

自动选择跟踪窗尺度的 Mean-Shift 算法

钱惠敏 茅耀斌 王执铨

(南京理工大学自动化学院, 南京 210094)

摘要 实用的跟踪系统要求能实时地适应运动目标的外观变化, 尺度固定不变的跟踪窗口不能有效地跟踪存在明显尺度变化的目标。本文将多尺度图像的信息量度量方法引入到运动目标跟踪中, 提出了一种跟踪窗口自动更新算法, 并用此算法改进了基于颜色直方图的 Mean-Shift 跟踪方案。实验结果表明, 改进的跟踪算法对尺寸逐渐减小和逐渐增大的目标都能自动选择合适的跟踪窗口大小。

关键词 目标跟踪 信息度量 Mean-Shift

中图分类号: TP391 **文献标识码**: A **文章编号**: 1006-8961(2007)02-0245-05

Mean Shift Tracking with Self-updating Tracking Window

QIAN Hui-min, MAO Yao-bin, WANG Zhi-quan

(School of Automation, Nanjing University of Science & Technology, Nanjing 210094)

Abstract A practical tracking system is required to update the appearance changes of moving objects in real-time. The system with fixed-size tracking window could not trace an object effectively when scale of the object has distinct changes, therefore it is important to select the scale of tracking window automatically. The information measure of multi-scale image in scale space has been used to differentiate the scale and is introduced into moving object tracking in this paper. Automatic updating method of tracking window is proposed, and is integrated into the classical Mean-Shift tracking algorithm based on color histogram. Experimental results demonstrate that the improved algorithm could select the proper size of the tracking window in the scenarios that not only of increasing scale but of decreasing scale.

Keywords object tracking, information measure, Mean-Shift

1 引言

目标跟踪一直是计算机视觉研究领域的热点问题, 其应用领域有视频监控、人机用户接口、虚拟现实等等。如何对视频序列中的目标进行稳健、有效的跟踪是目标跟踪的研究重点。因此, 一个实用的跟踪系统要求能够实时适应目标因各种运动导致的外观变化, 然而现在大多数的跟踪算法在跟踪存在明显尺度变化的目标(尤其是当目标尺寸逐渐增大以至超出跟踪窗口范围)时, 常常会导致目标的丢失。目前, 还没有系统的方法来解决跟踪窗口尺寸的自动调整问题。

Mean-Shift^[1]算法作为一种高效的模式匹配算法, 已经成功地应用在对实时性要求较高的目标跟踪

算法中。文献[2]~[6]对 Mean-Shift 算法核窗宽(即跟踪窗口尺寸)的选取提出了改进。文献[2]分别以核窗宽确定的原始跟踪窗口、核窗宽做正负10%增量修正后的跟踪窗口进行跟踪, 并选择较大 Bhattacharyya 系数对应的核窗宽作为最佳核窗宽。实验发现, 该方法只能对尺寸逐渐缩小的目标有较好的跟踪效果。文献[3]利用视频前后两帧中目标的角点匹配估计刚性物体的仿射模型参数, 并由此参数来更新核窗宽, 由于依赖于角点匹配, 不适合非刚性物体的跟踪。文献[4]、[5]对目标核函数直方图进行 Kalman 滤波来更新目标模型, 从而使算法对外观变化、遮挡和目标尺度的变化等稳健, 但其尺度更新方法主要还是基于文献[2]的方法。Collin^[6]通过对采样加权图像进行 DOG 滤波引入额外尺度 σ (σ 为

基金项目: 江苏省自然科学基金项目(BK2004421)

收稿日期: 2006-10-13; 改回日期: 2006-11-10

第一作者简介: 钱惠敏(1980~), 女, 南京理工大学博士研究生。主要研究方向为图像处理、视频跟踪、行为理解。E-mail: qhmin0316@163.com

Gauss 函数参数),在尺度空间中进行 Mean-Shift 迭代,并收敛到 3 维空间(由 2 维空间位置和尺度 σ 构成)中的局部极值点,因此计算量大;并且使用 Epanechnikov 核等价于在尺度空间中求各尺度的平均。

本文将多尺度图像的信息量度量方法^[7]引入到运动目标跟踪中,根据信息量的变化自动选取跟踪窗的大小。文献[7]中的信息量度量方法表明尺度空间中图像的一类视觉细节信息与尺度成反比;类似地,观察者与物体距离的远近也可以由此信息量反映,距离远时对应的信息量大,距离近时对应的信息量小。因此在跟踪尺度变化的目标时可以根据信息量的变化来自动扩大或缩小跟踪窗口尺寸。本文首先在文献[7]的基础上,从微分算子的局部极值角度统一度量信息量的两类特征点;并提出了跟踪窗口尺度更新算法,改进了基于彩色直方图的 Mean-Shift 目标跟踪算法,使其能跟踪存在明显尺度变化的目标。

2 多尺度图像信息量度量

众所周知,人在不同的距离观察同一图像获得的感受是不一样的,如远距离仅看到图像轮廓,近距离时能看到更多的细节,这就是尺度效应。而视频跟踪中尺度不断变化的被跟踪目标也满足尺度效应。因此,如果能找到有效的信息量度量方法来区分目标的不同尺度,就可以根据信息量的变化来自动选择跟踪窗口的尺度。

2.1 图像的两类特征点

Marr 视觉理论^[8]认为,视觉处理得到的初始简图中要素的数量可以作为一个十分好的信息度量标准。基于这一思想,文献[7]定义了两类特征点,并用图像中两类特征点的数量来度量尺度空间中图像的信息量。

从微分算子的局部极值角度统一这两类特征点。假设在 2 维图像平面内,点 P 为图像 $f(x, y)$ 内某像素点,1~8 为点 P 的 8 邻接像素点 $N_8(P)$,如图 1 所示。现以 P 为极点,建立极坐标系,设极角为 $l, l \in L, L = \{k\pi/4, k = -4, \dots, 0, \dots, 3, k \in \mathbb{Z}\}$,则对 P 的 8 邻接像素点进行如下的数学描述:

$$f_l(x, y) = \begin{cases} f(x - \sin(l), y + \cos(l)) & l = \frac{n\pi}{2} \\ f(x - \sqrt{2}\sin(l), y + \sqrt{2}\cos(l)) & l = \frac{n\pi}{4}, l \neq \frac{n\pi}{2} \end{cases} \quad (1)$$

在极角 l 方向上,取步长 $\Delta h = 1$,定义 $f(x, y)$ 的

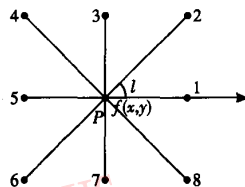


图 1 像素点 $P(f(x, y))$ 及 $N_8(P)$ 位置关系

Fig. 1 Location relationship of pixel $P(f(x, y))$ and $N_8(P)$

方向微分算子:

$$\nabla_l f(x, y) = \frac{f_l(x, y) - f(x, y)}{\Delta h} = f_l(x, y) - f(x, y) \quad (2)$$

下面给出文献[7]中两类特征点在式(2)定义的方向微分算子下的定义:

定义 1 设 2 维离散平面上点 $P(x, y)$, 若 $\forall l_1, l_2 \in L, L = \{k\pi/4, k = -4, \dots, 0, \dots, 3, k \in \mathbb{Z}\}$, 当 $l_1 - l_2 = m\pi, m = \pm 1$ 时, $\nabla_{l_1} f(x, y) \cdot \nabla_{l_2} f(x, y) > 0$, 则点 P 称为第 1 类特征点。

定义 2 设 2 维离散平面上点 $P(x, y)$, 若 $\forall l_1, l_2 \in L, L = \{k\pi/4, k = -4, \dots, 0, \dots, 3, k \in \mathbb{Z}\}$, 当 $l_1 - l_2 = m\pi, m = \pm 1$ 时, $\nabla_{l_1} (f_{l_1}(x, y) - f_{l_2}(x, y)) \cdot \nabla_{l_2} (f_{l_1}(x, y) - f_{l_2}(x, y)) > 0$, 则点 P 称为第 2 类特征点。

以企鹅图为例,给出图像对应的特征点图,图 2 为原图(penguin.jpg)及其对应的两类特征点图。可见,特征点图反映了原图的重要细节特征,特别地,特征点图能从视觉上反映感兴趣对象——企鹅的重要信息。



(a) 原图



(b) 第 1 类特征点



(c) 第 2 类特征点

图 2 原图(penguin.jpg)及其对应的特征点图

Fig. 2 Image and the relevant feature points figures

2.2 图像的信息量度量方法

由于第 1、2 类特征点都是视觉上的细节重要特征点,因此这两类特征点的集合从视觉上能反映图像的重要信息。下面根据这两类特征点来定义 2 维灰度离散图像 $f(x, y)$ 的信息量,而彩色图像的信息量等于 3 个颜色分量的独立信息量之和。

定义 3 设 2 维灰度离散图像 $f(x, y)$, 图像的第 1 类特征点个数为 I_1 , 第 2 类特征点个数为 I_2 , 则图像的信息量 I 定义为 I_1 与 I_2 之和。

在视频跟踪中, 用视频中的各帧图像构成一个多尺度图像空间, 尺度参数选为观察者与被观察场景的距离 d , 则尺寸小的跟踪目标对应大的尺度参数 d , 尺寸大的跟踪目标对应小的尺度参数 d 。文献[7]已证明了定义 3 给出的信息量度量方法在 Gauss 尺度空间中的有效性, 本文把它推广到由视频图像构成的多尺度图像空间中。

用以下实验来说明此推广的可行性。选择如图 3(a) 所示仙人球作为跟踪实验对象, 并从正上方对仙人球由远及近拍摄 45 帧图像; 然后对每帧图像内的仙人球手动划定矩形框使仙人球正好全部包含在框内, 并统计矩形框内信息量; 最后根据统计得的信息量, 给出随着仙人球尺寸的增大, 信息量的变化曲线。如图 3(b) 所示, 横轴表示当前帧的归一化矩形框高度 (即物体的高度), 纵轴是归一化的信息量值。由图可知, 当物体尺寸由小变大, 即物体与观察者距离由远及近时, 本文给出的信息量的值由小变

大。在此基础上, 分别求得前后两帧的信息量比值 S_t 及矩形框高度的比值 S_H , 并求得误差 $e = 1 + \lg(S_t) - S_H$, 误差曲线图如图 3(c) 所示。由图可知, 误差的值很小, 因此信息量的比值和物体尺寸存在如下的关系: $H \approx H_0(1 + \lg(S_t))$, 其中 H_0 指前一帧的物体尺寸, H 指当前帧的物体尺寸。

由此可见, 在由视频图像构成的尺度空间中, 对尺寸变化的被跟踪目标而言, 本文定义的信息量的大小与尺度参数 d 成反比, 尺度参数 d 较大, 即人眼从远处观察物体时, 对应的信息量小; 尺度参数 d 较小, 即人眼从近处观察物体时, 对应的信息量大。因此, 我们可以根据信息量大小的变化来自动更新跟踪窗口的大小。

3 自动选择跟踪窗尺度的目标跟踪算法

跟踪窗口的大小在目标跟踪的迭代过程中起关键作用。跟踪窗口过大, 跟踪窗内过多的背景区域会影响跟踪结果; 另外, 在多目标跟踪中, 一个跟踪窗内就可能会包含多个跟踪目标等等。跟踪窗口过小, 跟踪窗口内过少的目标信息不能有效跟踪目标, 从而可能使得目标丢失。因此, 在跟踪过程中自动更新跟踪窗口是改进跟踪算法的关键。根据前面描述的信息量与跟踪目标尺寸的关系, 本文在目标跟踪算法中加入了一个尺度更新项, 设计了跟踪窗口尺度自动更新算法; 并改进了基于颜色直方图的 Mean-Shift^[1] 算法, 得到了跟踪窗尺度自动更新的 Mean-Shift 算法, 既能跟踪尺寸逐渐减小的目标, 也能跟踪尺寸逐渐增大的目标。

跟踪窗口尺度更新算法的主要思想是根据实际视频中目标的尺度变化速度, 每隔 N 帧统计一次被跟踪目标信息量, 然后根据信息量的变化对跟踪窗口尺寸做相应更新。假设第 n 帧视频对应的跟踪窗口内图像的信息量为 I_1 , 对跟踪窗乘以 $1 \pm \alpha (0 \leq \alpha \leq 0.5)$ 后的区域的信息量为 I_2, I_3 (前者对应“+”, 后者对应“-”), 类似地, 对第 $n + N$ 帧统计得的信息量分别为 I_4, I_5, I_6 (此时的跟踪窗与第 n 帧的跟踪窗尺寸相等); 并设尺度变化比例为 S 。图 4 给出了信息量与窗口大小对应关系的示意图, 图中没有给出背景信息。由图可见, 通过判断最外层矩形框内信息量的变化可以进一步判断目标尺寸的增大或减小。

根据 1.2 节得到的信息量的比值和物体尺寸存在的关系, 以及实验验证, 本文得到如下的尺度更新

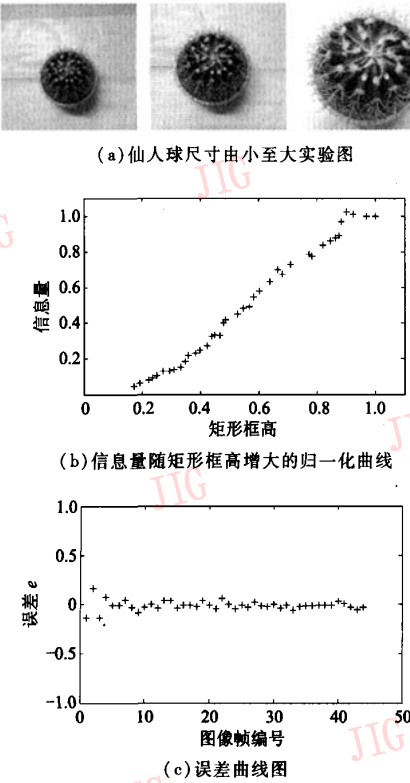


图 3 仙人球实验结果

Fig. 3 Result of experiment of cactus

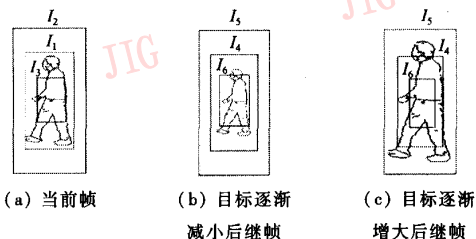


图 4 信息量与窗口大小的对应关系简单示意图

Fig. 4 Sketch map of relation between information and the size of window

方法。假如满足 $I_3 - I_2 \geq 0$, 判断目标尺度可能增加, 按式(3)更新尺度变化比例 S ; 否则, 判断目标尺度可能减小, 按式(4)更新尺度变化比例 S 。

$$S = \begin{cases} \lg\left(\beta \cdot \frac{I_3 - (I_2 - I_1)}{I_1}\right) & I_3/I_4 > I_2/I_1 \\ 0 & \text{其他} \end{cases} \quad (3)$$

$$S = \begin{cases} \lg(\beta \cdot I_4/I_1) & I_1 \cdot 0.95 > I_4 \\ 0 & \text{其他} \end{cases} \quad (4)$$

其中, 参数 β 的作用是消除背景对信息量的影响, 其值接近于 1, 背景越复杂与 1 的差值越大, 且目标尺寸增加时, $\beta \geq 1$; 目标尺寸减小时, $\beta \leq 1$ 。计算得当前帧的尺度变化比例 S 后, 就可以按式(5)更新当前帧跟踪窗口的高度 H 和宽度 W 。

$$\begin{aligned} H &= H(1 + S) \\ W &= W(1 + S) \end{aligned} \quad (5)$$

跟踪窗口尺度更新算法的步骤如下:

(1) 初始化帧数 $n = 0$, 尺度变化比例 $S = 0$, 设定跟踪窗口尺度更新间隔帧数 N ;

(2) 读取一帧图像, $n = n + 1$, 如果 $n = 1$, 手动划定跟踪窗口, 并计算初始高度 H 和宽度 W , 否则进入下一步;

(3) 如果帧数 n 是 N 的整数倍时, 进入步骤 4; 否则, $S = 0$, 进入步骤 5;

(4) 对第 n 帧图像的跟踪窗口对应区域及 $1 \pm \alpha$ 乘以跟踪窗口对应的区域分别统计信息量 $I_1(n)$,

$I_2(n), I_3(n)$, 判断目标尺度可能增加还是减小, 从而按式(3)和式(4)更新 S ;

(5) 按式(5)更新窗口的高度和宽度。若被跟踪视频未结束返回步骤 2; 否则结束算法。

4 实验结果

文献[2]分别以核窗宽确定的跟踪窗口、核窗宽做正负 10% 增量修正后确定的跟踪窗口进行 Mean-Shift 跟踪, 最后选择较大 Bhattacharyya 系数对应的核窗宽作为最佳核窗宽。此方法对尺寸逐渐减小的目标有很好地跟踪效果; 而本文方法对尺寸逐渐减小和增大的目标都有很好的跟踪效果。选取行人视频为实验对象来说明本文跟踪算法的有效性, 跟踪窗口都选取矩形窗, 且初始跟踪窗都在第 1 帧手动选择。

图 5 给出了本文方法对尺寸逐渐减小的行人的跟踪结果, 由图可见本文方法能自动选择逐渐减小的跟踪窗跟踪行人。实验中更新尺度时选择的参数为 $\alpha = 0.5, \beta = 0.97$ 。图 6 给出了采用文献[2]方法和采用本文方法的两种跟踪算法对目标逐渐增大的行人视频的跟踪结果比较。

由图可知, 文献[2]方法在跟踪尺寸逐渐增大的目标时, 跟踪窗口尺度不能随目标尺寸的增大逐渐自适应扩大; 而用本文方法能自适应扩大跟踪窗口, 实验中更新尺度时选择的参数为 $\alpha = 0.5, \beta = 1$ 。两种跟踪方法的差别在于: 文献[2]方法只是简单地跟踪窗尺寸做正负 10% 的增量, 在目标尺寸变化时没有对其模型做相应的变更; 而本文方法在跟踪过程中根据目标的视觉细节信息的变化来更新跟踪窗口, 从而也更新了目标模型的核函数直方图, 得到了有效的跟踪结果。另外, 本文方法中对视觉信息的统计的计算复杂度较小, 因为信息统计部分只有简单的比较和加法运算, 并且每隔 N 帧才对跟踪窗口区域(它是整幅图像的一小部分)做一次信息

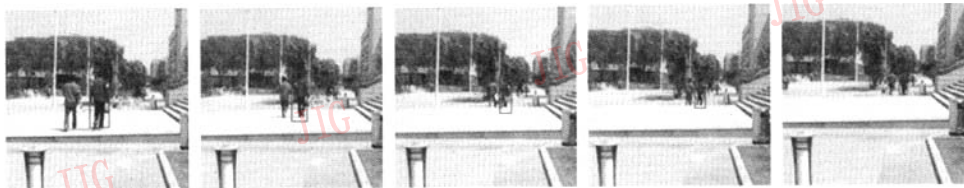


图 5 本文方法对逐渐减小目标的跟踪结果

Fig. 5 Tracking results for decreasing object of the proposed method

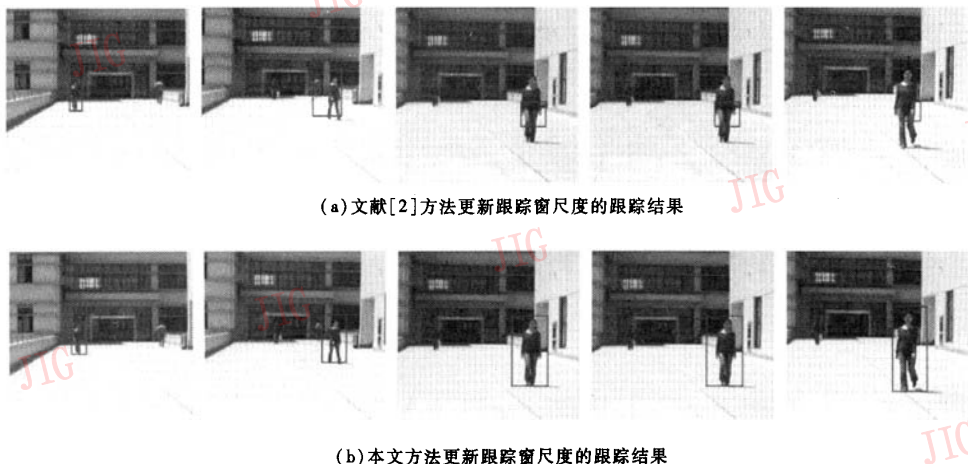


图6 Mean-Shift 算法对逐渐增大目标的跟踪结果

Fig.6 Tracking results for increasing object of methods based on Mean-Shift

量统计,能够满足实时性要求。

这里仅给出了将行人作为实验对象时的几个实验结果,而实验过程中选取了20个不同背景下的行人视频,另外还选取了其余的实验对象(如骑摩托车、自行车的人,人脸等),大量实验结果表明本文方法对各种实验对象都能自动选择合适的跟踪窗口大小。

5 结 论

跟踪窗口尺度固定不变的跟踪算法,在跟踪尺寸存在明显变化的目标时常会产生各种问题,如跟踪逐渐减小的目标时,一个跟踪窗口内可能会包含多个目标;跟踪逐渐增大的目标时,可能会导致目标的丢失;而目前还没有公认的方法来解决跟踪窗口的自动选取问题。本文将尺度空间中的信息量度量方法^[7]引入到运动目标跟踪中,可以根据跟踪窗内信息量的变化跟踪尺寸变化的目标。本文算法仍按原有算法在2维空间迭代,仅对跟踪窗口的尺寸更新,因此该算法有计算上的优势;并且该算法对光照变化是鲁棒的。本文跟踪窗口尺寸的选取依赖信息量的变化,而信息量仅与2.1节定义的两类特征点有关,特征点都是式(2)定义的方向微分算子的局部极值,因此光照的变化并不影响信息量的变化趋势,也就不会影响算法的有效性。另外,本文的尺度更新算法具有通用性,它也可用来改进其他跟踪算法。进一步的研究方向为寻找参数 β 与背景复杂度之间的明确的统计关系。

参考文献 (References)

- Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift[A]. In: IEEE International Proceeding on Computer Vision and Pattern Recognition[C], Stoughton: Printing House, 2000, (2):142~149.
- Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5):564~575.
- Peng Ning-song, Yang Jie, Liu Zhi, et al. Automatic selection of kernel-bandwidth for Mean-Shift object tracking[J]. Journal of Software, 2005, 16(9):1542~1550. [彭宁嵩, 杨杰, 刘志等. 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.]
- Peng Ning-song, Yang Jie, Liu Zhi. Mean shift blob tracking with kernel histogram filtering and hypothesis testing[J]. Pattern Recognition Letters, 2005, 26(5):605~614.
- Peng Ning-song, Yang Jie, Zhou Da-ke, et al. Mean-Shift tracking with adaptive model update mechanism[J]. Journal of Data Acquisition & Processing, 2005, 20(2):125~129. [彭宁嵩, 杨杰, 周大可等. Mean-Shift跟踪算法中目标模型的自适应更新[J]. 数据采集与处理, 2005, 20(2):125~129.]
- Collins R T. Mean-Shift blob tracking through scale space[A]. In: IEEE International Conference on Computer Vision and Pattern Recognition[C], Baltimore: Victor Graphics, 2003, (2):234~240.
- Wang Zheng-yao, Cheng Zheng-xing, Tang Shao-jie. Information measures of scale-space based on visual characters[J]. Journal of Image and Graphic, 2005, 10(7):922~928. [王郑耀, 程正兴, 汤少杰. 基于视觉特征的尺度空间信息量度量[J]. 中国图象图形学报, 2005, 10(7):922~928.]
- Milan Sonka, Vaclav Hlavac, Roger Boyle. Image Processing, Analysis, and Machine Vision(Second Edition). London: Brooks/Cole Publishing, 1999. [Milan Sonka等著. 图像处理、分析与机器视觉(第二版)[M]. 艾海舟, 武勃等译. 北京:人民邮电出版社, 2003.]