短期的な記憶に基づく環境変化の検出・記録が可能な 広範囲環境記述法*

香山 健太郎^{*1}, 長嶋 功一^{*2}, 近野 敦^{*3}
稲葉 雅幸^{*4}, 井上 博允^{*4}

Panoramic-Environmental Description as Robots' Visual Short-Term Memory

Kentaro KAYAMA, Koichi NAGASHIMA, Atsushi KONNO,
Masayuki INABA and Hirochika INOUE

In this paper, we propose an environmental description for robots to memorize temporality and spatiality of wide-range environment. The description is region-based and the environment is described as a mosaic of regions. The description consists of three components: Panoramic Labelled Image that is a segmented color panoramic image, Region Database that has spatial and temporal information about each region, and Regions Relation Network that describes adjacency state between regions. The description is constantly updated in the background of other visual processing. The description enables robots to detect changes in the environment, to grasp their characteristics, and to memorize them. The description corresponds to human visual short-term memory which is called working memory or visuospatial sketch pad in cognitive psychology. Moreover, we implement the environmental description on the humanoid robot Saika and show how it expands the robot's behavior.

Key Words: Mechatronics and Robotics, Robot, Image Processing, Humanoid Robot, Environmental Description, Panoramic Image, Short-Term Memory

1. 緒 言

人間型ロボットは、人間が生活している環境で人間と協調して行動することが要求される。そして、人間型ロボットが適切な行動を選択し、その行動を速やかに開始し、そして円滑に遂行するためには、広範囲の状況を把握しておく必要がある。

例えば、既に見たことのある、そして現在の視野の中にはない物体を探索する際には、広範囲の画像を記憶しておくことで、改めて首や目を動かして探す手間が省ける。また、物体認識の際にも、隣接関係、上下関係などの空間的な文脈を考慮することが不可欠である。

また、以下の理由から、環境の時間的変化も把握しておくことが望ましい。

- 人間の行動の意図を理解するためには、その行動

の時間的変化を観察する必要がある。

- 物体の時間的変化を観察することで、物体の性質や物体間の空間的な関係(物体Aの一部または全部が物体Bにより隠されている, など)がより明らかになる。
- 自らの行動の結果として起こる環境の変化を観察することによって、その成功・失敗を判定したり、あるいは自発的に環境を理解したりできるようになる。

広範囲の環境を記憶する方法として、Zhengらはパノラマ画像を提案した⁽¹⁾。これは、ある地点の周囲の画像をつなぎ合わせて円筒形の画像としたものであり、移動ロボットの位置決めのために用いられる。また、Hermanらは複数の画像から抽出した特徴を統合することによって、段階的に市街地の鳥瞰図を作成した⁽²⁾。しかし、これらの方法は、静的な環境記述を作成するものであり、変化が起こることを考慮していない。さらに、環境記述を作成するための行動が必要であり、その間は他のタスクが実行できない。

そこで、本論文では、広範囲の環境の空間的構造とその時間的変化の理解を容易にするための広範囲環境

* 原稿受付 1998年5月1日。

^{*1} 准員, 東京大学大学院 (〒113 8656 東京都文京区本郷7-3-1)。

^{*2} 東京大学大学院。

^{*3} 正員, 東北大学大学院工学研究科 (〒980-8579 仙台市青葉区荒巻字青葉01)。

^{*4} 正員, 東京大学大学院工学系研究科。

E mail: kaya@jsk.t.u-tokyo.ac.jp

記述法を提案する。これは、領域(ラベル)をベースとした環境記述法であり、環境は領域をはりあわせたものとして記述される。そして、各領域について様々な時空間的情報が記録される。なお、記録された変化情報は、記憶容量が足りなくなると、古いもの、重要でないものから消去される。これは、認知心理学において、作業記憶、視空間的記録メモ⁽³⁾などと呼ばれるような、人間の視覚の短期的な記憶に相当するものである。

この環境記述法の構成要素とその特徴について、2章で論じる。また、3章では、この環境記述法の作成方法について述べる。

さらに、本論文では、提案する環境記述法の有効性を示すため、人間型ロボット「才華」⁽⁴⁾において、記憶に基づく行動実験を行なった。この実験について4章で述べる。

2. パノラマラベル画像を用いた環境記述法

2・1 パノラマラベル画像 ここでは、カメラによりカラー画像を得ることができ、またステレオ視で物体のおおよその三次元位置が計測可能な視覚処理システムを対象とする。

提案する環境記述法では、広範囲の環境を領域の集合として記述した画像を核として環境を記述する。本論文では、これを「パノラマラベル画像」と呼ぶことにする。パノラマラベル画像は次のような手順で作成される。

1. 各瞬間にカメラに映っている画像をカラー情報を用いて領域分割する。これを「入力ラベル画像」と呼ぶことにする。
2. 入力ラベル画像およびカメラの姿勢に関する情報をもとに広範囲の環境についてのパノラマラベル画像(図1のPanoramic Labelled Image)を更新する。
3. 1, 2を繰り返す。

原画像やエッジ画像などではなくラベル画像を用いるのは、

- 入力ラベル画像とパノラマラベル画像間の対応付けや変化の検出が行いやすい
- 色情報を用いるため、様々な処理がよりロバストになる
- それぞれの領域と物体との対応関係や領域間の関係を記録、利用しやすい
- ロボットが移動しても、画像の変換、再構成が容易である

ためである。

Fig. 1 Environmental description using Panoramic Labelled Image

2・2 環境記述の構成要素 本論文で提案する環境記述法は、図1に示すように、

- ロボットが首やカメラ、身体を動かして見ることのできる範囲全体のパノラマラベル画像(Panoramic Labelled Image)
- 各領域の情報を保持する領域データベース (Region Database)
- 領域を節点、領域間の隣接関係を枝とする領域記述ネットワーク (Regions Relation Network)

の三つの要素からなる。

パノラマラベル画像は、ロボットが首やカメラ、身体を動かして見ることのできる範囲全体の画像を連結したパノラマ画像をカラー情報によって領域分割したものとみなすことができる。これは、ロボットのカメラから得られる画像に基づいて、常に更新され続ける。また、それに従い、領域データベースや領域記述ネットワークも更新される。なお、この一連の処理は、そのためにカメラを動かす必要はなく、他の視覚処理のバックグラウンドジョブとして行えばよい。

領域データベースには、各領域の色・面積・形状・距離という空間的情報と、出現・成長・縮小・消滅の履歴という時間的情報とが含まれる。また、物体との対応が取れた領域については、その情報も保持される。

領域記述ネットワークは領域間の隣接関係をあらわすものであり、その隣接関係の性質(隣接する領域が同一物体に含まれるのか、違う物体なのかなど)がわ

Fig. 2 Procedure of making the environmental description using Panoramic Labelled Image

かった場合には、その性質が枝の情報として付加される。また、どちらの領域がロボットにとって手前に存在するかという空間的な前後関係も記述される。

2・3 本環境記述法の特徴 この広範囲環境記述法は、全画面的な前処理と記号処理との中間に位置する表現であり、領域分割までのボトムアップ処理の統合と、高次の認識結果の保存を行うことができる。また、広範囲の環境の空間的構造と時間的変化を記憶することができる。

新規性、有用性は以下の4点にまとめられる。

- パノラマラベル画像は何枚もの画像を使って作成・洗練されるため、情報の部分性、不確実性を低減でき、ロバストな処理が可能である。
- 領域分割処理を行っているため、物体の探索などの高次の処理が行いやすい。
- 領域の時間的変化を記憶しているので、領域の重なり方の推定や同じ物体に属する領域かどうかの推定が容易である。
- 能動的にカメラを動かすのではなく受動的な処理であるため、トラッキングなどの他の視覚処理と並行して行なえる。

3. 環境記述の作成方法

本章では環境記述の作成方法を述べる。処理のながれは図2のようになる。

3・1 前処理 本環境記述法は領域をベースとした記述であるため、まずカメラから入力される画像に領域分割処理を施す必要がある。これは、次のような手順で行われる。

(a) raw image (b) segmented image

Fig. 3 Segmentation

1. 画像をHSI表色系によるカラー表現に変換する。
2. 明度が50以下の部分を「黒」とみなす。
3. 彩度が15以下の部分を「白」とみなす。
4. それ以外の部分を、判別分析法を繰り返し用いて単色領域に分割する。

HSI表色系とは、色相(H:hue)、彩度(S:saturation)、明度(I:intensity)からなる表色系で、色相がいわゆる色の違い、彩度が色の鮮やかさ、明度が色の明るさを示す⁽⁵⁾。RGB空間からHSI空間への変換方法は様々なものが提案されているが、本システムでは六角錐カラーモデルを用いている。

また、黒、白領域以外の領域分割は、繰り返ししきい値処理法⁽⁶⁾を用いており、その際の動的なしきい値決定には判別分析法を用いている⁽⁷⁾。

領域分割された画像を図3に示す。なお、黒い部分は、ノイズ、あるいは微小領域と見なされた部分である。

さらに、右目画像と左目画像を比較して、対応づけがとれた領域については視差を求め、その情報を付加する。

3・2 パノラマラベル画像の更新と環境の変化の検出 このように作成したラベル画像をもとに、パノラマラベル画像が更新される。パノラマラベル画像の例を図4に示す。これは、4章の実験を行った実験室の環境である。なお、黒い部分は、前述したノイズや微小領域、あるいは未観察の部分である。

パノラマラベル画像の更新は以下に行う。

まず、カメラの姿勢に関する情報をもとに、新しく入力された画像をパノラマラベル画像のどこにはりこむべきかを決定する。なお、入力画像とパノラマラベル画像の既観察の部分とに重なる部分があれば相関演算による微調整を行なう。

次に、入力画像の各領域とパノラマラベル画像の各領域との対応づけを調べる。このとき、対応づけがとれないものについては、

Fig. 4 An example of Panoramic Labelled Image

Fig. 5 Uncertainty in segmentation: (a) appearance/disappearance of boundary, (b) difference in boundary

- 領域分割の不確実性によるもの
- 実際の環境の変化によるもの

の二種類があり、これらは区別されなければならない。

前者については、図5のように、一方で二つの領域であるものが片方では一つの領域となっているものや、両方とも二つの領域であるが、その境界線が異なっているものなどがある。このようなものについては、複数の入力画像に基づいた投票によって、どちらの表現をとるべきかを決定する。

また、環境の変化によると思われる部分では、各領域について、出現・成長・縮小・消滅の判定を行う。さらに、視差から求めた各領域のカメラからの距離をもとに、各領域の前後関係を推定する。

3・3 領域データベースにおける環境の変化履歴の記述 パノラマラベル画像の更新の際、変化していると判定された領域については、領域データベースにその変化情報が記録される。なお、その情報がある一定数を越えた場合は、古いものから消去される。これは、認知心理学において、作業記憶、視空間的記憶メモ⁽³⁾などと呼ばれるような、人間の視覚の短期的な記憶に相当するものである。

この領域データベースに記録される変化情報は次の通りである。

- その時点での重心、面積などの量的情報

Fig. 6 Description of change history in environment and spatial relation

- 領域の出現・成長・縮小・消滅などの質的情報
- どの領域からどの領域に変化したかという情報

例えば、図6のように時刻 $t-1$ から t にかけて、領域Aが消滅し、その位置に領域Aよりもロボットにとって手前にある領域Bが現れたとする。すると、領域データベースの領域Aの性質が存在から隠れに変更され、さらに、領域Aには「時刻 t に領域Bによって隠される」という履歴が、領域Bには「時刻 t に領域Aを隠す」という履歴が加えられる。また、各領域の重心や面積などの量的情報も更新される。

3・4 空間的前後関係の記述 空間的な前後関係は、環境の変化を手がかりとして判定される。すなわち、環境の変化が検出された場合には、距離情報を元にして、変化の性質が判定される。例えば、図6のような変化が起こった場合、領域Bが領域Aを隠した、および領域Aがなくなってその後方に領域Bが出現した、の二種類の解釈が可能であるが、そのどちらであるかを各領域の距離情報を用いて判定する。

そして、この判定結果が領域記述ネットワークに反映され、領域間に空間的な前後関係を含む枝が張られる。図6の場合は、領域Bの方が領域Aより手前にあるので、領域A-B間に、Bを前、Aを後ろとする性質を持つ枝が張られる。

4. 実験

4・1 本環境記述法の実装 本環境記述法は、人間型ロボット「才華」⁽⁴⁾(図7)の視覚システム上に実装

Fig. 7 Humanoid robot Saika

Fig. 8 System configuration

されている。現在のシステム構成を図8に示す。

カメラからの入力画像を領域分割する処理は、DSP TMS320C40を搭載したISHTAR-II⁽⁸⁾上で行われる。領域分割する画像は32 pixel × 32 pixelであり、一枚の画像を領域分割するのにかかる時間は1/30秒である。これに、カラー変換、左右画像のマッチングと視差計算、ワークステーションへの取り込み時間を加えると4/30秒となる。

また、パノラマラベル画像の大きさは512 pixel × 512 pixelであり、この更新および領域データベース、領域記述ネットワークの更新はSun SparcStation上で行われ、1枚の画像につき1/30秒かかる。

ただし、この一連の処理はパイプライン処理で行われるため、更新間隔は1/30秒である。

4・2 実験の概要 環境の変化履歴の記述、および空間的前後関係の記述を利用した行動として、「ロボットに一度環境を見せた後、ある物体を隠す。そして、それをロボットに見つけさせる」という実験を、才華で行なった。

以下では、まず実験の手順を述べた後、ロボットがどのような処理を行ってこの行動を実現するかを述べる。

4・3 実験の手順とロボットの処理 実験の手順は以下の通りである。

1. 机の上にいくつか物体をおいておく。
2. ロボットが机の上を見ていない瞬間に、机の上の物体の上に別の物体をかぶせて隠す。
3. ロボットに、隠した物体を発見してつかむよう命令する。

上記の3の命令をロボットが受け取ったとき、環境を記憶できないシステムでは、ロボットは首やカメラを少しずつ動かして全環境(少なくとも256点)を探索することとなる。しかし、この探索には時間がかかる上、物体が隠されて見えなくなっているため、物体を発見することができない。

一方、本システムの場合は、以下のような処理を行って最短では一点の探索で物体を発見することができる。

1. 領域データベースの情報を利用して、指定された物体に対応する候補領域を求め、その領域の近辺に視点を動かす。
2. 1で求めた領域の近辺を観察し、その領域がまだ存在するのか、あるいは何らかの変化が起こっているのかを調べる。そして、領域が存在する場合は、つかむために邪魔なものがないかどうか、変化が起こっている場合は、その領域は消滅したのか、それとも隠されたのかを判定する。
 - a. 領域が存在していて、障害物もない場合 その領域が物体に対応することを確認し、物体をつかんで終了。
 - b. 障害物が存在する場合や領域が隠された場合 その物体をつかむために邪魔になる物体、あるいはその領域を隠している物体をまずつかんで移動する。そして1に戻る。
 - c. 領域が消滅した場合 次の候補領域を探索する。それがなければ、改めて全環境を探索する。

4・4 実験結果 実験の様子を図9に示す。

まず、才華に環境を観察させ、環境記述を作成する(図9(a))。そして、才華が机の上を見ていない時に、机の上にあった物体を別の物体で隠し(b)、その上で才華にその物体(ただし、左側にあったものとする)をつかむよう命令する。

Fig. 9 Scenes from the experiment of the delayed response

すると、才華はまずその物体があった場所を見に行く。しかし、物体は見つからず、かわりにその場所に別の物体があることを発見する(c)。そこで、才華は現在そこにある物体とそこにあったはずの物体との距離関係から目標物体はその物体によって隠されていると推論し、その物体を取り除きにかかる(d)。そして、才華は目標物体を発見し(e)、それをつかんで命令を遂行する(f)。なお図9(a)の環境記述作成に30秒、図9(c)の物体があるべき場所に首を動かし環境変化を認識するまでに2秒要した。実験全体は約4分で終了した。

このように、短期記憶を利用した環境記述を用いれば、タスクが速やかに開始できる、その瞬間の視覚情報だけでは達成できないタスクが達成できる、という利点がある。

5. 結 論

本論文では、環境の時間的、および空間的な理解を容易にするための表現としてパノラマラベル画像を用いた広範囲環境記述法を提案し、これによって環境の変化履歴、および空間的な前後関係を記述できることを示した。また、それを利用した行動として、人間型ロボットにおいて隠された物体を見つけ出すという行動を実現し、短期記憶の利用によってロボットの行動の幅が広がることを示した。

今後は、変化履歴の記述を用いてロボットの行動の成功判定が行えることを利用し、外界環境の観察によ

る行動の学習を行うことを考えている。

謝 辞

本研究の一部は文部省科学研究費補助金重点領域研究「知能ロボット」(課題番号:07245105)、および日本学術振興会未来開拓推進事業(課題番号: JSPS-RFTF 96P00801)の補助を受けた。

文 献

- (1) Zheng, J.Y. and Tsuji, S., Panoramic Representation for Route Recognition by a Mobile Robot, *Int. J. Computer Vision*, **9**-1(1992), 55-76.
- (2) Herman, M. and Kanade, T., Incremental reconstruction of 3D scenes from multiple, complex-images, *Artificial Intelligence*, **30**(1986), 289-341.
- (3) Baddeley, A.D., Working Memory, *Science*, **255**(1992), 556-559.
- (4) Konno, A., Nagashima, K., Furukawa, R., Nishiwaki, K., Noda, T., Inaba, M. and Inoue, H., Development of a Humanoid Robot Saika, *Proc. IEEE/RSJ Int. Conf. IROS*, (1997), 805-810.
- (5) Smith, A.R., Color Gamut Transformation Pairs, *SIGGRAPH-ACM Computer Graphics*, **12**-3(1978), 12-19.
- (6) Ohta, Y., Kanade, T. and Sakai, T., Region segmentation by recursive thresholding, *Computer Graphics and Image Processing*, **13**-3(1980), 222-241.
- (7) Otsu, N., A threshold selection method from gray-level histograms, *IEEE Trans. Systems, Man, and Cybernetics*, **SMC**-9-1(1979), 62-66.
- (8) 白井良明・佐々木繁・塩原守人・河口博紀・間宮章, 動画像処理システム, 第1回重点領域研究「知能ロボット」シンポジウム予稿集, (1996), 1-11.