

基于特征匹配滤波的自适应模板跟踪算法^①

周晨辰 黄 昶 王小明 (华东师范大学 信息科学与技术学院 上海 200241)

摘要: 针对头部目标跟踪的应用背景,在 Mean Shift 的框架下提出一种基于特征匹配滤波的自适应模板更新策略。该方法根据被跟踪目标在不同状态下所呈现出颜色的统计特征信息,采用颜色空间核密度估计匹配滤波的方法,对模板中局部区域像素的更新作出决策。实验结果表明,该算法有效解决了因头部旋转以及物体遮挡而导致模板不匹配的问题,实现了头部的连续跟踪,取得了很好的跟踪效果,提高了跟踪算法的整体稳定性和鲁棒性。

关键词: 模板更新;核密度估计; Mean Shift; 目标跟踪

An Auto-Adaptive Model Update Tracking Algorithm Using Feature Matching Filter

ZHOU Chen-Chen, HUANG Chang WANG Xiao-Ming

(Department of Information Science and Technology, East China Normal University, Shanghai 200241, China)

Abstract: This paper proposes a novel template adaptive updating strategy as part of Mean Shift framework which is based on feature matching filter in the scenario of head tracking. In accordance with the color space statistical information of tracked objects in different states, by using the color space kernel density estimation matching filter, this method decides on the specific pixels that need to be updated in template. The experimental results show that the algorithm is an effective solution to the rotating head and object block which leads to the mismatch of template. It also achieves the continuous tracking of the head, and has very good tracking results. The overall stability and robustness of the tracking algorithm are improved.

Keywords: template update; kernel density estimation; mean shift; object tracking

1 引言

Mean shift 算法是一种快速搜索算法,具有特征稳定、鲁棒性强、实时性好等诸多优势^[1-4]。但是,当场景中目标外观变化较大或非目标干扰较为严重时,仍可能因目标模型的偏移而导致跟踪失败。所以在较长时间的跟踪中,目标模板的更新势在必行。本文提出了一种 Mean shift 框架下的核密度估计匹配滤波的自适应更新策略。改进方法的基本思路就是通过模板特征匹配滤波识别场景变化情况,并对模板的局部区域进行自适应模板更新。在本文以下章节中,首先对 Mean shift 算法框架进行介绍;在此基础上,提出 Mean shift 框架下的基于核密度估计匹配滤波的模板更新策略及其算法;最后是本文的实验和结论。

1.1 基本介绍

在拥有标志寄存器的处理器中,绝大多数的运算指令都会对标志位定值,虽然这些定值并不一定会使用到。如果对源指令中标志位的每次定值都进行翻译,那么所产生的目标代码的效率就会大幅度下降。平均而言,在翻译所产生的代码中,对每一个标志位的定值所需要的代码与翻译一条指令基本功能的代码数量差不多,而每条指令都会对多个标志位进行修改,那么,如果完全把标志位定值翻译出来,而这些定值并没有被使用,那么浪费就显而易见了。

2 Mean Shift跟踪算法

Mean shift 是一种基于核概率密度估计的快速

① 收稿时间:2009-08-22;收到修改稿时间:2009-11-04

模式匹配算法,它采用某几种特征(如灰度、颜色、纹理等)的核函数直方图来描述目标^[5, 6],在每帧中对目标模板和候选搜索模型进行相似性度量,并沿着核直方图相似性的梯度方向迭代搜索目标位置,最终实现模式匹配和目标跟踪。

2.1 对象特征提取

一般而言,目标对象在特征空间的模式可由它的概率密度函数 q 表示。定义目标对象模式坐标 0 位置,则后一帧搜索对象模式为坐标 y 处,由概率密度函数 $p(y)$ 表示。为了满足实时处理的要求,采用离散概率密度以减少计算复杂度^[2]。函数 $b: R^2 \rightarrow \{1..m\}$ 表示在 X_i^* 位置的像素在特征空间中直方图上的区域索引值 $b(X_i^*)$ 。则可通过下式获得目标对象模式中特征 $u=1..m$ 的概率密度估计^[2]:

$$\hat{q}_u = C \sum_{i=1}^n k(\|X_i^*\|^2) \delta[b(X_i^*) - u], \quad (1)$$

其中, δ 是冲击函数。归一化常数 C 保证了 $\sum_{u=1}^m \hat{q}_u = 1$ 。

2.2 相似性度量

相似度评价函数定义了目标对象模式和搜索对象模式在特征空间上的距离。定义离散概率密度函数估计的 Bhattacharyya 系数来衡量 p 与 q 之间的相似程度

$$\hat{\rho}(y) = \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (2)$$

公式(2)的统计特性限于篇幅在这里就不一一赘述,有兴趣的读者请参阅文献^[2]。

当前帧搜索过程从前一帧图像的目标对象所在位置 y_0 开始搜索,初始搜索对象模式为 $\{p_u(y_0)\}_{u=1..m}$ 。将 $\hat{p}_u(y_0)$ 进行泰勒展开,并将公式(1)代入公式(2)化简后可近似表示为

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right), \quad (3)$$

其中,

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(x_i) - u]. \quad (4)$$

只要满足条件 $\{\hat{p}_u(y)\}_{u=1..m}$ 与 $\{\hat{p}_u(y_0)\}_{u=1..m}$ 间没有突变,则近似造成的误差就可以被忽略。在图像跟踪

应用背景下,视频图像连续两帧之间通常容易满足上述条件。因此,为了求解距离的极小值,公式(3)中的第一项与 y 无关,所以第二项必须要取到最大值。显然第二项表示的是当前帧中 y 位置的核函数分布 $k(x)$ 概率密度估计,并且被 w_i 式(4)加权。我们可以使用 Mean Shift 在 y 位置的附近搜索该项的最大值。在搜索中的每一次迭代过程是将核函数窗口从 y_0 移动到 y_i 。

2.3 目标跟踪

核函数分布的形式适合采用 Epanechnikov 核函数^[3],因为该核函数分布求导后的 $g(x)$ 是常数,得到迭代向量的像素坐标的加权平均形式(质心)

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (5)$$

由于核函数的作用,一次 Mean Shift 迁移的最大范围限制在核函数的范围内,所以如果当前帧的目标对象中心点位置在当前搜索对象核函数窗口范围内,则通过上述算法经一次迁移可以找到该 Bhattacharyya 系数的局部最大值点。当然其前提是该目标的特征提取方法提供了足够的辨析度,并且 Bhattacharyya 系数在该范围内只有一个局部最大值。若当前真的目标对象中心点不在搜索对象的核函数窗口范围内,则需要多次 Mean Shift 迭代迁移,最终将搜索对象的中心收敛到目标对象中心上。

3 自适应模板建立与匹配

对于一个鲁棒的跟踪算法,模板更新方法非常重要,该更新方法应能适应目标表现的缓慢和快速变化^[7],具体表现为:①目标的姿态或光照条件的变化;②目标由于遮挡,部分可见或完全不可见。大部分基于模板的跟踪算法,对模板的选取采用三种方法:一是模板来自于序列的第一帧图像^[8];二是模板取自于刚处理的前一帧图像^[9];三是来自于从开始图像到当前帧的所有序列^[10, 11]。本文提出的方法基于上述第三种思想,旨在提高模板匹配的鲁棒性,对于遮挡、目标姿态和光照条件的变化有较强的适应性。

3.1 模板对象特征模式

本文算法在场景中采用椭圆形区域表示目标,为

简化运算, 将目标区域归一化为后的单位圆定义为模板。不妨设为 $\{X_i^*\}_{i=1..n}$ 模板像素集合, 其中心点坐标为 0 。我们认为越靠外围的像素的可信度越低, 故采用一个外凸单调递减的核函数 $k(x)$, 对位于离中心较远距离的像素赋予较小的权值, 以增强模板核密度估计匹配的鲁棒性。模板模式可由(6)式给出。

$$\hat{T}_u = C_h \sum_{i=1}^n k\left(\left\|\frac{x_i}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (6)$$

C_h 是归一化常数。带宽 h 定义了搜索对象的尺度, 以调整归一化处理区域范围。

3.2 自适应模板更新

由上述分析可知, 鲁棒的自适应模板更新方法需能区分目标遮挡和目标渐变这两种情况。对于第一种情况, 系统应对模板进行维持(如遮挡部分不进行更新等), 当遮挡结束后, 还能重新捕捉到目标; 对于第二种情况, 系统应该及时更新模板, 以适应物体外在表面的变化。由此, 本文提出应用 **Bhattacharyya** 系数评价和核密度估计匹配滤波相结合的自适应更新策略。

首先采用 **Bhattacharyya** 系数衡量跟踪结果的有效性, 如果 ρ 很小, 说明本次跟踪结果受到了剧烈干扰, 不应更新模板而仍采用前一时刻的目标模板; 如果 ρ 大于阈值 thd_B , 则将当前模板与跟踪结果的归一化核密度估计相比较得到核密度估计波动评价, (7)式给出了其评价函数形式,

$$\Delta H_{u_{max}} = MAX\left(\left|\hat{T}_u - \hat{q}_u(y_{irg})\right|_{u=1..m}\right), \quad (7)$$

其中, \hat{T}_u 为模板的模式, $\hat{q}_u(y_{irg})$ 为跟踪结果的模式。若 $\Delta H_{u_{max}}$ 很小, 说明模板与跟踪结果非常接近, 为提高算法效率, 不对模板进行更新。若 $\Delta H_{u_{max}}$ 大于阈值 thd_H , 则将核密度估计波动最大处所对应的跟踪结果区域的像素更新到相应模板区域:

$$X^{tempnew} = \{x_i^{irg} \delta(b(x_i^{irg}) - u_{max}) + x_i^{tempold} (1 - \delta(b(x_i^{irg}) - u_{max}))\}_{i=1..n} \quad (8)$$

其中, $X^{tempnew}$ 为更新后的模板像素集, $x_i^{tempold}$ 为更新

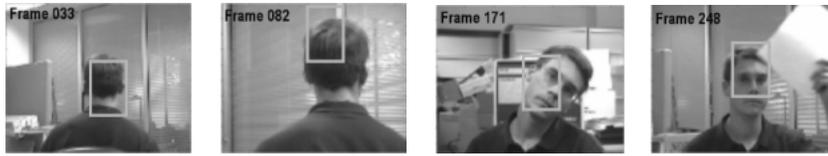
前模板中的像素, x_i^{irg} 为跟踪结果区域集中的像素, u_{max} 为核密度估计最大不匹配处的相应直方图位置。可见, 基于本文提出的特征匹配滤波模型更新策略, 可对目标模型中的每个像素进行有选择性的更新, 不再是同时更新或不更新整个目标模型。

4 实验结果

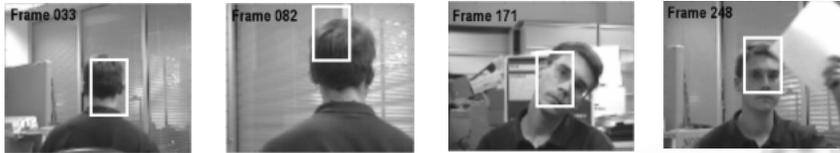
为验证本文算法, 采用 Stan Birchfield (<http://www.ces.clemson.edu/~stb/research/headtracker/>) 的两段人脸运动视频序列进行 matlab 仿真实验。两序列图像大小均为 128 像素 × 96 像素, 分别针对无模板更新(Non-model update, NMU)和本文提出的核密度估计模板更新(KDE model update, KMU)这两种机制进行测试。本文 KMU 算法采用参数 $thd_B=0.72$, $thd_H=0.44$ 。实验中初始目标手动选取, 采用量化等级为 $13 \times 13 \times 13$ 的 RGB 颜色特征, 核函数为 Epanechnikov 核。为便于对比和分析, 且结论也不失一般性, 两种方法均采用文献[2]提出的基本 Mean shift 算法框架, 搜索框固定, 其他改进措施(如背景滤除、目标旋转等)均未添加。

实验 1: 从 Stan Birchfield-man 视频序列中选取图像共 500 帧, 跟踪对象为人头部, 图 1 为两种模板更新算法的部分跟踪结果。由图 1 可见, 1)两模板更新算法对人脸部跟踪情况良好(171 帧)并且对遮挡具有一定鲁棒性(248 帧), 但是对脸部转动情况跟踪精度一般; 2)虽然场景中头部发生了转动, 但是由于对象头发的金色与肤色接近, 导致 KMU 方法未进行模板更新, 即 KMU 在模板不发生更新的情况下等同于 NMU 算法。

实验 2: 从 Stan Birchfield-woman 视频序列中选取图像共 500 帧, 跟踪对象为人脸, 图 2 为两种模板更新算法的部分跟踪结果。由图 2 可见, 1)KMU 算法在深色头发情况下对于人头部转动跟踪情况明显好于 NMU 算法(108 帧); 2)NMU 算法对于色彩特征相似的物体遮挡干扰敏感(424 帧), 甚至出现了跟踪失败的情况(500 帧), 但是本文提出的 KMU 算法对于特征相似物体遮挡体现了较好的鲁棒性。



(a) NMU

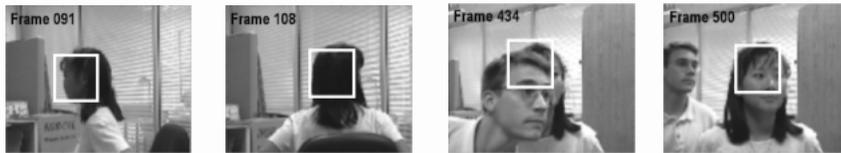


(b) KMU

图 1 实验 1 Stan Birchfield-man 视频序列 33、82、171 和 248 帧



(a) NMU



(b) KMU

图 2 实验 2 Birchfield-woman 视频序列 91、108、434 和 500 帧

5 结论

针对 Mean shift 在变化场景中的目标跟踪问题, 本文提出了一种基于核密度估计匹配滤波的自适应模板更新算法。该方法克服了更新算法对变化场景无法动态适应的不足; 同时, 对于目标遮挡具有较好的适应性。实验结果显示其增强了 Mean Shift 算法在头部跟踪过程中的自适应性、有效性和鲁棒性。该算法可以很好的解决由目标表现特征变化而引起目标跟踪性能下降, 甚至丢失的问题, 从而提高了经典 Mean Shift 算法跟踪精度和鲁棒性。

参考文献

1 Comanicu D, Meer P. Mean Shift: A Robust Approach Toward Feature Space Analysis. Pattern Analysis and

Machine Intelligence, IEEE Transactions on, May 2002,24(5):1-18.

2 Comanicu D, Ramesh V, Meer P. Kernel-based object tracking. Pattern Analysis and Machine Intelligence, IEEE Transactions, 2003,25(5):564-577.

3 Comanicu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift. Computer Vision and Pattern Recognition, 2000. Proc. IEEE Conference on, 2000,2:142-149.

4 Bradski GR. Computer Vision Face Tracking For Use in a Perceptual User Interface. Intel Technology Journal, 1998, 2nd Quarter' 98.

5 Eleftheriadis A, Jacquin A. Automatic Face Location Detection and Tracking for Model-Assisted Coding of

(下转第 76 页)

(上接第 166 页)

- Video Teleconference Sequences at Low Bit Rates. Signal Processing – Image Communication, 1995,7, (3): 231 – 248.
- 6 赵瑶,常发亮,郝洪霆.基于改进的均值漂移算法的非刚性目标跟踪.计算机工程与科学, 2007,29(12):71–73.
- 7 黄凤贤.彩色序列图像中的人脸检测与跟踪[硕士学位论文].西安:西安电子科技大学, 2004.
- 8 Li B, Chellappa R. Simultaneous Tracking and Verification via Sequential Posterior Estimation. Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, 2000,2:110 – 117.
- 9 Black M, Yacoob Y. Recognizing Facial Expressions in Image Sequences Using Local Parameterized Methods of Image Motion. Computer Vision, 1997,25(1):23 – 48.
- 10 沈志熙,杨欣,黄席樾.均值漂移算法中的目标模型更新方法.自动化学报. 2008, 35(5):478 – 483.
- 11 Tao H, Sawhney H, Kumar R. Dynamic Layer Representation with Applications to Tracking. Computer Vision and Pattern Recognition, 2000. Proceedings IEEE Conference on, 2000,2:134 – 141.

76 研究开发 Research and Development