

From a different perspective

形態覚、運動覚、色覚、明暗覚などの総称である視覚は、我々人間の最も重要な能力の一つですが、近年、コンピューターによりその幾つかを複製する技術が着目され、その研究と実用化が進んでいます。今回、TESマガジン向けに特別に寄稿をお願いしたRoberto Cipolla教授は、この分野の第一人者であり、ケンブリッジ、ペンシルベニア、東京、オックスフォード各大学の学位を持ち、現在ケンブリッジ大学情報工学の教授として、映像認識・ロボティクスのリサーチ部門を担当されています。また、2年間、東芝の研究部門で勤務される等、合計5年間に及ぶ在日経験が有る日本通で、2003年10月からは、トヨタとも、映像認識に関する共同研究を進められています。今回は、ルネッサンスの芸術家が用いた遠近法から説き起こし、最新の映像認識技術及び自動車含め多岐に渡るその活用法について、紹介します。きっと皆さんの知覚を刺激することでしょう。

原文:ケンブリッジ大学 教授 Roberto Cipolla
監修:TMEM先端技術部 Chief Technologist 柳原 弘道
Jonas Ambeck
協力:TESヨーロッパ 委員長 阪田 勝利
TMEM調達部 水嶋 正純

視覚—一人の最もパワフルな感覚の Artificial Intelligence化 ●●●●●●

視覚—見るという行為でモノ・場所を判別する感覚—はおそらく最もパワフルな感覚でしょう。我々の研究はコンピューターを使って映像を解析したり、対象を静止画像として検出したり、更にそれらの3次元像やある背景の中での空間的配置を再現したりすることで視覚が持つ幾つかの能力を再現することを目的としています。ここではケンブリッジ大学で行われているその研究活動の一端を紹介いたします。

2次元イメージから対象物の3次元イメージ情報を入手する為には様々な手法が利用されています。最も一般的な手法は線—遠近法—といわれる幾何学的な手法であり、これは15世紀以来、ルネッサ

ンスの芸術家達が背景の中での3次元配置を把握する為に使われて来た手法です。フロレンスの芸術家であるマサッチオは遠近法を正確に使用した最初の画家の一人でした。1427年に完成した彼のフレスコ画であるフロレンス、サンタマリアノベラ教会の'The Trinity'、がその良い例としてあげられます(図1)。それを見ると線—遠近法と数学との関係がよく理解できます。平行線と垂直線の延長上の消失点を背景の中で復元することで、容易に絵の投影像を2次元から3次元モデルに転換することができるのです。

(参照: <http://www-g.eng.cam.ac.uk/125/noflash/now/perspective.html>) .



図1

Roberto Cipolla教授



Roberto Cipolla氏の種々の遠近法

3D models of architectural scenes 建築物イメージの3次元モデル

2次元モデルから3次元モデルを作り出す為の線一遠近法のような数学的処理の考え方を発展させるとある建築物の写真からその3次元モデルを自動構築する事ができます。これは一つの建築物の写真を補正無しのカメラで複数枚撮影し、それを利用する手法です。二つの写真の中から同じ点、同じ線の投影がそれぞれどれに対応しているかを見つけて出すことがポイントとなります。対応する点と線からカメラの位置と向きが復元される事で、それぞれの点の3次元位置を見つける事が可能になるのです。

図2に示したJesus College 教会の3次元モデルは、二枚の写真からカメラの視点を復元し、そこから3次元VRML(仮想現実モデル言語)モデルを構築することによって、得られたものです。この技法は半自動的であり、更にバーチャルツアーでの使用を目的とした街並の構築へ適用することができます。

(<http://www.jesus.cam.ac.uk/virtualtour/hires/models/models.htm> でバーチャルツアーを楽しんでみてください)

また、この技術により、カーナビゲーションシステムに更なる先進性を与えることができます。ドライバーの位置と視点とを関連づけることにより、現実感を与えたり、あるいは3次元モデルの写真を入力することで車の位置を把握する助けとすることができます。



Curved Surfaces 曲面の認識

線一遠近法より更にチャレンジングな課題は、形、輪郭から曲面情報を見出すことです。ここでも最も一般的に使われる手法は幾何学であり、外形や形状が手がかりとなります。例えば、我々が彫刻を見る場合、その形の視覚情報を得るために回りを歩き、色々な角度からそれを見るのはよくある事です。我々は自動的にカメラでそのような情報を得るシステムを開発しました。他のシステムとは異なり、カメラやイメージの補正は必要ありません。固定カメラの前で対象物

を回転させ、その外形、輪郭を記録するだけでよいのです。'B splines' (滑らかな曲線多項式)により抽出された輪郭を表現し、それをトラックし、再構築する事で近似モデルを作り出します。これらの技法は、屋外彫刻のような不特定対象物の3次元モデルの復元にまで使用が拡大されています。

以下の例(図3)は、ハンディカメラで撮影された、ブロンズの馬の写真13枚のうちの2枚を示しています。

外形を検知、トラックし、カメラの動きを予測した上で、その右に示されたメッシュ構成の3次元形状が計算により得られます。最新の研究では対象物に穴やくぼみがある場合でも完全なモデルを構築することや反射する物体を復元することを目指しています。これはトヨタとの共同研究プロジェクトの候補になると思われれます。撮影された対象物の凡そ3次元復

製が可能ですので、例えば、色々な角度から見た車の3次元モデル復元ができます。また、この技術により、ゲーム中に使われるイメージを構成したり、ゲームプレイの3次元イメージを作ったりする事が可能である為、コンピューターゲーム業界やグラフィック業界から多くの関心が寄せられています。

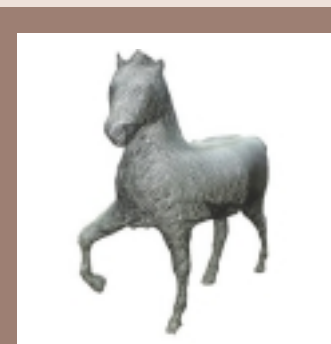
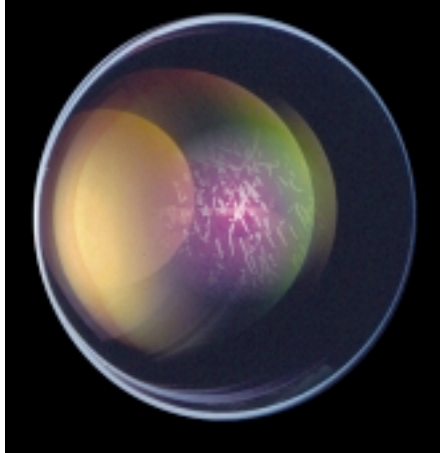


図3





我々の研究領域の3つ目は、ロボットの視覚追跡です。現在、部品を自動的にセットしたり、溶接接合部を視覚追跡するシステムに取り組んでいます。CADモデルにより、対象をロボットに指示すれば、ロボットは部品の位置・周囲環境にかかわらず、自動的にビデオカメラの映像から対象を見つけ出すことができます。

図4に示す映像は複雑な3次元構造をリアルタイムで追跡するシステムのデモンストレーション映像です。図5は、溶接追跡システムの視覚支援(視覚的サーボ)を受けたロボットの作動を示しています。システムの利用者が溶接箇所を教え込むと、この視覚支援システムは、溶接プロセスの完全自動化をも可能にします。部品のセット中、あるいは溶接中に起こるかもしれない動きを考慮しつつ、リアルタイムで自動的に溶接箇所の位置が決められます。このシステムは既にデンマーク、オーデンセの造船所で試験的に使われています。このシステムを用いれば既知の対象形状の3次元物体追跡をする事も可能となります。



Roberto Cipolla氏の種々の遠近法

Sign of the times 今望まれていること

このリアルタイム追跡システムは3次元情報である手話を認識し、それを通常の会話や、あるいは文章へと変換するシステムの構築という、新しい研究分野に拡大されてきています。更に情報処理環境と結びつけるハンドジェスチャー検知への応用も行われています。

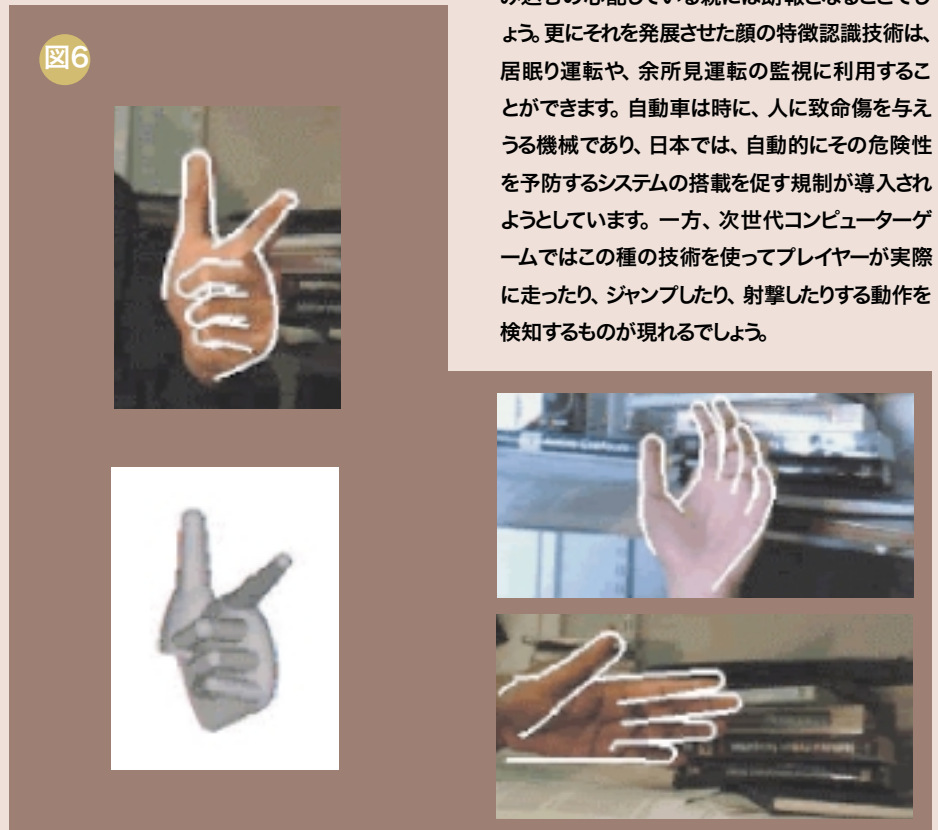
さて、コンピューターが人間の手の動きを検知し、その3次元形状を決定するのは容易に見えるかもしれませんが、しかしこれを不特定の背景内でリアルタイムで行うのはかなり困難なことです。コンピューターは示されるそれぞれの像から考える何千もの解釈を評価しなければならないからです。コンピューターの性能が向上し、比較的安価に利用できるようになった今だからこそ、このシステムが現実的なものになってきたと言えます。

また、これを実行する為には視覚的幾何学だけでなくBayesian Statistics(注:不確かな事象の連鎖が考えられる場合に、相互に起こり得る確率を求める手法)の領域にまで踏み込む必要があります。その目的は視覚による解釈を安定させること

です。これは'予測'の概念(例えば、手の形や動きについての予測)及び'可能性'の概念(イメージの色、影、輪郭のような特徴の測量)を考慮する事で、最も可能性の高い解釈(手の位置、形、動き)を決める手法です。27通りの手の動きは、一つの視点から見た場合、幾百万通りのイメージに膨れ上がります。我々のアルゴリズムを使えば、短時間でその中の最も可能性が高いイメージが決定することができるのです。(図6)

イギリス王立郵便はタイプされた文字を手話に変換できる自動ブースを既に備えていて、この研究にも関心を寄せています。彼らは、耳や言葉の不自由な顧客が'手話'を直接機械に読み取らせる事で彼らの要求を認知できるシステムを望んでいるのです。この技術の応用例としては他に、ドライバーの簡単な動きだけで自動車のラジオ、ライト、ワイパー等のスイッチを入れるようなインターフェースがあげられます。

また、この技術は車両前方を歩く歩行者の検知に使うこともできます。子供が自動車の前かがみ込むの心配している親には朗報となることでしょう。更にそれを発展させた顔の特徴認識技術は、居眠り運転や、余所見運転の監視に利用することができます。自動車は時に、人に致命傷を与える機械であり、日本では、自動的にその危険性を予防するシステムの搭載を促す規制が導入されようとしています。一方、次世代コンピューターゲームではこの種の技術を使ってプレイヤーが実際に走ったり、ジャンプしたり、射撃したりする動作を検知するものが現れるでしょう。



On the horizon 将来への展望

今まで紹介してきた新技術は既に実用化又はその途上にあるものですが、最新の取組みとして、未知の物体、即ち事前にデータ設定されていない物体を工場やラボの中で認識できるシステムの構築という究極的挑戦がなされています。例えば、任意の照明下で、リアルタイムで、ごちゃごちゃした背景の中で顔や歩行者を検出することは既に可能となってきています。もちろん、このようなシステムでは多くの習熟用データ(対象物が特定され、細区分化されているイメージ群)と習熟の為のコンピュータ処理日数が必要となります。従って、この次世代認識システムでは、数個ではなく、多くの異なる対象物を、ある背景の中の任意の場所で認識できるようになると思われますが、同時に自己学習機能を持つシステムとなることでしょう。

このようにコンピュータを用いた映像認識技術は、自動車産業を含め広範な産業に活用することができ、その発展性は無限とも言えます。是非一度その可能性に思いを馳せてみてください。

