

移動ロボットの視覚



白井良明*

Vision for Moving Robot

Key Words : Robot, Vision, Imageunderstanding

私は学生時代から、人間の眼の働きをする機械を作りたいという夢をもっていた。この夢は変わらないが、自分の置かれた環境によって具体的な夢は変化してきた。就職した電気試験所（現在の電子技術総合研究所）は通産省に属する国立研究所であったため、産業に役に立つことをやらなければならないと思った。そこでは、機械部品を認識する産業用ロボットや原子力発電所の中で働く極限作業ロボットの眼に関する研究が中心であった。

大学に移って6年目に入るが、大学では研究対象を大きく変えた。そのきっかけの一つは10年ほど前に始まった無人走行車の開発調査である。それは普通の道路に自動運転の乗用車を走らせる技術の調査と、段階的な開発計画を作ることである。第1段階として、専用道路を1台だけ走らせる実験が行われた。しかし、そこで用いられている技術をそのまま延長しても普通の道路を走らせることができない。根本的に研究しなければならない課題がいくつかあることがわかり、大学ではその課題に長期的に取り組むことができそうであった。

第1の課題は2台のテレビカメラを用いた両眼立体視によって距離情報を得ることである。

これによって、近くにある障害物などを認識できる。両眼立体視は航空写真から地図を作る場合にも応用できるため、比較的古くから研究されている。しかし、信頼性が十分でなく、計算量も多いため実用化していない。われわれは、人間に近い信頼性の高いステレオの研究を開始した。2つの画像の対応点を求めるのに、明るさや色の変化に関するできるだけ多くの情報を利用すること、画像の中のわかりやすい部分から順番に求めていくという戦術によってある程度の成果を得ている。もちろん計算量がさらに増えるため、現状では実時間処理ができない。もう少しハードウェアが進歩するのを待っている。

第2は動きの検出である。比較的遠くにあるものは相対運動によって3次元位置を知ることが多い。もちろん、視野内にある動くものを検出することも大切である。自動車から見る場合には背景も動くので、動く物体の検出は簡単でない。現在、一定範囲に一つだけ動いているものがある場合にはその動きを高速に得ることができる。この技術を用いた振れ防止ビデオカメラも実用化している。画面の多くの点のそれぞれの動きを高速に求めることはまだ困難であるが、対象を限れば研究的には実現されている。例えば図1には、ある時刻の入力画像と得られた動きベクトルが重ねて表示されている。

画像の多数の点で動きが得られただけでは十分でない。背景の領域と、背景と相対運動している領域を抽出し、動いている物体のそれぞれの3次元位置を求めなければならない。現在、



*Yoshiaki SHIRAI
1941年8月3日生
1969年東京大学大学院博士課程機械2学科修了
現在、大阪大学工学部電子制御機械工学科、教授、工学博士、コンピュータビジョン
TEL 06-877-5111(内線4706)

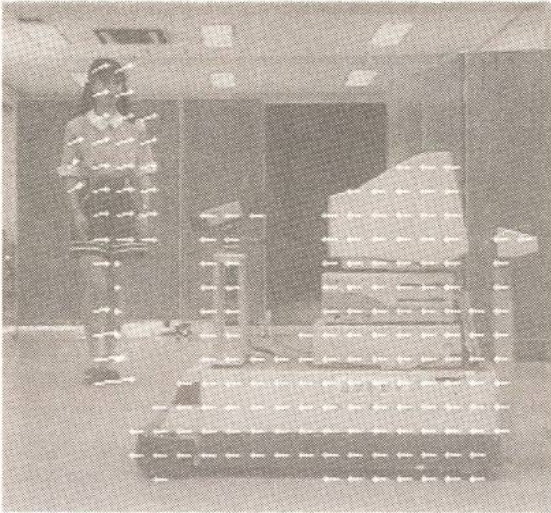
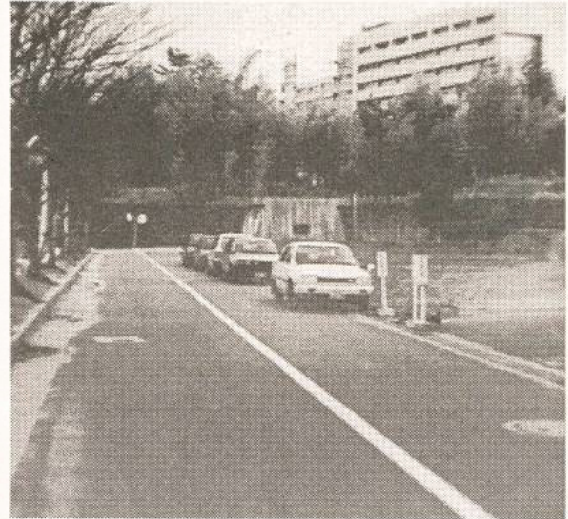


図1 動きの検出

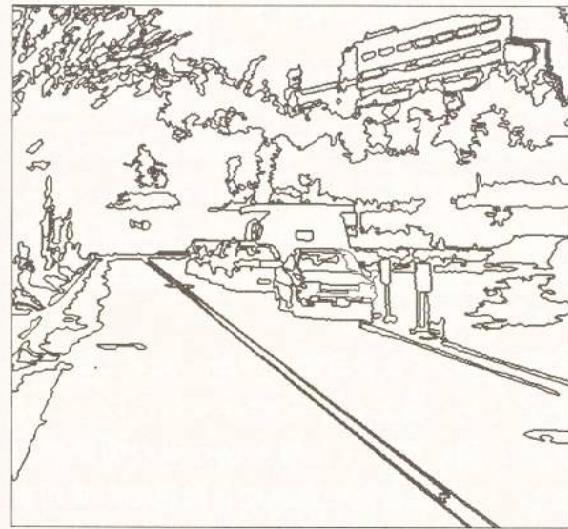
多くの動画処理の研究者がこの問題に取り組んでいる。我々も含め、いずれの方法も膨大な計算を伴うため、実時間にほど遠い。人間は実時間処理をしているので、何とかなるはずである。

第3はカラー画像からほぼ一様な色をもつ領域を求めることである。カラー画像を処理するための基礎となるため、いくつかの研究室でそれぞれの方法を開発している。普通の風景では、色が明確についていない部分が多い。そこで、色が使え部分とそうでない部分に分け、使える部分では色と明るさを用いて領域に分割し、使えない部分では明るさだけを用いて分割する。さらに、色や明るさが急変するエッジを探し、エッジによっても分割する。この方法でも完全でないが、ある程度満足できる結果が得られる。基礎的な手法は、安定した結果が得られ、その結果が後の処理に利用しやすいことが重要であろう。

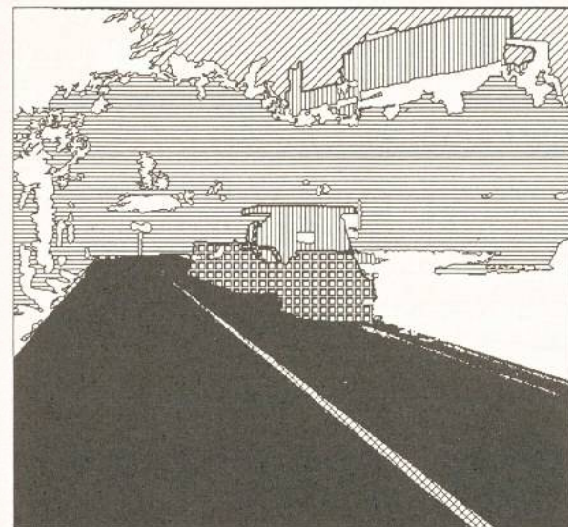
第4は得られた視覚情報を解釈することである。自動車でどこかに行く場合、道路に沿った要所要所を確かめるとともに、目的の場所を認識しなければならない。認識するためには対象に関する知識が必要である。例えば、道路は色が薄い細長い領域で、センターラインなどが書かれていて、領域内に自動車があるというような知識である。つぎに、このような知識を利用して画像を処理することが必要になる。現在、限られた範囲ではあるが、解釈ができるように



(a) 入力画像



(b) 一様な色の領域



(c) 解釈結果

図2 カラー画像の解釈例

なった。図2 (a) のカラー画像から (b) のような一様な色の領域を求め、それを解釈した結果を (c) に示す (道路, 自動車, 建物, 木, 空が異なる領域として示されている)。一つの対象に関する知識を蓄え、それを利用するプログラムを作るためにかなりの労力がかかることが問題である。言語の辞書に相当するものが利用できれば便利であろう。

第5は多数の情報の統合によってより多くの情報を得ることである。例えば、両眼視によって距離が得られる点は必ずしも多くない。距離が既知の点の密度が低ければ、どこに面があるかわかりにくい。一方、カラー画像は密な情報をもっているが、3次元情報がないため、2つの領域が離れているか同じ面上にあるかわからない。両方の情報を統合すれば、どこにどのような面があるかを知ることができる。異なるセンサ情報を統合することはセンサヒュージョンと呼ばれ、いま盛んに研究されている。動き情報と立体視情報の統合もおもしろいのである。さらに、動きながら両眼視でカラー画像を入力

し、すべてを統合すれば、信頼の高い解釈ができるであろう。

以上は無人行車にヒントを得た課題である。昨年発表されたミシガン大学が行ったデルファイ調査によれば、人と同じ能力の無人運転の実現は50年以上先であるとされている。しかし、そのための研究を途中でいろいろ応用できる。例えば、画像内にある動きや色の一様な領域が求められれば、映像の編集が楽になる。指定した人を別の風景に置いたり、その部分だけを詳しく伝送することができる。また、屋内走行車に応用することはより簡単である。現在は、すべての機能をもたせることはできないので、それぞれ焦点を絞った研究がなされている。我々も、廊下の自動走行、地図の自動作成、走行計画の自動作成などの研究を行っている。

幸いなことに、ここで述べたような夢をもつ人は少なくない。様々な方法が完成して実時間処理のシステムができてほしい。それらを組み合わせられて夢が現実に近づくことを楽しみにしている。

