

交互式多模型算法的 MT-R 机器人视觉跟踪*

弋英民, 黄莹

(西安理工大学自动化与信息工程学院, 西安 710048)

摘要: 针对机器人的视觉跟踪问题, 提出了一种基于交互式多模型算法的视觉跟踪方法。该方法采用 HSV 颜色模型, 通过交互式多模型算法进行滤波估计, 获得目标的运动属性。利用交互式多模型算法对目标的位置进行一步或者 N 步预测, 在获得下一帧的观测值后, 在预测位置的区域进行目标搜索, 可减少搜索区域, 节省搜索时间, 增加了跟踪的实时性。使用该方法对机器人足球比赛中的红色球进行搜索跟踪。实验结果表明, 应用该方法可获得目标的运动属性, 能快速搜索到运动目标, 并能进行准确跟踪。

关键词: HSV 颜色模型; 交互式多模型算法; 视觉跟踪

中图分类号: TP242.6+2

文献标志码: A

文章编号: 1001-3695(2014)05-1574-03

doi:10.3969/j.issn.1001-3695.2014.05.071

Visual tracing algorithm based on interacting multiple model for MT-R robot

YI Ying-min, HUANG Ying

(Faculty of Automation & Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: For the visual tracing problem, this paper presented a visual tracing method based on interacting multiple model (IMM) algorithm using HSV color model. It acquired the moving characteristics of the target by means of the filter function of IMM algorithm, and predicted the location of the target by IMM algorithm with one-step or N -steps. After obtaining the next observation frame, it searched the target in the predictive area. Thus in this way, it reduced the search area, saved the search time and increased the ability of real-time tracking. This method was put into practice to search and track the red ball in RoboCup. Experiment results show that the moving target can be searched quickly and tracked correctly. The moving characteristics of the target can be obtained with the proposed algorithm.

Key words: HSV color model; interacting multiple model algorithm (IMM); visual tracing

基于视觉的运动目标跟踪是近年来计算机视觉领域中备受关注的前沿方向,它在机器人、交通监控、医学研究、轮船导航等领域有着广泛的应用。人们要求在获得运动物体的观测值后,能对其快速进行搜索并准确跟踪,从而正确和实时地获得运动物体的位置、速度等运动属性。为此,人们提出了多种方法来解决这些问题^[1-6]。在这些方法中,基于区域的跟踪方法受到重视。基于区域的跟踪方法可产生稳定的跟踪结果,但是为找到运动区域的聚类中心,算法的迭代次数多,降低了跟踪的实时性。为了提高跟踪的实时性,可采取不对整幅图像进行处理,而只对包含目标的区域进行处理的方法。这就要求对目标的图像未来位置进行搜索。目前一般使用 SSD (sum-of-squared difference) 方法进行搜索^[7],它的缺点是累计图像处理区域大。本文提出了一种基于交互式多模型 (IMM) 算法的视觉跟踪方法。该方法利用 IMM 的滤波功能获得目标的运动属性,并利用 IMM 的预测功能对目标的位置进行一步或者 N 步预测,且在预测位置的区域搜索目标。与 SSD 方法相比较,该方法的优点是利用 IMM 的预测功能可大致确定目标的未来位置,不像 SSD 法那样重复多次使用图像处理对较大区域进行搜索才能确定目标的未来位置,这样就可以减少图像处理的区域,提高图像处理的速度和跟踪速度。本文将该方法用于机器

人足球比赛中红色球的搜索跟踪。实验结果表明,应用该方法能获得目标的运动属性,可快速搜索到运动目标,并能准确地实现跟踪。

1 交互式多模型算法

设一个使用确定模型集合 $M = \{m_i | i = 1, 2, \dots, m\}$ 的混合系统,状态方程和观测方程的离散形式为

$$\begin{aligned} x(k+1) &= \Phi_i(k)x(k) + w_i(k) \quad i = 1, 2, \dots, m \\ z(k) &= H_i(k)x(k) + v_i(k) \quad i = 1, 2, \dots, m \end{aligned} \quad (1)$$

其中: $w_i(k)$ 和 $v_i(k)$ 为状态噪声和观测噪声,协方差矩阵分别为 $Q_i(k)$ 和 $R_i(k)$ 。

已知 m 个模型的初始状态为 $x^i(0) = \hat{x}^i(0)$, 初始模型概率为 $u_i(0) = P(m_i(0)/z(0))$, 模型转移符合先验的 Markov 过程,其转移概率矩阵为 $[P_{ij}]$, P_{ij} 表示模型 m_i 向 m_j 转移的概率。设从 $k-1$ 时刻开始到 k 时刻,一个采样周期的 IMM 算法的实现步骤如下^[8,9]。

1.1 输入交互

对于模型 $m_j (j = 1, 2, \dots, m)$, $\forall m_j \in M$, 有

$$\text{模型预测概率: } \bar{c}_j = \sum_i P_{ij} u_i(k-1) \quad (2)$$

模型转移概率:

收稿日期: 2013-07-05; **修回日期:** 2013-08-25 **基金项目:** 国家自然科学基金资助项目(51275405); 陕西省教育厅自然科学基金资助项目(2013JK1078)

作者简介: 弋英民(1976-),男,陕西西安人,副教授,博士,主要研究方向为机器人 SLAM 研究、信息融合与控制理论(yiym@xaut.edu.cn); 黄莹(1989-),女,陕西渭南人,硕士研究生,主要研究方向为机器人运动及地图构建。

$$u_{ij}(k-1/k-1) = P(m_i(k-1)/m_j(k), z^{k-1}) = \frac{1}{c_j} P_{ij} u_i(k-1) \quad (3)$$

混合输入:

$$\hat{x}^{0j}(k-1/k-1) = \sum_i \hat{x}^i(k-1/k-1) u_{ij}(k-1/k-1) \quad (4)$$

$$P^{0j}(k-1/k-1) = \sum_i u_{ij}(k-1/k-1) \{ P^i(k-1/k-1) + [\hat{x}^i(k-1/k-1) - \hat{x}^{0j}(k-1/k-1)] [\hat{x}^i(k-1/k-1) - \hat{x}^{0j}(k-1/k-1)]^T \} \quad (5)$$

1.2 模型条件滤波

对于模型 $m_j (j=1, 2, \dots, m)$, $\forall m_j \in M$, 以 $\hat{x}^{0j}(k-1/k-1)$ 和 $P^{0j}(k-1/k-1)$ 为输入, 代入基于模型 m_j 的滤波器(一般选用 Kalman 滤波器), 进而获得状态估计 $\hat{x}^j(k/k)$ 及其估计的协方差 $P^j(k/k)$ 。

1.3 模型概率更新

计算 k 时刻的模型 $m_j (j=1, 2, \dots, m)$, $\forall m_j \in M$ 的似然函数:

$$A_j(k) = P(z(k)/m_j(k), z^{k-1}) = P(z(k)/m_j(k), \hat{x}^{0j}(k-1/k-1), \quad (6)$$

$$P^{0j}(k-1/k-1)) = N((z(k) - z^j(k/k-1)) | 0, S^j(k))$$

这里, $A_j(k)$ 服从均值为 0、方差为 $S^j(k)$ 的正态分布, $S^j(k)$ 为新息协方差矩阵。模型 $m_j (j=1, 2, \dots, m)$, $\forall m_j \in M$ 的校正概率为

$$u_j(k) = P(m_j(k)/z^k) = \frac{1}{c} \frac{P(z(k)/m_j(k), z^{k-1}) P(m_j(k)/z^{k-1})}{P(z(k)/z^{k-1})} = \frac{1}{c} A_j(k) \bar{c}_j \quad (7)$$

其中: $\bar{c}_j$ 由式(2)给出, c 由式(8)求得。

$$c = \sum_j P(z(k)/m_j(k), z^{k-1}) P(m_j(k)/z^{k-1}) = \sum_j A_j(k) \bar{c}_j \quad (8)$$

1.4 输出

$$\hat{x}(k/k) = \sum_j \hat{x}^j(k/k) u_j(k) \quad (9)$$

$$P(k/k) = \sum_j u_j(k/k) \{ P^j(k/k) + [\hat{x}^j(k/k) - \hat{x}(k/k)] [\hat{x}^j(k/k) - \hat{x}(k/k)]^T \} \quad (10)$$

2 基于交互式多模型算法的视觉跟踪

2.1 目标检测与跟踪

目标检测即在图像中搜索是否有目标, 给出目标在图像中的位置特征点。目前比较常用的目标检测技术有模板匹配、边缘检测、色彩分类等。在本文的视觉系统中, 由于实时性很高且目标是用颜色来区分的, 因此采用基于色彩的图像分割技术, 它不仅快捷、有效, 而且对光照强度的适应性强。本文算法采用的是 HSV 模型, 它将采集的颜色信息分为色度 H 、饱和度 S 和亮度 V 三种属性。由于这种模型具有线性伸缩性, 可感知的色差与颜色分量所相应的样值的欧氏距离成比例, 较之 RGB 模型更适合用于检测及识别处理。

目标的跟踪是确定同一物体在不同帧中所在位置的过程, 当运动目标被检测出来后, 对相邻帧中检测出的目标进行匹配。由于目标有一定的大小和形状, 为了对目标定位, 必须在图像中选取目标上对应点的图像位置。目标的质心点具有不随平移、旋转与比例的改变而变化的特点, 故选取目标的质心点作为目标点。根据目标的质心点在本帧以及相邻的连续前几帧的位置, 可以利用 IMM 算法的预测功能, 预测出目标的质心点在下一帧的位置。在获得下一帧的观测值后, 在目标质心点的预测值位置的区域进行目标搜索, 可快速搜索到运动目标并进行跟踪, 提高了目标跟踪的实时性。

2.2 目标的运动属性

由于图像序列前后两帧的时间间隔 T 很小, 因此可用二

阶微分方程来描述目标的运动轨迹。定义状态矢量 $x(t) = [x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t), \ddot{x}(t), \ddot{y}(t), \ddot{z}(t)]^T$, 状态方程和观测方程为

$$x(k+1) = \Phi(k)x(k) + w(k) \quad (11)$$

$$z(k) = H(k)x(k) + v(k) \quad (12)$$

其中: $w(k)$ 为零均值的高斯白噪声, 协方差矩阵为 $Q(k)$; $v(k)$ 为零均值的高斯白噪声, 协方差矩阵为 $R(k)$ 。由上可见, 目标的运动符合 IMM 算法的应用条件, 因此可通过 IMM 算法的滤波功能, 获得运动目标的位置、速度和加速度等运动属性。

2.3 基于 IMM 算法的跟踪

基于 IMM 算法的视觉跟踪主要步骤如下: a) 获取第一帧的观测值, 并作为 IMM 算法中滤波的初始值; b) 隔行隔列全局搜索, 搜索到目标内一点即跳出, 记录找到的目标内第一点的列坐标, 再沿着该列逐行向左向右搜索, 直到目标外; c) 对目标进行定位, 得到目标的质心点, 作为 IMM 算法的观测值; d) 利用 IMM 算法由第一帧的观测值进行初始化状态估计, 第一帧的观测值作为第二帧的状态预测; e) 利用 IMM 算法进行滤波估计, 得到目标的运动属性(位置、速度和加速度); f) 利用 IMM 算法对目标的位置进行一步或者 N 步预测, 在获得下一帧的观测值后, 在预测位置的区域进行目标搜索。

3 实验结果

实验采用 MT-R 机器人, MT-R 机器人为三层结构, 分动力底盘、计算机系统层、传感及运动控制层三层结构, 整机三点支撑, 两个驱动轮, 一个万向轮平衡。MT-R 机器人配有带云台的摄像机, 可以采集视频图像, 获取视频数据, 用户可对该视频数据进行处理。MT-R 机器人开发平台如图 1 所示。以机器人足球比赛中对红色球的搜索跟踪进行实验。



图 1 MT-R 机器人

图 2 给出了跟踪过程中获得的目标运动属性。由于仿真的相似性, 只给出了在 x 方向的实验结果。根据计算, 在相同条件下, 使用全局搜索与本文算法的每一帧图像处理平均时间如表 1 所示, 预测目标在下一帧所在的区域所需时间如表 2 所示。在 IMM 算法中, 使用了四个模型, 分别为: CV (constant velocity) 模型: $Q_1(k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R_1(k) = 0.2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; CA (constant acceleration) 模型: $Q_2(k) = 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R_2(k) = 0.2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; 常转弯模型 1 ($\omega_1 = 0.0349$ (rad/s)), $Q_3(k) = 10 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R_3(k) = 0.1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; 常转弯模型 2 ($\omega_2 = -0.0349$ (rad/s)), $Q_4(k) = 10 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R_4(k) = 0.1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, 模型转移概率 $P = \begin{bmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.97 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.97 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.97 \end{bmatrix}$, 采样时间 $T = 0.04$ s。

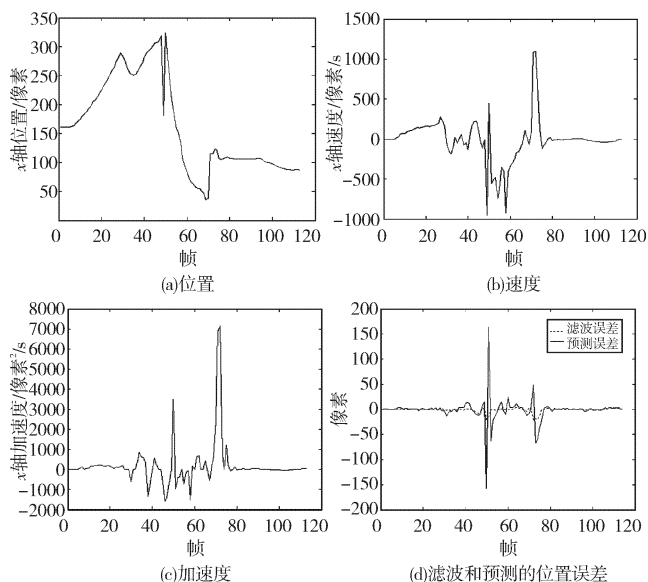


图2 x轴向仿真

由图2可见,在整个跟踪过程中,通过 IMM 算法进行滤波可以获得目标的运动属性(位置、速度和加速度)。由图2(d)可以看出,滤波获得的位置误差很小,预测获得的位置误差在整个跟踪过程中会发生大的波动,并且由图2(a)(b)和(c)中可以看出,发生波动的时刻即为目标发生机动的时刻。通过 IMM 算法的预测功能,可准确反映出运动目标发生机动的情况。由表1可知,本文算法相比全局搜索所花费的时间大幅度地减少。在表2中,相比利用单一模型的 Kalman 滤波的预测时间,本文算法效果要差一点,其原因是 IMM 算法是一个多模型的并行 Kalman 滤波算法,利用 Kalman 滤波预测目标所在的位置区域,必须要建立目标准确的运动模型,而 IMM 算法可应用于目标集所覆盖的所有目标运动模式。运用基于 IMM 算法进行目标搜索跟踪并获得目标的运动属性,不仅处理速度提高、应用范围广,而且还能准确反映目标的运动特征。

表1 平均时间

时间	算法	
	全局搜索	本文算法
处理时间/s	2.2	0.9

表2 预测时间

时间	算法		
	Kalman 滤波	本文算法	SSD
时间/s	0.32	0.37	0.48

(上接第1561页)等会导致训练时间过长,达到数周的量级。本文提出了采用消息传递接口技术并在小型集群服务器上运行训练程序的方案,有效地解决了存储空间不够和速度过慢的问题。结合算法特点提出了对存储和带宽进行优化,使训练速度提高了30倍,从而能更好地进行随机森林算法的训练,可以及时调整训练参数来多次训练。实验测试表明像素分类准确率能够达到平均80%以上,人体骨架节点实际误差比微软效果更好,继而在后续的人体节点估计模块中也能获得较好的结果。

参考文献:

- [1] AGARWAL A, TRIGGS B. 3D human pose from silhouettes by relevance vector regression [C]//Proc of IEEE Conference on CVPR. 2004:882-888.
- [2] HU Zhi-lan, WANG Gui-jin, LIN Xing-gang, et al. Recovery of upper body poses in static images based on joints detection [J]. *Pattern Recognition Letters*, 2009, 30(5):503-512.
- [3] GANAPATHI V, PLAGEMANN C, KOLLER D, et al. Real time motion capture using a single time-of-flight camera [C]//Proc of IEEE Conference on CVPR. 2010:755-762.

4 结束语

本文提出了一种基于 IMM 算法的视觉跟踪方法。该方法利用 IMM 算法的滤波功能获得目标的运动属性,利用 IMM 算法的预测功能对目标的位置进行一步或者 N 步预测,在获得下一帧的观测值后,在预测位置的区域搜索目标。将该方法用于机器人足球比赛中的红色球搜索跟踪实验,其结果表明了该方法的有效性。

参考文献:

- [1] CHO H, RYBSKI P E, ZHANG Wen-de. Vision-based 3D bicycle tracking using deformable part model and interacting multiple model filter [C]//Proc of IEEE International Conference on Robotics and Automation. [S.l.]: IEEE Press, 2011: 4391-4398.
- [2] STIERLIN S, DIETMAYER K C J. Scale change and TTC filter for longitudinal vehicle control based on monocular video [C]//Proc of the 15th International IEEE Conference on Intelligent Transportation Systems. [S.l.]: IEEE Press, 2012: 528-533.
- [3] GRINIAS I, TZIRITAS T. A semi-automatic seeded region growing algorithm for video object localization and tracking [J]. *Signal Processing: Image Communication*, 2001, 16(10):977-986.
- [4] KIM W, LEE C Y, LEE J J. Tracking moving object using Snake's jump based on image flow [J]. *Mechatronics*, 2001, 11(2):199-226.
- [5] 陈凤东, 洪炳镕. 基于动态阈值背景差分算法的目标检测方法 [J]. *哈尔滨工业大学学报*, 2005, 37(7): 883-955.
- [6] 刘并, 司秉玉. 基于图像区域搜索法的彩色球目标识别与跟踪 [J]. *仪器仪表学报*, 2003, 24(4):225-232.
- [7] STEPHEN J R, VIKRAMADITYA B, NELSON B J. Micropositioning of a weakly calibrated microassembly system using coarse-to-fine visual servoing strategies [J]. *IEEE Trans on Electronics Packaging Manufacturing*, 2000, 23(2): 123-131.
- [8] BLOM H A P, BAR-SHALOM Y. The interacting multiple model algorithm for systems with Markovian switching coefficients [J]. *IEEE Trans on Automatic Control*, 1988, 33(8):780-783.
- [9] MAZOR E, AVERBUCH A, BAR-SHALOM Y, et al. Interacting multiple model methods in target tracking: a survey [J]. *IEEE Trans on Aerospace and Electronic Systems*, 1998, 34(1):103-123.
- [4] WILSON J L. Microsoft Kinect for Xbox 360 [EB/OL]. (2010). <http://www.pcmag.com/article/0,2817,2372069,00.asp> Mag. Com.
- [5] SHOTTON J, FITZGIBBON A, COOK M, et al. Real-time human pose recognition in parts from single depth images [C]//Proc of IEEE Conference on CVPR. 2011:1297-1304.
- [6] ROGEZ G, RIHAN J, RAMALINGAM S, et al. Randomized trees for human pose detection [C]//Proc of IEEE Conference on CVPR. 2008:1-8.
- [7] BREIMAN L. Random forests [J]. *Machine Learning*, 2001, 45(1):5-32.
- [8] KATH R. Managing memory-mapped files in Win32 [EB/OL]. <http://msdn.microsoft.com>.
- [9] GROPP W. MPICH2: a new start for MPI implementations [M]//Recent Advances in Parallel Virtual Machine and Message Passing Interface. Berlin: Springer, 2002.
- [10] 都志辉, 李三立, 陈渝, 等. 高性能计算之并行编程技术——MPI 并行程序设计 [M]. 北京: 清华大学出版社, 2001.
- [11] 吕治国, 李焱, 贺汉根. 基于 Poser 模型的三维人体建模方法 [J]. *计算机工程*, 2008, 34(13): 256-258.
- [12] Smith Micro Software. Poser 7 [EB/OL]. <http://Poser.smithmicro.com>.