

ステレオ視覚を持つ移動ロボットの目標到達行動におけるオクルージョン回避行動の獲得

○中村 恭之 浅田 稔 細田 耕
大阪大学工学部

Acquiring Occlusion Avoidance Behavior for Stereo-vision-based Mobile Robot

○Takayuki NAKAMURA Minoru ASADA Koh HOSODA
Osaka University

1 はじめに

人間は行動するために必要な情報の多くを視覚により得ているが、逆に「見る」ためにもさまざまな動作を行っている。例えば、視点の位置、視線方向を変える事によって対象を極力正面で観測しようとするオクルージョンの回避行動によって、得られる情報の質を向上させていると考えられる。

本研究では、このようなオクルージョン回避行動を学習により獲得する事を目的とする。ここでは、視覚センサーとしてステレオカメラを用い、左右画像間の非対応領域の情報を基に、強化学習を用いて、目的行動を獲得する。一般的に強化学習の際に、状態空間内に「隠れ状態」が存在すると、目的の行動が学習により獲得できないが、ここでは、状態遷移確率を推定し、これと実際に行動した時の状態遷移の履歴から、「隠れ状態」を発見し¹⁾、その状態を新たに状態空間に加え再学習することにより、目的行動を獲得する。本研究の特徴は、障害物に関する記述を陽に求めることなく、視野内の目標物体のオクルージョンの回避行動を獲得することで、カメラがロボット本体に固定されているので、結果的に障害物回避行動が実現できるという点である。提案した手法の有効性を、シミュレーションによって検証する。講演会当日には、実機での実験結果について示す予定である。

2 タスクと仮定

ステレオ視覚を有する移動ロボットが目標物体に到達する過程で、目標物体の一部が視野内で隠された時、それを回避(オクルージョン回避行動)して到達するタスク(Fig.1(a)参照)について考える。

ここでは、静止環境内に静止目標物体とそれを隠す静止障害物が存在し、初期状態においては静止障害物は完全に見えていると仮定する。ロボットは自身の幾何学的パラメータや動的特性などに関する知識はもっていない。また、2つの独立のモータで駆動される独立2動輪操舵方式を採用しており、左右のモータに対して、 $-\tau_{max} \leq \tau \leq \tau_{max}$ の間で、 n 段階のトルクを発生でき、これらの組合せにより、 n^2 通りの行動を生成できる。

ロボットの機構としてはステレオカメラがロボット上に固定されており、ステレオカメラのベースラインの長さとしてロボットの幅はほぼ同程度の大きさとする。

3 ステレオ視に基づくオクルージョン回避行動の獲得

オクルージョン回避行動を獲得するためには、対象の見え方の評価基準を決める必要がある。ここでは、ステレオ視における非対応領域の有無を、その評価基準として用いる。従って、その評価基準を検出する視覚行動を実現する。

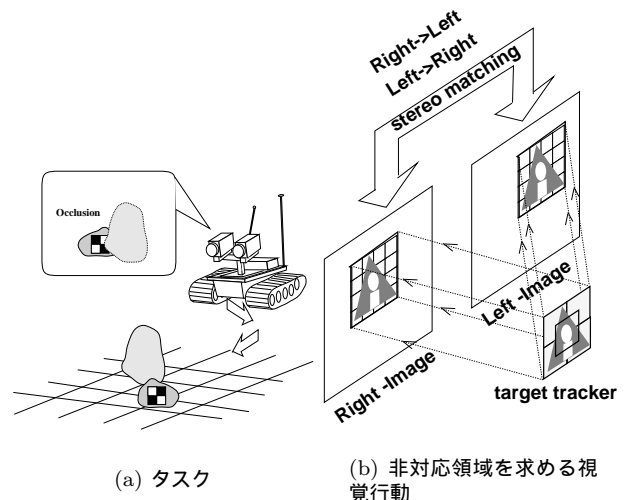


Fig.1 タスクと視覚行動

3.1 ステレオ視による非対応領域を求める視覚行動

本研究では、ロボット視覚として運動推定プロセッサ(MEP)を使用することにより、実時間(ビデオレート)で約140個のウィンド(各ウィンドは 8×8 の大きさ)を、追跡することができる²⁾。ここで、1個のウィンドを追跡する処理を1つの視覚追跡ルーチンとする。5個の視覚追跡ルーチンによりある1つの目標物体を安定して追跡するもの(T_c)を準備し、それら5個の追跡ウィンドウが占める矩形領域内に、さらに視覚追跡ルーチン $m \times m$ 個を一樣に配置させる(T_f)。Fig.1(b)参照。まず T_c で粗い対応を求め T_f でより細部の対応を求める粗密戦略により、左から右、右から左への対応点探索を行って、各時刻において画像間での非対応領域を、左右各画像において求める。対応・非対応の判定は以下のように行う。MEPによる相関演算では、テンプレートと最も相関の高い領域の位置が常に得られる。その近傍領域でのエッジ方向の相関値分布を求め、分布が単峰性になっていれば、対応していると判断し、それ以外の分布形態では、非対応と判断する。

3.2 強化学習における状態空間の構成

人間のステレオ視覚システムでは、左右画像間での非対応領域の存在によって、ある領域の距離(距離の不連続領域)を知覚しているという心理学の実験結果が示されている³⁾。従って、目標物体の領域が求められている状態で、その領域内の距離の不連続領域が求められると、目標物体領域とそれ以外の領域(障害物領域)の識別が可能になると考えられる。そこで、視差の不連続性を利用して、環境の状態を表す状態空間を以下のように構成する。

前述した視覚行動により、目標物体の左右の画像における対応領域、非対応領域を、環境の状態として知覚するこ

とができるので、左右のステレオ画像における目標物体の見え方を5段階(目標物体領域の右側, 左側, 両側がそれぞれ非対応領域, 目標物体領域の全体が対応領域, 全く見えていない)に量子化する。そして、それらの領域が左右各画像内のどのような位置に出現するかを3段階(左, 中央, 右)に量子化する。これらから、各画像において目標物体が見えている時には12通りの見え方が存在する。また、各画像において目標物体領域が見えなくなったときの状態として(左に消えた, 右に消えた)の2通りを考慮する。従って、全体として $148(12 \times 12 + 2 \times 2)$ 個の状態によって、環境の状態を知覚する。

3.3 隠れ状態の発見とオクルージョン回避行動の獲得

ロボットは、自身に搭載されたステレオカメラだけから環境の状態を識別しなければならない。しかし、センサの能力には限界があり、現在の環境の状態を完全に把握することはできず、知覚の見せかけ問題4)が発生する(環境の状態とエージェントの知覚が1対1に対応せず、“隠れ状態”が存在する)。強化学習を適用する場合、隠れ状態が存在すると、学習が困難になり、合目的行動が獲得できない。そこで、ここでは、一度学習を行い、その後に、環境状態の遷移に関するモデルとしての状態遷移確率を推定し、それを用いて隠れ状態を発見する。状態遷移確率 $P_{ij}(a)$ の推定は、以下の最尤推定によって求める。

$$P_{ij}(a) = \frac{n_{ij}^a}{\sum_{j \in S} n_{ij}^a},$$

ここで、 n_{ij}^a は、状態 i で行動 a をとったときに状態 j に遷移した回数を示す。学習中にこれらの履歴を保存し、ある程度学習を行った後に、状態遷移確率 $P_{ij}(a)$ を推定する。そして、この状態遷移確率を基にして、各状態 i について状態遷移確率の分布が双峰性になっていないかを調べ、双峰性になっている状態を隠れ状態の候補として識別する。この様にして識別された隠れ状態の候補それぞれに、2個のメモリを持たせる。そして、一度最適政策に従って行動し、隠れ状態の候補状態への状態から遷移したかを記憶する。その際、以前に記憶された状態と、異なる状態からの遷移がある場合、その候補状態を、隠れ状態として識別する。隠れ状態における動作は、識別された隠れ状態を新たに状態空間内に入れ学習を行って、新たに得られた政策に従って決定される。

4 シミュレーション結果および考察

移動ロボットは、左右各モータに対して5段階のトルクを発生でき、全部で25通りの行動がとれる。シミュレーションでは、簡単な為、目標物体をボールとし、ボールを十分に隠すことの出来る大きさの直方体を障害物とした。目標物体と障害物との配置をランダムに変えて、それらの相対的な配置に関して偏りがないようにして、目的の行動を学習した。ある程度学習を行った後に、本手法により、目的の行動を実行する時に、隠れ状態を特定し、その状態における準最適な動作を生成してタスクを達成させた。

例えば、隠れ状態として左右各画像内から目標物体が右に消えたという状態 S_{lost} を考える。これは、それに先立つ状態が、1) 障害物によって隠されている領域が目標物体領域の右側に存在する状態 S_{obs} からの遷移と、2) カメラの視野の限界によって隠されている領域が目標物体の右側に存在する状態 S_{lmt} からの遷移があり、それらを同一視することによって生じる隠れ状態になっている。

Fig.2は、隠れ状態 S_{lost} においてある行動を取った場合の状態遷移確率の分布 $P_{ij}(a)$ $a \in A$ の例を表わしている。

Fig.3は、隠れ状態が状態 S_{lost} となっている場合に本手法によってオクルージョンが回避され、目標物体に到達す

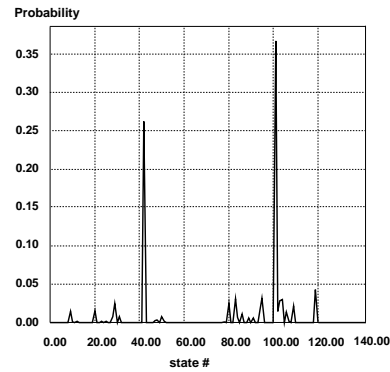


Fig.2 状態遷移確率の分布

る様子を示している。紙面の都合上、各図の上側は、ロボット上に搭載されたステレオカメラによって得られた画像の右側半分を表し、右画像, 左画像(点線で示されている)を重ねて表示している。本手法では、全く何も見えない状態が続く場合には、障害物に衝突してしまう可能性があるが、それ以外の場合では、障害物の記述を陽にすることなく、オクルージョンを回避し目標物体に到達する行動を実現することが可能である。

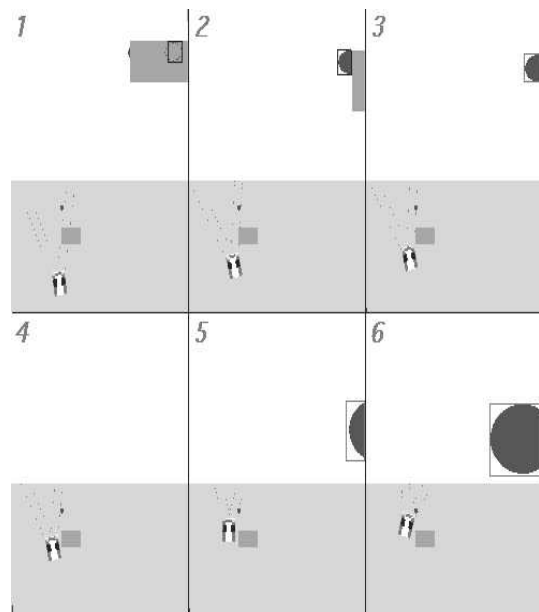


Fig.3 オクルージョン回避行動

参考文献

- 1) R.A. McCallum. "Instance-Based Utile Distinctions for Reinforcement Learning with Hidden State". In *Proc. of the 12th Int. Conf. on Machine Learning*, pp. 387-395, 1995.
- 2) H. Inoue, T. Tachikawa, and M. Inaba. "Robot vision system with a correlation chip for real-time tracking, optical flow and depth map generation". In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pp. 1621-1626, 1992.
- 3) K. Nakayama and S. Shimojo. "Da vinci stereopsis: depth and subjective occluding contours from unpaired image points". *Vision Res*, Vol. 30, No. 11, pp. 1811-1825, 1990.
- 4) S. D. Whitehead and D. H. Ballard. "Active Perception and Reinforcement Learning". In *Proc. of Workshop on Machine Learning-1990*, pp. 179-188, 1990.