

基于视觉显著性的均值漂移跟踪算法^{*}

樊祥锰, 尚振宏, 刘 辉, 钱 谦

(昆明理工大学 信息工程与自动化学院, 云南 昆明 650500)

摘 要: 为解决突变运动下的目标跟踪问题, 提出了一种基于视觉显著性的均值漂移跟踪算法。将视觉注意机制运用到均值漂移跟踪框架中, 利用时空显著性算法对视频序列进行检测, 生成视觉显著图, 从视觉显著图对应的显著性区域中建立目标的颜色特征表示模型来实现运动目标跟踪。实验结果表明: 该算法在摄像机摇晃等动态场景下可以较准确检测出时空均显著的目标, 有效克服了在运动目标发生丢失和遮挡等情况下跟踪不稳定的问题, 具有较强的鲁棒性, 从而实现复杂场景下目标较准确的跟踪。

关键词: 运动目标跟踪; 均值漂移; 视觉显著性; 显著图

中图分类号: TP212

文献标识码: A

文章编号: 1000-9787(2017)06-0130-04

Meanshift tracking algorithm based on visual saliency^{*}

FAN Xiang-meng, SHANG Zhen-hong, LIU Hui, QIAN Qian

(College of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: In order to solve target tracking problem under circumstances of abrupt motion, a meanshift tracking algorithm based on visual significance is proposed. The visual attention mechanism is applied to the framework of mean shift tracking, the video sequence is detected by using the spatial and temporal saliency algorithm, and the visual saliency map is generated. According to the visual figure corresponding significant area, from which target color characteristics of the saliency model is established to realize moving target tracking. Experimental results show that the proposed algorithm can accurately detect obvious spatial and temporal targets in the dynamic scene, and the algorithm effectively overcomes the problem of tracking instability in the case occlusion and loss of moving objects, it has strong robustness, so as to realize accurate tracking of target under complex scene.

Key words: moving object tracking; meanshift; visual saliency; saliency map

0 引 言

运动目标跟踪是在连续变化的图像序列中找出目标的位置和状态的过程^[1], 而目标跟踪实现的稳定性和鲁棒性则是要处理的主要问题。运动目标跟踪算法框架大体分为 3 种: 基于粒子滤波的目标跟踪算法^[2] 框架, 基于卡尔曼滤波的目标跟踪算法^[3] 框架, 基于均值漂移的运动目标跟踪算法^[4] 框架。均值漂移(MeanShift)法在单一环境下对目标的跟踪能够较准确地实现, 由于现实环境的复杂多样性, 算法对运动目标大小和形状的改变适应性较差, 目标发生突变后不能自动恢复, 很难实现目标跟踪的准确性和稳定性。

运动发生突变主要分为跟踪目标发生遮挡、目标移动太快以及目标丢失等, 针对上述问题, 基于 MeanShift 的多模板跟踪算法^[5] 通过构建并更新多样性模板库, 提高了算法的稳定性和鲁棒性; 基于时空上下文跟踪算法利用局部

上下文区域时空关系和跟踪的目标在贝叶斯框架中进行建模, 效果较鲁棒; 基于压缩感知目标跟踪算法^[6] 对多尺度图像特征运用合适压缩感知 RIP 条件的随机感知矩进行降维, 然后采用简单的朴素贝叶斯分类器在降维后的特征上进行分类。在复杂场景下能够对目标较准确地跟踪。

视觉注意机制^[7] 能自动过滤掉非注意场景区域。利用计算机模仿人类视觉机制构建视觉注意模型, 并把视觉注意模型运用到运动检测^[8]、运动识别和运动跟踪^[9-11] 框架中, 能够提高算法的准确性。本文利用人类视觉注意机制的研究成果, 提出了一种基于视觉显著性的 MeanShift 跟踪算法, 将时空显著性算法融合到 Meanshift 跟踪框架中, 利用时空显著算法检测定位运动目标。从显著性区域中搜索与运动目标模板最为相似的候选模板进行跟踪, 屏蔽非显著性区域对目标的干扰, 有效克服了在运动目标丢

收稿日期: 2016-06-24

^{*} 基金项目: 国家自然科学基金资助项目(61462052, 31300938)

失和遮挡等情况下跟踪不稳定的问题,具有较强的鲁棒性。

1 显著性检测算法

视觉显著性检测包括2种,即数据推动的自底向上显著性区域突显和任务推动的自顶向下目标突显。通过采用两种策略注意机制相结合的方式计算获取视频序列中的时空显著性。

本文利用4元数傅里叶相位谱时空变换的显著性检测算法^[12](PQFT)检测得到视频序列的时空显著图,并运用到均值漂移跟踪框架中。显著性检测算法分为2个阶段:1)构建包含两个颜色通道、一个强度通道和一个运动通道的4元数图像,运动通道可以获取具有运动目标的时间显著性,将视觉注意机制的视角从静态序列转变为运动目标在时间和空间上均显著的动态序列;2)通过计算获得视频序列的时空显著图。采用4元数模型可实现对图像特征的并行处理,并且能够满足算法的实时性。

1.1 构建4元数图像

假设 $F(t)$ 表示时间 t 时刻的输入图像, $t=1,2,\dots,T$, T 为所有序列帧的总数。 $F(t)$ 分为红、绿、蓝3个颜色通道,表示为 $r(t)$ 、 $g(t)$ 、 $b(t)$,那么可以将3个颜色通道扩展为4个广义的颜色通道

$$R(t) = r(t) - \frac{g(t) + b(t)}{2} \quad (1)$$

$$G(t) = g(t) - \frac{r(t) + b(t)}{2} \quad (2)$$

$$B(t) = b(t) - \frac{g(t) + r(t)}{2} \quad (3)$$

$$Y(t) = \frac{r(t) + g(t)}{2} - \frac{|r(t) - g(t)|}{2} - b(t) \quad (4)$$

类似于人类视觉系统,当信号传入大脑中心区域时,神经元对一种颜色的信号所刺激而兴奋,当另外一种颜色的信号传入时,会对其进行抑制,因此,颜色通道定义为

$$RG(t) = R(t) - G(t) \quad (5)$$

$$BY(t) = B(t) - Y(t) \quad (6)$$

强度通道和运动通道定义为

$$I(t) = \frac{r(t) + g(t) + b(t)}{3} \quad (7)$$

$$M(t) = |I(t) - I(t - \tau)| \quad (8)$$

式中 τ 为延迟系数,通常设定 $\tau=3$ 。

通过上述计算得到了4个通道的图像:2个颜色通道,1个强度通道和1个运动通道。因此,4元数图像可以表示为

$$q(t) = M(t) + RG(t)\mu_1 + BY(t)\mu_2 + I(t)\mu_3 \quad (9)$$

式中 μ_i 为权重系数, $i=1,2,3$,需满足 $\mu_3 = \mu_1\mu_2$ 。 $q(t)$ 为

$$q(t) = f_1(t) + f_2(t)\mu_2 \quad (10)$$

$$f_1(t) = M(t) + RG(t)\mu_1 \quad (11)$$

$$f_2(t) = BY(t) + I(t)\mu_1 \quad (12)$$

1.2 获得时空显著图

将视频序列图像中每一个像素点表示为 $q(n, m, t)$, (n, m) 为空间坐标, t 为时间坐标。4元傅里叶图像变化为

$$Q[u, v] = F_1[u, v] + F_2[u, v]\mu_2 \quad (13)$$

$$F_i(u, v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} e^{-\mu_1 2\pi i (mu/M + nu/N)} f[n, m] \quad (14)$$

式中 (u, v) 为频域坐标, N, M 为图像维度。4元逆傅里叶变化为

$$f_i(n, m) = \frac{1}{\sqrt{MN}} \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} e^{\mu_1 2\pi i (mu/M + nu/N)} F_i[u, v] \quad (15)$$

将 $q(t)$ 表示为 $Q(t)$ 的极坐标形式为

$$Q(t) = \|Q(t)\| e^{i\Phi(t)} \quad (16)$$

式中 $\Phi(t)$ 为 $Q(t)$ 的相位谱。设定 $\|Q(t)\|=1$,则只剩下相位信息 $\Phi(t)$ 。利用式(15)计算 $Q(t)$ 重建为 $q'(t)$

$$q'(t) = \rho_0 + \rho_1(t)\mu_1 + \rho_2(t)\mu_2 + \rho_3(t)\mu_3 \quad (17)$$

时空显著图 $S(t)$ 为

$$S(t) = g * \|q'(t)\|^2 \quad (18)$$

式中 g 为二维高斯平滑滤波。采用上述时空显著性算法得到的图像时空显著图如图1所示。



图1 时空显著图

2 基于视觉显著性的MeanShift算法

视觉显著性能够有效抑制非显著目标的干扰,在显著性区域中进行目标跟踪能够强化候选目标信息的描述。通过视觉显著性可有效筛选显著信息,从而提供给运动目标跟踪算法,提高了目标跟踪算法的准确性。MeanShift算法^[4]利用相似函数度量函数评价前一帧目标模型和候选模板的相似性,通过分别计算目标区域和候选区域内像素的颜色直方图得到关于目标模型和候选模型的描述,选择使相似度量函数值最大的候选模型得到目标模型的MeanShift向量。运动目标的最终位置通过多次迭代得到。利用第1节计算视频序列中的显著图,引入到MeanShift跟踪框架中,首先对视频序列计算显著图,然后从显著图对应的显著性区域中搜索与运动目标模板最为相似的候选模板进行跟踪,使颜色直方图更真实反映出目标的特征。

2.1 目标模型描述

从当前帧中确定目标区域,假设其中有 n 个像素用 x_i

表示其位置 $i = 1, \dots, n$ 均匀划分选取区域的灰度空间,得到由 m 个相等的区间构成的灰度直方图。目标模型 q_u 表示为

$$q_u = c \sum_{i=1}^n k \left[\left\| \frac{x_i - y_0}{h} \right\|^2 \right] \delta[b(x_i) - u] \quad \mu = 1, \dots, m \quad (19)$$

式中 y_0 为目标中心位置; x_i 为目标中心为原点的归一化位置; h 为带宽; k 为核函数,常选用 Epanechnikov 核函数; $\delta[b(x_i) - u]$ 为判断目标区域中像素 x_i 处的灰度值是否属于直方图中第 u 个单元,属于,则为 1; 否则,为 0; c 为归一化系数。

2.2 候选模型描述

从显著性区域中得到候选目标的中心位置坐标 f 为根据上一帧的目标中心位置 f_0 为搜索窗口中心得到,则候选目标模型为

$$p_u(y) = c \sum_{i=1}^n k \left[\left\| \frac{x_i - y}{h} \right\|^2 \right] \delta[b(x_i) - u] \quad (20)$$

式中 y 为候选区域目标中心。

2.3 相似性度量

利用相似性函数表示目标模型和候选目标模型的相似水平,通过目标和候选目标颜色直方图的 Bhattacharyya 系数得到

$$\rho(y) = \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y) q_u} \quad (21)$$

2.4 MeanShift 迭代过程

为得到最大的相似函数,对式(21)进行泰勒展开,得到系数的近似表达

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(y_0) q_u} + \frac{c_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (22)$$

式中 权值 w_i 为

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q_u}{p_u(y_0)}} \quad (23)$$

式(22)中第二项随 y 变化,其目标中心位置可以通过候选区域中心向真实区域中心的迭代完成,在一帧中找到目标位置需要迭代若干次才能得到最终的目标中心位置。

2.5 整体算法流程描述

1) 利用 1 节时空显著性检测算法检测图像序列的时空显著性区域;

2) 在当前帧中以 y_0 为起点,利用式(20)计算显著性区域中候选目标的特征模型 $P_u(y_0)$ $\mu = 1, \dots, m$;

3) 利用式(21)计算目标与候选目标的相似度 $\rho[p(y_0), q]$;

4) 利用式(23)计算权值 w_i $i = 1, \dots, m$;

$$5) \text{ 计算目标的新位置 } y_1 = \frac{\sum_{i=1}^m x_i w_i g \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^m w_i g \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)},$$

式中 $g(x) = -k'(x)$;

6) 若 $\|y_1 - y_0\| < \varepsilon$, 停止; 否则, 转步骤(2)。

3 实验结果分析

3.1 实验仿真

为验证本文基于视觉显著性的 MeanShift 算法的稳定性和准确性,采用目标在运动过程中有遮挡发生的国际公共视频序列^[13]进行运动目标跟踪遮挡测试,旨在运动目标发生丢失、遮挡等突变运动情况下能够准确地跟踪目标,并与传统 MeanShift 跟踪算法^[4]、基于 MeanShift 多模板跟踪算法^[5]、基于时空上下文跟踪算法以及基于压缩感知目标跟踪算法进行比较。

图 2 和图 3 为当目标发生遮挡、丢失时各算法的跟踪结果。(a) (c) (d) 都难以处理突变运动,跟踪的鲁棒性和准确性较差,对目标进行跟踪时容易跟踪到其他物体上造成目标丢失。(b) 在目标发生严重遮挡时可以恢复跟踪,但是还有待提高跟踪的稳定性和准确性。而(e) 本文算法可以对目标准确定位,跟踪效果较鲁棒。将时空显著性融合到均值漂移跟踪框架中,从显著区域中提取目标特征模型,屏蔽非显著区域对目标区域的干扰,跟踪结果鲁棒于上述其他运动目标跟踪算法。



图 2 视频序列 Girl 跟踪结果

3.2 定量分析

为进一步验证本文算法的准确性和稳定性,利用中心位置误差对实验结果进行定性分析。中心位置误差指的是预测目标中心位置和真实目标中心位置像素的偏差,图 4



图 3 视频序列 Boy 跟踪结果

给出了本文跟踪算法和其余 4 种跟踪算法在不同视频集上中心位置误差曲线的比较。

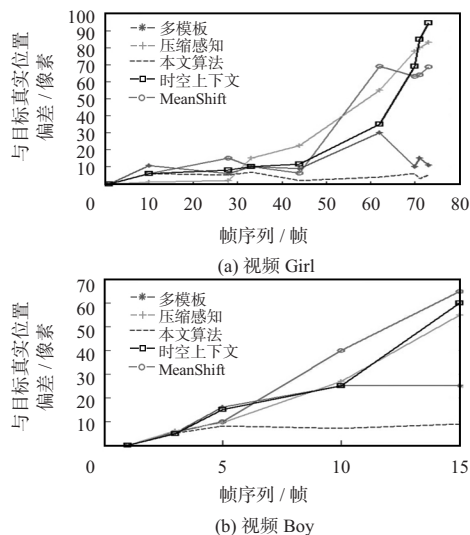


图 4 各跟踪算法的中心位置误差曲线

图 4 可以看出: 4 种对比算法在运动目标发生遮挡时目标中心都基本上偏离了真实目标中心位置, 而本文算法的目标中心位置的偏差较小。传统 MeanShift 目标跟踪算法、基于时空上下文跟踪算法、基于压缩感知目标跟踪算法在未发生遮挡等突变运动前跟踪的结果较准确, 但之后目标发生了严重的遮挡, 当前帧获取不到完整的信息, 造成了跟踪目标的丢失。而基于 MeanShift 多模板目标跟踪算法能够有效的在目标发生严重遮挡时恢复跟踪, 但在跟踪过程中目标跟踪的鲁棒性还有待完善。本文算法将时空显著性融合到 MeanShift 跟踪算法中, 从最显著的区域中建立目标的特征表示模型, 能够保持与真实目标位置较小的误差, 目

标跟踪的效果较鲁棒。

运动目标跟踪成功的帧数与视频图像序列的总帧数的比值定义为运动目标跟踪成功率。当前帧跟踪成功的标志就是目标的真实中心位置在预测的运动目标区域内。由表 1 可以看出本文算法的跟踪成功率优于其余 4 种算法。

表 1 5 种算法在视频序列中的跟踪成功率 %

算法	视频序列 boy	视频序列 girl	平均值
本文	90.0	95.8	92.9
均值漂移	20.3	63.0	42.3
多模板	33.3	75.3	54.0
压缩感知	34.0	49.3	41.0
时空上下文	33.0	52.0	42.5

4 结束语

为解决突变运动下的目标跟踪问题, 本文提出了基于视觉显著性的 MeanShift 跟踪算法, 利用时空显著性检测算法对视频序列进行检测, 生成显著图, 从显著图对应的显著性区域中搜索与运动目标模板最为相似的候选模板进行跟踪, 屏蔽非显著性区域对目标的干扰。实验证明: 当运动发生突变时, 本文算法不仅在动态场景下可较准确检测出时间和空间均显著的运动目标, 还能够有效解决目标发生遮挡后的目标跟踪问题, 在复杂场景下可实现目标较准确的跟踪, 具有较强的鲁棒性。下一步将考虑融合更多的特征对目标进行描述, 来增加目标跟踪的鲁棒性。

参考文献:

- [1] 张亚红, 杨欣, 沈雷, 等. 基于视觉显著性特征的自适应目标跟踪[J]. 吉林大学学报: 信息科学版, 2015, 33(2): 195-200.
- [2] Zhu Mingqing, Wang Zhiling, Chen Zonghai. Human visual intelligence and particle filter based robust object tracking algorithm[J]. Control and Decision, 2012, 27(11): 1720-1724.
- [3] 万顷浪, 张殿福. 基于卡尔曼粒子滤波的目标跟踪算法[J]. 电子科技, 2013, 26(8): 7-9, 12.
- [4] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shiftspace[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2000: 142-149.
- [5] 丁晓凤. 基于 MEAN SHIFT 的多模板目标跟踪算法的研究[D]. 昆明: 昆明理工大学, 2016.
- [6] 安宁, 闫斌, 熊杰. 基于压缩感知的多尺度绝缘子跟踪算法[J]. 传感器与微系统, 2016, 35(3): 140-143.
- [7] 黎万义, 王鹏, 乔红. 引入视觉注意机制的目标跟踪方法综述[J]. 自动化学报, 2014, 40(4): 561-576.
- [8] Li Z, Xu J, Wang Y, et al. Saliency detection based on protoobjects and topic model[C]//2011 IEEE Workshop on Applications of Computer Vision(WACV), [S.l.]: IEEE, 2011: 125-131.

(下转第 137 页)

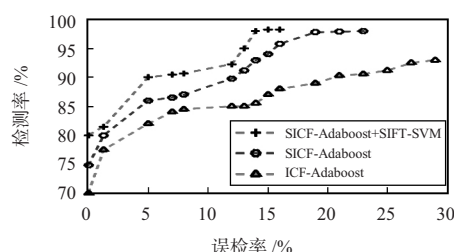


图4 3种算法的ROC曲线

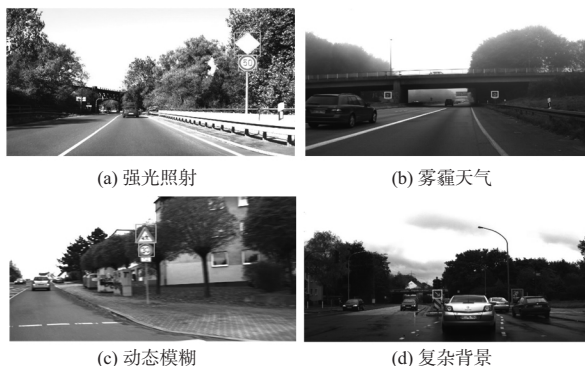


图5 不同实景下的交通标志检测结果示例

志位置并将其切割出来,同时,对光照、天气变化等因素带来的影响有较好的抑制作用。另外,对比4种形状类别的检测结果,发现倒三角形的检测率较低,如何提高倒三角形交通标志的检测率,将是下一步研究工作的重点。

参考文献:

- [1] Wang G, Ren G, Wu Z, et al. A robust coarse-to-fine traffic sign detection method [C]// International Joint Conference on Neural Networks, IEEE, 2013: 1-5.
- [2] Yang Y, Wu F. Real-time traffic sign detection via color probabi-

lity model and integral channel features [C]// Pattern Recognition, 2014: 545-554.

- [3] Ohn-Bar E, Trivedi M M. Learning to detect vehicles by clustering appearance patterns [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5): 2511-2521.
- [4] Dollár Piotr, Tu Zhuowen, Perona Pietro, et al. Integral channel features [C]// Proceedings of 2009 British Machine Vision Conference, BMVC 2009, London, UK, 2009.
- [5] 李文杰, 廖晓伟, 束仁义, 等. 基于 ZigBee 和模糊 C 均值聚类方法在火警中的应用 [J]. 传感器与微系统, 2012, 31(10): 143-145.
- [6] Fleyeh H, Biswas R, Davami E. Traffic sign detection based on AdaBoost color segmentation and SVM classification [C]// Eurocon, IEEE, 2013: 2005-2010.
- [7] Pizer S M, Amburn E P, et al. Adaptive histogram equalization and its variations [J]. Computer Vision Graphics & Image Processing, 1987, 39(3): 355-368.
- [8] 王程冬, 程筱胜, 崔海华, 等. SIFT 算法在点云配准中的应用 [J]. 传感器与微系统, 2012, 31(2): 149-152.
- [9] Moranduzzo T, Melgani F. A SIFT-SVM method for detecting cars in UAV images [C]// Geoscience and Remote Sensing Symposium, IEEE, 2012: 6868-6871.

作者简介:

郑天宇 (1992-) 男, 硕士研究生, 研究方向为机器学习与模式识别, E-mail: 780876068@qq.com.

王年 (1966-) 男, 教授, 博士生导师, 主要从事数字信号处理工作。

(上接第129页)

- [13] Park J W, Oh H D, Tahk M J. UAV collision avoidance based on geometric approach [C]// SICE Annual Conference, IEEE, 2008: 2122-2126.
- [14] 罗臻, 刘宏立, 徐琨. 无线传感器在协同攻击环境中的安全定位研究 [J]. 传感器与微系统, 2014, 33(7): 38-41.

(上接第133页)

- [9] Zhang Q, Li B. Artificial intelligence and computational Intelligence [M]. Berlin Heidelberg: Springer, 2011: 433-438.
- [10] Li Z, Wang W, Wang Y, et al. Visual tracking by proto-objects [J]. Pattern Recognition, 2013, 46(8): 2187-2201.
- [11] 聂晓华, 张夫鸣. 针对临近空间目标跟踪的 AMCKF 算法 [J]. 现代雷达, 2016, 38(5): 49-53.
- [12] Guo C, Ma Q, Zhang L. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recog-

作者简介:

刘鑫 (1993-) 男, 硕士研究生, 主要研究方向为军事航空通信, E-mail: kdyliuxin@163.com.

杨霄鹏 (1973-) 男, 副教授, 硕士生导师, 主要从事宽带无线通信及信号处理技术工作。

2008: 1-8.

- [13] Choi W, Shahid K, Savarese S. The 9th International Workshop on Visual Surveillance in Conjunction with ICCV [EB/OL]. (2012-05-03) [2015-03-08]. <http://wwwweb.eecs.umich.edu/vision/activity-dataset.html>

作者简介:

樊祥猛 (1990-) 男, 硕士研究生, 主要研究方向为计算机视觉、模式识别、图像处理。

尚振宏 (1975-) 通讯作者, 博士, 副教授, CCF 会员, 主要从事计算机视觉、图像处理研究工作, E-mail: shangzhenhong@126.com.