

基于多任务辅助学习特征的瞳孔中心检测



贾静¹, 付高波¹, 赵歆波^{1,2}, 邹晓春^{1,2}, 张宝尚²

1. 西北工业大学, 陕西 西安 710072

2. 光电控制技术重点实验室, 河南 洛阳 471009

摘 要: 眼动跟踪技术在航空领域极具潜力, 其作为关键的信息获取和人机交互手段, 可用于精确的目标瞄准、飞行员状态监测, 对飞行安全和作战精准意义重大。而瞳孔中心检测作为核心技术, 决定了眼动追踪系统的鲁棒性和准确性, 本文提出一种基于多任务辅助学习特征的瞳孔中心检测算法, 模拟人类视觉系统的多任务辅助学习特性, 引入多任务模块, 实现从粗到精的瞳孔中心检测, 并通过试验验证了本文方法的有效性和先进性, 从而提高了眼动跟踪的精度和准确性, 为航空领域的各项任务和操作提供了更精确、高效和安全的手段。

关键词: 瞳孔中心检测; 深度学习; 辅助特征; 多任务学习; 多尺度分类

中图分类号: TP391.4

文献标识码: A

DOI: 10.19452/j.issn1007-5453.2023.09.014

近年来, 随着计算机视觉、图形学、人工智能等新兴技术的发展以及应用需求的增加, 学术界和工业界对眼动跟踪技术进行了大量深入研究。人工智能技术与装备的有机结合, 将极大地提升装备作战效能^[1], 眼动技术应用于飞行训练方面的研究也不断发展, 从早期的眼动仪系统在飞行模拟器中的测试, 到现代眼动仪用于真实空中驾驶舱的当代试验, 利用眼动仪收集眼动数据, 监测飞行员的注意状态^[2], 眼动追踪技术在航空航天领域的应用正蓬勃发展。在瞄准具设计中, 眼动跟踪技术提供实时的目标识别和定位以改进瞄准精度, 可提高战斗人员的操作效率和打击效果。在飞行员行为分析中, 利用眼动追踪技术来评估注意力分配和疲劳程度, 以采取相应的措施进行提醒或干预。在飞行器界面设计中, 眼动跟踪技术被用来记录、评估飞行员的眼动情况, 以减少操作复杂性并量化用户体验。总之, 眼动追踪技术通过采集分析眼动数据, 为航空领域提供了实时精准的量化信息, 为瞄准打击、飞行安全和人机交互提供了数据基础。

而鲁棒准确的瞳孔中心定位是眼动追踪的前提和核心技术。在复杂环境下, 现有瞳孔检测方法的准确性和鲁棒

性不够稳定, 这与人类视觉相比是不可接受的, 即人类具有很强的鲁棒性。根据认知心理学的相关研究^[3], 人脑具有多任务辅助学习的特点。凭借这一特性, 人类在运动模糊、光照变化、噪声等复杂环境下仍能轻松完成视觉定位任务。因此, 本文认为人类视觉系统的这一特性可以促进瞳孔定位的准确性。

在瞳孔检测任务中, 现有的瞳孔定位方法可分为传统的瞳孔定位方法和基于深度学习的瞳孔定位方法。现有的瞳孔定位方法从瞳孔定位任务本身出发, 忽视了人类视觉系统在处理瞳孔定位问题上的优越性。因此, 无论是基于传统的图像处理还是基于深度学习, 在复杂的环境下都很难取得令人满意的效果。

为了解决上述问题, 受多任务辅助学习特性的启发, 本文提出了一种用于瞳孔中心检测的由粗到精的神经网络架构, 如图1所示, 其中, 蓝色背景代表粗分类阶段, 红色背景代表精细回归阶段。在网络的不同分支中定义了两个不同的坐标维度。粗分类阶段和精细回归阶段是级联的。该架构主要分为粗分类和精细回归两个阶段。在粗分类阶段, 将坐标值划分为不同的类别。这种分类在相应的尺度空间中执行多次。精

收稿日期: 2023-04-27; 退修日期: 2023-07-03; 录用日期: 2023-08-07

基金项目: 国家自然科学基金(61871326); 航空科学基金(20185453031, 20200051053003)

引用格式: Jia Jing, Fu Gaobo, Zhao Xinbo, et al. Pupil center detection inspired by multi-task auxiliary learning characteristic[J]. Aeronautical Science & Technology, 2023, 34(09): 121-126. 贾静, 付高波, 赵歆波, 等. 基于多任务辅助学习特征的瞳孔中心检测[J]. 航空科学技术, 2023, 34(09): 121-126.

细回归阶段进一步细化粗分类结果,计算出准确的瞳孔中心位置。粗分类阶段的主要任务用来辅助回归任务,以提高准确率。瞳孔中心的横坐标和纵坐标在两个独立的分类和回归模块中同时计算。它们共享相同的全图像卷积特征映射,这也在训练过程中引入了类似于多任务学习(MTL)^[4]的优势。此外,在开放标准和具有挑战性的数据集上,将本文方法与大量优秀的EISE^[5]、ExCuSe^[6]、SET^[7]、DeepEye^[8]等方法进行了比较,得到了最先进的结果。

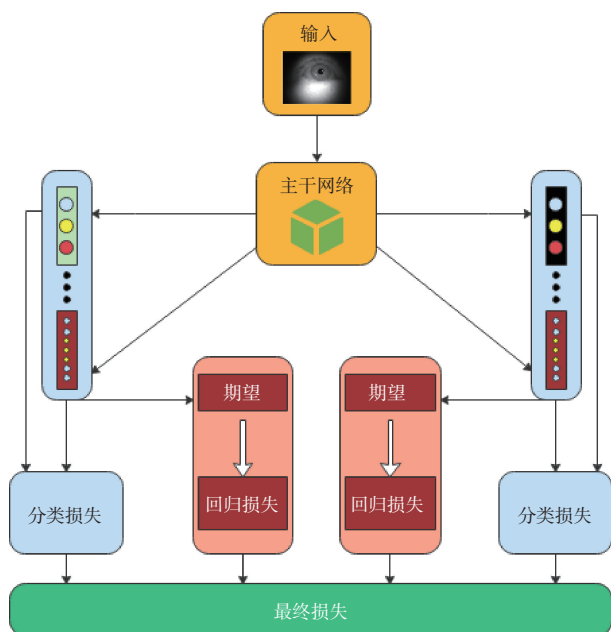


图1 本文提出的由粗到精网络的结构

Fig.1 The architecture of the proposed coarse-to-fine network

1 方法介绍

1.1 多任务辅助学习特征建模

为了提高瞳孔中心检测的鲁棒性和准确性,本文在瞳孔中心检测任务中引入多任务模块来模拟人类视觉系统的多任务辅助学习特性。多任务模块主要由两部分组成:一部分是在粗分类阶段使用多个具有不同尺度因子的子任务辅助主任务,以优化瞳孔中心检测的鲁棒性。另一部分通过对分类任务和回归任务的联合优化得到瞳孔中心坐标,从而提高瞳孔中心的检测精度。

假设输入图像为 I ,通过基本网络得到的输出为 $g(I)$; Tasks表示不同尺度空间中的分类任务,定义为

$$\text{Tasks} = \begin{cases} \text{subTask}_i(g(I)), & i = \{1, 2, 3, 4, 5, 6, 7, 8\} \\ \text{mainTask}_i(g(I)), & i = \{9\} \end{cases} \quad (1)$$

式中, i 表示分类任务的个数。在尺度空间中,将强约束分类和弱约束分类相结合,强约束分类表示主任务,弱约束分类表示子任务。

在获得多尺度分类节点的基础上,一方面,计算不同尺度空间中多任务单元的交叉熵损失。多任务分类损失定义为

$$L = \sum_{i=1}^9 \alpha_i \times C_i \quad (2)$$

式中, α_i 为不同尺度分类任务的权重系数; C_i 表示每个任务的交叉熵损失。

另一方面,在主要任务节点的基础上建立了回归模型。首先通过主要任务节点求解最大期望值 $E(\text{mainTask}(g(I)))$,定义为

$$E(\text{mainTask}(g(I))) = \sum_{i=1}^n K_i \times \text{Softmax}(\text{nodes}(i)) \quad (3)$$

式中, n 为主要任务分类节点个数;Nodes(i)表示主任务第 i 个节点的输出; K_i 表示节点 i 对应的分类类别。

因此,Softmax(nodes(i))表示第 i 个节点对应瞳孔中心坐标的分类概率。

在得到瞳孔中心坐标的最大期望值后,通过均方误差计算回归损失。将多任务分类损失和回归损失相结合对模型进行训练,以优化瞳孔中心坐标。最终损失 L 定义为

$$L = L_{\text{subTask}} + L_{\text{mainTask}} + \beta \times L_{E(\text{mainTask}(g(I)))} \quad (4)$$

式中; β 为回归损失函数的权值; $L_{E(\text{mainTask}(g(I)))}$ 是真实值与期望值之间的均方误差。

为了保证提取的特征尺度的一致性,Vera-Olmos等提出的DeepEye要求输入图像必须具有固定的大小,否则很难获得满意的结果。由于成像特性的不同,图像尺寸也会有所不同,这阻碍了深度学习网络鲁棒性的提高。在本文中,通过比例尺映射来解决这个问题。其思想是将不同大小的图像映射到固定大小的比例尺空间进行计算。

1.2 粗分类

粗分类主要包括点分类和多尺度约束。

1.2.1 点分类

N. Ruiz等^[9]首先使用姿态角范围(bin)分类提出了用于头部姿态估计的HopeNet。现有的bin类函数为 $B(\text{point}_i)=[a, b]$ 。如果 point_i 属于 $[a, b]$,则 $B(\text{point}_i)=a$,最大极值误差为 $E=[0, |b-a|]$ 。对于精确的回归阶段,这个误差太大而无法接受。为解决这一问题,本文在粗分类阶段提出一种新的点分类方法。在尺度空间中定义不同的等距点,每个点代表一个范畴。在点分类中,输出值属于最近的

点。当距离相等时,它们属于具有较高节点的点。与bin分类方法相比,本文所提出的点分类方法可以根据不同的任务改变距离测量,扩展了高维数据空间。引入点分类,为粗分类阶段提供了一种相对公平稳定的分类策略。

1.2.2 多尺度约束

在定义分类尺度空间时,通过设置分类点的个数来确定每个类别在尺度空间中的范围。在粗分类阶段,首先,定义了一个包含513个分类点的尺度空间作为主要任务。其次,通过减少分类点的数量,定义了257、129、65、33、17、9、5、3个分类点,同时进行训练。

可以选择骨干网络进行多次试验,并扩展多个完全连接的层来预测瞳孔中心坐标。每一个全连通层代表一个不同的分类尺度空间。每个尺度空间都有一个确定的交叉熵损失。分类点越少,损失的尺度空间对应的权重越大。当分类范围较大时,错误的分类对应更高的惩罚。每个维度的分类损失定义为

$$\text{ClassificationLoss} = \sum_{i=1}^9 \alpha_i \times C(y_i, y_i^*) \quad (5)$$

式中: $C(y_i, y_i^*)$ 表示交叉熵损失; α_i 表示不同尺度空间的权重。

1.3 精细回归阶段

经过粗分类阶段的计算,在一个较小的范围内确定各个维度的坐标值,该范围代表一个类别。而在实际应用中,瞳孔位置是一个准确值。为得到更精确的瞳孔中心坐标,本文扩展了粗分类阶段之后的细回归阶段。

在精细回归阶段,只使用粗分类阶段的主要任务,它具有更精细的分类范围,因此精细回归任务更容易收敛。用softmax层输出的数学期望来表示回归方程,定义为

$$y^* = \sum_{i=1}^{i=513} i \times \frac{e^{Z_i}}{\sum_{m=1}^{m=513} e^{Z_m}} \quad (6)$$

式中: Z 表示来自全连接层的矢量输出和softmax层的输入; i 表示分类类别; 分母 $\sum_{m=1}^{m=513} e^{Z_m}$ 起归一化作用。

回归损失定义为

$$\text{RegressionLoss} = \text{MSE}(y, y^*) \quad (7)$$

在定义了回归方程之后,选择均方误差损失作为回归损失。最后,生成两个最终损失,分别对应瞳孔定位任务中图像的两个不同维度。每个最终损失被定义为分类损失和回归损失的线性组合。每个维度的最终损失定义为

$$\text{Loss}_{\text{final}} = \text{ClassificationLoss} + \beta \times \text{RegressionLoss} \quad (8)$$

式中, RegressionLoss 表示均方误差损失函数; β 表示回归损失权值。

1.4 训练

在训练过程中,首先按照8/2的比例对LPW数据集进行分割,最终生成104685张训练图像和26171张测试图像。在将图像输入网络之前,将其大小调整为64×64。使用Resnet-50作为骨干网络,训练了30个epoch。在反向传播中使用Adam优化器。初始学习率设为0.001,每10次学习率减小,其中 $\beta_1=0.9$, $\beta_2=0.999$, $\eta=10^{-8}$ 。此外,在粗分类阶段定义了多个分类点,分类点为257的尺度空间的分类范围最小,对整体任务的贡献最大,而分类点为3的尺度空间贡献最小,根据Wang Haofan^[10]所提出的方案以及在训练测试过程中调整参数确定了权重因子。 $\alpha_i=\{1, 2, 3, 5, 7, 9, 11, 13, 15\}$, $\beta=0.15$ 。

2 试验

2.1 试验装置

2.1.1 数据集

为了测试和训练所提出的网络,在室内和室外场景中使用不同的传感器采集瞳孔图像,它们包括三个已发表的具有挑战性的红外瞳孔图像数据集。

(1) LPW: 该数据集包含66个瞳孔区域视频。该数据集由22人使用头戴式眼动仪收集,共生成130856个视频帧。该数据集涵盖了室内和室外的照片条件。

(2) ElSe: 该数据集包含55712张图像。数据集I-V通过眼动追踪设备在道路驾驶环境中记录。数据集VI-VII在室内试验中收集。

(3) PupilNet: 该数据集包含41217张图像。数据集PNI-PNV包括更快的变化、不利的光源和眼睛生理结构的破坏。

2.1.2 指标

为了评估和公平地比较本文提出的网络与其他算法,针对不同的阶段定义了不同的评价指标。

(1) 检测率: 如果预测坐标与真实标签之间的欧氏距离小于5px,则认为检测成功。检测成功的图像与所有图像的比值被定义为检测率。

(2) 分类检测率: 对于分类任务的每个维度,定义输出结果与真实标签之间的误差小于等于3px,表示分类成功。成功图像的比例定义为分类检测率。

(3) 分类精度: 当两个不同的维度同时分类成功时,就定义了分类精度。

(4) 绝对误差: 每个维度的预测值与真实值之间的像素误差的绝对值。

(5) 欧氏距离误差: 预测坐标与实际坐标之间的像素

误差的欧氏距离。

2.1.3 实现细节

本文收集了有代表性的算法,将 SET、Swisski、ElSe、ExCuSe、DeepEye 等算法与本文方法进行了比较。所有试验均在一台台式计算机上进行。

2.2 试验分析

2.2.1 分类网络与分类方法评价

为了获得更稳定的分类模型,考虑到不同的代表性模型 AlexNet、VGGNet、ResNet-50,在现有研究绝大多数任务上的性能,ResNet-50 具有最高的性能。因此,本文选择 Resnet-50 作为骨干分类网络。

此外,为了测试点分类的准确性,本文将其与 Ruiz 等在 LPW 数据集上提出的 bin 分类方法进行了比较。同样,在粗分类阶段,使用之前定义的评价指标来计算两个维度的分类检测率。在精细回归阶段,使用不同维度的平均绝对误差和标准差作为评价指标。最后,计算了不同分类方法的平均欧氏距离误差。

试验结果见表 1,分类阶段以分类精度作为评价指标。用平均绝对误差(MAE)、平均欧氏距离误差(MEDE)比较回归阶段。由表 1 可见,与 bin 分类相比,本文的点分类方法在两个维度的平均分类精度上提高了 0.77%,在回归精度上平均欧氏距离误差降低了 16.58%。与 bin 分类相比,点分类更加公平、合理,不会带来数据偏移。

表 1 bin 分类与本文在 LPW 数据集上的点分类在分类和回归阶段的对比

Table 1 Comparison between bin classification and proposed point classification on LPW data sets in classification and regression stages

阶段	方法	LPW		
		X	Y	
分类	bin	94.77%	92.71%	
	本文	95.55%	93.48%	
回归		X	Y	MEDE
		MAE	MAE	
	bin	1.38	1.04	1.93
	本文	1.11	0.89	1.61

2.2.2 多任务辅助特征有效性评价

为了验证架构的有效性,在 LPW 数据集上进行了分析试验,设置了两个对照组。在对照组 1 中,去除多尺度空间约束,只保留一个尺度空间与主任务。在对照组 2 中,删除主任务的分类损失,保留其他子任务的分类损失,结合最高尺度空间的回归损失对网络进行训练。

试验结果见表 2。对照组 1 的检出率相对较低,因为删除了很多分类子任务。仅使用主分类任务,网络难以稳定收敛。对照组 2 的分类精度较低,因为去掉了最高尺度的分类损失,很难得到准确的分类结果。

表 2 多任务辅助效能控制试验(通过分类准确率和检测率对结果进行可视化)

Table 2 Multi-task auxiliary effectiveness control experiment (The results are visualized through classification accuracy and detection rate)

方法	分任务	主任务	准确率/%	检测率/%
对照组 1	×	√	73.36	92.20
对照组 2	√	×	34.51	95.82
本文	√	√	92.82	97.21

2.3 试验对比

为了充分证明模型的性能,建立了对比试验。首先,将该模型与传统的瞳孔定位算法 Swiski、SET、ExCuSe 和 ElSe 进行了比较,将检出率作为评价指标。选择 LPW 作为测试数据集,因为它具有更多的连续帧图像和更高的图像分辨率。为了进行更详细和全面地比较,将每个参与者采集的图像视为一个独立的数据集,并对每个分割的数据集执行不同的算法。

表 3 给出了本文模型和传统算法在 LPW 数据集上的检测率。与传统算法相比,本文提出的模型有了显著的改进。还计算了每种方法在所有数据集上的平均检出率。本文方法的检出率为 97.2%。在传统的瞳孔定位方法中,ElSe 方法的定位效果最好。然而,本文方法仍有大约 28% 的改进。

图 2 描述了在 LPW 上小于特定欧氏距离的预测和手工标记的地面真值的像素误差的平均检测率。当允许像素误差小于 1 时,本文方法的检测率接近 60%。当允许像素误差小于 3 时,对应的检测率大于 90%。与传统的最佳算法相比,本文方法提高了 30% 以上。因此,本文方法适用于高精度瞳孔定位。这也表明本文模型具有更高的检测率和更好的鲁棒性。

本文选择了在 LPW 数据集上表现较好的两种传统算法 ExCuSe 和 ElSe 以及两种深度学习算法 PupilNet 和 DeepEye 进行比较。

对于所有 12 个数据集,本文模型在 8 个数据集上实现了更高的检测率(见表 4)。与 PupilNet 相比,所提出的方法将所有数据集的平均检出率提高了约 20%。与 DeepEye 相比,平均检出率略有提高 6%。ElSe 在传统方法中表现出更好的性能,但本文方法在平均检出率上仍然提高了 30%。

表3 本文方法与传统方法在LPW数据集上的检测率比较
Table 3 Comparison of detection rate between proposed method and traditional methods on LPW dataset

数据集	Swirski	SET	ExCuSe	ELSe	本文
1	0.84	0.57	0.64	0.88	0.99
2	0.42	0.49	0.30	0.70	0.99
3	0.31	0.28	0.35	0.58	0.96
4	0.17	0.08	0.25	0.37	0.98
5	0.08	0.07	0.19	0.23	0.96
6	0.63	0.11	0.53	0.84	0.94
7	0.66	0.44	0.66	0.74	0.98
8	0.78	0.42	0.75	0.81	0.94
9	0.56	0.36	0.60	0.62	0.97
10	0.71	0.10	0.59	0.73	0.94
11	0.32	0.31	0.50	0.71	0.99
12	0.72	0.55	0.73	0.90	0.98
13	0.27	0.15	0.45	0.52	0.94
14	0.76	0.30	0.59	0.71	0.99
15	0.38	0.27	0.45	0.54	0.97
16	0.74	0.23	0.73	0.82	0.99
17	0.69	0.21	0.42	0.73	0.99
18	0.61	0.51	0.66	0.79	0.99
19	0.25	0.12	0.22	0.54	0.98
20	0.41	0.2	0.12	0.83	0.98
21	0.56	0.3	0.47	0.89	0.96
22	0.06	0.42	0.31	0.70	0.98
平均值	0.497	0.295	0.478	0.690	0.972

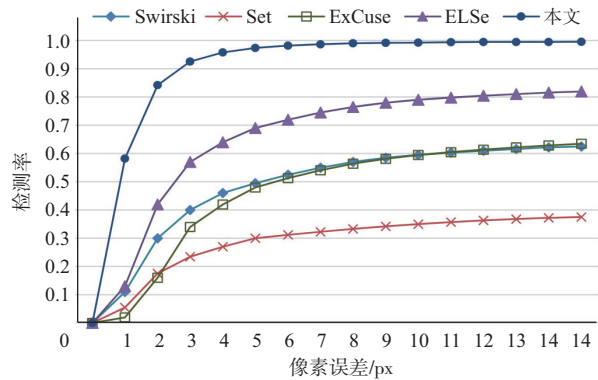


图2 所测试方法在LPW数据集上的像素误差和平均检测率分析
Fig.2 Pixel error and average detection rate analysis of tested method on LPW dataset

为比较不同方法的时间效率,进行了一个对比试验。首先,在LPW上随机选择100张图像。其次,每种方法执行100次,计算每张图像的平均处理时间,试验结果见表5。与DeepEye和ExCuSe相比,本文方法分别提高了1.34倍和3.04倍。

表4 本文方法与最先进的方法在混合数据库上的检测率比较
Table 4 Comparison of detection rate between proposed method and the best-advanced method on the hybrid dataset

数据集	ELSe	ExCuSe	PupilNet	DeepEye	本文
I	0.57	0.24	0.62	0.74	0.80
II	0.33	0.23	0.37	0.78	0.75
III	0.78	0.58	0.79	0.92	0.80
IV	0.47	0.52	0.83	0.88	0.78
V	0.53	0.26	0.58	0.80	0.84
VI	0.94	0.93	0.90	0.96	1.00
VII	0.53	0.46	0.55	0.55	0.99
PN I	0.62	0.22	0.69	0.64	0.89
PN II	0.26	0.16	0.45	0.83	0.94
PNIII	0.39	0.34	0.49	0.79	0.90
PN IV	0.54	0.48	0.82	0.90	0.84
PN V	0.75	0.59	0.81	0.83	0.83
Average	0.559	0.418	0.658	0.802	0.863

表5 不同方法在LPW数据集上的时间效率比较
Table 5 Comparison of time efficiency between different methods on LPW data sets

方法	Else	ExCuSe	DeepEye	本文
时间/ms	171.89	46.49	26.86	11.50

3 结论

本文受人类视觉系统特点的启发,探索了人类视觉系统的多任务辅助学习特征与瞳孔中心检测任务之间的隐藏关系,将多任务辅助学习特征引入瞳孔定位任务,提出了一种从粗到精多任务协同优化瞳孔定位方法。本文方法在LPW和ELSe、ExCuSe上取得了最先进的结果,分别获得了97.2%和86.3%的检测率,距离误差小于5px。试验结果验证了多任务辅助学习特征的有效性。在未来的工作中,本文提出的网络架构可以应用到其他类似的任务中,并且可以根据不同的需求灵活更换相应的骨干网。

AST

参考文献

[1] 吴明忠,刘永志. 人工智能技术在直升机领域的应用及发展展望[J]. 航空科学技术, 2021, 32(1):29-34.
Wu Mingzhong, Liu Yongzhi. Application and development prospect of artificial intelligence technology in helicopter field [J]. Aeronautical Science & Technology, 2021, 32(1):29-34. (in Chinese)
[2] 张益凡,王宇超,张琴喻,等. 基于眼动指标的飞行员注意状

- 态识别可行性研究[J]. 航空科学技术, 2022, 33(4):39-46.
- Zhang Yifan, Wang Yuchao, Zhang Qinyu, et al. Feasibility study on pilot attention state recognition based on eye movement index [J]. Aeronautical Science & Technology, 2022, 33(4):39-46.(in Chinese)
- [3] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the National Academy of Sciences, 2017, 114(13):3521-3526.
- [4] Zhang Yu, Yang Qiang. An overview of multi-task learning[J]. National Science Review, 2018, 5(1):30-43.
- [5] Fuhl W, Santini T C, Thomas C K, et al. ElSe: Ellipse selection for robust pupil detection in real-world environments[C]// Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. ACM, 2016.
- [6] Fuhl W, Kübler T C, Sippel K, et al. Excuse: Robust pupil detection in real-world scenarios[C]. Computer Analysis of Images and Patterns 16th International Conference, 2015.
- [7] Javadi A H, Hakimi Z, Barati M, et al. SET: A pupil detection method using sinusoidal approximation [J]. Frontiers in Neuroengineering, 2015, 8:4
- [8] Vera O, Pardo F J, Melero E, et al. DeepEye: Deep convolutional network for pupil detection in real environments[J]. Integrated Computer-Aided Engineering, 2019, 26(1):85-95.
- [9] Ruiz N, Chong E, Rehg J M. Fine-grained head pose estimation without keypoints[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2018.
- [10] Wang Haofan, Chen Zhenghua, Zhou Yi. Hybrid coarse-fine classification for head pose estimation[EB/OL]. (2020-10-15). <https://arxiv.org/abs/1901.06778>.

Pupil Center Detection Inspired by Multi-task Auxiliary Learning Characteristic

Jia Jing¹, Fu Gaobo¹, Zhao Xinbo^{1,2}, Zou Xiaochun^{1,2}, Zhao Baoshang²

1. Northwestern Polytechnical University, Xi'an 710072, China

2. Science and Technology on Electro-optic Control Laboratory, Luoyan 471009, China

Abstract: Eye tracking technology has great potential in the aviation field. As a key means of information acquisition and human-computer interaction, it can be used for accurate target aiming and pilot condition monitoring, which is of great significance to flight safety and combat accuracy. As the core technology, pupil center detection determines the robustness and accuracy of the eye tracking system. This paper proposes a pupil center detection algorithm based on multi-task assisted learning features, simulates the multi-task assisted learning characteristics of the human visual system, and introduces a multi-task module to achieve pupil center detection from coarse to fine. Experiments have verified the effectiveness and advanced nature of the proposed method. In turn, it improves the precision and accuracy of eye tracking, providing a more accurate, efficient and safe means for various tasks and operations in the aviation field.

Key Words: pupile center detection; deep learning; auxiliary characteristic; multi-task learning; multiscale classification

Received: 2023-04-27; Revised: 2023-07-03; Accepted: 2023-08-07

Foundation item: National Natural Science Foundation of China(61871326);Aeronautical Science Foundation of China(20185453031, 20200051053003)