

自主机器人视觉与行为模型及避障研究

梁 冰, 洪炳, 曙 光

(哈尔滨工业大学计算机科学与技术学院, 黑龙江哈尔滨 150001)

摘 要: 主动的障碍物探测与避障行为控制一直是自主机器人的重点研究课题. 在以往的研究中, 机器人的视觉系统与电动机驱动系统往往是各自独立的构成控制体系, 人们很少关注它们之间的关系, 这样就从本质上割裂了两者之间的紧密耦合关系, 减少了获取更多信息的途径. 本文主要研究如何建立机器人视觉与电动机驱动系统关系的视觉与行为模型, 通过这种紧密的耦合关系为机器人运动提供更丰富的信息来源. 此外, 在该模型的基础上引入了强化学习, 引导机器人进行动态避障. 实验表明该方法是可靠的, 能够满足实时性要求.

关键词: 视觉与行为模型; 动态避障; 强化学习; 光流场

中图分类号: TP2. 24 **文献标识码:** A **文章编号:** 0372-2112 (2003) 12A-2197-04

Research on Vision-Action Model of Autonomous Robot and Obstacle Avoiding

LIANG Bing, HONG Bing-rong, SHU Guang

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: Active obstacle detecting and avoiding for autonomous robot are significant research fields. In the past, the vision system and motor controlling system of the robot have been developed independently. Most people ignored the relationship of these two systems, separated the close coupling, and reduced the ways to obtain more information. Here, a Vision-Action model is developed to describe the relationship between the robot vision and motor action, in which the coupling provides more information resources of the environment. In addition, reinforced learning is used for autonomous robot to avoid the dynamic obstacle based on this Vision-Action model. Our experiments in the method were carried out stably and in real time.

Key words: vision-action model; obstacle avoiding; reinforcement learning; optical flow fields

1 引言

自主机器人作为一个能够在复杂环境下自主决策的智能体应具备三个基本功能: 感知功能、思维功能以及行动功能. 在机器人上安装车载摄像机的局部视觉系统具有信号探测范围宽、获取信息完整等优点, 成为近来的主要研究方向^[1,2]. Brooks 开创性地提出了机器人分层控制的感知-行为模式^[3], 他将复杂的任务分解成简单的可以并发执行的单元, 每个单元有自己的感知器和执行器, 构成感知动作. 这一理论重点强调了每个单元的交互, 并没有将感知与执行器之间进行协调关联, 只是将感知和执行作为一个简单的反馈环来封装. 此外, 有研究人员通过 CMAC 小脑模型神经网络来对传感器和执行器进行关联协调的, 但没有采用视觉传感器^[4].

本文提出一种视觉与行为模型, 该模型将机器人视觉系统与电动机驱动系统通过图像光流联系起来, 并且在标准环境下求得它们之间的关系, 即: 对机器人不同的电动机驱动命令所引发的光流矩阵进行数学分析, 构建能够线性表示光流场的基向量空间, 最终形成了机器人视觉光流与电动机驱动行为的关系图. 这样的视觉与行为模型有以下优点: 机器人在执行任务时, 只需计算该模型中相应参数即可, 无需进行摄像机校准和环境三维重建; 利用这样的视觉与行为模型驱动, 通

过强化学习方法, 能够压缩强化学习的搜索空间, 加速机器人与环境之间的信息交互, 实时地产生躲避障碍物的动作序列. 实验表明该方法是可靠的, 可以满足实时性要求.

2 自主机器人视觉与行为模型

机器人的视觉与行为模型是将视觉行为与电动机行为关联起来. 在以往的研究中, 机器人通过传感器感知外界环境, 并产生行为作用于外界环境, 传感器与电动机驱动系统往往作为两个相对独立的部分, 很少考虑到它们之间的联系. 在本文所构建的视觉与行为模型正是处于这两个系统之间中, 将这两者紧密的联系在一起. 如图 1 所示, 我们把机器人获得外界环境信息的过程称为视觉行为; 把机器人动作系列称为电动机行为. 视觉行为包括图像光流计算、动态目标跟踪、障碍物动态检测等, 光流计算是其中的基础部分. 在视觉与行为模型中, 视觉传感信息和电动机控制信息不仅仅是从环境到传感器, 从决策系统到电动机驱动系统这样单向流动的, 电动机驱动系统也同时相应的给视觉系统提供信息, 从而影响到视觉系统的光流理解和感知性能; 此外, 该模型也为高层决策系统提供机器人本体所固有的视觉与行为规律, 辅助机器人进行高层决策 (如机器人的强化学习).

在电动机行为模块中, 将机器人左右两个轮子的电动机

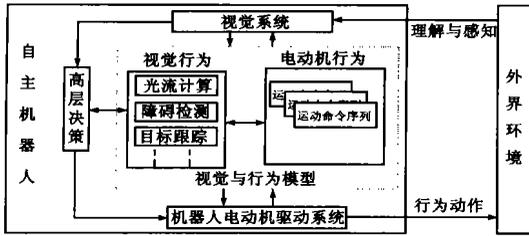


图1 视觉与行为模型图

控制命令 $r(t)$ 作 5 个级别的离散化处理,分别为快速前进 (qf);慢速前进 (sf);停止 (s);慢速后退 (sb);快速后退 (qb) 五种.这样,机器人在运动时两个轮子就有 25 种不同的动作组合,可选任一个动作 $i(v_{li}, v_{ri})$,其中 v_{li}, v_{ri} 分别代表自主机器人的左右两个驱动轮在第 i 个动作时分别接受控制命令所采用的速度,且 $v_{li}, v_{ri} \in \{qf, sf, s, sb, qb\}$.

2.1 机器人视觉行为的光流计算

图像光流场的计算是由美国学者 Horn 和 Schunck 提出的^[5].他将二维速度场与灰度相联系,导出了光流计算的基本等式.在构建视觉与行为模型过程中,我们采用区域匹配法对图像块计算光流:首先将机器人获取的图像(140 × 140 像素)按照 7 × 7 网络栅格进行区域划分,这样每个栅格区由 20 × 20 个像素构成,针对图像序列中相邻的两幅图像进行光流计算,将上一帧栅格区域的图像作为模板在当前帧图像的各个区域进行匹配计算,求最小误差绝对值和.计算过程如下:

$$D[i, j] = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} |R[k, l] - M[i+k, j+l]|, \quad 0 \leq i, j \leq 19(1)$$

其中 $R[k, l]$ 表示图像中坐标为 (k, l) 处的参考模板; $M[i, j]$ 为当前图像的匹配区域, $D[i, j]$ 为图像在该处的累计误差总和.利用该匹配过程搜索整个图像栅格区域,得到光流图.

2.2 机器人线性光流场主成份分析

要分析机器人视觉与行为的关系,首先将机器人放在理想的标准环境中(地面平坦,无障碍物),这样就排除了环境因素可能带来的干扰.图 2(a) 表示实验室平坦无障碍物的地面环境,机器人在这样的环境下选取动作 $i(qb, qb)$,则上节所述方法计算的光流图如图 2(b) 所示.

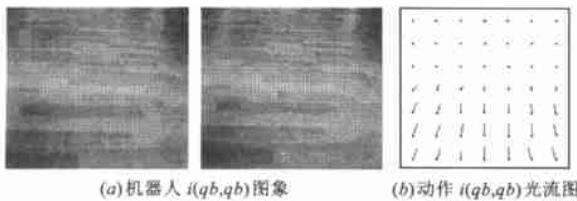


图2 机器人不同动作产生的光流图

多次计算机器人各个电动机驱动命令组合所产生的光流图,取均值后得到了平均光流场.要分析机器人电动机驱动与光流之间的关系,首先将每个动作 $i(v_{li}, v_{ri})$ 所对应光流图中的各个光流向量从左到右排列表示成一个行向量 p_i ,用每个动作 p_i 作为矩阵的行来构造矩阵 $P(25 \text{ 行}, 98 \text{ 列})$,矩阵的行代表机器人不同动作所产生的光流场,矩阵的列代表每个光流场各个光流向量(每个向量由 x, y 轴两个方向上像素偏移

量构成,共 49×2 个元素),将此矩阵 P 进行奇异值分解:

$$P_{m \times n} = U_{m \times m} S_{n \times n} E_n^T \quad (2)$$

其中 S 是矩阵 P 的奇异值为对角元素的对角矩阵, E^T 为主成份特征向量,是正交向量基. $m = 25$ 表示机器人电动机不同动作组合, $n = 49 \times 2$ 为矩阵 P 列数. U 是与列有关的正交阵.当取 K 个主成份的时候,每个电动机动作的光流场 p_i 可以用公式(2)中的方法表示为 K 个 E^T 的线性组合:

$$p_i = \sum_{k=1}^K u_{ik} s_k e_k^T \quad (3)$$

用线性关系表示 p_i 必定存在误差,但是如果误差很小,不影响实际应用的效果,则可以忽略这种偏差.为了在不增加误差的基础上尽可能地减少主成份分量的个数,下面计算主成份 p_i 的真值与由 E^T 的线性组合近似值之间的误差 $E(K)$,以确定当选取较少主成份分量时的最优 K 值:

$$E(K) = M \sum_{i=1}^m |p_i - p_i(K)| \quad (4)$$

其中: $|p_i - p_i(K)| = \sqrt{\sum_{j=1}^{n/2} (v_{ij}^j - v_{ij}^j(K))^2 + (v_{ij}^i - v_{ij}^i(K))^2}$,

式中 M 为常数系数, $v_{ij}^j, v_{ij}^j(k)$ 和 $v_{ij}^i(k)$ 分别为真实值与线性近似值在图像光流平面上的象素坐标.计算误差的约束条件为:

$$E(K) = \begin{cases} E(1) \\ E(K-1) - E(K) \end{cases} \quad (5)$$

实验中,依 25 个不同光流场构成矩阵 P 通过奇异值分解得到 25 个主成份.图 3 表示 $E(K)$ 与主成份个数 K 之间的关系,可以看出,当选取 3 个或者 3 个以上的主成份个数也不能显著降低误差,而且误差都在 2 个像素的精度范围之内,因此只要取到第二个主成份时就能较为有效的表示了.由此,将公式(4)简化展开用两个主成份分量的线性组合表示 p_i 为:

$$p_i = u_{i1} s_1 e_1^T + u_{i2} s_2 e_2^T \quad (6)$$

从机器人步进电动机系统理论角度上分析,机器人实际运动是由平移和旋转两种基本动作组合而成.而光流分析 $K = 2$ 的结果所得到的含有两个分量的基向量所对应的正是这两种运动,图 4(a) 对应的是机器人的旋转运动,图 4(b) 对应的是机器人的前后平移运动.其它的所有动作(25 个)都可以用此基向量的不同系数线性组合来加以表示.

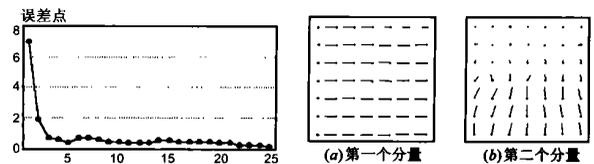


图3 主成份分量个数与误差分析

由此,已经求得两个主成份分量的线性组合表示各个光流场 p_i 的系数,这样就可以得到机器人所有电动机动作与光流系数之间的关系.定义 a_k^i 为线性组合的系数, i 为 25 个电动机动作空间, k 为两个主分量,根据公式(3)则有:

$$a_k^i = u_{ik} s_k, \text{ 其中 } i = 1, \dots, 25 \quad (7)$$

2.3 视觉与行为关系拓扑关系图

把机器人每个动作产生的光流映射到另外一个正交空

间,然后用一系列的系数进行线性组合就可以表示出每个电动机动作所引发的光流图.采用 ISODATA^[6]方法将这些系数点 a_k 进行二维空间的聚类分析,可以比较清晰地看出他们之间的分布关系,如图 5 所示.

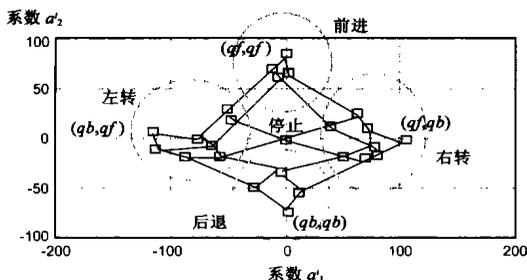


图 5 光流与电动机命令动作的关系图

图中坐标轴分别表示构成电动机动作光流的两个组合系数,聚类结果用曲线标出,具有相同的属性的系数点被聚类在一起,这些点与电动机命令一一对应.另外,这几个聚类分别代表了机器人左转向,右转向,前进,后退和停止的动作(其中停止动作单独聚成一个类).从这里可以看出光流系数与电动机驱动之间的关系.我们将每个系数点根据它与周围点距离的关系,用直线进行彼此互连,形成了机器人各种电动机动作变化的轨迹图.利用这样的拓扑关系,为以后的强化学习提供了方便.因为最初的学习是对每一个分类的代表向量进行,然后对 5 个类同时进行学习,能够压缩学习的规模.

3 基于强化学习的机器人动态避障

3.1 基于 Q-学习的自主机器人动态避障

Q-学习是强化学习中的一种重要算法,算法中 S 为环境状态空间的集合, A 为机器人针对环境所能采取的动作集合.机器人以一定的迁移概率依据当前状态而采用某种动作,如:当前状态为 s ,动作 a ,迁移到状态 s' 的概率为 $T(s', a, s)$.此时,环境给机器人报酬为 $r(s', a)$.我们的目的就是发现能够使报酬衰减总和达到最大的策略,因此策略 f 是状态空间 S 到行动集合 A 之间的一个映射^[7].

如前面所述,在标准环境下机器人的动作与其所产生的光流场 p_i 之间的关系已经得到.当机器人处于工作环境的时候,若机器人前方存在障碍物,那么在运动过程中视觉产生的光流则与标准环境下的光流场有所不同,利用这个特性能够检测动态障碍物.比较标准环境下产生的光流场 p_i 与有障碍物环境内的光流场 p_i^{obs} ,在归一化后计算均值误差把光流向量有明显差异的区域分离出来,超出误差阈值范围的被认为是障碍物潜在区.如图 6 所示,当机器人运动时前方出现了障碍物,通过光流场比较,分离出了误差值偏差较大区域(阴影区域),该区域被认为是障碍物区域,机器人应根据该区域的属性利用 Q-学习进行动作规划.图中(a)表示机器人在标准环境下采用当前电动机命令应该产生的光流,(b)表示在工作环境中实际所产生的光流,(c)为出现横向运动的障碍物图像.可以看出通过对比光流场的不同,能够检测出可能存在障碍物的区域.

当检测到前方存在障碍物时,机器人根据障碍的情况产生躲避动作序列.根据 Q-学习算法,首先对存在障碍物的环境进行状态划分:根据障碍物区域在图像中出现的面积大小可划分为三类:大、中、小;依据障碍物区域在图像中水平方向的位置分为:左、中央、右三种状态.另外,为了让机器人适应动态变化的环境,根据图像序列中障碍物在前后两帧的位置变化定义了九种不同的状态:水平方向表现为障碍物向左移动、向右移动和不动三种;以及障碍物前进和后退的位置变化:表现为面积的增大、不变和减小三种状态.此外,为了区分由于机器人本体运动所引起的图像变化和自身环境的变化,我们把当前机器人自身的动作也包含到状态空间内(25 种),因此集合 S 内有 $9 \times 9 \times 25$ 种不同状态组合.

在开始 Q-学习时,由于不知道转移概率 T 和报酬分布 r 的初始值,就从 0 开始不断累计更新 Q 值,其中 r 是执行动作 a 后获得的报酬, s' 是下一个状态, α 为学习率(在实验中取 0.8).具体步骤如下:

初始化环境:将 Q-学习的行为价值评测函数及环境变量都置为零;然后循环执行下面过程:

- (1) 将当前机器人状态置为 s .
- (2) 根据策略 f 选择行动 a (可以随机选择)执行
- (3) 得到状态 s' 以及报酬 r .
- (4) 更新 $Q(s, a)$:

$$Q(s, a) \leftarrow (1 - \alpha) Q(s, a) + (\alpha (r(s', a) + \max_{a'} Q(s', a))) \quad (8)$$

- (5) 更新策略 f 即: $f(s) \leftarrow a$, 当

$$Q(s, a) = \max_{a'} Q(s, b) \quad (9)$$

在视觉与行为模型的基础上进行 Q-学习规划,可降低系统的搜索空间,提高运行速度.文中状态空间的大小为 $N_s \times N_a$, N_s 为障碍物及环境状态数, N_a 为自身动作个数.由它们决定的状态空间 T 的大小为 $|T| = N_s \times N_a \times N_a$.通过前面的视觉与行为模型构建了拓扑关系图(如图 5 所示).设动作空间被分割成 c 个类,每个类包含的电动机动作个数为: N_{c_1} ,

$$N_{c_2}, \dots, N_{c_c}, \text{ 那么有: } N_a = \sum_{i=c_1}^c N_i \quad (10)$$

分割以后最初的学习是对每一个类的代表向量进行,然后对 c 个类同时进行学习,因此搜索空间状态个数变成:

$$|P| = N_s \times (N_a^2 + N_{c_1} \times N_{c_1} + N_{c_2} \times N_{c_2} + \dots + N_{c_c} \times N_{c_c}) \quad (11)$$

因此对于分 5 个动作类的机器人来说,由于:

$$N_a^2 \geq N_a^2 + N_{c_1} \times N_{c_1} + N_{c_2} \times N_{c_2} + \dots + N_{c_c} \times N_{c_c} \quad (12)$$

可看出 $|T| \geq |P|$,表明采用视觉与行为模型的拓扑结构图,能在一定程度上压缩搜索空间,加快系统的实时反应能力.

3.2 自主机器人动态避障实验

本实验使用美国 ActiveMedia 公司 Pioneer DX2 型自主机器人,视觉系统配备 Sony D30/31 单目彩色镜头,图像采集频率 30 帧/秒.图 7 为机器人在运动过程中,前方右侧突然出现横向运动物体,机器人当检测到障碍物之后,通过 Q-学习选择适当的动作序列,产生主动规避的行动过程.

实验中利用视觉与行为模型的光流场检测动态障碍物,

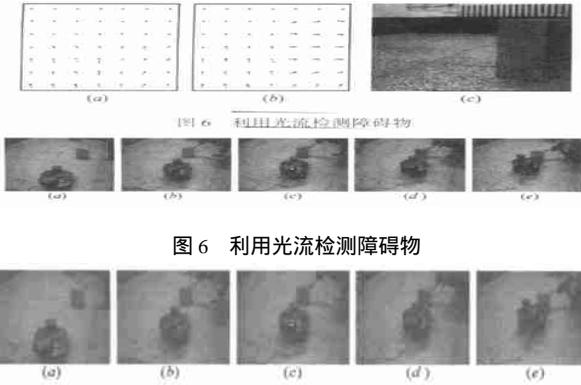


图 6 利用光流检测障碍物

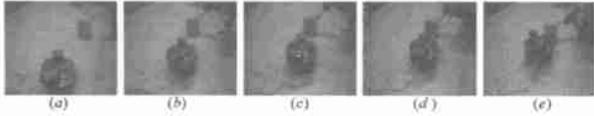


图 7 机器人避障过程环境示意图

本质上说是基于光流块区域的障碍物检测,相对于像素级光流计算而言,降低了计算量,同时也提高了抗干扰性能。从实验中可以看出,在机器人运动过程中,只有较近的物体才有较为明显的光流向量异常,这也符合机器人躲避近处障碍物的常理,这时机器人能够迅速产生躲避动作。图 8 表示了在机器人运动过程中,检测光流异常区域的示意图。机器人在检测到前方障碍物后,开始减速并产生躲避动作,从障碍物上的光流能够看到这个变化。

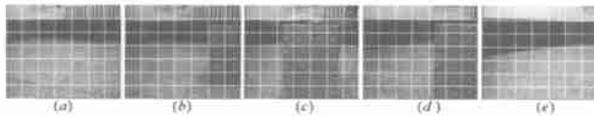


图 8 机器人障碍物检测及动态避障过程示意图

图 9 表示了机器人避障过程中两个驱动电动机的命令序列。可以看出,当机器人面前近距离突然出现障碍物时,机器人两个电动机转速相反,急速躲避障碍物。机器人在进行动作规划的过程中,根据动作的分类属性,在一定范围内选择最佳电动机驱动动作,图 10 表示机器人在避障过程中候选动作类范围以及所选择电动机驱动命令的变化路径。

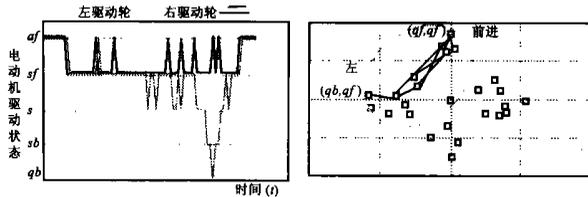


图 9 电动机命令序列

图 10 电动机命令序列路径

实验中,机器人所配备的图像采集卡图像采集频率为 30 帧/秒,在获取图像后计算图像块的光流场,由于已经将图像分解成 7×7 的栅格区域,因此匹配计算量非常小。此外,利用视觉与行为关系模型所构造的拓扑关系图在进行 Q-学习的过程中也简化了处理,减少了搜索空间。在实验中,最终机器人视觉系统能达到 9 帧/秒左右的信息处理速度。机器人的电动机作了离散化处理,在实验室中设定机器人最高速度为 0.6m/s,机器人可以根据前方障碍物存在的具体情况调节运动速度(也就是每个轮子的电动机可以在 0m/s、 $\pm 0.3m/s$ 和 $\pm 0.6m/s$ 的速度上切换),能够适应对环境实时性的要求。

4 结论

以往的机器人应用中,视觉系统与电动机驱动系统都作为两个独立的模块进行工作,很少考虑到它们之间的联系,仅仅是通过命令来进行参数传递,降低了两个本来具有紧密关系系统的耦合程度。本文所建立的机器人视觉与行为关系模型,能够将这两个系统更加密切的关联在一起,丰富了机器人获取信息的手段。在此基础上,该模型也为强化学习提供了思路,通过分析不同电动机驱动动作的聚类分布,能够减少强化学习的搜索空间,提高机器人实时性能。实验表明这种方法具有稳定的性能,能够满足实时性要求。

参考文献:

- [1] Boley D L, Sutherland K T. A rapidly converging recursive method for mobile robot localization [J]. International Journal of Robotics Research, 1998, 17(10) : 1027 - 1039.
- [2] Ohya A, Kosaka A, Kak. Vision-based navigation by a mobile robot with obstacle using single-camera vision and ultrasonic sensing [J]. IEEE Transactions on Robotics and Automation, 1998, 14(6) : 969 - 978.
- [3] Brooks R A. A robust layered control system for a mobile robot [J]. IEEE Journal of Robotics and Automation, 1986, RA-2(1) : 14 - 23.
- [4] Yao Shu, Zhang Bo. The learning convergence of CMAC in cyclic learning [J]. Journal of Computer Science and Technology, 1994, 9(4) : 320 - 328.
- [5] Horn B, Schunch B. Determining optical flow [J]. Artificial Intelligence, 1981, 17: 185 - 203.
- [6] 黄岗岗, 宋克欧. 模式识别 [M]. 哈尔滨: 哈尔滨工程大学出版社, 1998.
- [7] Watkins C J C H, Dayan P. Q-learning [J]. Machine Learning, 1992, 8: 279 - 292.

作者简介:



梁冰 男, 1974 年出生于河北省石家庄市, 博士生, 于 1999 年在哈尔滨工业大学攻读博士学位, 主要研究领域为机器人视觉, 智能机器人导航。



洪炳男, 1939 年出生于吉林, 教授, 博士, 现为哈尔滨工业大学教授, 博导, 主要研究领域为分布式人工智能, 虚拟现实技术, 智能机器人。



曙光 男, 1970 年出生于内蒙古呼和浩特市, 博士生, 主要研究领域为图像处理, 多智能体合作。