

王德铖, 张著洪. 基于果蝇视觉神经网络的单目标运动检测与跟踪[J]. 智能计算机与应用, 2024, 14(4): 1-11. DOI: 10.20169/j.issn.2095-2163.240401

基于果蝇视觉神经网络的单目标运动检测与跟踪

王德铖¹, 张著洪^{1,2}

(1 贵州大学 大数据与信息工程学院, 贵阳 550025; 2 贵州大学 贵州省系统优化与科学计算特色重点实验室, 贵阳 550025)

摘要: 运动目标跟踪因对实时输入信息的处理、跟踪模型的性能有较高要求,使得对应用性较强的目标跟踪模型的构建仍然是较为活跃的研究重点。针对单目标跟踪问题,基于果蝇视觉信息处理机制及目标跟踪具有的固有特征,本文提出不同于现有目标跟踪模型的新型模型。模型设计中,基于果蝇视觉信息处理机制建立改进型前馈果蝇视觉神经网络,进而借助其输出的运动方向量矩阵及运动目标的固有运动特性,构建运动目标的运动方向检测以及位置、速度、偏航角估计模型,由此获得计算复杂度由输入图像的分辨率确定的前馈果蝇视觉目标跟踪模型。比较性的实验表明,相较于经典的和基于深度学习的目标跟踪模型,所获新型目标跟踪模型在多种指标下具有实时处理能力强、跟踪效果好且有较好应用潜力的优点,为目标跟踪研究提供了又一新的解决方案。

关键词: 运动方向检测; 单目标跟踪; 果蝇视觉信息处理机制; 前馈果蝇视觉神经网络

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 2095-2163(2024)04-0001-11

Fly visual neural network based single target motion detection and tracking

WANG Decheng¹, ZHANG Zhuhong^{1,2}

(1 College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China; 2 Guizhou Provincial Characteristic Key Laboratory of System Optimization and Scientific Computation, Guizhou University, Guiyang 550025, China)

Abstract: Due to the high requirements of real-time input information processing and target tracking performance, the construction of target tracking models with strong applicability is still a relatively active research focus. To address the single target tracking problem, a novel target tracking model, which completely differs from existing target tracking models, is proposed based on the visual information processing mechanism of the fly and the intrinsic characteristics of the target tracking problem. In the design of the model, a feedforward fly visual target tracking neural network is developed, in which the computational complexity is decided by the input image's resolution. Herein, on the one hand, an improved feedforward fly visual neural network model is acquired based on the fly visual information processing mechanism, on the other hand a computational model is constructed to perform the attitude estimation of a moving object, which involves the target's motion direction detection, velocity estimation, and yaw angle estimation. Comparative experiments have validated that, compared to classical and deep learning-based object tracking models, the new model has significant advantages over the compared models based on the evaluation indicators. It exhibits the strong abilities of real-time processing, target tracking performance, and good application potential. This provides another new scheme to handling object tracking problems.

Key words: motion direction detection; single target tracking; the visual information processing mechanism of the fly; feedforward fly visual neural network

0 引言

运动目标跟踪是一种依据目标位置和速度进行跟踪的研究分支,涉及的经典跟踪模型或算法可概括为两种类型,即生成式模型与判别式模型。生成

式模型首先通过 Harris 角点检测或 SURF(Speed Up Robust Features)算法提取运动目标的特征,然后在跟踪过程中通过特征匹配获取当前运动目标;陈薇等^[1]针对 Meanshift 算法在干扰场景下目标跟踪难的问题,将 MeanShift 与自适应卡尔曼滤波结合,获

基金项目: 国家自然科学基金(62063002)。

作者简介: 王德铖(1998-),男,硕士研究生,主要研究方向:图像处理。

通讯作者: 张著洪(1966-),男,博士,教授,博士生导师,主要研究方向:数据科学与计算智能,深度学习。Email:zhzhang@gzu.edu.cn

收稿日期: 2023-11-13

得自适应卡尔曼滤波模型,进而利用三帧差分法确定目标所在区域以及利用候选目标的概率密度函数估计目标的位置;夏斯维等^[2]针对传统目标跟踪算法因目标遮挡导致跟踪精度低的问题,借助跟踪失败判别机制及运动目标在图像中每个像素点处的光流与图像的平均光流之差检测目标跟踪成功与否,提出一种基于光流法与分块思想的目标跟踪模型。尽管生成式模型具有设计原理简单且不依赖于先验知识,但跟踪效果差和环境自适应能力弱。判别式模型主要有相关滤波跟踪法和深度学习算法,其中相关滤波跟踪法是一种借助图像处理方法提取目标与背景信息的分类模型,其将运动目标与背景分离,把训练后的分类器用于下一帧中寻找最优预测区域,实现目标定位。首先,在第一帧图像的跟踪区域内利用循环移位等方式提取用于训练相关滤波器的特征;其次,通过后续帧与前一帧作相关运算,获得最大相应点和预测目标的像素点。高帅华等^[3]针对尺度变化及遮挡问题,融合 KCF (Kernelized Correlation Filters)、彩色空间特征与梯度直方图,引入尺度池与卡尔曼滤波器,获得能提升运动目标表征能力以及能抵抗部分遮挡的改进型 KCF 模型。尽管相关滤波法的目标跟踪精度高和跟踪速度快,但因其模型中特征提取与跟踪器的训练是单独处理的,导致模块的输出与期望输出的偏差较大,模型的训练效果难以满足期望的要求。随着深度学习的兴起,越来越多的网络框架在计算机视觉领域已得到应用,使得目标跟踪的研究已上升到新的高度^[4]。Bertinetto 等^[5]通过离线训练得到相似度匹配模型,在跟踪过程中寻找运动目标的最大响应位置,首次获得一种基于孪生网络的目标跟踪模型 SiamFC (Siamese Fully Convolutional)。随后,多种基于深度学习的目标跟踪模型陆续被报道,针对目标跟踪中目标尺度变化导致目标跟踪失败的问题, Li 等^[6]将 SiamFC 与区域候选网络 RPN 结合,提出适用于多尺度目标的 SiamRPN (Siamese Region Proposal Network); Li 等^[7]基于 SiamRPN,引入空间感知采样策略提出 SiamRPN++; Guo 等^[8]将跟踪任务分解为像素类别的分类与该像素处的边界框回归两个子问题,提出 SiamCAR (Siamese Fully Convolutional Classification and Regression); Zhang 等^[9]利用对象感知无锚框网络,以无锚框的方式直接预测目标的位置和尺度,提出新的跟踪器 Ocean。尽管这些模型的运行效率高、目标跟踪效果好、实时性强,但模型训练需有标记的数据集,模型结构复杂,硬件与软

件资源消耗大,且目标跟踪效果极大依赖于已训练的模型,可扩展能力弱。

自然界中,生物视觉系统是生物连接外界的重要纽带,经由视觉神经信息处理机制感知外部环境的光信号强度,产生能刻画外在环境变化的行为量,进而对环境中运动目标的运动方向、位置、速度、轮廓与尺度等特征的变化作出响应,这为构建能解决如目标跟踪问题的前馈神经网络提供了重要的生物启发。随着神经生理学理论的快速发展,前馈视觉、视脑神经网络已得到足够重视, Reichardt 等^[10]基于甲壳虫视觉特性,提出一种对运动方向进行检测的模型 EMD (Elementary Motion Detector); 英国莱斯特大学的 Yue 教授等^[11]基于蝗虫视觉神经系统中小叶巨型运动检测器的响应特性,提出一系列解决小车运动碰撞检测的蝗虫视觉神经网络;姜真^[12]基于蛙眼的视觉信息处理机制,借助 4 种视神经节细胞对环境明暗的敏感性,提出一种运动目标检测与跟踪模型;崔光照^[13]基于鹰视觉系统的生理特点,依据视网膜的分层结构与高级神经中枢的分层投射,构建一种基于鹰视觉响应特性的弱小运动目标检测模型;余翔^[14]受人类的生物视觉系统启发,基于人类视觉系统的长短时程突触互补特性,提出一种边缘检测方法。果蝇作为一种经典模式生物,因体积小、结构简单和响应敏捷,其视脑神经响应特性已初步应用于运动目标的检测, Joesch 等^[15]对果蝇的运动视觉进行研究,证明了感光信号在果蝇的视觉神经中分离为 ON 通道与 OFF 通道。特别地, Missler 等^[16]基于果蝇视觉处理机制及 EMD,获得仅能解决黑白背景下小目标运动检测与跟踪神经结构模型 ANNFPT (A Neural Network for Pursuit Tracking),但模型的设计复杂,较难应用于非黑白噪声环境下运动目标的位置、速度与姿态估计; Eichner 等^[17]基于 EMD 以及果蝇视觉神经层的 ON、OFF 通道的特性,提出一种四象限的运动方向检测器 4QMD (4-Quadrant-Motion-Detector); Shen 等^[18]针对弱小目标的运动检测问题,基于果蝇视野神经纤维网中 ON 和 OFF 通道及小目标运动检测神经元 LC11,提出一种小目标运动检测与跟踪模型。另外, Zhang 等^[19]在 Missler 等的研究基础上,于 2015 年提出一种适用于碰撞检测的前馈人工果蝇视觉神经网络 (Artificial Fly Visual Neural Network, AFVNN)。这些为本文的目标跟踪研究已奠定较好基础。

综上,鉴于经典目标跟踪模型与深度学习方法在目标跟踪中的研究与应用现状,克服传统目标跟

踪方法的建模难及基于深度学习的模型训练代价大的问题, 探寻新型目标跟踪模型与算法仍属于目标检测与跟踪研究面临的科技难题。为此, 本文在已有研究基础上, 基于果蝇视觉神经的信息处理机制, 提出一种响应速度快、运动目标检测与跟踪能力强且具有较好应用潜力的前馈果蝇视觉目标跟踪神经网络 (Feedforward Fly Visual Target Tracking Neural Network, FVTTNN)。

1 果蝇视觉神经信息处理与神经网络

果蝇有两只复眼, 每只复眼由约 3 500 个晶状体小眼组成, 每个小眼包含 8 个光感受器 (R1-R8)。其视觉信息处理主要依次经由感光 (Photoreceptor)、视网膜 (Retina)、薄膜 (Lamina)、髓质 (Medulla)、小叶板 (Lobula) 5 个神经层完成, 各层之间的信息处理机制有显著差异, 但又紧密相连, 如图 1 所示^[20]。光感受器接收外界环境的亮度信息后, 将信息传递给后续神经层, 经信息处理后, 视觉信息变成由视觉神经元输出的膜电位。

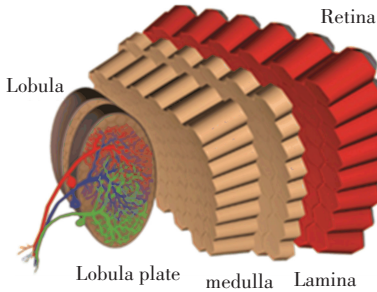


图 1 果蝇视觉神经的信息处理结构图

Fig. 1 The fly visual system's information-processing diagram

感光层中感光细胞接收来自外界环境中的光强度信息, 并将视觉信息转化为视神经信号; 视网膜层从感光层中接收视神经信号, 对同一光轴接收的视神经信号作极化处理, 消除冗余噪声; 薄膜层由两个子层构成, 第一子层为墨盒节点 (Cartridge, cart) 层, 其每个节点接收来自视网膜层中相应节点及其邻近节点的输入, 通过投影或滤波处理产生输出行为量; 第二子层为开关节点 (On-Off, oo) 层, 接收 cart 层中对应节点的输出信号, 并由侧抑制或分流抑制得到输出行为量; 髓质层由 m_1 和 m_2 节点子层构成, m_1 节点层接收来自 oo 层中对应节点及其邻近节点输出的信息, 产生运动目标的局部运动方向量, 而 m_2 节点层中的节点则接收并整合 m_1 节点层中对应节点处邻近节点的输出行为量, 产生该节点在水平与垂直方向上的局部运动方向行为量; 小叶板层中小叶板切向细胞

(Lobula Plate Tangential Cells, LPTCs) 接收 m_2 节点输出的运动方向量, 产生刻画视觉场景中运动目标运动变化状态的全局运动方向量。

基于以上果蝇的视觉响应特性和视觉神经生理学理论, 张著洪等于 2015 年获得适用于碰撞检测问题的果蝇视觉神经网络 (Artificial Fly Visual Neural Network, AFVNN) 的拓扑结构示意图如图 2 所示。

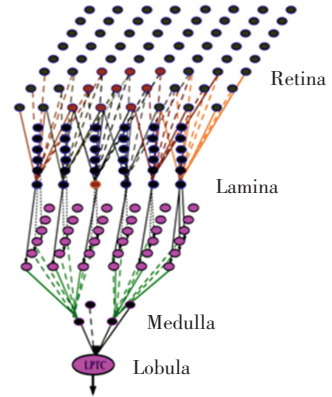


图 2 AFVNN 拓扑结构示意图

Fig. 2 Schematic architecture illustration on AFVNN

其各层的设计简要描述如下:

(1) Retina 层: 由 $M \times N$ 个节点按矩阵排列表示, $M \mid 8, N \mid 8$; 将此矩阵均匀划分为 $m \times n$ 个 8×8 子块 A_{pq} , $1 \leq p \leq m, 1 \leq q \leq n, m = \frac{M}{8}, n = \frac{N}{8}$; 在此, 每个节点 (i, j) 接收当前输入图像 L 中对应像素点 (i, j) 处的灰度值 L_{ij} , 并由式 (1) 对灰度值作平滑处理, 消除噪声信号, 由此获得子块 (p, q) 的输出行为量矩阵 A_{pq} 。

$$p_{ij} = \frac{1}{2}(L_{ij} + g(L_{ij})), 1 \leq i \leq M, 1 \leq j \leq N \quad (1)$$

其中, $g(\cdot)$ 为一帧延时函数。

(2) Lamina 层: 由 cart、oo 节点子层构成, cart 节点子层由 $m \times n$ 个子块按顺序排列而成, 其子块 (p, q) 与 Retina 层中子块 (p, l) 对应, 经由式 (2) 卷积产生大小为 6×6 的行为量矩阵 B_{pq} :

$$B_{pq} = A_{pq} \times W, 1 \leq p \leq m, 1 \leq q \leq n \quad (2)$$

其中, W 为如下 3×3 的高斯卷积核:

$$W = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

所有行为量矩阵按顺序排列后形成矩阵 $B, B = (b_{rs})_{6m \times 6n}$ 。

oo 节点子层由 $6m \times 6n$ 个节点构成。对 cart 节点子层作扩边处理后, 每个节点 (r, s) 首先接收矩

阵 \mathbf{B} 中对应位置 (r, s) 处的值 b_{rs} , 然后经由分流抑制产生兴奋或抑制量 x_{rs} , 如式(3)所示:

$$\dot{x}_{rs} = -A_l x_{rs} + (B_l - x_{rs}) e_{rs}, 0 \leq r < 6m, 0 \leq s < 6n \quad (3)$$

其中, A_l 是正的衰减系数; B_l 是给定的最大兴奋量; e_{rs} 为 cart 节点层中节点 (i, j) 的行为偏差量, 如公式(4)所示:

$$e_{rs} = b_{rs} - \frac{1}{8} \sum_{-1 \leq u, v \leq 1g} (b_{r+u, s+v}), (u, v) \neq (0, 0) \quad (4)$$

(3) Medulla 层: 由 $2m \times 2n$ 个节点按矩阵形式排列而成, 在将 oo 节点子层按大小为 3×3 均匀划分为 $2m \times 2n$ 个子块后, 节点 (p, q) 接收子块 (p, q) 中各节点输出的行为量, 进而借助熟知的运动检测器 EMD, 获局部运动方向行为量 M_{pq} , 如式(5)所示:

$$M_{pq} = \frac{1}{1 + \exp(-m_{pq})} - 0.5, 0 \leq p < 2m, 0 \leq q < 2n \quad (5)$$

其中,

$$m_{pq} = (x(3p+1, 3q+1)f(x(3p+1, 3q+1) - x(3p+u, 3q+v)f(x(3p+1, 3q+1)))^2 \quad (6)$$

(4) LPTC 层: 收集 Medulla 中 $2m \times 2n$ 个节点的输出行为量, 并经由式(7)产生输出行为量:

$$\dot{Y}_{lob} = -A_l Y_{lob} + (B_l - Y_{lob}) \cdot \left(\sum_{p=0}^{2m-1} \sum_{q=0}^{2n-1} M_{pq} - f \left(\sum_{p=0}^{2m-1} \sum_{q=0}^{2n-1} M_{pq} \right) \right) \quad (7)$$

AFVNN 是依据果蝇视觉信息处理机制而设计, 能实时检测运动目标的变化行为, 且适用于刻画运动目标的逼近行为。其与碰撞检测方案结合, 可解决复杂背景下的碰撞检测问题。为了使其能适应于目标跟踪问题的求解, 对其网络结构、各神经层的设计加以修改和扩展, 提出刻画视觉场景中单个运动目标的运动方向变化状态的改进型前馈果蝇视觉神经网络 (Improved Artificial Fly Visual Neural Network, IAFVNN)。

2 改进型人工果蝇视觉神经网络

基于果蝇视觉信息处理的神经层结构关系, IAFVNN 仍由 Retina 层、Lamina 层、Medulla 层和 Lobula 层顺次连接而成, 如图 3 所示。Retina 层、Lamina 层与 Medulla 层均由 $M \times N$ 个节点构成; Retina 层是信号接收及信号平滑处理层; Lamina 层由 cart、oo 节点子层构成, 每个子层均由 $M \times N$ 节点

矩阵表示, 接收来自 Retina 的输出信号后, 经卷积滤波进行平滑处理以及利用分流抑制模型产生输出行为量; Medulla 层由 m_1, m_2 节点子层构成, 每个子层也由 $M \times N$ 个节点构成, 经由四象限运动检测器模型 4QMD 获取方向偏差量, 进而经由投影产生在水平与垂直方向的方向量。另外, Lobula 由 $m \times n$ 个小叶板切向细胞 (Lobula Plate Tangential Cells, LPTCs) 按顺序排列形成的矩阵表示, $m = M/2, n = N/2$; 每个 LPTC 被视为一个节点, 每个节点汇聚 Medulla 中特定区域内节点的输出, 经侧抑制机制产生运动目标的运动方向行为量。

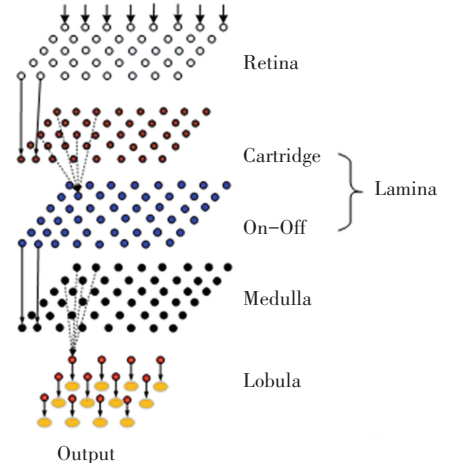


图3 IAFVNN 拓扑结构示意图

Fig. 3 Schematic architecture illustration on IAFVNN

各神经层的设计如下:

1) Retina 层: 与 AFVNN 中 Retina 层的设计相同, 每个节点输出行为量 p_{ij} 。

2) Lamina 层: 由 cart、oo 节点子层构成, 经由滤波处理滤除 Retina 层中节点输出的噪声信息且增强有用信息。具体设计如下:

(1) cart 节点子层: 对 Retina 层作扩边处理后, 每个 cart 节点接收 Retina 层中对应节点处 3×3 个节点的输出行为量, 通过滑动窗口大小为 3×3 、移动步长为 1 的方式进行卷积滤波。为使运动目标提取过程中边缘信息不受滤波的影响, 利用双边滤波对 Retina 层输出的行为量进行平滑处理, cart 节点 (i, j) 的输出行为量由式(8)计算:

$$x_{\text{cart}}(i, j) = \frac{1}{\sum_{u, v=0}^2 w_{uv}(i, j)} \sum_{u, v=0}^2 w_{uv}(i, j) p_{i+u, j+v-1} \quad (8)$$

其中, $1 \leq i \leq M, 1 \leq j \leq N, w_{uv}(i, j)$ 是在 cart 节点 (i, j) 处 3×3 动态权重矩阵 $\mathbf{W}(i, j)$ 中位置 (u, v) 处的元素, 如式(9)所示:

$$w_{uv}(i, j) = w_{d, uv}(i, j) \times w_{r, uv}(i, j) \quad (9)$$

$w_{d,uv}(i, j)$ 是以 cart 节点 (i, j) 为中心的空间域高斯核函数在 (u, v) 处的值,如式(10)所示:

$$w_{d,uv}(i, j) = \exp\left(-\frac{(u-i)^2 + (v-j)^2}{2\sigma_d^2}\right) \quad (10)$$

其中, σ_d 是空间域标准差。

类似地, $w_{r,uv}(i, j)$ 是以 cart 节点 (i, j) 为中心,在 (u, v) 处的值域高斯核函数值,其经由式(11)计算:

$$w_{r,uv}(i, j) = \exp\left(-\frac{\|L_{i+u-1, j+v-1} - L_{ij}\|^2}{2\sigma_r^2}\right) \quad (11)$$

其中, σ_r 是值域标准差。

以上 cart 节点子层的模型设计中,利用非线性的滤波模型对图像的空间邻近度与像素值相似度进行折中处理,结合图像的空域位置与灰度信息,在滤除噪声的同时,保留图像的边缘信息。

(2) oo 节点层:每个节点 (i, j) 接收 cart 节点层中对应节点处输出的行为量,进而由式(12)计算其兴奋或抑制行为量:

$$\begin{aligned} \dot{x}_{oo}(i, j) = & -A_l x_{oo}(i, j) + \\ & [E_{\text{cart}}(i, j)(B_l - x_{oo}(i, j))]^+ + \\ & [E_{\text{cart}}(i, j)(B_l + x_{oo}(i, j))]^- \end{aligned} \quad (12)$$

其中, $1 \leq i \leq M, 1 \leq j \leq N, A_l$ 为式(3)中对应的符号, B_l 为当前所有 cart 节点输出的值中的最大值, $[n]^+ = \max(0, n), [n]^- = \min(0, n)$; 式(12)被用于模拟果蝇视觉神经的兴奋与抑制的响应特性;当 $x_{oo}(i, j)$ 较大时, oo 节点 (i, j) 处于兴奋状态,此暗示视觉场景中相应像素点处含有运动目标信息;反之,则无运动目标。另外, $E_{\text{cart}}(i, j)$ 为 cart 节点层中节点 (i, j) 的行为偏差量,如式(13)所示:

$$E_{\text{cart}}(i, j) = x_{\text{cart}}(i, j) - g(x_{\text{cart}}(i, j)) \quad (13)$$

3) Medulla 层:由 m_1, m_2 两个子层构成,检测运动目标的局部运动方向以及产生运动目标在水平、垂直方向的局部运动方向量。

(1) m_1 节点子层:4QMD 是一种受生物视觉启发的四象限运动检测器模型,由 4 个关联检测器模型构成,如图 4 所示。每一个关联检测器由两个相同的 EMD 对称排列而成,通过检测相邻两个感光细胞接收的亮度信号对两点间的运动方向进行判断,能检测运动目标的局部运动方向并抑制静止图像的响应;通过设置一个并行的信号时延与乘法结构,提取运动目标的局部运动信息。图 4 中, oo_c, oo_a 分别表示中心与邻近 oo 节点;从左至右的 4 个关联检测器依次对应 on-on、on-off、off-on、off-off 通道;由汇

集 4 个并行通道的信息,获取 m_1 节点的输出行为量,即方向偏差量。

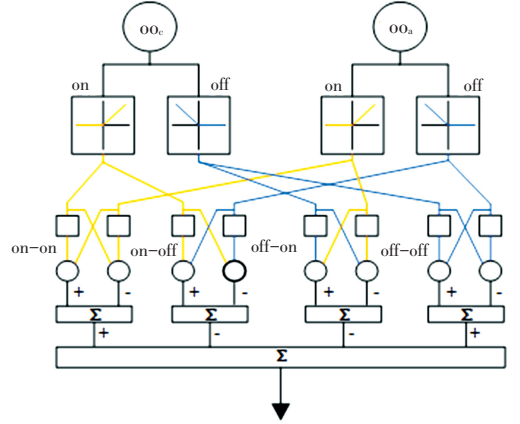


图 4 四象限运动检测器模型

Fig. 4 4-Quadrant-Detector model

每个 m_1 节点接收 oo 节点层中对应节点的输出行为量后,经由半波整流将此行为量分解为 ON 和 OFF 通道的分量,由以上 4 条并行通道产生 4 种行为量;随后,4QMD 汇集如此 4 种行为量,获得 m_1 节点的输出行为量,其计算过程描述如下:

对 oo 节点子层作扩边处理后, m_1 节点 (i, j) 首先接收 oo 节点子层中对应节点输出的行为量 $x_{oo}(i, j)$ 及其 3×3 邻域内其余 8 个节点输出的行为量 $x_{oo}(i+k, j+l)$, $-1 \leq k, l \leq 1, (k, l) \neq (0, 0)$; 当 (k, l) 依次取 $-2 \sim 2$ 时,此 8 个节点的编号依次记为 1 ~ 8。于是,对于节点 (i, j) 及第 1 个节点,即节点 $(i-1, j-1)$,经由 on-on 通道及式(14)获得行为量:

$$m_{11}^{\text{on-on}}(i, j) = [f(x_{oo}(i+k, j+l))]^+ [x_{oo}(i, j)]^+ - [f(x_{oo}(i, j))]^+ [x_{oo}(i+k, j+l)]^+ \quad (14)$$

类似地,经由 on-off、off-on、off-off 通道可获得如下行为量:

$$m_{11}^{\text{on-off}}(i, j) = [f(x_{oo}(i+k, j+l))]^+ [-x_{oo}(i, j)]^+ - [f(x_{oo}(i, j))]^+ [-x_{oo}(i+k, j+l)]^+ \quad (15)$$

$$m_{11}^{\text{off-on}}(i, j) = [-f(x_{oo}(i+k, j+l))]^+ [x_{oo}(i, j)]^+ - [-f(x_{oo}(i, j))]^+ [x_{oo}(i+k, j+l)]^+ \quad (16)$$

$$m_{11}^{\text{off-off}}(i, j) = [-f(x_{oo}(i+k, j+l))]^+ [-x_{oo}(i, j)]^+ - [-f(x_{oo}(i, j))]^+ [-x_{oo}(i+k, j+l)]^+ \quad (17)$$

经以上四象限运动检测器可获 m_1 节点 (i, j) 与第 1 个节点确定的运动方向行为量,如式(18)所示:

$$\hat{m}_{11}(i, j) = m_{11}^{\text{on-on}}(i, j) + m_{11}^{\text{off-off}}(i, j) - m_{11}^{\text{on-off}}(i, j) - m_{11}^{\text{off-on}}(i, j) \quad (18)$$

类似地可获 oo 节点 (i, j) 与邻近的其余 7 个

节点产生的行为量 $\hat{m}_{1r}, 2 \leq r \leq 8$ 。

(2) m_2 节点子层: 每个 m_2 节点 (i, j) 接收 m_1 节点子层中对应节点处产生的 8 个按图 5 所示排列的输出行为量 $\hat{x}_{m_{1k}}(i, j), 1 \leq k \leq 8$, 由式 (19) 和式 (20) 计算其在水平、竖直方向的局部运动方向量:

$$x_{m_{2h}}(i, j) = \hat{x}_{m_{11}}(i, j) - \hat{x}_{m_{13}}(i, j) + \hat{x}_{m_{18}}(i, j) - \hat{x}_{m_{14}}(i, j) + \hat{x}_{m_{17}}(i, j) - \hat{x}_{m_{15}}(i, j) \quad (19)$$

$$x_{m_{2v}}(i, j) = \hat{x}_{m_{11}}(i, j) - \hat{x}_{m_{17}}(i, j) + \hat{x}_{m_{12}}(i, j) - \hat{x}_{m_{16}}(i, j) + \hat{x}_{m_{13}}(i, j) - \hat{x}_{m_{15}}(i, j) \quad (20)$$

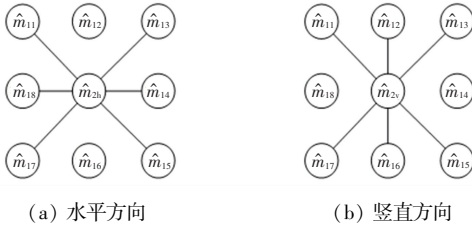


图 5 m_2 节点的结构

Fig. 5 The architecture of m_2 node

4) Lobula 层: 由 $m \times n$ 个神经元按矩阵排列, 每个神经元被视为一个节点; 将 m_2 节点子层按 2×2 划分为 $m \times n$ 个子块, Lobula 中节点 (r, s) 接收对应子块 M_{rs} 中各节点在水平、竖直方向的行为量, 由式 (21) ~ 式 (22) 可获 Lobula 层中 $m \times n$ 个节点在水平、竖直方向的行为量, 用于检测视觉场景中运动目标的运动方向变化状况。

$$\hat{x}_h(r, s) = -A_l x_h(r, s) + [(B_l - x_h(r, s)) M_{h,r}]^+ + [(B_p + x_h(r, s)) M_{h,r}]^- \quad (21)$$

$$\hat{x}_v(r, s) = -A_l x_v(r, s) + [(B_l - x_v(r, s)) M_{v,r}]^+ + [(B_p + x_v(r, s)) M_{v,r}]^- \quad (22)$$

其中, $1 \leq r \leq m, 1 \leq s \leq n, M_{h,r}$ 和 $M_{v,r}$ 分别是子块 M_{rs} 中各节点在水平、竖直方向上的行为量之和。

当视觉场景中没有运动目标时, 节点在水平、竖直方向的行为量均为 0, 反之则视觉场景中运动目标。

3 运动目标跟踪模型与算法

AFVNN 作为一种前馈视觉神经网络, 借助 Lobula 层实时获知视觉场景中单运动目标的运动方向, 但不能获知运动目标的整体运动状态。以下给出单运动目标的运动位置、速度和偏航角估计模型。

3.1 目标检测模型

IAFVNN 中输出层 (Lobula) 的各节点在水平、竖直方向的方向量被用于计算该节点处的运动方向

量, 节点 (r, s) 处的运动方向量用式 (23) 计算:

$$\hat{x}_g(r, s) = \frac{1}{1 + \exp(-|x_h(r, s) + x_v(r, s)| - T_g)} - 0.5 \quad (23)$$

其中, T_g 是给定的运动量阈值。

若 $\hat{x}_g(r, s) > 0$, 则视 Lobula 层中节点 (r, s) 处含有运动目标的运动信息; 反之, 则视此点处不存在运动目标。借助式 (24) 得到的 0 或 1 按所在位置排列, 构成 0-1 矩阵 \mathbf{X}_l :

$$x_l(r, s) = \begin{cases} 1, \hat{x}_g(r, s) > 0 \\ 0, \hat{x}_g(r, s) \leq 0 \end{cases} \quad (24)$$

若 \mathbf{X}_l 中值为 1 的位置形成连通区域, 则此区域定义为运动目标所在区域, 且此区域在水平、竖直方向的最长连通长度分别取为运动目标的宽度与高度; 将由此区域所在矩形框的中心取为运动目标的重心。由此, 运动目标的全局运动向量由式 (25) 计算:

$$x_d = \sum_{r=0}^m \sum_{s=0}^n x_l(r, s) |x_g(r, s)| \quad (25)$$

3.2 速度与偏航角估计模型

根据运动目标的位置信息矩阵 \mathbf{X}_l , 确定是否存在运动目标; 若存在运动目标, 则其运动速度由式 (26) 计算:

$$v = \frac{\sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2}}{\Delta t} \quad (26)$$

其中, (x_1, y_1) 和 (x_2, y_2) 分别为上一帧、当前帧运动目标的重心位置, Δt 为相邻两帧的时间间隔。

另外, 运动目标偏离水平方向的偏航角 α 由式 (27) 计算:

$$\tan \alpha = \frac{\sum_{r=0}^m \sum_{s=0}^n x_l(r, s) |x_v(r, s)|}{\sum_{r=0}^m \sum_{s=0}^n x_l(r, s) |x_h(r, s)|} \quad (27)$$

4 前馈果蝇视觉目标跟踪神经网络

FVTNN 由 IAFVNN 和目标检测、运动速度与偏航角估计模型按顺序连接而成, 实时地由 IAFVNN 中 Retina 层接收视觉灰度图像, 然后由 Lamina、Medulla 神经层对视觉信息进行处理, 由 Lobula 层输出刻画运动目标的运动方向变化状态的运动方向量。在此基础上, 利用目标检测模型估计运动目标的重心和运动目标所在区域的矩形框, 借助速度与偏航角估计模型估计运动目标的运动速度和偏离航向的程度。FVTNN 算法流程如图 6 所示。

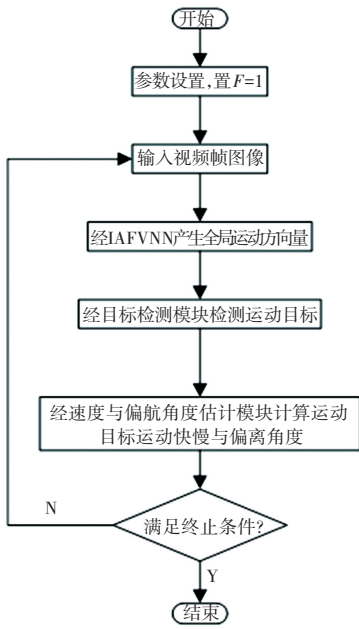


图6 FVTNN的算法流程图

Fig. 6 Algorithm flowchart on FVTNN

FVTNN的算法描述如下:

- 步骤1** 参数设置: $M \times N$, 空间域标准差 σ_d , 值域标准差 σ_r , 衰减系数 A_l , 运动量阈值 T_g
- 步骤2** 置 $t \leftarrow 1$
- 步骤3** 执行 IAFVNN
- 步骤3.1** Retina 层接收第 t 帧灰度图 G_t
- 步骤3.2** 由式(8)~式(11)、式(12)~式(13)分别计算 Lamina 层中 cart、oo 节点的行为量 x_{cart} 和 x_{oo}
- 步骤3.3** 由式(14)~式(20)计算 Medulla 层中 m_1 、 m_2 节点的行为量
- 步骤3.4** 由式(21)~式(22)计算 Lobula 层中各节点处在水平和垂直方向的行为量
- 步骤4** 由式(23)~式(25)确定运动目标的位置及所在位置的矩形框
- 步骤5** 由式(26)~式(28)计算运动目标的运动速度 v 和偏航角 α
- 步骤6** $t \leftarrow t + 1$
- 步骤7** 返回步骤3,直到目标跟踪结束

由以上算法描述可知,FVTNN的计算复杂度主要由 IAFVNN 确定。在一次循环中,IAFVNN 的 Retina 层执行 $2MN$ 次算术运算;Lamina 层中的 cart 节点子层执行 $46MN$ 次运算, on-off 节点子层执行 $11MN$ 次运算, Lamina 层共执行 $57MN$ 次运算; Medulla 层的 m_1 节点子层执行 $248MN$ 次运算, m_2 节点子层执行 $10MN$ 次运算,因此其共执行 $258MN$ 次运算; Lobula 层执行 $7MN$ 次运算。由此可知,

FVTNN 的计算复杂度为 $O(324MN)$, 即只要适当设置 M 和 N 的取值,其执行效率能得到保障。

5 数值实验

在配置为 Windows10/CPU/R5-5600X 3.7 GHz/RAM/16 GB/Matlab R2016a 环境下展开数值实验。选取适用于单目标跟踪的果蝇视觉神经网络 ANNFPT, 相关滤波算法 KCF, 深度学习中孪生网络 SiamFC、SiamRPN 与 SiamRPN++ 参与比较, 其参数设置源于相应文献; FVTNN 的参数设置为 $M = 128$, $N = 128$, $\sigma_d = 3$, $\sigma_r = 0.1$, $A_l = 100$, $T_g = 0.25$ 。测试事例包括实际场景下摄像机拍摄的1个物体滚动视频序列以及2个行人正常行走、奔跑情形下的基准视频序列。

5.1 评价指标

选取平均重叠率 (Average Overlap Rate, AOR)、重叠阈值分别为 0.5 与 0.75 时的成功率 $SR_{0.5}$ 与 $SR_{0.75}$ 以及精确度 (Precision) 对模型的性能进行定量分析。

重叠率是指预测框与真实框交集内像素点个数与并集内像素点个数之比, 公式(28):

$$IoU = \frac{B^{pr} \cap B^{gt}}{B^{pr} \cup B^{gt}} \quad (28)$$

其中, B^{pr} 表示预测框, B^{gt} 表示真实框。

进一步, 获得 m 帧图像的平均重叠率, 用于刻画预测框锚定目标的效果, 公式(29):

$$AOR = \frac{1}{m} \sum_{k=1}^m \frac{B^{pr}(k) \cap B^{gt}(k)}{B^{pr}(k) \cup B^{gt}(k)} \quad (29)$$

AOR 越大, 则锚定目标的效果越好; 反之, 则锚定目标的效果越差。

$SR_{0.5}$ 与 $SR_{0.75}$ 分别指在重叠阈值为 0.5 与 0.75 情况下的重叠率不小于给定阈值 T_s (取 0.5, 0.75) 的帧数占总帧数的比例, 用于刻画正确锚定目标程度的高低, 如式(30)所示:

$$SR = \frac{1}{m} \sum_{k=1}^m I_k \quad (30)$$

其中,

$$I_k = \begin{cases} 1, & IoU(k) \geq T_s \\ 0, & \text{else} \end{cases} \quad (31)$$

若 SR 越大, 则正确锚定目标的程度越高; 反之, 则越低。

精确度 (Pr) 是指真实框与预测框的重心之间的欧氏距离不大于给定误差阈值的帧数占总帧数的比例, 用于刻画预测框与真实框之间的接近程度, 由

式(32)计算:

$$Pr = \frac{1}{m} \sum_{k=1}^m I_k \quad (32)$$

其中

$$I_k = \begin{cases} 1, & d_k \leq T_s \\ 0, & \text{else} \end{cases} \quad (33)$$

d_k 为算法在第 k 次运行获得的预测框与真实框的重心之间的距离; Pr 越大, 则预测框越接近目标; 反之, 则预测框越远离目标。

5.2 实验结果与分析

情形1 行人行走

行人在最右侧位置以平均运动速度 4.06 pixels/s 沿水平方向向左步行移动, 生成的视频经解帧后得到由 84 帧构成的图像序列, 每帧的分辨率为 128×128 , 如图 7 所示, 箭头和矩形框分别为 FVTNN 获得的运动目标的运动方向和锚定运动目标的矩形框。通过 FVTNN 获得的每帧的速度与偏航角见

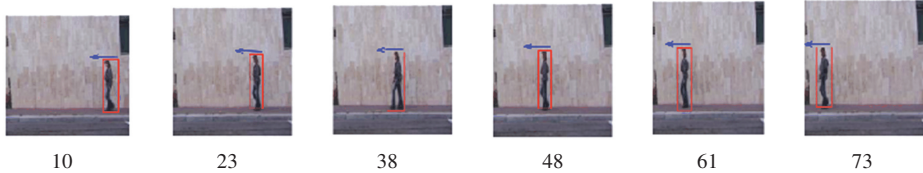


图7 行人行走及 FVTNN 的目标锚定与方向检测

Fig. 7 FVTNN's target anchoring and direction detection in pedestrian walking motion

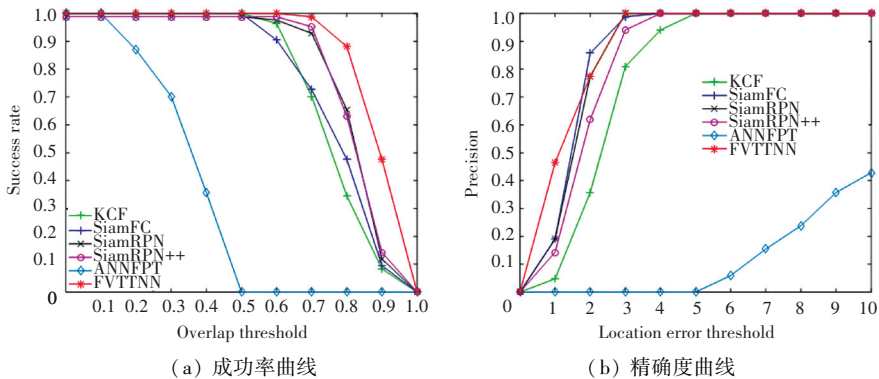


图8 在情形1下各模型获得的成功率曲线与精确度曲线比较

Fig. 8 Comparison of success rate curves and accuracy curves of different methods in condition 1

由表 2 可知, 在行人行走场景下, FVTNN 相较于参与比较的模型, 获得的平均重叠率 AOR 、成功率 $SR_{0.5}$ 与 $SR_{0.75}$ 均明显较高, 同时 AOR 值说明其获得的预测框能较好地锚定运动目标; 另外, $SR_{0.5}$ 、 $SR_{0.75}$ 及 Pr 值说明生成的预测框锚定运动目标的程度较高且能较好地接近运动目标。因此, FVTNN 能很好地跟踪图像序列中的运动目标, 跟踪效果与参与比较模型获得的跟踪效果相比具有显著优势。在参与比较的模型中, ANNFPT 的跟踪性能最差, 其

获得的平均覆盖率最低, 且获得关于 $SR_{0.5}$ 、 $SR_{0.75}$ 及 Pr 的值均为 0, 仅适用于黑白背景下的 ANNFPT 并不适用于情况 1 图像序列的目标跟踪。表 2 中数值表明 KCF 作为一种相关滤波模型并不能有效处理图像中行人的姿态变化, 导致平均覆盖率仅有 76%, 且当阈值为 0.75 时, 锚定运动目标的程度较低; 其次, SiamFC、SiamRPN 及 SiamRPN++ 作为深度学习方法, 得到的平均覆盖率介于 77%~82% 之间, 同时随着阈值的增大, $SR_{0.5}$ 、 $SR_{0.75}$ 值表示其不能在

表1 情形1: 运动偏航角检测结果

Table 1 Yaw angle detection results in condition 1

帧号	10	23	38	48	61	73
速度 (pixels · s ⁻¹)	4.8	7.6	7.2	7.6	7.6	5.4
偏航角/(°)	181.6	181	181.9	181.3	181.3	180.3

表2 情形1 各模型的性能测试比较

Table 2 Comparison of the models' performance test results in condition 1

方法	AOR / %	$SR_{0.5}$ / %	$SR_{0.75}$ / %	Pr / %
KCF	76.0	100	58.3	94.0
SiamFC	77.1	100	64.3	100
SiamRPN	81.3	100	79.8	100
SiamRPN++	81.4	100	84.5	100
ANNFPT	33.9	0	0	0
FVTNN	83.6	100	97.6	100

高重叠阈值情形下有效地锚定运动目标,也说明样本的规模对其模型的训练效果有重要影响。

图 7 中箭头和矩形框表示 FVTTNN 不仅能跟踪运动的运动方向,还能较为准确地跟踪运动目标的位置。由于行人在水平方向行走时会有微小踮脚动作,导致在竖直方向上产生微小运动且运动幅度远小于在水平方向上的运动幅度,因此 FVTTNN 得到的行人运动方向与水平方向的夹角较小,此与实际场景下行人运动方向的变化一致。图 8 表明重叠阈值低于 0.5 时,与 KCF 和 ANNFPT 相比,FVTTNN 的目标跟踪成功率较高;与 SiamFC、SiamRPN 及 SiamRPN++相比,有相同的成功率;当重叠阈值高于

0.5 时,与比较的模型相比,FVTTNN 的跟踪成功率较高,相较于其它模型能更好地锚定运动目标。

情形 2 行人快速奔跑

行人在最右侧位置沿着水平方向以平均运动速度 7.8 pixels/s 向左方向快速奔跑,由此生成的视频解帧后得到由 42 帧构成的图像序列,每帧的分辨率为 128×128 ,如图 9 所示。通过 FVTTNN 获得的每帧的速度与偏航角见表 3;将参与比较的模型与 FVTTNN 均作用于图 8 中的图像序列,获得各自在此图像序列下的统计结果见表 4;各模型在对运动目标跟踪过程中形成的成功率、精确度曲线如图 10 所示。



图 9 行人奔跑运动及 FVTTNN 的目标锚定与方向检测

Fig. 9 Target anchoring and direction detection in pedestrian running motion

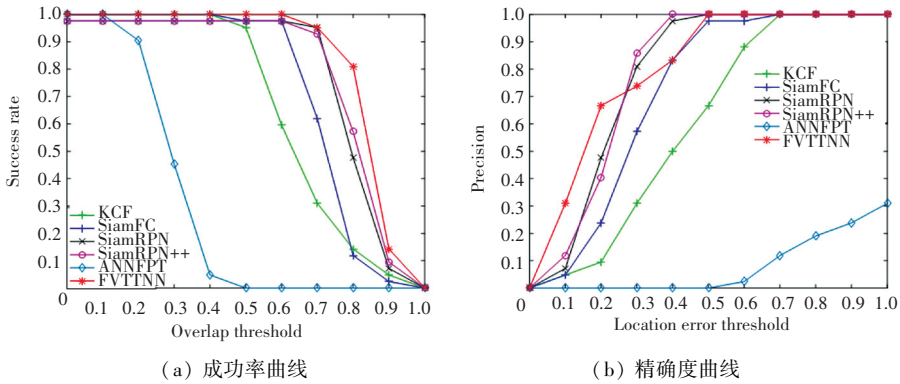


图 10 在情形 2 下各模型的成功率曲线与精确度曲线比较

Fig. 10 Comparison of the models' curves on success rate and precision in condition 2

表 3 情形 2:运动偏航角的检测结果

Table 3 Yaw angle detection results in condition 2

帧号	6	15	23	27	32	40
速度/(pixels · s ⁻¹)	9.9	14.0	14.4	7.6	13.6	9.9
偏航角/(°)	180.3	184.2	182.4	182.5	180.1	180.3

表 4 情形 2:各模型的性能测试比较结果

Table 4 Comparison of the models' performance test results in condition 2

方法	AOR /%	SR _{0.5} /%	SR _{0.75} /%	Pr /%
KCF	64.8	95.2	19.0	50.0
SiamFC	72.9	97.6	42.9	83.3
SiamRPN	78.5	97.6	83.3	97.6
SiamRPN++	79.3	97.6	83.3	100
ANNFPT	29.7	0	0	2.4
FVTTNN	82.6	100	88.1	83.3

由表 4 可知,在行人快速奔跑场景下,FVTTNN

相较于参与比较的模型,获得的平均重叠率 AOR、成功率 SR_{0.5} 与 SR_{0.75} 均明显提高,这些值与 Pr 值表明其得到的预测框锚定运动目标的程度较高且能较好地接近运动目标,因此能很好地跟踪图像序列中的运动目标,其跟踪效果与参与比较模型的跟踪效果相比具有明显优势。另外,在参与比较的模型中,ANNFPT 的跟踪性能最差、获得的平均覆盖率最低、得到的 SR_{0.5}、SR_{0.75} 及 Pr 值均为 0,说明该模型不适宜图 8 图像序列的目标跟踪;表 4 中数值表明 KCF 作为一种相关滤波模型,并不能有效处理图像中行人快速运动下姿态的变化,导致平均覆盖率仅有 65%,且当阈值为 0.75 时,锚定运动目标的程度较低。SiamFC、SiamRPN 及 SiamRPN++作为深度学习方法得到的平均覆盖率介于 72%~80%之间,同时

随着阈值的增大,其 $SR_{0.5}$ 、 $SR_{0.75}$ 的值表明其不能在高重叠阈值情形下有效地锚定运动目标,但与 FVTNN 相比,跟踪精确度较高,表明在快速运动情形下深度学习算法获得的预测框能较好地接近运动目标,但计算资源开销较大。

图9中箭头和矩形框表明 FVTNN 不仅能跟踪运动目标的运动方向,且能较为准确地跟踪运动目标的位置。由于行人在水平方向奔跑时会产生小幅度的踮脚动作,导致在竖直方向上产生运动且运动幅度小于在水平方向上的运动幅度,因此 FVTNN 得到的行人运动方向与水平方向的夹角较小,这与实际场景下行人的运动方向变化一致。图10(a)表明在所有重叠阈值区间内,FVTNN 的跟踪成功率相较于其它模型的成功率较高;图10(b)表明在距离误差阈值所在区间[2,5]内,FVTNN 的跟踪精度低于 SiamRPN 与 SiamRPN++ 的跟踪精度,但整体跟踪精确度较高。

情形3 物体滚动

实验室环境下,一物体在工作台面板上以平均运动速度 5.2 pixels/s 由左向右运动,其尺度在视觉区内发生变化,并逐渐向右上方运动,随后逐渐偏离

摄像机。在此情形下,视频解帧后生成由53帧构成的图像序列如图11所示;每帧在分辨率为 128×128 下的速度与偏航角见表5。参与比较的模型与 FVTNN 的统计结果见表6;各模型产生的成功率、精确度曲线如图12所示。

表5 情形3:运动偏航角检测结果

Table 5 Yaw angle detection results in condition 3

帧号	7	17	24	30	40	47
速度 ($\text{pixels} \cdot \text{s}^{-1}$)	9.6	4.8	7.6	4.8	4.8	2.4
偏航角 ($^{\circ}$)	181.3	185.5	184.6	182.4	181.6	180.1

表6 情形3:各算法的性能测试比较结果

Table 6 Comparison of the algorithms' performance test results in condition 3

方法	AOR / %	$SR_{0.5}$ / %	$SR_{0.75}$ / %	Pr / %
KCF	78.2	100	71.7	100
SiamFC	76.4	100	55.8	100
SiamRPN	76.2	100	62.3	100
SiamRPN++	76.5	100	66.0	100
ANNFPT	29.7	0	0	0.094
FVTNN	91.6	100	100	100

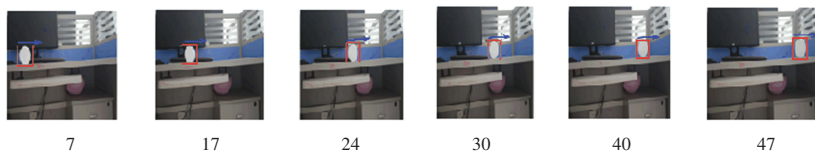


图11 物体滚动及 FVTNN 的目标锚定与方向检测

Fig. 11 Target anchoring and direction detection in a object scrolling

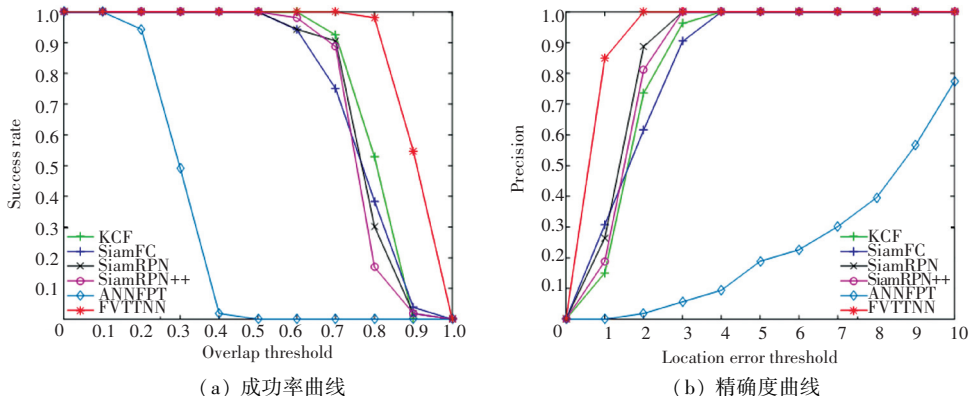


图12 在情形3下:各算法的成功率曲线与精确度曲线比较

Fig. 12 Comparison of success rate curves and accuracy curves of different methods in condition 3

由表6可知,在物体滚动场景下,FVTNN 相较于参与比较的模型,获得的 AOR 、 $SR_{0.5}$ 、 $SR_{0.75}$ 、 Pr 值均较高,获得的预测框能较好地锚定运动目标,跟踪效果具有显著优势。另外,在参与比较的模型中,ANNFPT 的跟踪性能最差,获得的平均覆盖率最低,

且获得关于 $SR_{0.5}$ 、 $SR_{0.75}$ 及 Pr 的值均为0,说明该模型也不适宜情形3图像序列中运动目标的跟踪,主要原因:一方面其仅适合于跟踪黑白背景下的运动目标,另一方面其缺乏有效的噪声处理方法,同时其运动目标检测模型的稳健性较弱;表6中的数值表

明 KCF 作为一种相关滤波模型,能较好地检测做规则运动的目标,得到的平均覆盖率为 78%,与其它参与比较的模型相比,能较好地锚定运动目标。SiamFC、SiamRPN 及 SiamRPN++作为深度学习方法得到的平均覆盖率介于 76%~77%之间,同时随着阈值的增大, $SR_{0.5}$ 、 $SR_{0.75}$ 值表明其不能在高重叠阈值情况下有效地锚定运动目标,也说明样本的规模对其模型的训练效果有重要影响。

图 11 中箭头和矩形框表明 FVTTNN 不仅能跟踪运动目标的运动方向,而且得到的矩形框说明其能较为准确地跟踪运动目标的位置;图 12 表明, FVTTNN 的跟踪成功率及跟踪精度均较高。图 12 成功率与精确度曲线表明 FVTTNN 能更好地锚定运动目标。因此,在物体滚动情形下, FVTTNN 在正常范围内不同重叠率阈值与距离误差阈值下,目标跟踪能力较强。

6 结束语

单目标运动检测与跟踪是一个经典且也是热点的研究问题。经典目标跟踪方法存在环境噪声消除难、目标位置估计精度低、跟踪实时性差的问题,基于深度学习的目标跟踪方法的模型训练成本高且目标跟踪性能极大依赖于样本的容量。为此,本文在深入研究已有适用于碰撞检测的前馈果蝇视觉神经网络(AFVNN)基础上,构建改进型前馈果蝇视觉神经网络并建立运动目标的位置、速度以及偏航角度估计模型,获得前馈果蝇视觉目标跟踪神经网络,此神经网络的计算复杂度由输入图像的分辨率决定。比较性的实验结果表明,本文的神经网络与参与比较的相关滤波模型、ANFPT 和基于深度学习的模型相比,目标跟踪性能具有明显优势,不仅能实时、准确地确定运动目标的运动方向,而且也能通过所获视觉神经网络有效抑制噪声对目标跟踪效果的影响,目标跟踪效果好。

参考文献

- [1] 陈薇,袁文定,方强,等.基于自适应卡尔曼滤波的 Meanshift 跟踪算法[J].制造业自动化,2021,43(6):16-20,48.
- [2] 夏斯维,陈迅.基于光流法与分块思想的目标跟踪算法[J].计算机工程与设计,2020,41(7):2063-2068.
- [3] 高帅华,张建花,王创海.基于改进 KCF 算法的加油软管锥套目标跟踪研究[J/OL].电光与控制:1-6[2023-07-12].
- [4] MA C, HUANG J B, YANG X, et al. Hierarchical convolutional features for visual tracking [C]// Proceedings of the IEEE

- International Conference on Computer Vision. IEEE, 2015:3074-3082.
- [5] BERTINETTO L, VALMADRE J, HENRIQUES J F, et al. Fully-convolutional siamese networks for object tracking [C]// Proceedings of the Computer Vision - ECCV 2016 Workshops. Amsterdam. 2016: 850-865.
- [6] LI B, YAN J, WU W, et al. High performance visual tracking with siamese region proposal network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8971-8980.
- [7] LI B, WU W, WANG Q, et al. Evolution of siamese visual tracking with very deep networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 15-20.
- [8] GUO D, WANG J, CUI Y, et al. SiamCAR: Siamese fully convolutional classification and regression for visual tracking [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 6269-6277.
- [9] ZHANG Z, PENG H, FU J, et al. Ocean: Object-aware anchor-free tracking [C]//Computer Vision-ECCV 2020: 16th European Conference. 2020:771-787.
- [10] REICHARDT W, ROSENBLITH W A. Autocorrelation, a principle for evaluation of sensory information by the central nervous system [C]//Proceedings of Symposium on Principles of Sensory Communication. 1959:303-317.
- [11] YUE S, RIND F C. Collision detection in complex dynamic scenes using an LGMD-based visual neural network with feature enhancement [J]. IEEE Transactions on Neural Networks, 2006, 17(3): 705-716.
- [12] 姜真.基于蛙眼视觉特性的运动目标检测与跟踪方法研究[D].长沙:湖南大学,2013.
- [13] 崔光照.基于生物视觉系统的小目标运动检测研究[D].重庆:重庆大学,2020.
- [14] 余翔.生物视觉启发的智能计算模型及其应用[D].杭州:杭州电子科技大学,2021.
- [15] JOESCH M, SCHNELL B, RAGHU S V, et al. ON and OFF Pathways in Drosophila Motion Vision [J]. Nature, 2010, 468: 300-304.
- [16] MISSLER J M, KAMANGARF A. A neural network for pursuit tracking inspired by the fly visual system [J]. Neural Networks, 1995, 8(3): 463-480.
- [17] EICHNER H, JOESCH M, SCHNELLB. Internal structure of the fly elementary motion detector [J]. Neuron, 2011, 70(6): 1155-1164.
- [18] SHEN K, YANG Y, LIANG Y, et al. Modeling Drosophila vision neural pathways to detect weak moving targets from cluttered backgrounds [J]. Computers and Electrical Engineering, 2022, 99: 107678.
- [19] ZHANG Z, YUE S, ZHANG G. Fly visual system inspired artificial neural network for collision detection [J]. Neurocomputing, 2015, 153: 221-234.
- [20] ALEXANDER B, MICHAEL D, MATTHIAS M. The neural network behind the eyes of a fly [J]. Current Opinion in Physiology, 2020, 16: 33-42.