

基于复杂特征融合的改进 mean shift 目标跟踪

杨欣^{1,2,3}, 费树岷³, 李刚¹, 周大可¹

(1. 南京航空航天大学自动化学院, 南京 210016; 2. 光电控制技术重点实验室, 河南 洛阳 471000; 3. 东南大学自动化学院, 南京 210096)

摘要: 提出一种融合 Gabor 小波纹理特征与颜色特征的改进 mean shift 目标跟踪算法. 首先, 提取移动目标的颜色特征和纹理特征直方图; 其次, 基于 mean shift 算法定义融合相似度系数, 对特征空间进行融合并得出目标中心位置; 再次, 通过定义特征自适应系数来融合基于颜色和纹理特征的目标位置; 最后, 对上述结果进行处理, 得到目标最终位置. 实验结果表明, 该算法在跟踪目标存在变形、噪声、遮挡时能够得到比较理想的跟踪效果.

关键词: 目标跟踪; 均值转移算法; Gabor 小波; 特征融合

中图分类号: TP391.41

文献标志码: A

Improved mean shift tracking algorithm based on complicated feature fusion

YANG Xin^{1,2,3}, FEI Shu-min³, LI Gang¹, ZHOU Da-ke¹

(1. College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; 2. Key Laboratory of Photoelectric Control Technology, Luoyang 471000, China; 3. School of Automation, Southeast University, Nanjing 210096, China. Correspondent: YANG Xin, E-mail: yangxin@nuaa.edu.cn)

Abstract: A novel improved mean shift algorithm based on texture and color features is proposed. Firstly, the moving object histograms of the color feature and texture feature are got respectively. Secondly, the fusion similarity coefficients are defined to fuse the different feature space, and the central location of moving object is calculated based on MS tracking. Thirdly, according to the color feature and texture feature, the object location is updated by using the feature adaptive coefficients. Finally, the above results are processed to get the final object location. Experimental results show that the proposed tracking algorithm exhibits good results in the presence of noise, deformation and occlusion.

Key words: object tracking; mean shift algorithm; Gabor wavelet; feature fusion

0 引言

当前, 目标跟踪技术已得到了广泛关注. 过去几十年中, 大量目标跟踪算法被相继提出, 其中 mean shift 算法属于经典算法, 是由 Fukunaga 等^[1]于 1975 年提出的并迅速引起了人们的重视. 徐琨等^[2]提出了基于自适应颜色空间算法, 具有较强的鲁棒性. Polate 等^[3]提出了结合运动特征和颜色跟踪目标的方法. Adam 等^[4]则把待测目标进行了分块处理, 对每一块小的区域进行搜索, 再对每个小区域最佳匹配位置进行加权处理, 最终根据投票原则确定目标所在位置. Ojala 等^[5]通过融合图像的灰度和纹理特性来描述目标以实现目标的跟踪. Zivkovic 等^[6]根据人类视觉的多信息融合特点, 提出融合算法来实现跟踪. 王永忠

等^[7]提出了融合多特性核函数跟踪法, 通过融合边缘特征和颜色特征来描述目标.

为了进一步提升 mean shift 算法在遮挡和干扰情况下的跟踪能力, 受上述文献的启发, 本文提出一种结合纹理特征和颜色特征的改进 mean shift 算法. 实验结果表明, 该方法能较好地适应复杂背景视频序列, 特别是颜色干扰和纹理干扰, 从而改进了经典算法的不足.

1 特征提取与描述

1.1 基于颜色特征的目标描述

Mean shift 算法是基于颜色特征来描述目标形状, 对 RGB 颜色空间中每个子空间进行划分, R、G、B 空间分为 k 个相等区间 bin, 这些 bin 构成了特征空

收稿日期: 2013-05-01; 修回日期: 2013-06-26.

基金项目: 国家自然科学基金项目(60905009, 61172135, 61101198); 航空基金项目(20115152026).

作者简介: 杨欣(1978—), 男, 副教授, 从事模式识别、图像处理的研究; 费树岷(1961—), 男, 教授, 博士生导师, 从事非线性控制、模式识别等研究.

间, 其中 bin 的个数为 $m_c = k^3$. 针对视频序列中的目标区域的像素点, 计算该特征空间中所有特征值的概率, 便可得到目标特征模型, 即

$$\hat{q}_{uc} = C \sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right) \delta[b_c(x_i) - uc]; \quad (1)$$

$$k(x) = \begin{cases} \frac{1}{2} C_d^{-1} (d+2)(1 - \|x\|^2), & \|x\| \leq 1; \\ 0, & \text{else.} \end{cases} \quad (2)$$

其中: $uc = 1, 2, \dots, m_c$; $\{x_i\}_{i=1,2,\dots,n}$ 表示目标区域中 n 个元素; x_0 为目标中心, 归一化后目标模型中心像素位置为 0; 含有字母 C 的所有形式都表示为归一化常数; n 为候选目标区域中像素总数; 单调递减凸函数 $k(\cdot)$ 是定义每个像素权值大小的核函数轮廓函数; $\|(x_0 - x_i)/h\|$ 的作用是为了消除不同大小目标计算时的影响, h 为核函数带宽; $\delta(\cdot)$ 为 Kronecker delta 函数, 其作用是判断目标区域中像素是否属于第 u 个 bin, 相等时为 1, 否则为 0; $b_c: R^2 \rightarrow \{1, 2, \dots, m_c\}$ 为像素点到像素特征的映射. 在序列图像第 2 帧及以后的每帧中可能包含目标的区域为候选区域, 候选模型特征为

$$\hat{p}_{uc}(y) = C_c \sum_{i=1}^n k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \delta[b_c(x_i) - uc], \quad (3)$$

其中 y 为候选目标区域中心位置.

上面所述的即为经典 mean shift 算法中基于颜色特征建模的一般步骤. 为了增强算法的跟踪效果, 本文将纹理特征融合到 mean shift 算法中, 其主要思路是颜色特征及纹理特征具有一定程度的互补性, 当其中的任一特征受到干扰带来跟踪干扰时, 另一个特征可以实现补充功能, 从而使得跟踪的准确率大大提高.

1.2 基于纹理特征的目标描述

纹理特征展现了图像中像素的灰度级空间分布特性, 是检测图像的一种重要参考属性. 本文选用 Gabor 小波对图像进行纹理特征的提取, 并结合统计直方图来适应 mean shift 算法. 一个二维 Gabor 函数空间表达的一般形式为

$$h_{mn}(x, y) = \frac{1}{2\pi\sigma^2} \exp \left(-\frac{1}{2\sigma^2} [(x \cos \theta + y \sin \theta)^2 + (y \cos \theta - x \sin \theta)^2] \right) \exp[2\pi j F(x \cos \theta + y \sin \theta)]. \quad (4)$$

其中: m, n 分别代表不同的尺度变换和角度旋转, F 为 Gabor 函数中心频率, σ 为尺度因子.

设 Gabor 滤波输入图像为 $I(x, y)$, 具体步骤如下.

1) 卷积运算. 需处理的图像 $I(x, y)$ 与 Gabor 图像的实部、虚部分别进行卷积运算, 再按照下式进行处理, 得到滤波后的图像 h'_{mn} :

$$h'_{mn} = ([h_{mn}(x, y)_r * I(x, y)]^2 + [h_{mn}(x, y)_i * I(x, y)]^2)^{\frac{1}{2}}. \quad (5)$$

2) 建立纹理特征直方图. 首先由下式得出 G_{mn} :

$$G_{mn} = \frac{1}{\sqrt{M}} \sum_{ROI} h'_{mn}(x, y); \quad (6)$$

然后得出像素灰度值分布信息, 一般分为 6~8 个相等的 bin 区间 (区间个数用 j 表示), 这些区间形成了特征空间, 特征值的个数为 $m_t = j^{s \cdot d}$, 统计 j 个区间的每个区间中对应的特征值数目; 最终完成纹理直方图的描述.

式 (6) 中: $M = sd$, s 表示 Gabor 滤波尺度的大小, d 表示滤波方向, 综合计算复杂度的考虑, 一般选择 3 组不同大小和不同尺度, 方向的选择一般为 0、 $\pi/4$ 、 $\pi/2$, 尺度大小分别为 3、4、5; ROI 表示感兴趣的区域, 即初始化时所标定的目标区域. 则基于纹理特征的目标特征模型为

$$\hat{q}_{ut} = C_1 \sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right) \delta[b_t(x_i - ut)], \quad (7)$$

基于纹理特征的候选模型为

$$\hat{p}_{ut}(y) = C_t \sum_{i=1}^n k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \delta[b_t(x_i) - ut]. \quad (8)$$

其中: $ut = 1, 2, \dots, m_t$; 其他变量符号的解释可参考上文, 意义与基于颜色的特征模型一致.

2 本文算法

2.1 融合相似度系数

在对目标区域和候选区域进行相应的特征描述之后, 利用融合的相似性函数来度量候选与目标区域的相似性, 利用对当前帧的搜索结果, 确定相似性函数的最大值, 得到目标位置.

本文提出了融合相似度系数 ρ_{csc} (combine similarity coefficient) 的概念, 即定义两个系数乘积的组合形式, 具体定义如下:

$$\rho_{csc} = \Delta \cdot \rho(\hat{p}_{uc}(y), \hat{q}_{uc}) \cdot \rho(\hat{p}_{ut}(y), \hat{q}_{ut}) = \Delta \cdot \sum_{uc=1}^{m_c} \sqrt{\hat{p}_{uc}(y)\hat{q}_{uc}} \sum_{ut=1}^{m_t} \sqrt{\hat{p}_{ut}(y)\hat{q}_{ut}}. \quad (9)$$

其中: Δ 为敏感系数, 取值区间为 (0, 1); ρ 为巴氏系数. 从数学定义可看出, 融合的相似度系数与巴氏系数大小成正比, 所以该值越大, 相似度越高, 跟踪效果越好.

计算式 (9) 中的 $\hat{p}_{uc}(y_0)$ 和 $\hat{p}_{ut}(y_0)$, 再将 (9) 中的两个巴氏系数分别在初始状态 y_0 处泰勒展开, 即

$$\rho(\hat{p}_{uc}(y), \hat{q}_{uc}) = \frac{1}{2} \sum_{uc=1}^{m_c} \sqrt{\hat{p}_{uc}(y_0)\hat{q}_{uc}} + \frac{C_c}{2} \sum_{i=1}^{n_h} \omega_{ic} k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right), \quad (10)$$

$$\rho(\hat{p}_{ut}(y), \hat{q}_{ut}) = \frac{1}{2} \sum_{ut=1}^{m_t} \sqrt{\hat{p}_{ut}(y_0) \hat{q}_{ut}} + \frac{C_t}{2} \sum_{i=1}^{n_h} \omega_{it} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right). \quad (11)$$

其中

$$\omega_{ic} = \sum_{uc=1}^{m_c} \sqrt{\frac{\hat{q}_{uc}}{\hat{p}_{uc}(y_0)}} \delta[b_c(x_i) - uc],$$

$$\omega_{it} = \sum_{ut=1}^{m_t} \sqrt{\frac{\hat{q}_{ut}}{\hat{p}_{ut}(y_0)}} \delta[b_t(x_i) - ut],$$

C_c 、 C_d 为颜色和纹理归一化系数。由式(9)~(11), 有

$$\rho_{csc} \approx \frac{1}{4} \sum_{uc=1}^{m_c} \sqrt{\hat{p}_{uc}(y_0) \hat{q}_{uc}} \sum_{ut=1}^{m_t} \sqrt{\hat{p}_{ut}(y_0) \hat{q}_{ut}} + C' \sum_{i=1}^n \omega_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right). \quad (12)$$

其中: C' 为颜色及纹理归一化系数的数学组合, 权重系数 ω_i 为

$$\omega_i = \sum_{ut=1}^{m_t} \sqrt{\hat{p}_{ut}(y) \hat{q}_{ut}} \sum_{uc=1}^{m_c} \sqrt{\frac{\hat{q}_{uc}}{\hat{p}_{uc}(y_0)}} \delta[b_c(x_i) - uc] + \sum_{uc=1}^{m_c} \sqrt{\hat{p}_{uc}(y) \hat{q}_{uc}} \sum_{ut=1}^{m_t} \sqrt{\frac{\hat{q}_{ut}}{\hat{p}_{uc}(y_0)}} \delta[b_t(x_i) - ut]. \quad (13)$$

基于均值转移算法, 通过对融合相似度系数函数 ρ_{fsc} 求最大值, 可以得到目标位置

$$y_{fsc} = \frac{\sum_{i=1}^n x_i \omega_i g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \omega_i g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}, \quad (14)$$

其中 $g(x) = -k'(x)$ 为核密度估计。

2.2 特征自适应系数

基于相似度函数可以求取目标区域的当前帧位置。为了提高跟踪的精度, 本文再作一次跟踪处理。相似度函数依旧选取巴氏系数, 分别对颜色特征和 Gabor 纹理特征进行一次求解巴氏系数过程; 然后基于 mean shift 算法, 为了获取最高相似度或者最大巴氏系数, 对当前帧的目标位置进行更新处理。当前帧目标位置为 y_0 , 更新后的目标新位置分别为基于颜色空间的位置 y_{1c} 和基于纹理空间的位置 y_{1t} , 即

$$y_{1c} = \frac{\sum_{i=1}^n x_i w_{ic} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_{ic} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}, \quad (15)$$

$$y_{1t} = \frac{\sum_{i=1}^n x_i w_{it} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_{it} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}. \quad (16)$$

考虑融合特征的特性, 本文对目标位置进行更新处理, 定义了两种基于特征的自适应系数, 称之为提升颜色系数 (ICC) 和提升文本系数 (ITC), 分别为 γ_c 和 γ_t , 有

$$y_{fac} = \gamma_c y_{1c} + \gamma_t y_{1t}, \quad (17)$$

$$\gamma_c = \left[1 - \frac{\rho[\hat{p}_{uc}(y_{1t}), \hat{q}_{uc}]}{\rho[\hat{p}_{uc}(y_{1c}), \hat{q}_{uc}] + \rho[\hat{p}_{uc}(y_{1t}), \hat{q}_{uc}]}\right], \quad (18)$$

$$\gamma_t = \left[1 - \frac{\rho[\hat{p}_{uc}(y_{1c}), \hat{q}_{uc}]}{\rho[\hat{p}_{uc}(y_{1c}), \hat{q}_{uc}] + \rho[\hat{p}_{uc}(y_{1t}), \hat{q}_{uc}]}\right], \quad (19)$$

其中 ρ 为巴氏系数。从式(18)和(19)可以看出, 当某一特定空间的巴氏系数增加时, 其对应的系数也相应增加, 这样通过式(17)便能更精确地得到当前帧目标所在的位置 y_f 。

最后基于位置 y_{fac} 和 y_{fsc} , 再作一次均值处理, 即可得到当前帧目标位置

$$\bar{y} = (y_{fac} + y_{fsc})/2. \quad (20)$$

3 仿真实验

为了直观检验出算法的效果, 基于 Matlab 环境, 本文对一些视频序列进行了实验, 硬件条件为 1.7 GHz CPU 和 1.2 G 内存, 视频序列从 PETS2000-2007 中挑选。

第1组视频序列中, 采用基于颜色特征的 mean shift 算法进行跟踪, 如图1所示。图1为某视频序列中的第1、9、20、35帧, 用矩形框来标定目标。从图1中可以看出, 由于受到地面颜色的干扰, 造成了目标跟踪结果的不准确, 质心发生了较大的偏移。



图1 基于颜色特征算法效果图

图2从数据上反映了两种算法的差别, 即传统 mean shift 算法(基于颜色特征的 mean shift 法)和本文算法。横坐标表示帧序列, 纵坐标表示实际目标中心位置与跟踪算法所得到的目标中心位置之间的距离。从图2中可以发现, 本文算法的距离差明显小于经典算法, 跟踪效果更好。

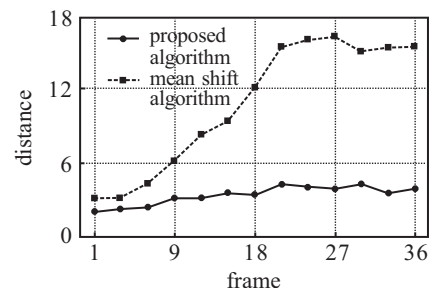


图2 实际目标中心与跟踪目标中心距离 (基于纹理特征算法和本文算法)

图3、图4分别为另一段视频序列的第1、20、30、45帧,同样用矩形框来标定目标.该组选用的视频序列同样存在背景干扰,采用的第1个算法是基于纹理特征的mean shift法,该方法基于纹理特征进行目标描述.从图3中可以发现,单一使用纹理特征只能完成对目标的小部分跟踪,随着视频帧的前进,跟踪效果越来越差.图4则显示了本文算法的优越性,其能够完成对目标的全部跟踪,明显优于传统的基于单一特征的算法.图5从数据上显示了本文算法的优越性.



图3 基于纹理特征算法效果图



图4 基于本文算法效果图

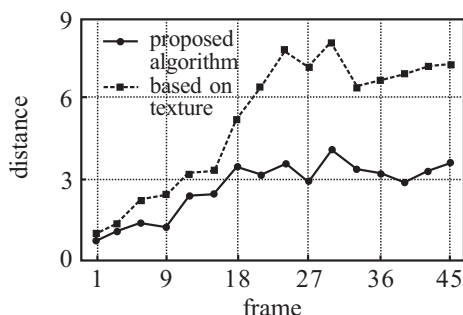


图5 实际目标中心与跟踪目标中心距离(基于颜色特征的mean shift算法和本文算法)

综上所述可以发现,本文算法不仅在颜色干扰下能够达到较好的跟踪效果,而且在纹理干扰下也能达到预期效果.为了提高算法的实时性,在纹理特征提取过程中,方向数和尺度数均选择为3,这是因为如果选择太大,会影响算法的速度,而过小则会使得效果提升不明显.由实验结果可知,本文在增加算法鲁棒性的同时,也带来了计算复杂度大的问题,由原来的 $O(m_c)$ 增加到 $O(m_c + m_t)$,平均每帧的耗时由原先的0.015 s增大到0.13 s.但是对于一般的工程项目而言,是能够满足实时性要求的.

4 结 论

本文提出了一种改进的mean shift算法,将Gabor小波应用于纹理特征的提取,并对巴氏系数作了进一步的改进,使得跟踪的准确性得到提高;同时对传统的mean shift算法中存在的不足作了较明显的改进,例如背景颜色与目标颜色相近的情况.从实验结果可以看出,本文算法能够准确、有效、实时地跟踪目标,并对复杂背景具有较好的适应性,增强了跟踪算法的鲁棒性.目标跟踪还存在许多技术难点,包括目标遮挡等问题,这将是下一步的研究内容.

参考文献(References)

- [1] Fukunaga, Housteler L D. 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.
- [2] 徐琨, 贺昱曜, 王卫亚. 基于CAMSHIFT的自适应颜色空间目标跟踪算法[J]. 计算机应用, 2009, 29(3): 757-760.
(Xu K, He Y Y, Wang W Y. Object tracking algorithm with adaptive color space based on CAMSHIFT[J]. J of Computer Applications, 2009, 29(3): 757-760.)
- [3] Polate, Ozden. A nonparametric adaptive tracking algorithm based on multiple feature distributions[J]. IEEE Trans on Multimedia, 2005, 8(6): 1156-1163.
- [4] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2006: 798-805.
- [5] Ojala T, Pietikainen M, Maenpaa T. Multi-resolution gray scale and rotation invariant texture classification with local binary patterns[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.
- [6] Zivkovic Z, Krose B. An EM-LIKE algorithm for color-histogram-based object tracking[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Washington DC: IEEE CS, 2002: 249-254.
- [7] 王永忠, 梁彦, 赵春晖, 等. 基于多特征自适应融合的核跟踪方法[J]. 自动化学报, 2008, 34(4): 393-399.
(Wang Y Z, Liang Y, Zhao C H, et al. Kernel-based tracking based on adaptive fusion of multiple cues[J]. Acta Automatica Sinica, 2008, 34(4): 393-399.)

(责任编辑: 李君玲)