

基于边缘多通道梯度模型的多运动目标检测

陈婕妤^{*}, 奚峥皓^{**}, 卢俊鑫^{***}

上海工程技术大学电子电气工程学院, 上海 201620

摘要 针对视频序列运动目标检测易受环境噪声干扰、提取目标轮廓困难的问题,提出了一种基于边缘多通道梯度改进模型的多运动目标检测算法。首先,利用 Canny 算子获取视频序列中目标的边缘信息,并根据人类视觉色彩的恒常特性,对目标边缘建立时间、空间、颜色多通道梯度模型;然后,利用该模型获取目标边缘像素点的运动状态描述信息,实现背景边缘和运动物体边缘的分离;最后,将间断边缘像素点与其邻域点的运动状态相关联,以连接目标间断边缘,实现运动目标轮廓的提取,并将连接后的轮廓进行形态学处理以分割出目标。实验结果表明,与同类型算法相比,本算法在运动目标检测中具有实时性、准确性和鲁棒性更好。

关键词 机器视觉; 运动目标检测; 多通道梯度; Canny 算子; 边缘连接

中图分类号 TP29

文献标志码 A

doi: 10.3788/LOP202158.0415002

Multi-Moving Object Detection Based on Edge Multi-Channel Gradient Model

Chen Jieyu^{*}, Xi Zhenghao^{**}, Lu Junxin^{***}

College of Electrical and Electronic Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

Abstract Aiming at the problems that moving object detection in video sequence is easily interfered by environmental noise and the object contour is difficult to extract, this paper proposes a multi-moving object detection algorithm based on the improved model of edge multi-channel gradient. First, the Canny operator is used to obtain the edge information of the object in the video sequence, and a multi-channel gradient model of time, space and color is established on the object edge according to the constant characteristics of human visual color; Then the model is used to obtain the motion state description information of the object edge pixel and achieve the separation of the background edge and the edge of the moving object; Finally, the discontinuity edge pixels are associated with the motion state of their neighboring points to connect the discontinuity edges of the object, which achieve the complete extraction of the contour of the moving object; The connected contour is morphologically processed to segment the object. Experimental results show that, compared with similar algorithms, the algorithm has superior real-time performance, accuracy, and robustness in moving object detection.

Key words machine vision; moving object detection; multi-channel gradient; Canny operator; edge connection

OCIS codes 150.0155;150.1135;150.4620

1 引言

目标检测作为机器视觉领域的研究热点,在智

能监控、智能交通、无人驾驶等方面具有极其重要的应用和发展前景。如何在目标检测过程中克服光照、异物、遮挡等环境噪声,提高检测效率,提取完整

收稿日期: 2020-06-15; 修回日期: 2020-07-17; 录用日期: 2020-08-07

基金项目: 国家自然科学基金(61801286)

^{*} E-mail: 15026621362@163.com; ^{**} E-mail: zhenghaoxi@sues.edu.cn; ^{***} E-mail: 792902802a@gmail.com

运动目标的轮廓,也成为这类问题的研究重点^[1-2]。

目前,主流的目标检测算法可分为基于经典模型和基于学习模型的检测算法,混合高斯法与视觉背景提取算法是基于经典模型的检测算法。Chen 等^[3]提出了一种在混合高斯改进模型中,利用码本和区间估计提高检测准确性的方法,一定程度上解决了前景像素移动缓慢造成的像素扰动和检测误判问题。但在光照、背景变化等噪声干扰情况下的鲁棒性较差,目标轮廓提取不完整,且无法满足实时性的要求。王旭等^[4-6]提出了多种针对目标检测问题的视觉背景提取算法,改进后的算法可以有效抑制视觉背景提取算法的鬼影现象,提升算法的实时性。但这些算法依赖背景模型的更新,受环境噪声干扰较大,无法获取准确的目标轮廓。

YOLO(You only look once)网络是目前比较热门的一种基于学习模型的目标检测算法^[7],相比区域卷积神经网络(RCNN)^[8]、Fast-RCNN^[9]算法在速度上得到了极大的提升,但对微小目标、尺度变化的目标,YOLO 网络的检测率和泛化能力都较弱。针对该问题,Liu 等^[10]提出了一种基于 YOLO 网络改进的多尺度单步多框检测(SSD)网络,大幅提升了对小目标的检测性能。基于 YOLO 网络,人们还提出了 YOLOv2^[11]和 YOLOv3^[12]网络算法,在确保算法准确性的同时改进了算法整体的运算速度。近年来,生物视觉特性与机器视觉结合的目标检测算法逐渐受到了人们的关注。如受人类视觉感知系统启发的 RFBNet(Receptive field block net)^[13]算法利用 Inception 与空洞卷积模拟人类视觉的感受野,并将 SSD 网络与感受野相结合,相比 YOLOv3 网络算法,该算法的检测精度和运行效率均有较大提升。基于学习模型的目标检测算法虽然在检测性能方面有较大优势,但算法的网络泛化和实用性依赖大量训练数据集,且耗时耗力。

针对上述问题,本文在文献[14]中的多通道梯度模型(McGM)基础上,结合生物视觉特性,提出了一种边缘 McGM(E-McGM),并将该模型用于多运动目标检测。针对多运动目标检测过程中出现的目标轮廓提取不完整问题,利用 E-McGM 设计一种运动目标间断边缘的连接方法。实验结果表明,相比其他算法,本算法在 3 个序列中的平均查全率、平均查准率和平均 F-Measure 均得到了一定的提高。

2 基于 E-McGM 的运动目标检测算法

文献[14]中的模型没有对图像信息进行筛选,

算法需要计算大量的冗余信息,时间效率低且空间复杂度高。针对该问题,选取图像中具有代表性的边缘信息,并参考文献[14]中的模型,提出了一种基于 E-McGM 的多运动目标检测算法。

2.1 E-McGM 的建立

生物学认为物体在人眼中的成像是该物体反射光能量在视觉单元中二维信号的反映。不同波长的反射光对应不同光谱能量值,通过光谱能量值得到物体在视觉系统中对应的颜色信息^[15]。由文献[15]中的视觉光学理论及图像学知识,得到图像边缘在图像 I 中各颜色通道的描述为

$$S_{\gamma}(\lambda_0) = \int_{\Omega} E(\lambda)G_{\gamma}(\lambda; \lambda_0, \sigma)d\lambda, \gamma \in (0, 1, 2), (1)$$

式中, λ 为波长, Ω 为可见光波长的范围,用均值为 λ_0 、标准差为 σ 的高斯函数 $G(\lambda, \lambda_0, \sigma)$ 表示视觉传感器灵敏度, γ 为 $G(\lambda, \lambda_0, \sigma)$ 的求导阶数。令 $\lambda_0 = 520 \text{ nm}$, $\sigma = 55 \text{ nm}$, 依据文献[16]提出的人类视觉彩色拮抗理论,将 $G(\lambda, \lambda_0, \sigma)$ 最高求导至二阶,用 $G_0(\lambda, \lambda_0, \sigma)$, $G_1(\lambda, \lambda_0, \sigma)$ 和 $G_2(\lambda, \lambda_0, \sigma)$ 分别表示人类视觉的亮度、黄-蓝和红-绿拮抗通道的加权函数, S_0 , S_1 和 S_2 分别为对应物体边缘的亮度、黄-蓝和红-绿拮抗通道的光谱响应。物体表面反射光的能量分布 $E(\lambda)$ 可表示为

$$E(\lambda) = e(\lambda)R(\lambda), (2)$$

式中, $e(\lambda)$ 为入射光的能量分布, $R(\lambda)$ 为物体表面光的反射系数。当物体运动时, $R(\lambda)$ 不发生变化, $e(\lambda)$ 变化比较缓慢,可认为短时间内的 $E(\lambda)$ 为不变量。因此, $S_{\gamma}(\lambda_0)$ 具有色彩恒常性,即对光照变化的鲁棒性较强。

将 $S_0(\lambda_0)$ 在 λ_0 邻域做二阶泰勒展开,得到图像 I 中物体边缘的光谱描述为

$$\hat{S}[\lambda_0 + s; E(\lambda), \sigma] = S_0[\lambda_0; E(\lambda), \sigma] + sS_1[\lambda_0; E(\lambda), \sigma] + \frac{s^2}{2}S_2[\lambda_0; E(\lambda), \sigma] + O(s^3), (3)$$

式中, s 为波长 λ 的偏移量, $O(s^3)$ 为泰勒公式的余项。将(3)式的光谱信息与图像 I 的物体边缘空间、时间信息结合,再次进行泰勒展开,可得到视觉皮层图像信息的 E-McGM 描述,即图像边缘像素 $P(x, y, t, \lambda)$ 对应的 E-McGM 为

$$\hat{S}(x + p, y + q, t + r, \lambda + s) = \sum_{i=0}^e \sum_{j=0}^f \sum_{k=0}^g \sum_{l=0}^h \frac{p^i q^j r^k s^l}{i! j! k! l!} \cdot \frac{\partial^{(i+j+k+l)} S(x, y, t, \lambda)}{\partial x^i \partial y^j \partial t^k \partial \lambda^l}, (x, y \in E^t), (4)$$

式中, x, y 分别为图像横轴、纵轴的坐标, E^t 为由 Canny 算子得到的第 t 帧边缘图像, (p, q, r, s) 为 (x, y, t, λ) 的偏移参数, e, f, g, h 为关于图像边缘像素 $P(x, y, t, \lambda)$ 在泰勒展开中的最高阶数, 实验取 $e=3, f=2, g=2, h=2$ 。

2.2 基于 E-McGM 的运动目标边缘检测

合并(4)式中偏导阶数相同的项, 构成向量 $k(x, y, t, \lambda)$, k 在空间 xy -时间 t -光谱 λ 上的差分可反映边缘像素点 P 的运动变化趋势。视觉皮层对运动物体的感知过程可看作单帧图像中每个像素点在空间梯度上的变化, 可表示为

$$\mathbf{A} = D[k(x, y, t, \lambda)] = (X, Y, T, \lambda), \quad (5)$$

式中, D 为差分算子, X, Y, T, λ 分别为向量 k 对 x, y, t, λ 的差分值。物体运动过程中的描述参数为

$$\mathbf{A}^T \mathbf{A} \rightarrow \begin{array}{cccc} XX & XY & XT & X\lambda \\ YX & YY & YT & Y\lambda \\ TX & YT & TT & T\lambda \\ \lambda X & \lambda Y & \lambda T & \lambda\lambda \end{array} \rightarrow \begin{array}{ccc} \frac{\lambda \cdot Y + T \cdot Y}{\lambda \cdot T + T \cdot T} & \frac{X \cdot Y}{X \cdot X} & \frac{\lambda \cdot T + X \cdot T}{\lambda \cdot X + X \cdot X} \\ \frac{\lambda \cdot T + Y \cdot T}{\lambda \cdot Y + Y \cdot Y} & \frac{X \cdot Y}{Y \cdot Y} & \frac{\lambda \cdot X + T \cdot X}{\lambda \cdot T + T \cdot T} \end{array} \quad (6)$$

生物学认为初级视觉皮层大部分的细胞具有方向选择性, 即表征初级视觉皮层细胞会在 m 个运动方向 θ 中选择梯度变化最强响应作为物体的运动方向^[17]。图像边缘点 $P(x, y, t, \lambda)$ 在视频序列三维空间中的正交方向速度 $v(v_x, v_y)$ 可表示为

$$\begin{aligned} v_x &= \frac{\lambda \times T + X \times T}{\lambda \times X + X \times X}, \\ v_y &= \frac{\lambda \times T + Y \times T}{\lambda \times Y + Y \times Y}, \end{aligned} \quad (7)$$

将(7)式作为初速度建立速度基坐标系, 通过旋转速度基坐标系可获得 m 个不同方向 θ 的速度响应, 令图像边缘点的正速度为 \hat{s} 、逆速度为 \check{s} , 则 θ 中任一个方向的运动可分为平行和垂直于基坐标的分速度, 即 $\hat{s} = (\hat{s}_{\parallel}, \hat{s}_{\perp})$, $\check{s} = (\check{s}_{\parallel}, \check{s}_{\perp})$ 。结合(7)式和文献[17]得到速度 \hat{s} 与逆速度 \check{s} 对 m 个不同方向 θ 的旋转公式为

$$\begin{aligned} \hat{s}_{\parallel} &= \sqrt{\frac{2}{m}} \left\{ v_x \left[1 + \left(\frac{X \cdot Y}{X \cdot X} \right)^2 \right] - 1 \right\} \\ \check{s}_{\parallel} &= \frac{1}{2} \sqrt{\frac{2}{m}} \left(\frac{\lambda \cdot X + T \cdot X}{\lambda \cdot T + T \cdot T} \right) \\ \hat{s}_{\perp} &= \sqrt{\frac{2}{m}} \left\{ v_y \left[1 + \left(\frac{X \cdot Y}{Y \cdot Y} \right)^2 \right] - 1 \right\} \\ \check{s}_{\perp} &= \frac{1}{2} \sqrt{\frac{2}{m}} \left(\frac{\lambda \cdot Y + T \cdot Y}{\lambda \cdot T + T \cdot T} \right) \end{aligned} \quad (8)$$

像素点 P 的速率为

$$S^2 = \frac{\begin{vmatrix} \hat{s}_{\parallel} \cdot \cos \theta & \hat{s}_{\parallel} \cdot \sin \theta \\ \hat{s}_{\perp} \cdot \cos \theta & \hat{s}_{\perp} \cdot \sin \theta \end{vmatrix}}{\begin{vmatrix} \hat{s}_{\parallel} \cdot \check{s}_{\parallel} & \hat{s}_{\parallel} \cdot \check{s}_{\perp} \\ \hat{s}_{\perp} \cdot \check{s}_{\parallel} & \hat{s}_{\perp} \cdot \check{s}_{\perp} \end{vmatrix}} \quad (9)$$

像素点 P 的运动方向 Θ 可表示为

$$\Theta = \arctan \frac{(\hat{s}_{\parallel} + \check{s}_{\parallel}) \sin \theta + (\hat{s}_{\perp} + \check{s}_{\perp}) \cos \theta}{(\hat{s}_{\parallel} + \check{s}_{\parallel}) \cos \theta + (\hat{s}_{\perp} + \check{s}_{\perp}) \sin \theta} \quad (10)$$

由(9)式和(10)式可得到像素点 P 的速度描述, 从而提取出第 t 帧视频图像的运动物体边缘 E_m^t 。

2.3 基于 E-McGM 的间断运动边缘连接

2.2 小节获得的运动目标边缘图 E_m^t 出现的边缘间断现象会影响运动目标的目标块获取。因此, 基于 E-McGM 设计了一种融合运动信息的边缘间断连接方法, 具体步骤如下。

1) 设动态像素点的标记 w 为 1, 静态像素点的标记 w 为 0。初始化当前第 t 帧 E_m^t 得到的目标边缘像素点集合 $\mathbf{P}^t = \{p_{x_1, y_1}^t, p_{x_2, y_2}^t, \dots, p_{x_c, y_c}^t\}$ 中所有元素的 w 为 1。

2) 在集合 \mathbf{P}^t 中按顺序选取第 n 个动态边缘点 p_{x_n, y_n}^t , 创建属于该点的动态数组 ξ_{x_n, y_n}^t , 将 p_{x_n, y_n}^t 存入 ξ_{x_n, y_n}^t , 即 $\xi_{x_n, y_n}^t = \{\xi_0\}$, $\xi_0 = p_{x_n, y_n}^t$ 。令 ξ_{x_n, y_n}^t 中边缘点个数 X_{count} 的初值为 1, 设 ξ_{x_n, y_n}^t 中包含元素点对应的速度均值为 $M(S_{x_n, y_n}^t)$ 、方向均值为 $M(\theta_{x_n, y_n}^t)$, ξ_0 的速度值和方向值作为 $M(S_{x_n, y_n}^t)$ 和 $M(\theta_{x_n, y_n}^t)$ 的初值。

3) 对于数组 ξ_{x_n, y_n}^t 中的元素 ξ_0 在 Canny 边缘图 E^t 上以自身为中心建立一个 3×3 的邻域为 e_{ϵ} , $\epsilon \in [0, 7]$ 。若邻域点 $e_{\epsilon} \in E^t$ 且 $e_{\epsilon} \notin E_m^t$, 则判定该点是否为缺失动态边缘点的方法为

$$w = \begin{cases} 1, & |S_{e_{\epsilon}}^t - M(S_{x_n, y_n}^t)| < T_s \text{ and} \\ & |\Theta_{e_{\epsilon}}^t - M(\theta_{x_n, y_n}^t)| < T_{\theta} \\ 0, & \text{else} \end{cases} \quad (11)$$

式中, $S_{e_{\epsilon}}^t$ 为该点的速度, $\Theta_{e_{\epsilon}}^t$ 为该点方向值, T_s 、 T_{θ} 分别为速度阈值和方向阈值, 分别设为 0.18 和 $\pi/36$ 。若 $w=1$, 则判定当前邻域点为动态边缘点并将其加入数组 ξ_{x_n, y_n}^t 和运动物体边缘图 E_m^t 中, 同时令 $X_{\text{count}} = X_{\text{count}} + 1$ 。

4) 更新 $M(S_{x_n, y_n}^t)$ 和 $M(\theta_{x_n, y_n}^t)$ 的方法可表示为

$$M(S_{xn,yn}^t) = \frac{\sum_{i=1}^{X_{count}} S_{\xi_i}^t}{X_{count}}, M(\theta_{xn,yn}^t) = \frac{\sum_{i=1}^{X_{count}} \theta_{\xi_i}^t}{X_{count}} \quad (12)$$

将 $p_{xn,yn}^t$ 从 $\xi_{xn,yn}^t$ 中移除,同时令 $X_{count} = X_{count} - 1$ 。若 $X_{count} \neq 0$,则从 $\xi_{xn,yn}^t$ 中提取新的 ξ_0 ,返回步骤 3);否则,令 $n = n + 1$ 。重复步骤 2),直到集合 P^t 中所有元素提取完毕。运动目标间断边缘连接完成后,通过形态学膨胀和物体轮廓连通方法,得到每个运动物体的完整边缘轮廓。然后利用矩形描述符分割出运动目标块,得到单帧图像中运动物体的检测区域。在公共视频序列 PES09-S2L1 中测试本算法,结果如图 1 所示。其中,图 1(a)为经过形态学处理后得到的检测目标,图 1(b)为外接矩形框得到的目标块结果。

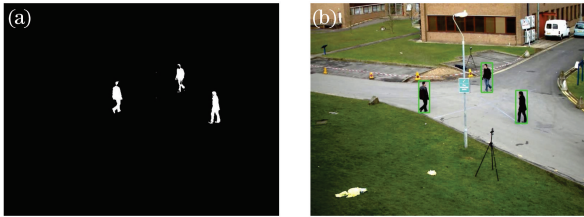


图 1 目标检测的结果。(a)形态学处理的结果;(b)目标块
Fig. 1 Result of object detection. (a) Result of morphological processing; (b) object block

3 实验结果分析与算法评估

本算法的实现与验证环境:操作系统为 64 位

Ubuntu16.04,CPU 处理器为 Intel Core Xeon E5-2603、内存为 16 GB 的个人电脑,编程语言为 C++ 和 OpenCV 库。

实验挑选 CDnet2014 数据集的 blizzard 和 PETS2006 视频序列以及 visual tracker benchmark 的 David 序列对本算法进行有效性验证。三个视频序列对应的参数如表 1 所示。将本算法与文献 [18]、文献[19]以及文献[20]中的算法在上述数据集中进行对比。

表 1 视频序列参数

Table 1 Parameter of the video sequence

Number	Video sequence	Frame	Size
1	PETS2006	1200	720×576
2	blizzard	7000	720×480
3	David	770	320×240

3.1 实验结果

首先在人流密集、天气恶劣、光照变化等场景下对不同算法进行多场景性能验证。视频序列 1 为室内地铁站人员流动场景,测试环境复杂,运动目标较多;视频序列 2 为室外下雪场景,在该恶劣天气下进行运动目标检测,会受到图像亮度灰暗、背景存在较多噪声和雪花轻微飘动的影响;视频序列 3 为从昏暗房间向明亮客厅移动的光照变化场景。图 2 为不同算法对 3 个不同场景视频序列的运动目标检测结果。

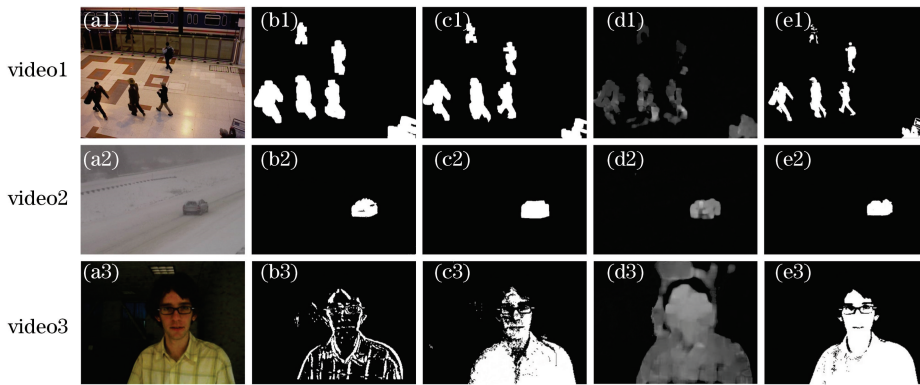


图 2 不同算法的检测结果。(a)原始图像;(b)文献[18]中的算法;(c)文献[19]中的算法;(d)文献[20]中的算法;(e)本算法

Fig. 2 Detection results of different algorithms. (a) Original image; (b) algorithm of Ref. [18]; (c) algorithm of Ref. [19]; (d) algorithm of Ref. [20]; (e) our algorithm

从图 2 可以看出,文献[18]中的算法在序列 1 中的检测结果较好;在序列 2 中,检测结果存在空洞,原因是该算法采用显色体积作为特征,对区域间

颜色差异敏感。序列中大幅的灰暗背景使目标某一小区域颜色与目标整体有差异时,该小区域会被误检为背景;在序列 3 中,对于光照变化不鲁棒,存在

明显的目标空洞和背景残留现象。文献[19]中的算法在序列 1 和序列 2 中均取得较好的检测结果,但该算法实际上是结合时空信息的背景减法,无法完全克服光照变化带来的干扰,在序列 3 的检测结果中出现空洞现象。文献[20]中的算法融合动态梯度和静态梯度特征构建时空梯度特征,会将产生梯度信息的动态背景误判为目标。本算法在 3 个序列中均取得了较好的检测结果,检测目标边缘清晰,在背景中存在缓慢运动干扰噪声以及光照变化等场景下均可以取得稳定的检测结果,鲁棒性较强。但本算法对于序列 1 中位于地铁围栏另一侧的运动目标检测结果不是特别完整,原因是存在围栏遮挡,算法无法提取到被遮挡物体完整的边缘轮廓,从而影响该运动目标的最终检测结果。

3.2 实验结果性能分析

为了进一步量化各算法的检测性能,用查全率

表 2 不同算法的检测性能

Table 2 Detection performance of different algorithms

Algorithm	Video1				Video2				Video3			
	R/%	P/%	$X_{F,M}/%$	FPS/frame	R/%	P/%	$X_{F,M}/%$	FPS/frame	R/%	P/%	$X_{F,M}/%$	FPS/frame
Ref. [18]	87.3	90.6	88.9	2	79.8	82.1	80.9	3	76.9	70.5	73.6	14
Ref. [19]	98.8	78.9	87.7	1	71.5	85.8	78.0	5	78.2	69.9	73.8	11
Ref. [20]	77.8	85.7	81.6	4	79.7	74.8	77.2	7	70.5	73.8	72.1	13
Ours	93.3	88.1	90.6	26	87.4	83.5	85.4	32	83.5	79.8	81.6	43

从表 2 可以看出,文献[18]中的算法在序列 1 中能获得稳定的区域显色体积特征,检测效果较好,但在具有噪声的序列 2 和光照变化序列 3 中,性能出现明显下降;文献[19]中的算法通过结合目标的时间和空间信息降低目标漏检率,在序列 1 中取得较高的查准率,但在背景干扰较多的序列 2 和序列 3 上漏检数量较大;文献[20]中的算法在 3 个序列上的检测精度最低,原因是该算法将动态与静态梯度图进行融合,背景中存在的微运动物体和光照变化会对图像梯度产生影响,降低目标检测精度,增大算法的误检率。相比其他算法,本算法在序列 1 中虽没有取得最佳的查准率和查全率,但能有效地平衡漏检率和误检率,在背景复杂的序列 2 和序列 3 中均取得了较好的性能。在实时性方面,其他算法对每帧图像的处理时间较长,无法达到实时检测要求。在视频序列 2 上对不同算法的检测误差率进行对比,结果如图 3 所示。可以看出,文献[18]与文献[20]中的算法在检测过程中性能不稳定,检测误

(R)、查准率(P)、F-Measure($X_{F,M}$)以及检测误差率(X_{Error_ratio})对检测结果进行量化。前三个指标值越大,表明对应算法的检测结果越好。

$$\begin{aligned} R &= \frac{X_{TP}}{X_{TP} + X_{FN}} \\ P &= \frac{X_{TP}}{X_{TP} + X_{FP}} \\ X_{F,M} &= \frac{2 \times R \times P}{R + P} \\ X_{Error_ratio} &= \frac{X_{FN} + X_{FP}}{X_{GT}} \end{aligned} \quad (13)$$

式中, X_{TP} 为正确检测的前景目标像素点数, X_{FP} 为被错误检测为前景目标的背景像素点数, X_{FN} 为被错误检测为背景的前景目标像素点数, X_{GT} 为前景目标像素总数。本算法与对比算法在视频序列的性能如表 2 所示,其中,FPS 为每秒检测帧数。

差率波动较大,原因是这两种算法不能克服恶劣天气的影响,检测结果有背景残留。相比其他算法,本算法具有明显优势,抗干扰能力较强,且性能稳定。

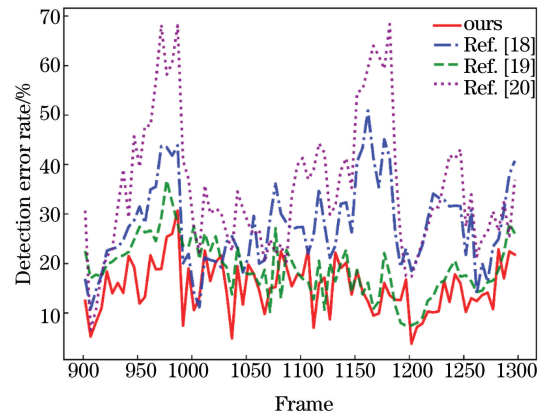


图 3 不同算法的检测误差率

Fig. 3 Detection error rates of different algorithms

3.3 数据优化分析

在目标检测中,图像中许多像素点为无用背

景点。而本算法只对目标边缘像素建立模型,大大减少了需要处理的像素点数,从而降低了算法的时间复杂度,提升了算法的运算速度。表 3 为 E-McGM 进行运动目标速度求解各阶段需要计算和存储的数据量,其中, N 为视频总帧数, B 为未处理的视频帧数, $n_x \times n_y$ 为视频帧中的像素点总数,

n_{CF} 、 n_{TF} 、 n_{SF} 分别为颜色滤波器、时间滤波器和空间滤波器个数, m_θ 为每个像素对应的 m 个响应角度。可以看出, $n_x \times n_y$ 数量影响 E-McGM 所有阶段的数据计算与内存消耗,这表明通过提取图像边缘可以减少 $n_x \times n_y$ 的数量,从而减少运算的时间与空间消耗。

表 3 各阶段计算和存储的数据量

Table 3 Amount of data calculated and store in each stage

Memory consuming and computational demanding stage	Data store
Establishment of the E-McGM	$(N - B) \times n_x \times n_y \times n_{CF} \times n_{TF} \times n_{SF} \times m_\theta$
Computing $\mathbf{A}^T \mathbf{A}$	$(N - B) \times n_x \times n_y \times m_\theta \times 6$
Velocity responses in m directions around a point	$(N - B) \times n_x \times n_y \times m_\theta \times 4$
Computing velocity magnitude and direction	$(N - B) \times n_x \times n_y$

4 结 论

针对视频序列中运动目标检测易受环境噪声干扰、目标轮廓提取困难的问题,在 McGM 模型基础上提出了 E-McGM,并利用所得的运动状态对运动目标轮廓实现完整提取。实验结果表明,相比其他算法,本算法在噪声干扰环境下的目标误检率和漏检率更低。通过结合 E-McGM 的运动物体间断边缘连接方法,提升了运动目标检测的准确性和鲁棒性,且具有更优的时间复杂度。本算法虽然可以解决视频序列的缓慢运动目标检测问题,但当背景中运动扰动较大时,检测结果并不理想。因此,在大扰动背景环境中的目标检测问题是未来需要研究的重点。

参 考 文 献

- [1] Fang L, Yu F Q. Moving object detection algorithm based on removed ghost and shadow visual background extractor [J]. *Laser & Optoelectronics Progress*, 2019, 56(13): 131002.
方岚, 于凤芹. 去除鬼影及阴影的视觉背景提取运动目标检测算法 [J]. *激光与光电子学进展*, 2019, 56(13): 131002.
- [2] Zhou B, Li R X, Shang Z H, et al. Object detection algorithm based on improved faster R-CNN [J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 101009.
周兵, 李润鑫, 尚振宏, 等. 基于改进的 Faster R-CNN 目标检测算法 [J]. *激光与光电子学进展*, 2020, 57(10): 101009.
- [3] Chen X R, Xi C L, Cao J H. Research on moving object detection based on improved mixture Gaussian model [J]. *Optik*, 2015, 126(20): 2256-2259.
- [4] Wang X, Liu Y, Li G Y. Moving object detection algorithm based on improved visual background extractor algorithm [J]. *Laser & Optoelectronics Progress*, 2019, 56(1): 011007.
王旭, 刘毅, 李国燕. 基于改进视觉背景提取算法的运动目标检测方法 [J]. *激光与光电子学进展*, 2019, 56(1): 011007.
- [5] Huynh-The T, Banos O, Lee S, et al. NIC: a robust background extraction algorithm for foreground detection in dynamic scenes [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(7): 1478-1490.
- [6] Ma Y J, Chen M L, Liu P P, et al. ViBe algorithm-based ghost suppression method [J]. *Laser & Optoelectronics Progress*, 2020, 57(2): 021007.
马永杰, 陈梦利, 刘培培, 等. ViBe 算法鬼影抑制方法研究 [J]. *激光与光电子学进展*, 2020, 57(2): 021007.
- [7] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [8] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [9] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision

- (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [10] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer Vision - ECCV 2016. Lecture Notes in Computer Science. Cham: Springer, 2016, 9905: 21-37.
- [11] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [12] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. [2020-10-25]. <https://arxiv.org/abs/1804.02767>.
- [13] Liu S T, Huang D, Wang Y H. Receptive field block net for accurate and fast object detection[EB/OL]. [2020-10-25]. <https://arxiv.org/abs/1711.07767>.
- [14] Liang X F, McOwan P W, Johnston A. Biologically inspired framework for spatial and spectral velocity estimations[J]. JOSA A, 2011, 28(4): 713-723.
- [15] Golland P, Bruckstein A M. Motion from color[J]. Computer Vision and Image Understanding, 1997, 68(3): 346-362.
- [16] MacLeod R B. Outlines of a theory of the light sense[J]. American Journal of Psychology, 1964, 80(1): 163.
- [17] Johnston A, McOwan P W, Benton C P. Robust velocity computation from a biologically motivated model of motion perception[J]. Proceedings of the Royal Society of London. Series B: Biological Sciences, 1999, 266(1418): 509-518.
- [18] Liu G H, Yang J Y. Exploiting color volume and color difference for salient region detection[J]. IEEE Transactions on Image Processing, 2019, 28(1): 6-16.
- [19] Boulmerka A, Allili M S. Foreground segmentation in videos combining general Gaussian mixture modeling and spatial information [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(6): 1330-1345.
- [20] Guo Y C, Li Z, Liu Y, et al. Video object extraction based on spatiotemporal consistency saliency detection[J]. IEEE Access, 2018, 6: 35171-35181.