

一种尺度自适应的机器人目标跟踪算法

成新田 唐振民

(南京理工大学计算机科学与技术学院 南京 210094)

摘要 Mean-Shift 算法是一种简单高效的目标识别算法,但是不能有效地识别被遮挡的目标和有尺度变化的目标。基于仿射变换,提出了一种尺度自适应的机器人目标跟踪算法。定义了转角点,并根据转角点匹配对目标进行区分,最后通过仿射变换识别出目标的尺度变化。与其它相关算法相比,该算法能有效地识别被跟踪目标的遮挡问题;当被跟踪目标的尺度发生改变时,该算法仍然能准确地对目标进行识别。分析表明,当视屏流中每秒的图像小于 25 帧并且目标的图像小于 2×10^4 个像素时,该算法可以用于目标的实时跟踪。

关键词 尺度,机器人,目标跟踪,图像处理

中图分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.12.060

Scale Adaptive Target Tracking Algorithm for Robot

CHENG Xin-tian TANG Zhen-min

(School of Computer Science and Engineering, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract Mean-Shift algorithm is a simple and efficient target tracking algorithm, but it can't recognize occluded target and the target of scale changes. This paper proposed a scale adaptive target tracking algorithm for robot based on affine transformation. We defined the corner points, recognized target according to the defined corner points, and recognized the scale changes of target using affine transformation. Compared with relative algorithms, the proposed algorithm can re-recognize the occluded target effectively, and when the scale of target changes, the proposed algorithm can also recognize the target accurately. The analysis shows that, when there is less than 2×10^4 pixels in an image and less than 25 frames per second in a video stream, the proposed algorithm can be used in real-time target tracking.

Keywords Scale, Robot, Target tracking, Image processing

实时的目标跟踪是机器人的重要研究内容之一。随着图像处理技术的发展,视频捕捉设备实时地捕捉目标信息的图像,并对目标进行识别^[1]。例如,用于服务的机器人需要跟踪目标用户并进行服务;用于救援的机器人可以跟踪有生命迹象的对象并获取更多的用于救援的信息^[2]。在这些实时的目标跟踪应用中,目标的尺度改变以及被遮挡给目标的实时识别提出了严峻的挑战。机器人对目标进行跟踪的算法有多种,Mean-Shift 跟踪算法^[3]是最有效的目标跟踪算法之一,能满足应用的实时性要求。然而,当被跟踪的目标的尺度改变或者被遮挡时,Mean-Shift 跟踪算法的性能会急剧下降。

本文重点研究了 Mean-Shift 跟踪算法在处理目标的尺度改变以及被遮挡时所带来的问题,提出了一种尺度自适应的 Mean-Shift 目标跟踪算法。首先,利用仿射结构得到尺度因子,并通过尺度因子调整 Mean-Shift 算法。然后,分别基于转角点匹配、Mean-Shift 算法,以及尺度自适应的 Mean-Shift 跟踪算法,通过高效的仿射转换获得 3 个候选目标,并选取最优的候选目标,使跟踪算法能更好地适应目标对象的尺度变化。

本文通过实验评估了上述算法的性能,并将本文提出的

算法与几种具有代表性的目标跟踪算法进行了对比。实验结果表明,与其它相关算法相比,本文提出的尺度自适应的 Mean-Shift 算法有相对较好的尺度自适应性和遮挡健壮性。此外,本文提出的算法具有很高的执行效率,可以用于基于视觉的机器人实时目标跟踪。

1 相关工作

本节介绍了基于视觉的机器人目标跟踪的相关研究。

目标跟踪算法在目标区域附近的背景对目标的识别产生干扰时,会大大降低目标的识别效率。为了解决这一问题,Comaniciu 等人^[3]在 Mean-Shift 算法中引入背景加权直方图来减小背景对目标的干扰。该方法通过对核的半径进行修改($\pm 10\%$),以不同的核尺度独立地应用 3 次 Mean-Shift 算法,并选取具有最大 Bhattacharyya 系数的尺度作为核尺度。

Ning 等人^[4]通过研究证明:背景加权直方图等价于共同的目标表示。在引入背景加权直方图时,Mean-Shift 算法并没有引入新的信息。Ning 等人通过对原始的目标模型而不是目标候选项进行转换,提出了一种基于 CBWH 的 Mean-Shift 算法。CBWH 模型通过增强目标模型的主特征来减小

到稿日期:2014-01-22 返修日期:2014-03-19 本文受国家自然科学基金课题(91220301)资助。

成新田(1985-),女,博士生,主要研究方向为计算机视觉、智能机器人自主导航、模式识别;唐振民(1961-),男,博士,教授,博士生导师,主要研究方向为智能机器人与目标识别、图象处理与模式识别、智能信息系统。

目标与背景之间的相似图像特征,明显地减小了背景区域对目标的识别产生的干扰。在CBWH的基础上,Mohammadi等人^[5]研究了被跟踪目标逐渐缩小情况下的Mean-Shift算法。当被跟踪的目标的尺度逐渐缩小时,该算法的性能很好;当被跟踪目标的尺度逐渐增大,尤其是目标的尺度超过跟踪窗口时,算法的Bhattacharyya系数^[6]总是收敛于小搜索窗口的局部最大值,因而算法的性能很低。

Peng等人^[7]提出了一种自动带宽选择的Mean-Shift算法。该方法通过后跟踪和对对象质心注册来提高Mean-Shift算法的性能。该方法的理论依据是:核内的目标尺度和目标位置的变化不会影响Mean-Shift算法对目标进行跟踪的准确性。然而,该方法不适用于目标的尺度逐渐减小的情况。为了解决上述问题,Bradski等人^[8]提出了一种基于不变距的Mean-Shift算法——CAMSHIFT。CAMSHIFT算法在计算二阶中心距时,计算复杂度很高,因此不适用于目标的实时跟踪。与CAMSHIFT算法相似,Hu等人^[9]通过将空间颜色特征空间和相似性函数相结合,提出了一种增强的Mean-Shift跟踪算法。该算法用特征值来监视目标对象的尺度。上述方法在目标跟踪时会产生很大的偏差,并且当目标被遮挡时有可能使目标不能识别。

文献^[10,11]基于仿射映射来解决目标的尺度变化问题。这些方法简单地引入仿射变换并能获得很好的尺度自适应性。然而,在这些方法中,兴趣点匹配并不能有效地获得正确的仿射结构。此外,当目标被遮挡时,其仍然会丢失对目标的识别和跟踪。

2 尺度自适应的目标跟踪算法

2.1 Mean-Shift 算法

对于某未知的密度函数 $f(x)$,令 $S = \{x_i\}_{i=1, \dots, n}$ 是该密度函数的一个独立同分布的采样。目标的大小为 $M \times N$ 像素,Mean-Shift目标跟踪算法定义的核密度估计为:

$$f(x) = \frac{1}{M \times N} \sum_{i=1}^{M \times N} k_h \left(\left| \frac{x - x_i}{h} \right|^2 \right) \quad (1)$$

其中,带宽 $h > 0$ 。

在Mean-Shift算法中,根据前一帧的目标中心一个接一个地构建目标候选项,计算当前帧的候选项与前一帧目标的Bhattacharyya系数。Bhattacharyya系数的定义如式(2)所示。

$$d(y) = \sqrt{1 - \rho(y)} \quad (2)$$

其中

$$\rho(y) = \rho(p(y), q) = \sum_{b=1}^B \sqrt{p_b(y)q_b} \quad (3)$$

式中, q 表示目标模型, $p(y)$ 为候选模型, B 是用来计算的容器个数。

2.2 尺度自适应的Mean-Shift算法

2.2.1 算法流程图

为了提高Mean-Shift算法的尺度适应性和遮挡健壮性,本文提出了尺度自适应Mean-Shift算法。在同一视频流中,相同的对象在邻接的帧中有着不同的视图。如果建立一个包含尺度变化的并且描述不同视图的目标模型,那么可以在视频中更准确地跟踪对象。仿射变换方法可应用于对象的旋转、变换等多种视图,也可以用于图像的缩放。然而,在仿射变换中,需要进行特征点匹配,本文应用转角点作为特征点。

本文提出的尺度自适应的Mean-Shift算法的工作流程如图1所示,其中 W_i 为第 i 帧的跟踪窗口, C_i 是 W_i 的中心, C_i' 是第 i 帧的目标中心, W_i' 是 W_i 经过参数 s_x 和 s_y 调整后得到的, W_i'' 是以 C_i' 为中心的跟踪窗口经参数 s_x 和 s_y 调整后得到的, N_i 是 W_i 和 W_{i-1} 之间的匹配对的个数。

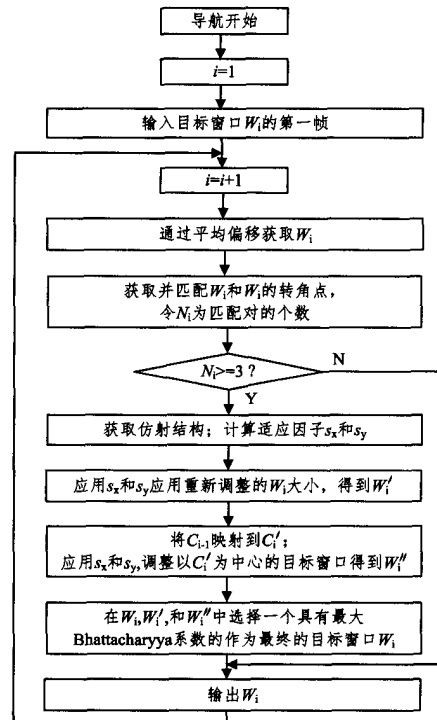


图1 算法流程

2.2.2 转角点匹配

为了应用仿射变换,需要寻找兴趣特征点对来计算仿射结构。实际上,具有不变图像特征的兴趣点是两个图像间的联系。错误的特征点对匹配将产生错误的仿射结构,并得到不正确的目标尺度。

转角点是从两个或多个边缘中形成的。边缘是区分两个不同的对象或者同一对象的不同部分的边界,因此转角点可以用来计算仿射结构。此外,基于相似彩色柱状图分布,转角点也可以用来在背景中区分出目标对象。Mean-Shift算法基于统计数据对比,即彩色柱状图对比,因此对目标和背景的差别比较敏感。当背景与目标的彩色柱状图分布相似时,Mean-Shift算法得到的目标中心经常偏离实际的位置,与此对比,转角点匹配能更准确地识别和跟踪目标。

本文采用Harris转角检测器^[8]用于目标的识别,在 W_i 中探测转角特征。令 I_x 和 I_y 分别是点 $I(x, y)$ 在水平和垂直方向的灰度变换,Harris转角检测器定义如下:

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (4)$$

其中

$$\begin{aligned} A &= I_x^2 \otimes \omega \\ B &= I_y^2 \otimes \omega \\ C &= I_x I_y \otimes \omega \end{aligned} \quad (5)$$

式中, \otimes 是卷积操作符, ω 是以 $I(x, y)$ 为中心的高斯窗口。本文采用 5×5 的高斯窗口。当点的转角测量大于阈值时,该点就是一个转角点,转角测量 R 的定义如下:

$$R = \det(M) - k(\text{trace}(M))^2 \quad (6)$$

其中, $k=0.04$ 。

接下来,采用均值为 0 的规范化交差相关在第 i 帧和第 $i-1$ 帧的转角特征之间寻找最优的匹配对。特征点对的相似性定义如下:

$$C(x, x') = \frac{1}{(2N+1)\sigma_{I_0}(x)\sigma_{I_1}(x')} \times \sum_{i=-N}^N \sum_{j=-N}^N [I_0(x-i, y-j) - \bar{I}_0] \times [I_1(x'-i, y'-j) - \bar{I}_1] \quad (7)$$

其中, \bar{I}_0 和 \bar{I}_1 分别是以 x 和 x' 为中心的、大小为 $2N+1$ 的窗口的平均灰度, σ_{I_0} 和 σ_{I_1} 为相应的方差。 $C(x, x')$ 的值域为 $[-1, 1]$ 。

2.2.3 仿射变换

当在连续的两帧中探测出转角点时,应用转角匹配计算仿射结构。在 2 维情况下,为了计算连续帧中的仿射结构,至少需要 3 个匹配点对^[9]。如果匹配点对不少于 3 个,那么可以用这些匹配对计算仿射结构 A_i 。根据仿射结构的原则,兴趣点需要满足如下条件:

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = A_i \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix} + b_i = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix} + \begin{pmatrix} e_x \\ e_y \end{pmatrix} \quad (8)$$

其中, $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$ 和 $\begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}$ 分别为匹配点在第 i 帧和第 $i-1$ 帧中的位置, A_i 是一个 2×2 的矩阵, s_x 和 s_y 分别为水平和垂直方向的弹性振幅, b_i 是一个 2×1 的矩阵, e_x 和 e_y 分别为水平和垂直方向上的平移。

在式(8)中,由于 $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$ 和 $\begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}$ 是已知的,可以计算出 s_x 和 s_y 。在本文提出的算法中, s_x 和 s_y 就是水平和垂直方向上的尺度因子。在应用 Mean-Shift 算法得到目标的候选项后,应用尺度因子 s_x 和 s_y 来调整目标的大小。

3 实验设计与结果分析

3.1 实验环境

实验采用的平台为一台 Pentium(R) 3.0 GHz、内存为 2Gb 的普通 PC,采用的数据集为 VIRAT 视频数据集。在该数据集中,图像的大小都被转换为 320×240 像素。

实验采用 3 种评价算法性能的度量标准。第 1 种是通过记录跟踪的结果来展示跟踪效果随着环境的变化(尺度变化和遮挡)。第 2 种和第 3 种评价标准分别是目标中心位置误差和目标相对尺度误差。目标中心位置误差是跟踪结果和真实位置之间的相对误差。目标相对尺度误差是将真实尺度规范化后的尺度,其定义如式(9)所示。

$$S = \frac{W \times H}{W_t \times H_t} \quad (9)$$

式中, W 和 H 分别是跟踪算法得到的当前目标尺度的宽度和高度, W_t 和 H_t 分别是真实目标尺度的宽度和高度。在完美的跟踪中,目标的中心位置误差接近于 0,相对尺度误差接近 1。

3.2 实验结果

将本文提出的算法记为 Proposed method,并且与 3 个具有带变形的跟踪算法进行对比:原始的 Mean-Shift 算法(MS)、具有 10% 自适应的 Mean-Shift 算法(MS with 10% adaption)和立体色 Mean-Shift 算法(Spatial color method)。

在实验中,上述算法均在 PC 中实现,每个算法的目标是

跟踪一个黑色的轿车。在数据集中,这辆轿车的尺度随着时间的变化而变化;当该轿车停在停车场时,它被其它的轿车所遮挡。图 2 分别给出了 4 种算法在目标被遮挡时的识别情况。从中可以看出,MS 与 MS with 10% adaption 算法都错误地将目标附近的黑色轿车当作了目标, Spatial color method 丢失了所跟踪的目标,而本文提出的算法可以正确地识别出目标测量。

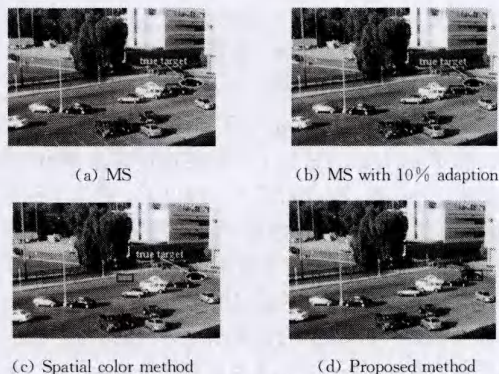


图 2 算法对遮挡的识别对比图

图 3 为 4 种算法对目标进行跟踪的中心位置误差。从图中可以看出, Spatial color method 算法的误差最大, MS 和 MS with 10% adaption 算法的误差接近,本文提出的算法的误差最小。这说明本文提出的算法在对目标进行跟踪时能很好地对目标进行识别。图 4 为 4 种算法对目标进行跟踪的相对尺度误差。本文提出的算法由于对目标进行了仿射变换,能更好地识别目标的尺度变换,其效果明显优于其它 3 种算法。

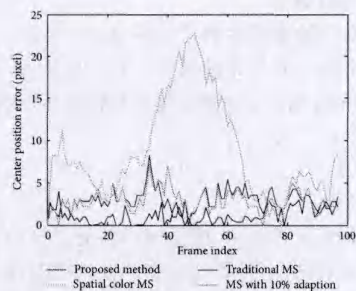


图 3 算法的中心位置误差对比图

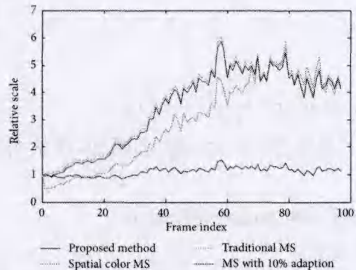


图 4 算法的目标相对尺度误差对比图

最后对本文提出的算法的执行时间进行了测试,实验结果如图 5 所示。随着目标的尺寸不断变大,算法的执行时间也相应地增大。然而,图中的曲线表明,当目标的像素个数小于 2×10^4 时,算法的运行时间不到 40ms。这表明,当视屏流

(下转第 292 页)

些共同的客观的评价准则。

参 考 文 献

[1] Avidan S, Shamir A. Seam carving for content aware image resizing [J]. *ACM Transactions on Graphics*, 2007, 26(3): 1-9

[2] Rubinstein M, Shamir A, Avidan S. Improved seam carving for video retargeting [J]. *ACM Transactions on Graphics*, 2008, 27(3): 1-9

[3] Shamir A, Avidan S. Seam carving for media retargeting [J]. *Communications of the ACM*, 2009, 52(1): 77-85

[4] Ariel S, Olga S. Visual media retargeting [C]// *ACM SIGGRAPH ASIA 2009 Courses*. New York, NY, USA, 2009, 11: 1-13

[5] Huang Hua, Fu Tian-nan, Rosin P L, et al. Real-time content-aware image resizing [J]. *Science in China Series F-Information on Sciences*, 2009, 52(2): 172-182

[6] Thilagam K, Karthikeyan S. Optimized image resizing using piecewise seam carving [J]. *International Journal of Computer Applications*, 2012, 42(14): 24-30

[7] Lin Xiao, Sheng Bin, Ma Li-zhuang, et al. Seamlet Carving for Shape-Aware Image Resizing [J]. *Science China: Information Sciences*, 2012, 55(5): 1073-1081

[8] Wang Y S, Tai C L, Sorkine O, et al. Optimized Scale-and-Stretch for Image Resizing [J]. *ACM Transactions on Graphics*, 2008, 27(5): 1-8

[9] Zhang Guo-xin, Cheng Ming-ming, Hu Shi-min, et al. A Shape-Preserving Approach to Image Resizing [J]. *Computer Graphics*

Forum, 2009, 28(7): 1897-1906

[10] 雷励星. 基于混合能量的内容敏感图像缩放新方法[J]. *计算机学报*, 2010, 33(10): 2015-2021

[11] 施美玲, 徐丹. 主体大小能控的内容感知图像缩放[J]. *计算机辅助设计与图形学学报*, 2011, 23(5): 915-922

[12] Niu Y Z, Liu F, Li X Q, et al. Image resizing via non-homogeneous warping [J]. *Multimedia Tools and Applications*, 2012, 56(3): 485-508

[13] Dong W M, Zhou N, Paul J C, et al. Optimized image resizing using seam carving and scaling [J]. *ACM Transactions on Graphics*, 2009, 28(5): 1-10

[14] Rubinstein M, Shamir A, Avidan S. Multi-operator media retargeting [J]. *ACM Transactions on Graphics*, 2009, 28(3): 1-11

[15] Dong W M, Bao G B, Zhang X P, et al. Fast multi-operator image resizing and evaluation [J]. *Journal of Computer Science and Technology*, 2012, 27(1): 121-134

[16] Luo S Q, Zhang J P, Zhang Q, et al. Multi-operator image retargeting with automatic integration of direct and indirect seam carving [J]. *Image and Vision Computing*, 2012, 30(9): 655-667

[17] Itti L, Koch C, Niebur E. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis [J]. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 1998, 20(11): 1254-1259

[18] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]// *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*. San Diego, CA, United States, 2005: 886-893

(上接第 282 页)

中每秒的图像小于 25 帧并且目标的图像小于 2×10^4 个像素时, 该算法可以用于目标的实时跟踪。

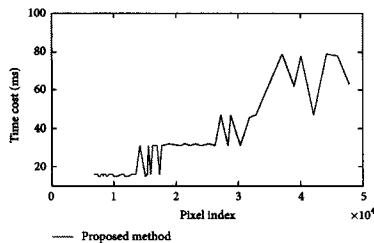


图 5 算法的运行时间评估

结束语 基于机器人视觉的目标跟踪是机器人研究领域的重要研究内容之一。经典的 Mean-Shift 算法虽然对目标跟踪有着很好的执行效率, 但是不能有效地应对被跟踪目标的遮挡以及尺度的变化问题。本文基于目标尺度的仿射变换, 提出了一种尺度自适应的机器人目标跟踪算法。与其它相关算法相比, 该算法能有效地识别被跟踪目标的遮挡问题; 当被跟踪目标的尺度发生改变时, 该算法仍然能准确地对目标进行识别。最后, 通过实验表明, 当视屏流中每秒的图像小于 25 帧并且目标的图像小于 2×10^4 像素时, 该算法可以用于目标的实时跟踪。

参 考 文 献

[1] Li X R, Jilkov V P. Survey of maneuvering target tracking. Part I. Dynamic models [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2013, 39(4): 1333-1364

[2] 邵文坤, 黄爱民, 韦庆. 目标跟踪方法综述 [J]. *影像技术*, 2006

(1): 17-20

[3] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(5): 564-577

[4] Ning J, Zhang L, Zhang D, et al. Robust mean-shift tracking with corrected background-weighted histogram [J]. *IET Computer Vision*, 2012, 6(1): 62-69

[5] Mohammadi S A, Amoozegar S, Jolfaei A, et al. Enhanced adaptive bandwidth tracking using mean shift algorithm [C]// *Proceedings of the IEEE 3rd International Conference on Communication Software and Networks (ICCSN'11)*. May 2011: 494-498

[6] Bhattacharyya A. On a measure of divergence between two multinomial populations [J]. *Sankhyā: The Indian Journal of Statistics*, 1946, 7(4): 401-406

[7] Peng N, Yang J, Liu Z, et al. Mean-Shift 跟踪算法中核函数窗宽的自动选取 [J]. *软件学报*, 2005, 16(9): 1542-1550

[8] Bradski G R. Computer vision face tracking for use in a perceptual user interface [C]// *Proceedings of the 4th IEEE Workshop on Applications of Computer Vision (WACV '98)*. 1998: 214-219

[9] Hu J, Juan C, Wang J. A spatial-color mean-shift object tracking algorithm with scale and orientation estimation [J]. *Pattern Recognition Letters*, 2008, 29(16): 2165-2173

[10] Zhao C, Knight A, Reid I. Target tracking using mean-shift and affine structure [C]// *19th International Conference on Pattern Recognition (ICPR 2008)*. IEEE, 2008: 1-5

[11] Leichter I, Lindenbaum M, Rivlin E. Tracking by affine kernel transformations using color and boundary cues [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(1): 164-171