博でも実演し、生活環境を想定して、家具や必要な対象物を認識するために VVV 技術が活躍している。最近のデモンストレーションでは、ビールが飲みたいとこのロボットに頼めば、家の中の障害物を自分でどかして道を開け、冷蔵庫を開けて、冷えたビールを取ってきてくれる。



©産業技術総合研究所
(手前のジュース缶を認識している HRP-2)

VVVは工場等での製造分野はもちろん、土砂中から露出した配管の検出といった土木・建設分野での応用、医療用システムでの応用等がある。



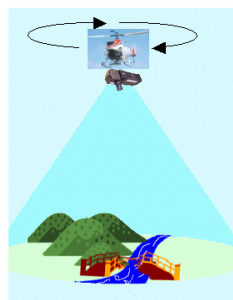
(a)



(c)



(b)



(d)

(a)物体操作システム

各種物体の各種操作の基本動作を体系的にモジュール化することにより、専門家によるプログラミングを必要としないマニピュレータ。

(b)自律移動システム

屋内、屋外の走行路を検出し、障害物を避けながら目標地点まで自律走行し、自動駐車できる車両や、パーソナルモバイルロボット。

(c)視覚障害者用視覚代行システム

視覚障害者をユーザとして、各種の視覚情報を聴覚情報に変換して、歩行支援、スポーツ支援のできるウェアラブルシステム。

(d)環境マップ生成システム

飛行体の空中撮影による広域な環境の3次元地図を自動構築し、単なる表示だけでなく、対象物の認識も可能な知的 GIS。

ベンチャー企業

電子技術総合研究所の時代から産業界での先端技術の普及を目的として、企業とも共同研究、技術指導を積極的に行ってきた。しかし企業からの全ての要望に応えられる状況にはなかった。

特に、共同研究を推進する際に企業側の人材の育成の時間がかかること、企業固有の課題まで対応できるほど、限られた研究者に余力がない等の問題があった。近年は、企業から、一定の性能を持った技術を一定の価額、一定の条件でモジュールとして導入し、企業側の人的負担は減らしたいという声も多くなってきた。

幸いなことに産総研の独立行政法人化後、技術移転のチャンネルとして、研究者自らがベンチャー企業を設立することが可能となっている。2004年にVVVの開発にタッチしてきた所外のスタッフも参加して、産総研発ベンチャー企業としてアプライド・ビジョン・システムズ(AVS)を設立することができた。

AVSの設立によって、産業界からの様々な要望に応えることが可能になっている。例えば、研究者の研究課題とするには個別すぎるが、各企業にとっては非常に重要な問題に、ベンチャーのスタッフが対応できるようにもなっている。

また、VVVのうちの要望の高い機能を個々の商品として普及する作業もすすめている。去年は回路パターン画像分析に関するソフトウェア、3次元ステレオの計測ソフトウェア等の出荷を開始した。AVSとパートナーシップを組んで商品を開発する事例も増えており、東京エレクトロニクスデバイスのステレオカメラボード(TD-BD-SCAM)とVVVを組み合わせた商品を出荷することができた。