



基于动态视觉传感器的铝基盘片表面缺陷检测

马居坡 陈周熠 吴金建

Dynamic Vision Sensor Based Defect Detection for the Surface of Aluminum Disk

MA Ju-Po, CHEN Zhou-Yi, WU Jin-Jian

在线阅读 View online: <https://doi.org/10.16383/j.aas.c240307>

您可能感兴趣的其他文章

基于轻量化重构网络的表面缺陷视觉检测

Visual Inspection of Surface Defects Based on Lightweight Reconstruction Network

自动化学报. 2022, 48(9): 2175–2186 <https://doi.org/10.16383/j.aas.c200535>

基于计算机视觉的工业金属表面缺陷检测综述

A Review of Metal Surface Defect Detection Based on Computer Vision

自动化学报. 2024, 50(7): 1261–1283 <https://doi.org/10.16383/j.aas.c230039>

基于事件相机的机器人感知与控制综述

A Survey of Robot Perception and Control Based on Event Camera

自动化学报. 2022, 48(8): 1869–1889 <https://doi.org/10.16383/j.aas.c210263>

基于事件相机的定位与建图算法: 综述

Event-based Visual Localization and Mapping Algorithms: A Survey

自动化学报. 2021, 47(7): 1484–1494 <https://doi.org/10.16383/j.aas.c190550>

基于深度学习的表面缺陷检测方法综述

A Survey of Surface Defect Detection Methods Based on Deep Learning

自动化学报. 2021, 47(5): 1017–1034 <https://doi.org/10.16383/j.aas.c190811>

基于自适应全局定位算法的带钢表面缺陷检测

Strip Surface Defect Detection Based on Adaptive Global Localization Algorithm

自动化学报. 2024, 50(8): 1550–1564 <https://doi.org/10.16383/j.aas.c210467>

基于动态视觉传感器的铝基盘片表面缺陷检测

马居坡^{1,2} 陈周熠¹ 吴金建^{1,2}

摘要 现有视觉缺陷检测技术通常基于传统电荷耦合器件 (Charge-coupled device, CCD) 或互补金属氧化物半导体 (Complementary metal-oxide-semiconductor, CMOS) 相机进行缺陷成像和后端检测算法开发。然而, 现有技术存在成像速度慢、动态范围小、背景干扰大等问题, 难以实现对高反光产品表面弱小瑕疵的快速检测。针对上述挑战, 创新性地提出了一套基于动态视觉传感器 (Dynamic vision sensor, DVS) 的缺陷检测新模式, 以实现对具有高反光特性的铝基盘片表面缺陷的高效检测。DVS 是一种新型的仿生视觉传感器, 具有成像速度快、动态范围大、运动目标捕捉能力强等优势。首先开展了面向铝基盘片高反光表面弱小瑕疵的 DVS 成像实验, 并分析总结了 DVS 缺陷成像的特性与优势。随后, 构建了第一个基于 DVS 的缺陷检测数据集 (Event-based defect detection dataset, EDD-10k), 包含划痕、点痕、污渍三类常见缺陷类型。最后, 针对缺陷形态多变、纹理稀疏、噪声干扰等问题, 提出了一种基于时序不规则特征聚合框架的 DVS 缺陷检测算法 (Temporal irregular feature aggregation framework for event-based defect detection, TIFF-EDD), 实现对缺陷目标的有效检测。

关键词 缺陷检测, 动态视觉传感器, 高反光表面, 不规则特征提取, 时序融合, 事件相机

引用格式 马居坡, 陈周熠, 吴金建. 基于动态视觉传感器的铝基盘片表面缺陷检测. 自动化学报, 2024, 50(12): 2407-2419

DOI 10.16383/j.aas.c240307 **CSTR** 32138.14.j.aas.c240307

Dynamic Vision Sensor Based Defect Detection for the Surface of Aluminum Disk

MA Ju-Po^{1,2} CHEN Zhou-Yi¹ WU Jin-Jian^{1,2}

Abstract Current visual defect detection technologies usually rely on conventional charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) cameras for defect imaging and the development of backend detection algorithms. However, these technologies encounter challenges such as slow imaging speed, limited dynamic range, and significant background interference, which hinder the rapid detection of minor defects on highly reflective product surfaces. To address these challenges, we innovatively propose a new defect detection mode based on dynamic vision sensor (DVS) to achieve efficient defect detection on the highly reflective surfaces of aluminum disks. DVS is a novel bio-inspired visual sensor with advantages such as fast imaging speed, high dynamic range, and excellent ability to capture moving objects. First, we conduct DVS imaging experiments for minor defects on the highly reflective surfaces of aluminum disk and analyze the characteristics and advantages of DVS on defect imaging. Then, we establish the first event-based defect detection dataset (EDD-10k) based on DVS, including three common defect types: Scratch, point and stain. Finally, to address the issues such as varying defect shapes, sparse textures, and noise interference, we propose a temporal irregular feature aggregation framework for event-based defect detection (TIFF-EDD), and realize the effective detection of defect targets.

Key words Defect detection, dynamic vision sensor (DVS), highly reflective surface, irregular feature extraction, temporal fusion, event camera

Citation Ma Ju-Po, Chen Zhou-Yi, Wu Jin-Jian. Dynamic vision sensor based defect detection for the surface of aluminum disk. *Acta Automatica Sinica*, 2024, 50(12): 2407-2419

收稿日期 2024-06-03 录用日期 2024-07-23

Manuscript received June 3, 2024; accepted July 23, 2024

国家重点研发计划 (2023YFA1008500), 陕西省自然科学基金基础研究计划 (2024JC-YBQN-0627), 中央高校基本科研业务费专项资金 (XJSJ23079, ZYTS24006) 资助

Supported by National Key Research and Development Program of China (2023YFA1008500), Natural Science Basic Research Program of Shaanxi (2024JC-YBQN-0627), and the Fundamental Research Funds for the Central Universities (XJSJ23079, ZYTS24006)

本文责任编辑 孟凡利

Recommended by Associate Editor MENG Fan-Li

1. 西安电子科技大学人工智能学院 西安 710071 2. 琶洲实验室 (黄埔) 广州 510555

1. School of Artificial Intelligence, Xidian University, Xi'an 710071 2. Pazhou Lab (Huangpu), Guangzhou 510555

工业品表面缺陷是指在制造过程中出现的影响产品性能、外观或安全性的表面瑕疵、损伤、裂纹、凹陷等产品问题^[1]。这些缺陷往往由制造过程中的加工工艺、机械损伤、环境污染等因素引起。发展缺陷检测技术对改良生产工艺、提升产品质量有重要的意义。

目前缺陷检测主要分为三大类: 人工检测方法, 利用介质的检测方法和视觉缺陷检测方法。人工检测方法效率低、准确性不高, 受人工经验和主观因素影响大, 不适应现代大规模工业生产。利用介质

的检测方法,如磁粉探伤、渗透探伤等,主要基于待检测产品与介质之间特定的物理或化学反应特性实现对缺陷的发现与定位.随着机器视觉的发展^[2],视觉缺陷检测技术^[3-4]由于其检测精度高、速度快、安全可靠等优势,成为当前的主流方法,已在纺织、陶瓷、半导体等行业广泛应用.

然而,视觉缺陷检测技术仍面临一些挑战.主流的视觉缺陷检测系统通常基于传统电荷耦合器件 (Charge-coupled device, CCD) 或互补金属氧化物半导体 (Complementary metal-oxide-semiconductor, CMOS) 相机进行缺陷成像和后端检测算法开发.传统相机在固定的曝光时间内进行能量积分,并以固定的帧率输出图像,在应用于工业品表面缺陷检测时存在以下局限性: 1) 由于能量积分以及模数 (Analog-to-digital, AD) 转换过程的存在,传统相机成像的速度慢,在快速生产流水线中拍摄高速目标易发生模糊^[5-7]; 2) 对前景和背景同等权重成像,造成信息冗余大、感兴趣目标不突出等问题,从而影响计算效率和目标的高效表征^[8]; 3) 由于器件基底噪声和满阱容量的限制,传统相机成像动态范围小,在强光或暗光条件下无法有效成像^[9-10].当其对于铝基盘片等高反光产品表面缺陷进行成像时,极易发生过曝或失真.高反光表面是指具有高亮、类镜面并含局部镜面反射特性的混合表面^[11].因此,当光线照射到表面时,会发生界面反射现象,在某些点产生聚焦效应^[12],从而导致过曝产生.此外,高反光表面易产生环境倒影,对缺陷检测造成极大干扰.上述问题导致现有的视觉缺陷检测系统在面对高反光表面的弱小缺陷快速检测任务时,难以有效发现与识别缺陷.

针对铝基盘片等高反光产品表面的弱小缺陷快速检测任务,本文创新性地提出了一种基于动态视觉传感器 (Dynamic vision sensor, DVS) 的缺陷检测新方法. DVS 是一种新型仿生视觉传感器^[13-15],它颠覆了传统相机的能量积分工作原理,采用能量差分的方式,仅对动态变化的光强作出响应.当 DVS 上某像素的感光强度变化超出预设阈值时,该像素会输出一个脉冲信号.本文基于 DVS 的工作原理,首先搭建了一套 DVS 缺陷检测实验平台,并以机械硬盘的主要原材料——具有高反光表面的铝基盘片为实验对象,开展了一系列的验证实验,来探究总结 DVS 的缺陷成像特性与优势.实验结果表明 DVS 能够有效应对强光导致的过曝、产品高速运动导致的模糊、高反光表面导致的环境倒影干扰等问题,由于其能敏感捕捉缺陷导致的光强变化信号,

因此具有弱小缺陷成像特征明显的优势.

随后,本文构建了第一个基于 DVS 的缺陷检测数据集 (Event-based defect detection dataset, EDD-10k),包含划痕、点痕、污渍三类缺陷,并人工标注了接近一万个标签.

最后,针对 DVS 缺陷成像的时空特性,本文提出了一种基于时序不规则特征聚合框架的 DVS 缺陷检测算法 (Temporal irregular feature aggregation framework for event-based defect detection, TIFF-EDD).其中,为了适应缺陷形态多变性、纹理稀疏性,设计了一种基于可变形卷积的多尺度不规则特征提取 (Multiscale irregular feature extraction, MIFE) 网络,实现了对不同大小、形态缺陷的自适应特征学习;为了挖掘事件流时序信息、缓解噪声干扰,提出了一种基于跨时间注意力的时序特征融合 (Cross-time attention aggregation, CT-AA) 方法,以充分利用时间上下文信息.此外,在后处理阶段提出了一种决策级时序预测融合 (Decision-level prediction fusion, DPF) 策略,来进一步降低噪声带来的虚警干扰.我们在 EDD-10k 数据集上进行了一系列的验证实验,结果表明,本文所提算法能够实现 DVS 缺陷目标的有效检测.

本文的主要贡献点总结如下:

1) 本文首次提出基于 DVS 的缺陷检测方法,并开展了相关的 DVS 缺陷成像实验,验证分析了 DVS 缺陷成像的特性与优势;

2) 构建第一个基于 DVS 的工业品表面缺陷检测数据集 EDD-10k,包含 116 段事件流,3 类常见缺陷类型,并人工手动标注了接近一万个标签;

3) 针对缺陷形态多变、纹理稀疏、噪声干扰等问题,创新性地设计了一种基于时序不规则特征聚合框架的 DVS 缺陷检测算法 (TIFF-EDD),并在 EDD-10k 数据集上进行一系列实验评估,验证了所提算法的有效性.

1 DVS 缺陷成像机理与特性分析

1.1 DVS 缺陷成像机理

DVS^[16]是一种通过模拟生物视觉系统 Where 通路对动态变化目标敏感的特性而设计出的新型仿生视觉传感器. DVS 像素阵列中的每个像素独立异步工作,输出非结构化的异步事件流^[17],因此 DVS 又被称为事件相机 (Event camera). DVS 事件信号产生过程表示如下:

$$|\Delta \ln I| = |\ln I(u, t) - \ln I(u, t + \Delta t)| \geq C \quad (1)$$

式中, $u = (x, y)$ 表示像素坐标位置, I 表示光强, t 表示上一事件的发生时刻, C 为事件产生阈值. 当光强变化超过阈值 C 时, 则产生事件信号.

DVS 对事件采用地址事件表示 (Address event representation, AER) 进行编码^[18], 每个事件 e 由一个四元组 (x_e, y_e, p_e, t_e) 组成. 其中, t_e 表示事件发生的时间戳; p_e 表示事件的极性, 反映了光强变化的趋势. 当光强变大, $p_e = +1$, 表明该事件为 ON 事件 (即正极性事件); 当光强变小, $p_e = -1$, 表明该事件为 OFF 事件 (即负极性事件). DVS 仅对光强的变化作出响应, 避免了能量积分过程和 AD 转换过程, 具有成像速度快、动态范围大等特性.

在拍摄待检产品时, 传统相机通常需要产品保持静止, 以便对缺陷区域进行成像. 不同于传统相机, 由于 DVS 仅捕捉变化的信号, 因此需要在产品运动的情况下对缺陷区域进行成像. DVS 对缺陷产生事件响应的机理如图 1 所示.

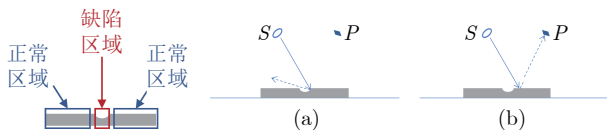


图 1 DVS 缺陷成像机理说明

Fig.1 Explanation of the mechanism of DVS-based defect imaging

图 1 中, S 表示光源, 入射光强度为 I_s , P 表示 DVS 感光阵列中的某一像素, 实线箭头表示入射光线, 虚线箭头表示反射光线. 从图 1(a) 到图 1(b), 光源和 DVS 相机均固定不动, 只有待测产品发生了移动. 假设待测产品平面为完全反射表面, 在平整的产品表面内部有一凹痕缺陷. 图 1(a) 表示 t_0 时刻入射光线照射在凹痕处的光路传播情况. 由于光源 S 的入射光线没有反射到像素 P 上, 此时像素 P 的感受光强为 $\ln(I_b)$, I_b 表示环境光强度. 图 1(b) 表示 t_1 时刻入射光线照射在产品平整表面处的光路传播情况. 由于光源 S 的入射光线完全反射到像素 P 上, 此时像素 P 的感受光强为 $\ln(I_b + I_s)$. 那么 t_0 到 t_1 时间内, 像素 P 的感受光强发生了变化, 其变化量为:

$$\Delta \ln I = \ln(I_b) - \ln(I_b + I_s) = \ln\left(\frac{I_b}{I_b + I_s}\right) \quad (2)$$

当 $|\Delta \ln I| \geq C$ 时, 则像素 P 产生了事件信号. 式 (2) 表明由于缺陷的存在, 会引起 DVS 感光强度的变化, 从而引发反映缺陷信息的脉冲事件. 此外, 图 2 展示了 DVS 某像素由于光强变化而产生的事

件数量情况^[16]. 由图 2 可知, 同样时间内光强变化越大, 产生事件越多. 因此, 当式 (2) 中入射光源 I_s 变大时, 缺陷移动时产生的光强变化变大, 产生的事件数量变多^[19]. 因此, 通过增大光源强度, 可充分利用 DVS 成像动态范围大的优势, 有效提升对缺陷的捕捉能力.

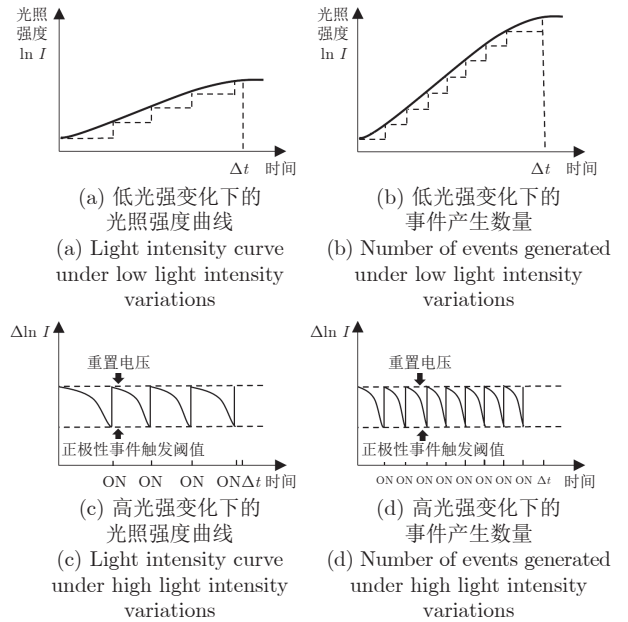


图 2 事件产生数量与光强变化大小的关系示意图

Fig.2 Relationship diagram between the number of generated events and the magnitude of light intensity variations

1.2 DVS 缺陷成像特性分析

本文以铝基盘片为检测对象, 搭建了一套面向 DVS 工作机理的缺陷检测实验平台, 并开展了一系列的成像验证实验. 铝基盘片是机械硬盘中用来存储信息的一种圆形铝基材料介质. 作为一种高精密度产品, 铝基盘片对工艺要求非常高, 缺陷判断的标准非常严格. 即使非常弱小的表面缺陷, 也会降低硬盘的存储性能和使用寿命. 例如, 合格产品要求表面划痕最大深度不能超过 1500 埃¹. 此外, 铝基盘片具有类镜面性质的高反光表面, 极易造成严重的光影干扰. 因此, 目前生产铝基盘片的主要厂商仍然采用人工检测的传统方式进行缺陷检测. 面对产线快速生产-检测需求, 实现对铝基盘片表面缺陷的高效自动化检测是一项十分困难的任务.

本文搭建的实验平台示意图如图 3 所示, 主要由成像模块、运动模块和照明模块组成. 成像模块负责对铝基盘片表面进行拍摄, 获取成像数据. 采

¹ 1 微米 = 10000 埃

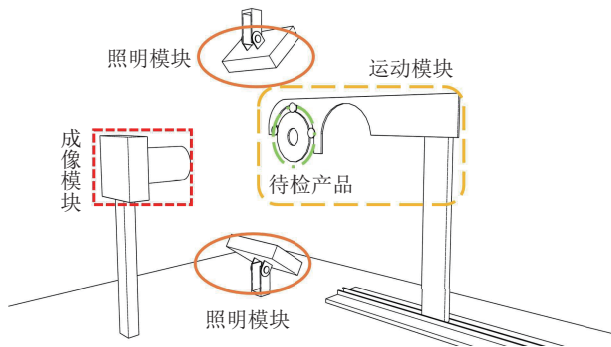


图 3 缺陷检测实验平台示意图

Fig. 3 Diagram of the defect detection experimental platform

用 CeleX-V^[14] 作为动态视觉传感器, 像素分辨率为 $1\ 280 \times 800$, 芯片大小为 $14.3\ \text{mm} \times 11.6\ \text{mm}$, 单个像元尺寸为 $9.8\ \mu\text{m} \times 9.8\ \mu\text{m}$ ^[7]. 铝基盘片与传感器的距离为 $240\ \text{mm}$ 左右, 相机的成像视场为 $215\ \text{mm} \times 175\ \text{mm}$ 左右. 运动模块负责夹持铝基盘片, 并做平行于盘片表面的直线平移运动. 照明模块则负责打光, 采用两个光强均匀的海康条形工业光源进行照明, 光源功率为 $7.8\ \text{W}$, 以确保整个盘片表面获得均匀且充分的光照, 提高缺陷区域的成像效果.

本文模拟工厂真实生产-检测环境, 对 DVS 的缺陷成像效果进行实验和分析. 为可视化事件数据, 采用第 2.2 节中式 (3) 所示的方法, 以时空堆叠的方式, 将时间窗口大小为 $10\ \text{ms}$ 的事件流转为一帧事件图像. DVS 成像特性总结如下:

1) DVS 弱小缺陷成像明显. DVS 具有敏感捕捉动态目标的优势. 由式 (2) 可知, 缺陷会引起 DVS 感光强度发生变化, 从而引发事件产生; 而对于平滑匀质的正常区域, 则不会引起 DVS 产生事件信号. 图 4 展示了对于深度约为 $2000\ \mu\text{m}$ 的划痕成像效果. DVS 可以实现对划痕的特征显著成像, 如图 4(a) 中红色矩形区域所示. 而传统相机由于对前景背景以相同的权重成像, 难以有效捕捉弱小缺陷以及缺陷区域和正常区域之间的差异, 因此对缺陷成像不明显, 如图 4(b) 中红色矩形区域所示.

2) DVS 受环境倒影干扰小. 铝基盘片的高反光表面容易产生环境背景的倒影, 如图 4(b) 中橙色圆形区域所示. 传统相机会对倒影成像, 从而给后端的检测算法带来干扰. 然而, 对于 DVS 而言, 由于在世界坐标系中, DVS 每个像素的感受野是固定的, 倒影出现的位置也是固定的. 因此 DVS 与倒影之间并没有发生相对运动. 这一特性使得 DVS 能够有效避免对倒影作出响应事件, 如图 4(a) 中橙色圆形区域所示.

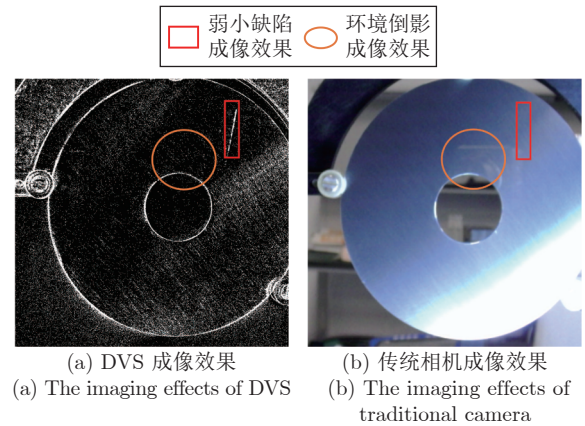


图 4 弱小缺陷、环境倒影的成像效果对比

Fig. 4 Comparison of imaging effects of small defects and environmental reflections

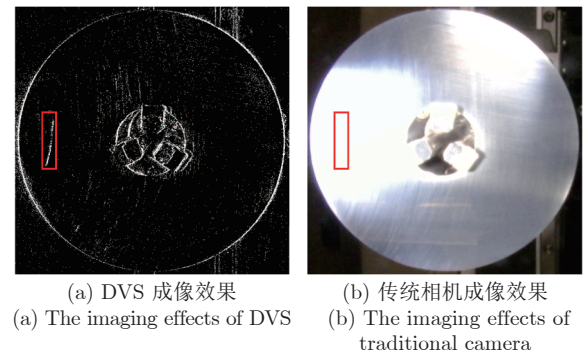


图 5 大动态范围成像效果对比

Fig. 5 Comparison of high dynamic range imaging effects

3) DVS 有效缓解强曝光影响. 在缺陷检测过程中, 往往需要额外的光源进行打光, 以提高缺陷和背景之间的对比度. 高反光表面由于强烈的光线反射, 容易导致传统相机曝光过度而无法成像. 如图 5(b) 中红色矩形区域所示, 传统相机过曝导致看不到缺陷. DVS 由于其大动态范围成像特性, 在强光下依然能够有效成像出缺陷, 如图 5(a) 中红色矩形区域所示.

4) DVS 延迟低, 响应速度快, 敏感捕捉快速运动目标. 图 6 展示了待测产品在不同运动速度下, DVS 和传统相机的缺陷成像情况. DVS 由于响应速度快, 无论铝基盘片是在低速运动还是高速运动情况下, 均能够有效捕捉缺陷信号². 然而, 传统相机由于成像速度慢, 在拍摄快速运动的缺陷目标时, 难以及时捕捉目标信号且易发生模糊, 造成缺陷难以识别, 如图 6(d) 中红色矩形区域所示.

² 由于在事件可视化时采用了同样大小的时间窗口, 快速运动的目标会产生更多的事件. 因此, 图 6(c) 比图 6(a) 具有更粗的边缘或拖影.

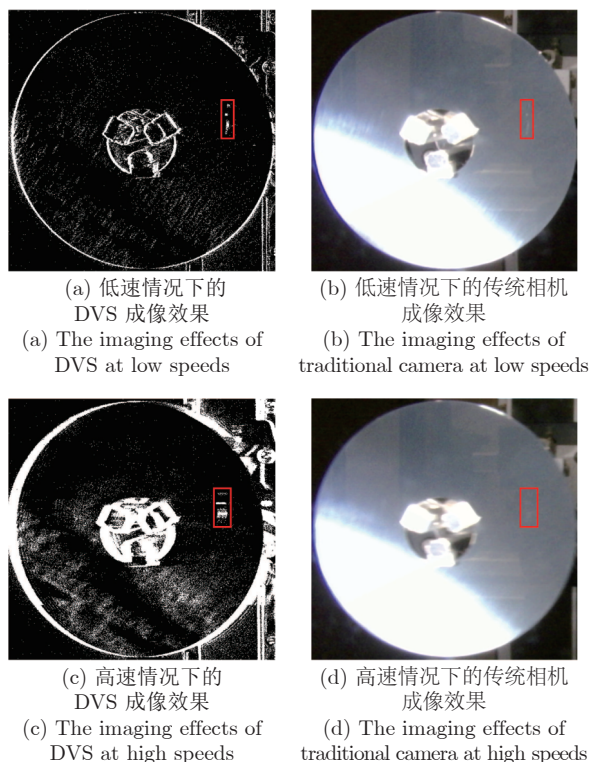


图 6 在不同运动速度下的缺陷成像效果对比

Fig.6 Comparison of defect imaging effects under different motion speeds

除了上述特性之外, DVS 的成像机理使其能够自动过滤静止不变的背景信息, 从而减少了数据冗余, 并降低了对计算资源的需求. DVS 仅对变化目标作出响应, 这种“特征成像”的特性为后端检测算法的目标特征提取提供了便利.

2 基于 DVS 的缺陷检测数据集

缺陷检测数据集是研究缺陷检测算法的基础工作^[20]. 现有主流缺陷检测数据集基于传统相机拍摄, 如 NEU-CLS^[21], KoletorSDD^[22], MVTecAD^[23] 和 MixedWM38^[24] 等. 目前尚无公开的 DVS 缺陷检测数据集. 因此, 为推动 DVS 在缺陷检测方向的应用研究, 本文构建了第一个基于 DVS 的缺陷检测数据集 EDD-10k. 接下来, 本节将会对 EDD-10k 的采集、标注过程和统计特性等进行介绍.

2.1 数据采集

本文采用第 1 节图 3 所示的缺陷检测平台作为数据采集装置, 使用 DVS 对收集的铝基盘片录制数据. 为保证缺陷真实性, 所有的产品样本均来源于工厂中实际生产的铝基盘片. 主要收集了 3 种常见的缺陷类型, 包括点痕、划痕和污渍. 此外, 为保证数据类型的丰富性, 也收集了一部分合格盘片.

每张盘片样本都经过了专业工艺人员的核查, 以确定盘片是否有缺陷或属于何种缺陷类型. 最终, 共收集了 116 张铝基盘片并对其进行了数据录制, 其中包括缺陷盘片 112 张, 合格盘片 4 张.

2.2 事件表征

DVS 的原始输出为异步非结构化的事件流^[25], 不便于人工可视化与标注. 为此, 本文采用了一种简单高效的事件表征方法将事件流转换为张量化的事件图像进行可视化^[26]. 该方法首先按照固定的时间窗口 ΔT 以无重叠的方式划分原始事件流; 随后, 将每个时间窗口内的事件进行时空堆叠得到一张二值事件图像 \bar{B} ^[27]. 在时空堆叠过程中, 按照式 (3) 对时间窗口内的每个事件 $e = (x_e, y_e, p_e, t_e)$ 进行处理, 从而实现对其 \bar{B} 的像素赋值.

$$\bar{B}(x, y) = 1, \text{ if } x = x_e \text{ and } y = y_e \quad (3)$$

即在某个时间窗口内, 只要某像素发生了事件, 则该像素值取 1; 否则, 取 0. 为了平衡事件图像可视化效果和时间分辨率之间的关系, 设置 $\Delta T = 10 \text{ ms}$. 通过上述过程, 可以将一段事件流 E 转化为一段事件图像序列 $(\dots, B_k, B_{k+1}, B_{k+2}, \dots)$.

2.3 数据标注

本文对每张事件图像采用人工标注的方法进行缺陷标注. 在标注之前, 标注人员会先观看若干段标注样例, 以建立对缺陷成像效果的主观认知. 由于某些微弱的缺陷在单张事件图像上难以识别, 比如微小的点痕、断断续续的细微划痕等, 标注人员必须以视频的方式观看事件图像序列, 通过时空关联性才能够有效识别出缺陷目标. 因此, 在标注过程中, 标注人员首先以视频的形式观看整个事件图像序列, 建立起图像之间的时空关联性; 随后, 再对每一张事件图像进行缺陷目标标注. 每一条标注包括缺陷目标的边界框信息和缺陷类别信息.

图 7 展示了不同缺陷样本的可视化结果. 每一行的事件图像来源于同一段事件流. 从图中可以看出, 在不同的缺陷类型之间, 缺陷表现出不同的形态. 比如划痕常常表现为细长条的直线或弧线; 点痕一般情况下表现为尺寸较小、形状不规则的点状区域; 污渍则表现为形状更加不规则、形态更加多样的块状区域. 在同一缺陷类型内, 不同的事件流之间, 缺陷也表现出不同的形态. 比如对于划痕缺陷, 既存在密度集中的连续划痕, 也存在断断续续的稀疏划痕, 如图 7(b) 中的第 1 行和第 2 行所示. 在同一段事件流内部, 由于在产品运动时进行拍摄,

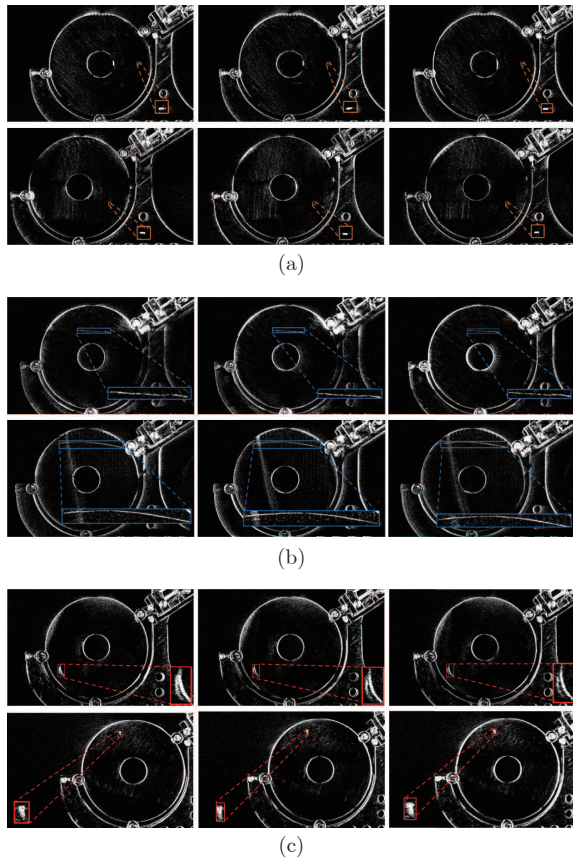


图 7 不同缺陷类别的事件图像以及标注框的可视化结果 ((a) 点痕; (b) 划痕; (c) 污渍)

Fig. 7 Visualization results of event images and annotation boxes for different defect categories ((a) Point; (b) Scratch; (c) Stain)

所以缺陷成像位置和形态也会发生变化, 如图 7(c) 第 2 行中所示的污渍缺陷成像效果. 此外, 由于在产品加工过程中存在打磨工艺, 事件图像中也会成像出打磨痕迹造成的纹路信息, 其通常表现为平行的纹理. 由于 DVS 电路的不稳定性和环境光线的影响, 事件图像中还存在大量的噪声干扰.

2.4 数据集统计特性分析

本文构建的基于 DVS 的缺陷检测数据集 EDD-10k 共包含 116 段采集自不同铝基盘片样本的事件流, 按照时间窗口 $\Delta T = 10 \text{ ms}$ 进行事件图像可视化和缺陷标注, 并人工标注了 9094 个标签. 将其中 89 段事件流作为训练集, 27 段事件流作为测试集, 并保证训练集和测试集之间没有重复的事件流. 表 1 给出了缺陷盘片的事件流和合格盘片的事件流分别在训练集、测试集中的数量分布³. 训练集中的点痕、划痕和污渍缺陷事件流数量分别占各自类别总量

的 77%、76% 和 77%. 表 2 给出了在训练集、测试集中不同类别的缺陷标签的数量. 训练集中的三类缺陷标签数量分别占各自类别总量的 82%、79% 和 76%.

表 1 事件流在训练集和测试集中的数量分布
Table 1 Quantity distribution of event streams in the training set and testing set

类别	事件流总数	训练集事件流数量	测试集事件流数量
点痕	35	27	8
划痕	39	30	9
污渍	44	34	10
合格	4	2	2

表 2 EDD-10k 数据集中每一类缺陷标签的数量
Table 2 The number of labels for each defect category in the EDD-10k dataset

类别	事件流总数	训练集事件流数量	测试集事件流数量
点痕	2865	2381	484
划痕	2159	1781	378
污渍	4070	3060	1010
合格	9094	7222	1872

为了进一步分析 DVS 缺陷成像特性, 图 8(a)、图 8(b) 和图 8(c) 分别展示了点痕、污渍、划痕缺陷目标边界框宽高比的分布. 点痕的宽高比主要集中在 0.5 到 2 之间; 污渍的宽高比主要集中在 0.5 到 3 之间; 而划痕的宽高比分布较广, 在 0 到 5 的区间内呈现不规则的分布. 图 8(d) 展示了所有缺陷目标边界框的面积分布. 图 8(d) 表明, 数据集中的缺陷目标以边界框面积小于 1000 像素的目标为主, 但也同样涵盖了面积在 1000 到 10000 像素范围内的目标. 由图 8(a) ~ 图 8(d) 可知, 缺陷目标的宽高比例差异大、尺寸分布范围广, 对缺陷检测算法的目标检测能力提出了很高的要求. 图 8(e) 展示了事件流持续时长的分布情况, 揭示了不同事件流在时间跨度上的差异. 图 8(f) 则展示了数据集中缺陷标签在训练集和测试集中的数量分布. 由图 8(f) 可知, 训练集和测试集中各类别的缺陷标签数量存在差异, 检测算法面临着类别不均衡的挑战.

3 面向 DVS 的缺陷检测算法

本文首次提出了基于 DVS 的缺陷检测新模式, 针对 DVS 缺陷成像特性, 本节进一步设计了一种基于时序不规则特征聚合框架的 DVS 缺陷检测算法 (TIFF-EDD), 实现了对形状多变、尺寸不一、干

³ 某些缺陷盘片表面可能同时具有多种缺陷, 因此表 1 中不同类别的总事件流数量相加之后大于 116.

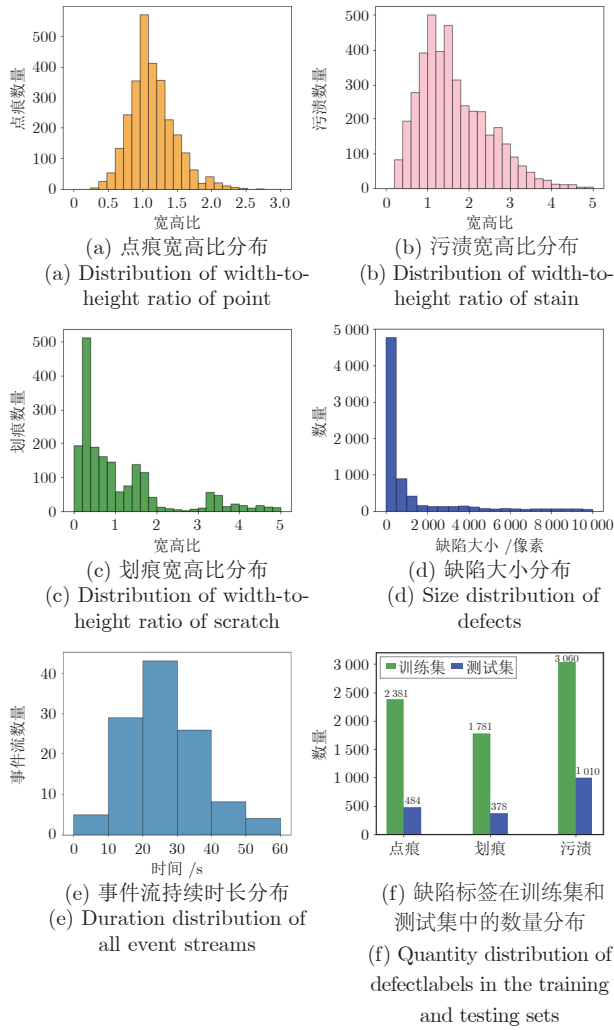


图 8 数据集 EDD-10k 统计特性

Fig.8 Statistical characteristics of the EDD-10k dataset

扰严重的弱小缺陷的准确检测。

3.1 缺陷检测定义和挑战

基于 DVS 的缺陷检测就是从 DVS 拍摄的事件流中检测缺陷的位置和类别。具体来讲,对一段事件流 $E = \{e_1, e_2, \dots, e_N\}$, 设检测器为 D , 在 T_j ($T_j > 0$) 时刻进行缺陷检测, 可表示为:

$$Box_j, Cls_j = D \left(\{e_i\}_{T_j - \Delta t_1 < t_{e_i} \leq T_j + \Delta t_2} \right) \quad (4)$$

式中, $\Delta t_1, \Delta t_2 > 0$, t_{e_i} 表示事件 e_i 的时间戳。 $Box_j = \{box_j^l = (x_j^l, y_j^l, w_j^l, h_j^l)\}_{l \geq 0}$ 表示预测目标框的位置集合, l 表示第 l 个预测目标框, x_j^l, y_j^l 表示预测目标框中心位置的坐标, w_j^l, h_j^l 表示预测目标框的宽和高。 Cls_j 表示预测目标的类别集合。

式 (4) 表示检测器 D 利用 $(T_j - \Delta t_1, T_j + \Delta t_2]$ 时间邻域内的事件完成 T_j 时刻的缺陷检测。根据

第 1 节和第 2 节对于 DVS 成像机理、缺陷成像特性、缺陷数据特性的分析可知, 开发 DVS 缺陷检测算法面临以下几个挑战:

- 1) DVS 事件数据的异步性和非结构性使得传统的图像处理技术无法直接应用于 DVS 缺陷检测;
- 2) 缺陷往往表现出形态多变、尺度不一、纹理稀疏、分布随机等特性;
- 3) 由于产品表面反光、材质变化、环境光影以及噪声事件等因素的影响, 事件流中包含大量噪声干扰。

3.2 方法概述

针对上述挑战, 本文提出了首个面向 DVS 的缺陷检测算法 TIFF-EDD, 其包含检测器部分和后处理部分。图 9 展示了检测器部分的主要网络框架, 图 10 展示了后处理部分的工作流程。

检测器部分主要负责时序不规则特征的提取与聚合, 并初步预测缺陷目标, 包含如下几个模块: 输入事件表征模块, 多尺度不规则特征提取模块, 基于跨时间注意力的时序特征融合模块以及预测头模块。后处理部分在检测器初始预测结果的基础上, 设计了一种决策级时序预测融合策略, 进一步优化预测结果。

输入事件表征模块采用与第 2 节中同样的事件表征方法, 按照固定时间窗口 $\Delta T = 10$ ms 将一段事件流 E 转化为事件图像序列。同时, 为了更充分地利用事件流中的时间上下文信息, 选择连续三帧事件图像, 简记为 $[B_1, B_2, B_3]$, 作为输入样本进行后续处理, 完成对中间帧 B_2 的缺陷目标检测。在预测头模块中, 采用了将目标位置定位和类别预测分开处理的解耦头 (Decouple head) 设计^[28] 和无锚框 (Anchor-free) 的目标框生成策略^[29]。其余的模块将在下文中进行详细介绍。

3.3 多尺度不规则特征提取模块

可变形卷积^[30] 在普通卷积的基础上, 添加了一个能够被学习到的偏移量 Δp_n , 允许采样点扩散成不规则形状, 其过程如下式所示:

$$y(p_0) = \sum_{p_n \in \mathbf{R}} w(p_n) x(p_0 + p_n + \Delta p_n) \quad (5)$$

可变形卷积拥有更灵活的感受野, 能够自适应处理各种不规则的缺陷目标。

受此启发, 本文提出了一个基于可变形卷积的多尺度不规则特征提取模块 MIFE。如图 9 所示, 该模块包含了三个权重共享的特征提取器, 分别接

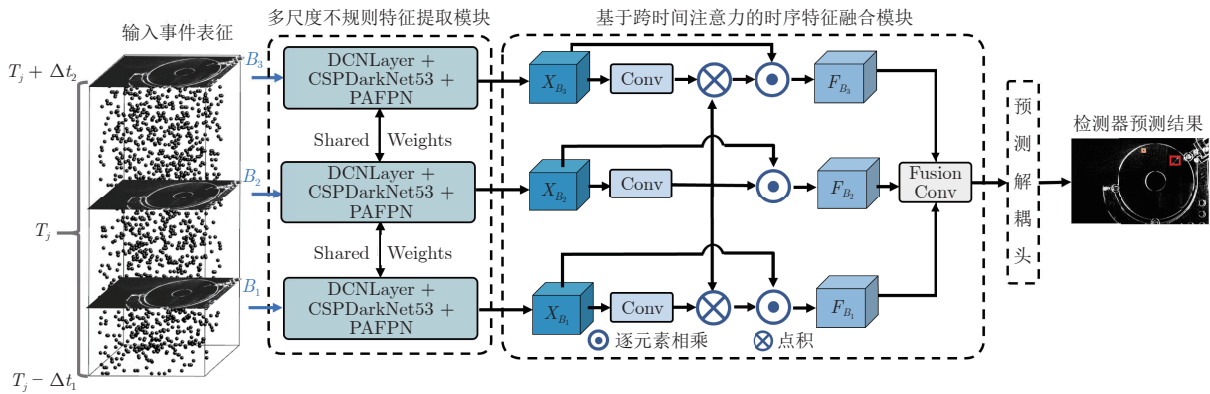


图 9 缺陷检测器的主要网络框架

Fig. 9 Main network framework of the defect detector

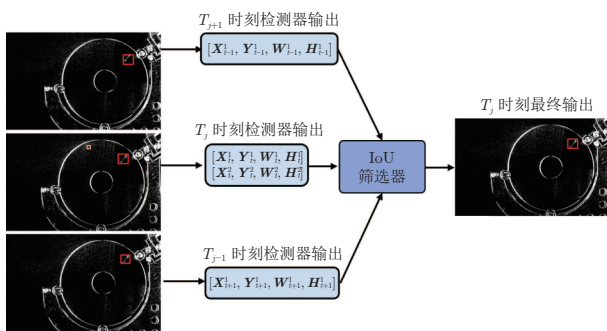


图 10 基于决策级时序预测融合的后处理模块工作流程图

Fig. 10 The workflow diagram of the post-processing module based on decision-level temporal prediction fusion

受单帧事件图像 (即 B_1 , B_2 或 B_3) 作为输入, 并输出相应的特征图 (即 X_{B_1} , X_{B_2} 或 X_{B_3}).

其中, 每个特征提取器主要由可变形卷积层 DCNLayer, CSPDarkNet53^[31] 与 PAFPN^[32] 三部分组成. DCNLayer 由 1×1 的普通卷积层、 3×3 的可变形卷积层以及 SiLU 激活层组成, 设置在网络的最前端, 以加强网络对形状不规则、纹理稀疏目标的特征提取能力. CSPDarkNet53 在 DarkNet53 基础上引入了 CSPNet (Cross stage partial network) 模块^[33], 利用残差连接^[34] 和特征融合机制提高了网络的信息传递和特征提取能力. 路径聚合特征金字塔网络 PAFPN 在特征金字塔网络 FPN^[35] 基础上增加自底向上的路径, 用浅层神经网络中的定位信息提升特征金字塔的性能, 通过聚合不同尺度的特征图以提升对多尺度缺陷特征的检测能力.

MIFE 通过可变形卷积和多尺度特征融合策略, 可以实现对不同形态、不同大小缺陷特征的有效提取, 从而提高模型精度和鲁棒性.

3.4 时序特征融合模块

事件流是在时域上的连续信号. 挖掘和利用事件流中的时序信息^[36-37], 对于 DVS 缺陷准确检测至关重要. 一方面, 在 DVS 缺陷检测过程中, 待检产品需要移动进行缺陷成像, 造成了光照、缺陷和 DVS 三者之间的角度会不断发生变化, 从而导致成像出的缺陷形态不断发生变化. 如图 7 所示, 对于同一段事件流, 在不同的事件图像中, 缺陷的形状、尺寸、位置、显著程度都会发生变化. 在一些情况下, 由于有用信息量不足, 仅靠单帧的事件图像难以判定是否存在缺陷^[38]. 通过建立相邻时间内事件图像之间的相互关联, 则可以有效地发现和判断是否存在缺陷. 另一方面, 由于电路不稳定性、光影干扰等影响, DVS 缺陷事件流中也会存在大量在时域上随机分布的噪声事件. 通过挖掘事件之间的时序关联性, 可以缓解噪声事件对缺陷检测的影响.

针对上述问题, 本文提出了一种基于跨时间注意力的时序特征融合 (CTAA) 模块, 旨在学习时间邻域内事件之间的相互依赖关系, 实现对噪声的抑制和有效信息的融合增强.

如图 9 所示, CTAA 主要通过计算相邻事件图像特征之间的相似性以融合时序特征. 其计算过程为:

$$S_{B_1} = \text{Conv}(X_{B_1}) \otimes \text{Conv}(X_{B_2}) \quad (6)$$

$$S_{B_2} = \text{Conv}(X_{B_2}) \otimes \text{Conv}(X_{B_2}) \quad (7)$$

$$S_{B_3} = \text{Conv}(X_{B_3}) \otimes \text{Conv}(X_{B_2}) \quad (8)$$

式中, S_{B_1} , S_{B_2} , S_{B_3} 分别代表连续输入的三帧图像与中间帧的相似度矩阵, \otimes 表示点积. 利用该相似度矩阵可进一步得到输入的三帧图像与中间帧的关联特征 F_{B_1} , F_{B_2} , F_{B_3} :

$$\begin{cases} F_{B_1} = X_{B_1} \odot S_{B_1} \\ F_{B_2} = X_{B_2} \odot S_{B_2} \\ F_{B_3} = X_{B_3} \odot S_{B_3} \end{cases} \quad (9)$$

式中, \odot 表示逐元素相乘. 接下来, 利用卷积操作融合关联特征, 得到增强后的特征图 $F_{B_2}^*$:

$$F_{B_2}^* = \text{Conv}(\text{Cat}(F_{B_1}, F_{B_2}, F_{B_3})) \quad (10)$$

最后, 将 $F_{B_2}^*$ 输入到预测头部分进行缺陷目标的预测.

CTAA 通过将相邻帧之间的特征相似性作为不同时刻的注意力权重, 实现时序特征的自适应融合, 充分利用事件图像序列中的时间上下文信息.

3.5 网络优化

本文以 30 ms 事件流累计成的连续 3 帧事件图像作为输入, 预测中间帧的目标检测结果. 因此, 在检测器网络训练过程中, 以中间帧的真值标签作为监督信号. 借鉴 YOLOX 框架^[28], 损失函数设计如下:

$$L_{all} = L_{obj} + L_{cls} + L_{reg} \quad (11)$$

式中, L_{cls} 表示用于类别预测的分类损失, L_{reg} 表示用于目标框位置预测的回归损失, L_{obj} 表示用于判断前景或背景的置信度损失. 本文采用交叉熵损失函数和 IoU (Intersection-over-union) 损失函数来构建分类损失 L_{cls} 和回归损失 L_{reg} , 如式 (12) 和式 (13) 所示:

$$L_{cls} = - \sum_{i=1}^n \bar{p}_i \log(p_i) \quad (12)$$

$$L_{reg} = 1 - \text{IoU} \quad (13)$$

式中, n 为正样本总数, \bar{p}_i 代表目标真实标签, p_i 代表模型预测类别概率, IoU 代表模型预测框与目标真实边界框的交并比.

在 DVS 缺陷数据集中, 正负样本极度不均衡, 其中正样本 (缺陷) 在整张事件图像中所占比例极小. 因此, 本文参考 Varifocal Loss^[39] 来构建置信度损失函数:

$$L_{obj} = \begin{cases} -\alpha p^\gamma \log(1-p), & \text{if } q = 0 \\ -q(q \log(p) + (1-q) \log(1-p)), & \text{if } q > 0 \end{cases} \quad (14)$$

其中, α 是防止过度抑制的超参数. 对于负样本, 通过将损失函数乘以一个尺度因子 p^γ , 降低易分类样本的权重, 减少它们对总损失的贡献, 从而缓解易分类负样本对模型训练的干扰, 提升训练稳定性.

对于正样本, 鉴于其相对稀缺, 因此使用 q 对其进行加权, 使模型在训练过程中更关注难以分类的高质量的正样本. 式 (14) 可以有效解决 DVS 缺陷数据集集中正样本稀缺问题, 提高模型的缺陷检测能力.

3.6 后处理部分

为充分挖掘时序信息, 同时减少由噪声造成的虚警, 本文进一步提出了一种基于决策级时序预测融合 (DPF) 的后处理策略. TIFF-EDD 方法的后处理部分主要负责对检测器部分预测的初始检测结果进行进一步优化, 如图 10 所示, 红色预测框预测目标为污渍, 橙色预测框预测目标为点痕. 由于橙色预测框与前后时刻同类别的预测框的交并比低, 因此被过滤掉. 红色预测框与前后时刻同类别的预测框的交并比高, 因此被保留.

根据式 (4), 假设 Box_{j-1} , Box_j , Box_{j+1} 分别为三个连续时刻 T_{j-1} , T_j , T_{j+1} 对应的检测器输出的初始预测目标框. 如果在 T_j 时刻检测到缺陷目标, 那么在其相邻时刻 T_{j-1} 、 T_{j+1} , 检测到该缺陷的概率依然很大. 基于这一思想, DPF 策略针对 Box_j 中的每一个预测框 box_j^l 分别进行如下操作: 1) 计算 box_j^l 与 Box_{j-1} , Box_{j+1} 中类别相同的每个预测目标框的交并比, 并记取得最大值的交并比为 IoU_{\max} ; 2) 如果 IoU_{\max} 小于预设阈值 IoU_{th} , 则删除该预测框. 本文设置 $\text{IoU}_{\text{th}} = 0.15$. 实验结果与在实践中的应用表明, 决策级融合策略可以进一步优化本文提出的 TIFF-EDD 算法的预测结果, 降低虚警率, 提高缺陷检测的精确度.

4 实验结果分析

在第 2 节构建的 DVS 缺陷数据集的基础上, 对本文所提的 DVS 缺陷检测算法 TIFF-EDD 进行了一系列实验验证. 具体实验分析如下.

4.1 实验设置

TIFF-EDD 网络模型采用 PyTorch1.12 深度学习框架、CUDA11.3 和 Ubuntu20.04 系统进行训练, 采用 SGD 优化器, 初始学习率设置为 0.01, 并使用余弦学习率衰减调整学习率, 最低学习率设置为 0.0001, 批次大小设置为 4. 总共的训练轮次为 100. 所有的训练和评估实验均在两张 RTX3090 显卡上进行.

由于铝基盘片表面存在的缺陷尺寸较小, 且实际工艺产线中往往只需要实现对缺陷大致定位即可, 因此本文使用了 mAP@0.4 (在 IoU 阈值为 0.4

时, 模型的平均精度) 作为评估指标.

4.2 与其他算法的对比实验

表 3 列出了本文所提算法与其他目标检测算法的实验结果. 用于对比的目标检测算法包含三类面向不同模态数据的算法, 分别为面向图像的目标检测算法、面向视频的目标检测算法和面向事件流的目标检测算法. 本文所对比的所有方法均在 EDD-10k 数据集上重新训练. 除 Single shot detector-event (SSD-event)^[40] 之外, 都采用了作者提供的源代码. 对于面向图像的目标检测算法与面向视频的目标检测算法, 采用第 2.2 节中式 (3) 转换得到的事件图像与事件图像序列作为模型输入. 对于面向事件流的目标检测算法, 采用事件流作为模型输入.

图像目标检测算法: Faster R-CNN^[41], YOLOv5, YOLOv7^[42] 均是图像目标检测领域里程碑式的工作. 其中, 前者是双阶段目标检测器, 后两者则是单阶段目标检测器. 它们在基于传统图像的各种目标检测任务场景中得到了广泛应用^[43-46], 并取得了优异的效果. 在 DVS 缺陷检测任务中, 由于缺陷弱小且形状不规则, Faster R-CNN 仅取得了 0.210 的 mAP@0.4. YOLOv5 和 YOLOv7 分别达到了 0.569 和 0.543 的 mAP@0.4. 本文提出的 TIFF-EDD 取得 0.617 的 mAP@0.4, 相比于 YOLOv5 和 YOLOv7 分别提升 8.43%, 13.62%.

表 3 与其他算法在 EDD-10k 数据集上的对比实验

Table 3 Comparison experiments with other algorithms on the EDD-10k dataset

方法	mAP@0.4	AP@点痕	AP@划痕	AP@污渍
Faster R-CNN ^[41]	0.210	0.000	0.536	0.095
YOLOv5	0.569	0.393	0.756	0.559
YOLOv7 ^[42]	0.543	0.471	0.644	0.514
RDN ^[47]	0.512	0.553	0.476	0.507
MEGA ^[48]	0.401	0.356	0.509	0.349
YOLOV ^[49]	0.537	0.112	0.628	0.670
SSD-event ^[40]	0.236	0.087	0.626	0.138
SODformer-event ^[25]	0.394	0.363	0.161	0.495
TIFF-EDD	0.617	0.512	0.701	0.639

视频目标检测算法: 本文选取了三种 SOTA 算法, 分别为 RDN^[47]、MEGA^[48] 和 YOLOV^[49], 它们各自取得 0.512、0.401 和 0.537 的 mAP@0.4. 相较于这三类方法, 本文所提算法分别取得 20.51%、53.86% 以及 14.90% 的 mAP@0.4 提升. 在单类缺陷的 AP 指标方面, TIFF-EDD 在点痕、划痕和污

渍类型上均取得前二的性能, AP 得分分别为 0.512、0.701 和 0.639.

事件流目标检测算法: 鉴于事件相机的优良特性, 近年来一些面向事件流的目标检测方法也被提出. SSD-event^[40] 便是其中的代表工作之一. SSD-event 首先将事件流转为 2D 结构化张量, 进而采用 SSD 算法^[50] 进行目标检测. 然而由于 SSD-event 中缺乏对浅层特征的保留^[40], 导致小目标检测能力弱, 因此仅取得 0.236 的 mAP@0.4. SODformer-event^[25] 是近两年来事件流目标检测的代表性工作, 该方法基于 Transformer 实现了对事件流中连续移动车辆的有效检测. 由于基于 Transformer 的目标检测方法通常需要大规模的数据集进行训练, 因此在 EDD-10k 测试集上的效果不佳, 仅取得 0.394 的 mAP@0.4. 本文所提的 TIFF-EDD 方法利用可变形卷积增强对不规则目标的提取能力, 采用 PAFPN 提升了对多尺度信息的表征能力, 同时基于 CTAA 充分挖掘了事件图像序列中的时序信息, 故取得了更好的性能.

4.3 消融实验

本节开展了一系列的消融实验, 以验证本文所提方法在 DVS 缺陷检测任务上的有效性. 表 4 列出了不同配置下的消融实验结果. TIFF-B 是由 CSPDarkNet53, PAFPN 和 DeCoupled Head 组成的基线模型, 采用单帧单通道的事件图像作为输入, 并基于式 (11) 采用 BCELoss 构建置信度损失函数, 最终取得 0.577 的 mAP@0.4. TIFF-MIFE 将 TIFF-B 模型中的 CSPDarkNet53 与 PAFPN 配置替换为本文提出的多尺度不规则特征提取模块 (MIFE). 相比于基线模型, TIFF-MIFE 取得 0.590 的 mAP@0.4, 性能提升 2.3%, 说明了 MIFE 模块在处理形态多变缺陷目标时的有效性. 进一步地, TIFF-CTAA 在 TIFF-MIFE 的基础上加入了基于跨时间注意力的时序特征融合模块, 取得 2.9% 的性能提升. 这一结果验证了本文所提 CTAA 方法能够有效实现对事件图像序列中时序特征的整合.

表 4 在 EDD-10k 数据集上的消融实验结果

Table 4 Ablation experimental results on the EDD-10k dataset

	基线	MIFE	CTAA	VFLoss	DPF	mAP@0.4
TIFF-B	✓					0.577
TIFF-MIFE	✓	✓				0.590
TIFF-CTAA	✓	✓	✓			0.607
TIFF-VF	✓	✓	✓	✓		0.612
TIFF-EDD	✓	✓	✓	✓	✓	0.617

TIFF-VF 模型在采用 MIFE 和 CTAA 模块的同时, 通过使用 VF Loss 对模型进行优化, 缓解正负样本分布不均衡问题, 并取得了 0.612 的 mAP@0.4. 最后, 通过在后处理阶段引入决策级时序预测融合策略, 得到本文所提算法 TIFF-EDD, 最终取得 0.617 的 mAP@0.4.

一系列的消融实验结果表明, 本文所提出的 DVS 缺陷目标检测方法能够有效应对缺陷形态多变、纹理稀疏、噪声干扰、样本不均衡等挑战, 实现对缺陷特征的鲁棒表征和时序信息的高效整合, 取得对缺陷目标的准确检测.

5 结束语

本文针对传统视觉检测技术在高反光产品表面缺陷检测能力不足的问题, 提出了一种基于动态视觉传感器的缺陷检测新模式. 首先分析了 DVS 的缺陷成像机理, 并以具有高反光表面的铝基盘片为实验对象, 验证总结了 DVS 缺陷成像的特性与优势. 接下来, 构建了首个基于 DVS 的缺陷检测数据集 EDD-10k, 并手动标注了接近一万个标签. 最后, 提出了一个基于时序不规则特征聚合框架的 DVS 缺陷检测算法 TIFF-EDD. 在 EDD-10k 数据集上的一系列实验结果表明, TIFF-EDD 能够有效应对缺陷形态多变、纹理稀疏、噪声干扰等挑战, 实现对缺陷目标的高效检测. 本文所提出的基于 DVS 的缺陷检测新模式也可推广应用于其他高反光产品表面缺陷检测任务, 为推动缺陷检测技术的发展提供了一种新方案.

References

- Jin Xia-Ting, Wang Yao-Nan, Zhang Hui, Liu Li, Zhong Hang, He Zhen-Dong. DeepRail: Automatic visual detection system for railway surface defect using Bayesian CNN and attention network. *Acta Automatica Sinica*, 2019, **45**(12): 2312–2327 (金侠挺, 王耀南, 张辉, 刘理, 钟杭, 贺振东. 基于贝叶斯 CNN 和注意力网络的钢轨表面缺陷检测系统. *自动化学报*, 2019, **45**(12): 2312–2327)
- Tao Xian, Hou Wei, Xu De. A survey of surface defect detection methods based on deep learning. *Acta Automatica Sinica*, 2021, **47**(5): 1017–1034 (陶显, 侯伟, 徐德. 基于深度学习的表面缺陷检测方法综述. *自动化学报*, 2021, **47**(5): 1017–1034)
- Chen Y Q, Pan J W, Lei J Y, Zeng D Y, Wu Z Z, Chen C S. EEE-Net: Efficient edge enhanced network for surface defect detection of glass. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 5029013
- Jiang W B, Liu M, Peng Y N, Wu L H, Wang Y N. HDCB-Net: A neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges. *IEEE Transactions on Industrial Informatics*, 2021, **17**(8): 5485–5494
- Zheng Y J, Zheng L X, Yu Z F, Shi B X, Tian Y H, Huang T J. High-speed image reconstruction through short-term plasticity for spiking cameras. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE, 2021. 6354–6363
- Xu X Y, Sun D Q, Pan J S, Zhang Y J, Pfister H, Yang M H. Learning to super-resolve blurry face and text images. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 251–260
- Li Jia-Ning, Tian Yong-Hong. Recent advances in neuromorphic vision sensors: A survey. *Chinese Journal of Computers*, 2021, **44**(6): 1258–1286 (李家宁, 田永鸿. 神经形态视觉传感器的研究进展及应用综述. *计算机学报*, 2021, **44**(6): 1258–1286)
- Pan L Y, Hartley R, Scheerlinck C, Liu M M, Yu X, Dai Y C. High frame rate video reconstruction based on an event camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(5): 2519–2533
- Han J, Yang Y X, Duan P Q, Zhou C, Ma L, Xu C, et al. Hybrid high dynamic range imaging fusing neuromorphic and conventional images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(7): 8553–8565
- Hu Y H, Liu S C, Delbruck T. V2e: From video frames to realistic DVS events. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville, USA: IEEE, 2021. 1312–1321
- Feng Wei, Xu Shi-Nan, Wang Heng-Hui, Xiong Zhi, Wang Xuan-Ze, Zhai Zhong-Sheng. Three-dimensional measurement method of highly reflective surface based on per-pixel modulation. *Chinese Optics*, 2022, **15**(3): 488–497 (冯维, 徐仕楠, 王恒辉, 熊芝, 王选择, 翟中生. 逐像素调制的高反光表面三维测量方法. *中国光学*, 2022, **15**(3): 488–497)
- Wang Ying, Ni Yu-Bo, Meng Zhao-Zong, Gao Nan, Guo Tong, Yang Ze-Qing, et al. Adaptive coding fringe projection profilometry on color reflective surfaces. *Acta Optica Sinica*, 2024, **44**(7): Article No. 0712001 (王颖, 倪育博, 孟召宗, 高楠, 郭彤, 杨泽青, 等. 彩色高反光表面自适应编码条纹投影轮廓术. *光学学报*, 2024, **44**(7): Article No. 0712001)
- Lichtsteiner P, Posch C, Delbruck T. A 128 × 128 120db 30mw asynchronous vision sensor that responds to relative intensity change. In: Proceedings of the IEEE International Solid State Circuits Conference-Digest of Technical Papers. San Francisco, USA: IEEE, 2006. 2060–2069
- Chen S S, Guo M H. Live demonstration: CeleX-V: A 1M pixel multi-mode event-based sensor. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach, USA: IEEE, 2019. 1682–1683
- Brandli C, Muller L, Delbruck T. Real-time, high-speed video decompression using a frame- and event-based DAVIS sensor. In: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS). Melbourne, Australia: IEEE, 2014. 686–689
- Gallego G, Delbruck T, Orchard G, Bartolozzi C, Taba B, Censi A, et al. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(1): 154–180
- Lagorce X, Orchard G, Galluppi F, Shi B E, Benosman R B. HOTS: A hierarchy of event-based time-surfaces for pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(7): 1346–1359
- Lazzaro J, Wawrzynek J, Mahowald M, Sivilotti M, Gillespie D. Silicon auditory processors as computer peripherals. In: Proceedings of the 5th International Conference on Neural Information

- Processing Systems. Denver, USA: ACM, 1992. 820–827
- 19 Ma Yan-Yang, Ye Zi-Hao, Liu Kun-Hua, Chen Long. Event-based visual localization and mapping algorithms: A survey. *Acta Automatica Sinica*, 2021, **47**(7): 1484–1494 (马艳阳, 叶梓豪, 刘坤华, 陈龙. 基于事件相机的定位与建图算法: 综述. *自动化学报*, 2021, **47**(7): 1484–1494)
- 20 Li Y Z, Wang H L, Yuan S H, Liu M, Zhao D B, Guo Y W, et al. Myriad: Large multimodal model by applying vision experts for industrial anomaly detection. arXiv preprint arXiv: 2310.19070, 2023.
- 21 Song K C, Yan Y H. Micro surface defect detection method for silicon steel strip based on saliency convex active contour model. *Mathematical Problems in Engineering*, 2013, **2013**: Article No. 429094
- 22 Tabernik D, Šela S, Skvarč J, Skočaj D. Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing*, 2020, **31**(3): 759–776
- 23 Bergmann P, Fauser M, Sattlegger D, Steger C. MVTEC AD—A comprehensive real-world dataset for unsupervised anomaly detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 9584–9592
- 24 Wang J L, Xu C Q, Yang Z L, Zhang J, Li X O. Deformable convolutional networks for efficient mixed-type wafer defect pattern recognition. *IEEE Transactions on Semiconductor Manufacturing*, 2020, **33**(4): 587–596
- 25 Li D Z, Tian Y H, Li J N. SODFormer: Streaming object detection with transformer using events and frames. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(11): 14020–14037
- 26 Nguyen A, Do T T, Caldwell D G. Real-time 6DOF pose relocalization for event cameras with stacked spatial LSTM networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach, USA: IEEE, 2019. 1638–1645
- 27 Kim J, Bae J, Park G. N-ImageNet: Towards robust, fine-grained object recognition with event cameras. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, USA: IEEE, 2021. 2126–2136
- 28 Ge Z, Liu S T, Wang F, Li Z M, Sun J. YOLOX: Exceeding YOLO series in 2021. arXiv preprint arXiv: 2107.08430, 2021.
- 29 Duan K W, Bai S, Xie L X, Qi H G, Huang Q M, Tian Q. CenterNet: Keypoint triplets for object detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019. 6568–6577
- 30 Xiong Y, Li Z, Chen Y, Wang F, Zhu X, Luo J, et al. Efficient deformable ConvNets: Rethinking dynamic and sparse operator for vision applications. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE, 2024. 5652–5661
- 31 Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv: 2004.10934, 2020.
- 32 Liu S, Qi L, Qin H F, Shi J P, Jia J Y. Path aggregation network for instance segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 8759–8768
- 33 Wang C Y, Liao H Y M, Wu Y H, Chen P Y, Hsieh J W, Yeh I H. CSPNet: A new backbone that can enhance learning capability of CNN. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, USA: IEEE, 2020. 1571–1580
- 34 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770–778
- 35 Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 936–944
- 36 Liu B D, Xu C, Yang W, Yu H, Yu L. Motion robust high-speed light-weighted object detection with event camera. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 5013113
- 37 Li J N, Li J, Zhu L, Xiang X J, Huang T J, Tian Y H. Asynchronous spatio-temporal memory network for continuous event-based object detection. *IEEE Transactions on Image Processing*, 2022, **31**: 2975–2987
- 38 Wang X T, Chan K C K, Yu K, Dong C, Loy C C. EDVR: Video restoration with enhanced deformable convolutional networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach, USA: IEEE, 2019. 1954–1963
- 39 Zhang H Y, Wang Y, Dayoub F, Sünderhauf N. VarifocalNet: An IoU-aware dense object detector. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE, 2021. 8510–8519
- 40 Iacono M, Weber S, Glover A, Bartolozzi C. Towards event-driven object detection with off-the-shelf deep learning. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 1–9
- 41 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: ACM, 2015. 91–99
- 42 Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE, 2023. 7464–7475
- 43 Zhu X M, Wang S, Su J J, Liu F, Zeng L. High-speed and accurate cascade detection method for chip surface defects. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 2506612
- 44 Yuan M H, Zhou Y B, Ren X Y, Zhi H, Zhang J, Chen H J. YOLO-HMC: An improved method for PCB surface defect detection. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 2001611
- 45 Wang Y R, Song X K, Feng L L. MCI-GLA plug-in suitable for YOLO series models for transmission line insulator defect detection. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 9002912
- 46 Zhu J, Pang Q W, Li S S. ADDet: An efficient multiscale perceptual enhancement network for aluminum defect detection. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 5004714
- 47 Deng J J, Pan Y W, Yao T, Zhou W G, Li H Q, Mei T. Relation distillation networks for video object detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE, 2019. 7022–7031
- 48 Chen Y H, Cao Y, Hu H, Wang L W. Memory enhanced global-local aggregation for video object detection. In: Proceedings of

the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 10334–10343

- 49 Shi Y H, Wang N Y, Guo X J. YOLOV: Making still image object detectors great at video object detection. In: Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington, USA: AAAI, 2023. 2254–2262
- 50 Jeong J, Park H, Kwak N. Enhancement of SSD by concatenating feature maps for object detection. arXiv preprint arXiv: 1705.09587, 2017.



马居坡 西安电子科技大学人工智能学院讲师. 主要研究方向为仿生动态视觉处理和视觉缺陷检测.

E-mail: majupo@xidian.edu.cn

(MA Ju-Po Lecturer at the School of Artificial Intelligence, Xidian University. His research interest

covers bionic dynamic vision processing and visual defect detection.)



陈周熠 西安电子科技大学人工智能学院博士研究生. 主要研究方向为仿生动态视觉处理和视觉缺陷检测.

E-mail: chenzhouyi@stu.xidian.edu.cn

(CHEN Zhou-Yi Ph.D. candidate at the School of Artificial Intelligence, Xidian University. His re-

search interest covers bionic dynamic vision processing and visual defect detection.)



吴金建 西安电子科技大学人工智能学院教授. 主要研究方向为高质量成像和图像智能处理. 本文通信作者.

E-mail: jinjian.wu@mail.xidian.edu.cn

(WU Jin-Jian Professor at the School of Artificial Intelligence, Xidian University. His research in-

terest covers high-quality imaging and intelligent image processing. Corresponding author of this paper.)