

文章编号: 1001-0920(2012)03-0399-04

基于 SIFT 特征的目标多自由度 mean-shift 跟踪算法

董蓉, 李勃, 陈启美

(南京大学电子科学与工程学院, 南京 210093)

摘要: 传统的 mean-shift 跟踪算法不能跟踪目标的旋转、缩放运动, 且常常因此造成定位不准. 鉴于此, 将尺度不变特征变换(SIFT)特征检测融入到 mean-shift 跟踪过程, 提出 SIFT 特征点的尺度变化与目标的尺度变化成正比, 特征点主方向变化与目标旋转角度一致, 给出了基于 SIFT 特征的自适应目标尺度、方向计算方法, 且利用带方向、可变带宽的椭圆核改进传统的 mean-shift 跟踪方法. 实验表明, 该算法能够较好地跟踪目标的旋转、缩放运动, 定位也更准确.

关键词: 尺度不变特征变换; 均值漂移; 特征点; 尺度空间; 目标跟踪

中图分类号: TP391

文献标识码: A

Multi-degree-of-freedom mean-shift tracking algorithm based on SIFT feature

DONG Rong, LI Bo, CHEN Qi-mei

(School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China. Correspondent: LI Bo, E-mail: liboee@nju.edu.cn)

Abstract: Traditional method can't follow the scale and orientation change of the object, which usually causes inaccurate localization. Therefore, this paper combines SIFT(scale invariant feature transform) feature detecting with the mean-shift tracking method, which proposes that the scale change of the SIFT keypoint is proportional to that of the object, and the dominant direction change of the SIFT keypoint is the same as that of the object. The algorithm for adaptively constructing the scale and the orientation information from SIFT features of the object is presented. And a variable-bandwidth and orientation ellipse kernel is used to improve the traditional mean-shift method. The experimental results show that the proposed algorithm provides good tracking of the scale and orientation change of the object, and the localization is more accurate.

Key words: scale invariant feature transform; mean-shift; keypoint; scale space; object tracking

1 引言

运动目标跟踪是智能视频监控的关键步骤, 在交通监测、公共安全防等系统中均有广泛应用. 跟踪时不仅需要获得准确的目标质心轨迹, 还希望得到目标缩放、旋转等参数, 即进行多自由度跟踪. mean-shift 算法由文献 [1] 首次提出, 是一种无参密度梯度上升算法, 用于寻找概率密度函数的极值. [2] 将其应用于目标跟踪, 以目标色彩直方图作为特征, 以巴氏系数作为相似性度量, 以 mean-shift 迭代方式定位目标, 实时性、有效性较好, 获得了广泛推崇. 但传统的 mean-shift 算法采用固定的核函数带宽, 不能自适

应跟踪目标的缩放、旋转等多自由度变化, 且容易导致定位不准. [3] 在二维空间的基础上增加一维尺度空间, 进行三维 mean-shift 迭代求取极值点, 计算量较大, 且本质上与 [2] 相同; [4-6] 用带宽矩阵替代单一的带宽变量后再进行迭代, 可以同时求取尺度、方向信息, 本质上与 [2] 相同; [7] 提出了一种后向跟踪、形心配准的尺度自动选取算法, 用于解决目标不断变大时 mean-shift 定位不准的问题, 但通常目标尺度变化未知, 如果目标尺度正当变小, 则用文中算法将导致后向跟踪定位不准, 从而给前向跟踪带来根本性误差.

本文在 mean-shift 算法中引入尺度不变特征变

收稿日期: 2010-09-27; 修回日期: 2010-12-17.

基金项目: 江苏省科技支撑计划项目(BE2009667); 江苏省自然科学基金项目(BK2010366); 中兴通讯高校合作基金项目.

作者简介: 董蓉(1984—), 女, 博士生, 从事机器视觉、匹配跟踪的研究; 陈启美(1949—), 男, 教授, 博士生导师, 从事网络通信、信号处理等研究.

换 (SIFT) 特征检测算法^[8], 利用 SIFT 特征检测时获得的特征尺度和特征点主方向信息跟踪目标缩放、旋转运动, 同时调整 mean-shift 算法中椭圆核函数的带宽和方向. 另外, mean-shift 算法为 SIFT 特征检测提供了预测位置, 可以提高效率.

2 mean-shift 跟踪算法及其问题

mean-shift 跟踪算法^[2]中目标中心位置由下式迭代收敛得到:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}. \quad (1)$$

其中: $g(x) = -k'(x)$, $k(x)$ 为所选取的核函数; h 为核函数带宽; x_i 为以 y_0 为中心的候选区域内 n_h 个像素的坐标; w_i 为权重, 且有

$$w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y_0)}} \delta(b(x_i) - u); \quad (2)$$

q_u 为目标模板的特征概率分布, 有

$$q_u = C \sum_{i=1}^n k(\|x_i^*\|^2) \delta(b(x_i^*) - u),$$

$$C = 1 / \sum_{i=1}^n k(\|x_i^*\|^2); \quad (3)$$

$p_u(y_0)$ 为当前帧中以 y_0 为中心的候选区域所构建的特征概率分布, 有

$$p_u(y_0) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) \delta(b(x_i) - u),$$

$$C_h = 1 / \sum_{i=1}^{n_h} k\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right). \quad (4)$$

由式 (1) 可见, mean-shift 算法本身并不具备 h 自动调节的功能, 同时算法中未考虑目标方向的旋转, 因而不能跟踪目标的多自由度变化. 此外, 定位准确性也将受到影响: 当所选择的 h 过大时, 提取的候选区域特征概率分布将包含背景干扰; 反之, 则只能获得目标局部的特征概率分布. 运动物体往往具有长短轴, 通常情况下需采用椭圆核, 如果椭圆核始终保持水平, 则当物体发生旋转时定位必受干扰.

3 SIFT 特征信息挖掘

SIFT 特征对于图像旋转、缩放、仿射变换、光照变化等均保持一定不变性, 其独特性、稳定性较好, 较小的物体也能产生较多的 SIFT 特征点, 适合于匹配跟踪. SIFT 算法依据其提出的 DOG (difference of gaussian) 算子^[8]提取图像在不同尺度上的特征点, 利用具有不同均方差参数 σ 的高斯函数 $G(x, y, \sigma)$ 将一帧图像扩展到尺度空间的一系列图像. 特征点通过对相邻尺度两幅图像的差分 $D(x, y, \sigma)$ 求局部极值得到,

即

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma))I(x, y),$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \quad (5)$$

其中 $I(x, y)$ 为原图像. 经过 SIFT 特征提取, 可以获得特征点坐标位置和特征点描述符信息, 因 SIFT 特征定位精度高、稳定性好, 常用于图像配准^[9]和图像检索^[10]. 另外, 特征检测过程中还能得到特征点所在尺度 (即 σ 参数) 以及特征点主方向这两个中间产物.

对于视频序列中的一帧图像, SIFT 检测中所谓的尺度参数实际代表该特征处于哪一个细节层次上: 利用较大的尺度参数进行高斯滤波将产生较为模糊的图像, 在这个层次上检测到的特征极值点通常是物体轮廓、概貌上的特征; 而较小的尺度参数则保留较多的细节信息, 这个层次上检测到的特征点往往是图像局部细微处的特征. 另一个值得关注的现象是, 对于视频序列中的一个运动目标, 当摄像机与其距离较远时, 目标成像较小, 此时利用较小的 σ 参数即可获得目标的轮廓概貌特征, 即轮廓概貌特征处于较小的尺度参数上; 随着距离减小, 目标成像越来越大, 细节信息也越来越多, 需要较大的 σ 参数才能获得轮廓概貌特征, 相对之前, 轮廓概貌特征位于一个更大的尺度参数上, 不仅是轮廓概貌特征, 随着距离拉近, 目标上所有特征所处的尺度都相应增大. 因此, 视频序列中目标的尺度变化与目标上某一特征所在尺度的变化应存在一定联系, 这种联系有助于利用 SIFT 特征点尺度变化进行目标尺度自适应检测.

SIFT 特征点描述符由特征点邻域内采样点的梯度信息组合而成. 如果采用 4×4 的子区域划分方式, 对每个子区域构建 8 个方向的梯度幅度信息, 则最终获得 128 维的特征向量. 特征点邻域梯度信息同时也用于构建特征点主方向, 通过构建特征点邻域内所有像素的梯度方向直方图, 寻找其峰值作为该特征点的主方向. 确定主方向后, 特征描述符中的梯度方向信息将以此为标杆, 从而保证 SIFT 特征的旋转不变性. 可见, 对于刚性物体, 主方向的变化其实也是物体旋转的标志.

4 目标多自由度 mean-shift 跟踪算法

文献 [11-12] 等曾利用 SIFT 特征点匹配的方法进行跟踪, 但仅利用 SIFT 特征描述符向量以及特征点位置信息, 忽略了特征点所在尺度和特征点主方向信息. 本文将特征点尺度和主方向信息用于目标缩放、旋转运动的多自由度跟踪, 同时改进 mean-shift 核函数尺度和方向自适应问题.

4.1 尺度自适应

实验证明, 目标上特征点所在尺度的变化反映了

相应的物体尺度变化, 如图 1 所示. 由图 1 可见, 物体尺度变化与特征点尺度变化成正比, 即

$$S_{\text{cur}}/S_{\text{model}} = \sigma_C/\sigma_M. \quad (6)$$

其中: 物体横向半径长度为物体尺度 S ; σ_C, σ_M 分别为 SIFT 特征检测中得到的两个匹配点 C, M 各自所在的尺度.

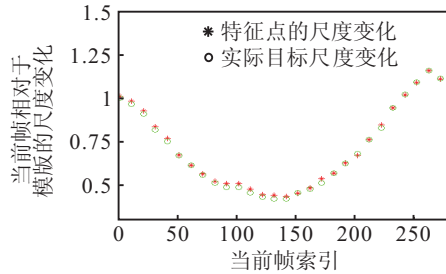


图 1 实际目标尺度变化与特征点尺度变化的关系

实验采用如图 2 所示的视频序列, 其中物体尺度经历由大变小再变大的过程, 物体上的箭头表示检测到的 SIFT 特征点, 箭头方向表示特征点主方向, 箭头长度是统一缩放后的特征点尺度.

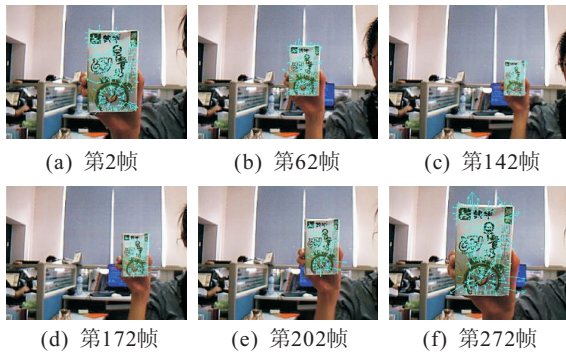


图 2 从视频序列中截取的 6 帧代表图像及 SIFT 检测结果

实验中利用所有匹配点尺度变化的统计值, 以获得稳定的目标尺度变化, 即

$$S_{\text{cur}} = S_{\text{model}} \times \frac{1}{n} \sum_{i=1}^n \frac{\sigma_{C_i}}{\sigma_{M_i}}, \quad (7)$$

其中 n 为所有匹配上的特征点数目.

4.2 方向自适应

对于匹配上的两个特征点 C 和 M , 设 SIFT 特征检测过程中得到其主方向分别为 R_C, R_M , 则物体旋转的角度 Δo 可通过两者差分得到, 即

$$\Delta o = R_C - R_M. \quad (8)$$

依据单个特征点得到的计算结果可能存在一定误差, 通过多个特征点主方向差分的统计结果得到

$$\overline{\Delta o} = \frac{1}{n} \sum_{i=1}^n (R_{C_i} - R_{M_i}), \quad (9)$$

其中 n 为当前帧和模板帧中所有匹配上的特征点数目. 当前帧物体的方向为

$$O_{\text{cur}} = O_{\text{model}} + \overline{\Delta o}. \quad (10)$$

4.3 抗噪处理

SIFT 特征提取与匹配的过程中可能出现误匹配, 导致依据该特征点计算出的尺度比例 σ_C/σ_M 或 Δo 偏离真实值, 成为干扰噪声. 由于 SIFT 特征的高独特性和稳定性, 这种干扰点偶尔才会出现且数目非常少, 同时由于实时视频处理的需求, 采用一种简便的自适应阈值控制法来剔除这些干扰点. 对于 $X_i = \sigma_{C_i}/\sigma_{M_i}, i = 1, 2, \dots, n$, 如果

$$|X_i - \bar{X}| > \sigma_X, \quad (11)$$

则剔除 X_i , 并使 $n = n - 1$, 再利用式 (7) 和 (9) 计算目标尺度和方向变化, 其中 \bar{X} 和 σ_X 分别为 X_i 的均值和均方差. 也可以采用各种更为高级但也更为复杂耗时的聚类算法^[13]替代式 (7) 和 (9), 以剔除可能存在的干扰; 或者在进行 SIFT 特征匹配时引入抗干扰机制, 比如在匹配每一个特征点描述符的同时检查其邻域特征点是否也匹配, 以降低误匹配率. 但实验证明, 采用式 (7), (9), (11) 已经能获得较好的效果, 采用更为复杂的算法性价比不高.

4.4 算法步骤

前文获得的尺度、方向变化参数, 一方面可以跟踪目标的缩放、旋转运动, 另一方可以调整 mean-shift 算法中核函数的带宽和方向. 同时, mean-shift 算法可用于目标的初步定位, 为 SIFT 特征检测提供初步方位和区域, 避免对大范围图像进行 SIFT 特征检测浪费时间, 也可以减少误匹配. 具体步骤如下:

Step 1: 利用上一帧获得目标尺度 S_{prev} 和方向 O_{prev} , 构建带方向的椭圆核. 假设目标模板的方向 $O_{\text{model}} = 0$, 模板横纵尺度分别为 h_x, h_y , 以物体横向半径长度代表物体尺度, $S_{\text{model}} = h_x$, 概率密度分布计算式改写为

$$p_u(y_0) = C_h \sum_{i=1}^{n_h} k(\|T(y_0 - x_i)\|^2) \delta(b(x_i) - u),$$

$$T = \frac{1}{S_{\text{prev}}} \begin{bmatrix} 1 & 0 \\ 0 & h_x/h_y \end{bmatrix} \begin{bmatrix} \cos O_{\text{prev}} & \sin O_{\text{prev}} \\ -\sin O_{\text{prev}} & \cos O_{\text{prev}} \end{bmatrix},$$

$$C_h = 1 / \sum_{i=1}^{n_h} k(\|T(y_0 - x_i)\|^2), \quad (12)$$

用 mean-shift 算法进行初步定位.

Step 2: 根据 mean-shift 定位结果和 $S_{\text{prev}}, O_{\text{prev}}$ 值, 确定 SIFT 特征检测区域, 将检测到的特征点与模板中的特征点匹配, 利用式 (7), (9)~(11) 计算新的尺度 S_{cur} 和方向 O_{cur} .

Step 3: 利用 S_{cur} 和 O_{cur} 再次进行 mean-shift 计算重新定位, 概率密度分布计算类似式 (12).

5 实验分析

实验在 Matlab 7.1 环境下进行. mean-shift 核函

数采用高斯函数, SIFT 特征检测采用 128 维的特征向量. 视频序列 1 为普通聊天摄像头获得, 分辨率为 320×240 , 帧率 30 帧/s; 视频序列 2 为 CIF 格式录像, 分辨率为 352×288 , 帧率 24 帧/s. 图 3(a) 和图 4(a) 展示了本文算法效果, 其中箭头表示跟踪过程中检测到的 SIFT 特征点, 椭圆表示跟踪到的目标. 另外, 将本文算法与经典的文献 [2] 算法进行对比, 结果如图 3(b) 和图 4(b) 所示. 图中, 采用 $h - dh, h, h + dh$ 三个带宽分别进行 mean-shift 迭代, 选取巴氏系数较大的为当前带宽, dh 为步长.



(a) 本文算法实验结果



(b) 文献[2]算法实验结果

图 3 视频序列 1 跟踪效果对比 (第 2, 172, 312 帧)



(a) 本文算法实验结果

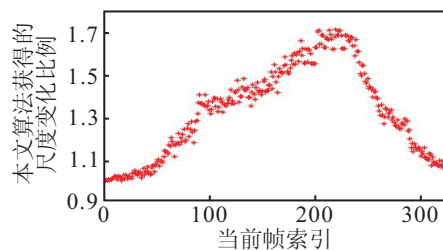


(b) 文献[2]算法实验结果

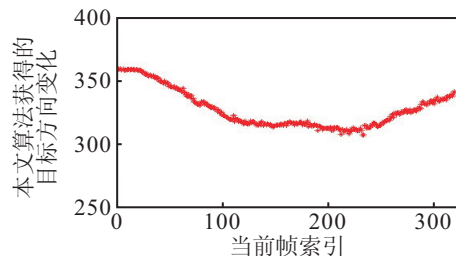
图 4 视频序列 2 跟踪效果对比 (第 22, 112, 212 帧)

视频序列 1 中花露水瓶子先向右旋转同时向前放大, 然后向后缩小并向左旋转. 由图 3(a) 可见, 无论是旋转角度还是目标尺度, 本文算法都能准确计算跟踪. 经典算法由于没有考虑方向变化, 在此次跟踪中几乎失效, 如图 3(b) 所示. 另外, 以此典型视频序列为例, 图 5(a) 和图 5(b) 给出了本文算法跟踪目标尺度和方向的数据结果, 图 5(c) 给出了文献 [2] 算法跟踪目标尺度的数据结果 (该算法不能进行方向跟踪), 其完全背离真实目标运动情况. 相比之下, 本文算法结果与真实目标运动情况吻合度较高, 这也验证了第 2 节对于传统 mean-shift 算法所存在问题的描述.

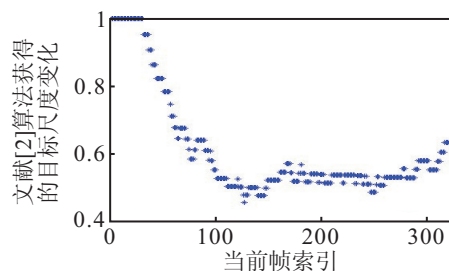
视频序列 2 是某十字路口的交通监控录像, 一辆公交车正左转弯通过, 期间尺度和方向均发生变化. 由图 4(a) 可见, 尽管图像质量不佳, 但仍能检测出较



(a) 本文算法获得的目标尺度变化



(b) 本文算法获得的目标方向变化



(c) 文献[2]算法获得的目标尺度变化

图 5 跟踪数据结果对比

多 SIFT 特征点, 在此基础上的尺度、方向计算也较准确. 图 4(b) 中文献 [2] 算法的跟踪逐渐偏离真实目标.

6 结 论

本文将 SIFT 特征检测融入到 mean-shift 跟踪算法框架中, 提出了一种位移、旋转、缩放多自由度的跟踪算法, 解决了原有 mean-shift 跟踪算法不能自适应调整尺度和方向且定位不准的问题. 利用 SIFT 特征检测算法高度的稳定性、独特性、高精度以及 mean-shift 算法的高效率, 本文算法取得了较好结果. SIFT 特征检测在灰度域进行, 实际上代表的是图像纹理、结构的信息, 而 mean-shift 跟踪算法采用目标的颜色直方图, 二者结合可以完整地描述目标. 下一步工作拟将融合 SIFT 特征点的位置信息, 用于解决遮挡干扰问题. 另外, 为了保证 SIFT 特征模板匹配时能够获得足够多的匹配对以确保计算结果稳定, 模板更新工作也较为重要.

参考文献(References)

- [1] Fukunaga K, Hostetler L. The estimation of the gradient of a density-function with applications in pattern recognition[J]. IEEE Trans on Information Theory, 1975, 21(1): 32-40.