中图法分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1006-8961(XXXX)XX-0001-12

论文引用格式:Zhang Runhua, Hao Xiaohui, Li Bo, She Rui. Complex-scene high-dynamic infrared adaptive processing and edge hardware acceleration[J/OL]. Journal of Image and Graphics, XXXX:1-12. DOI: 10.11834/jig.260471.(张润华, 郝晓辉, 李波, 厍睿. 复杂场景高动态红外图像自适应处理及端侧硬件加速[J/OL]. 中国图象图形学报,XXXX:1-12. DOI: 10.11834/jig.260471.)[DOI:10.11834/jig.260471]

复杂场景高动态红外图像自适应处理及端侧硬件加速

张润华[1],郝晓辉[2],李波[1],厍睿[1*]

1. 北京航空航天大学人工智能学院,北京市 100191; 2. 北京航空航天大学计算机学院,北京市 100191

摘　要:**目的**　面向低空无人机巡检、夜间搜救和应急监测中的机载红外感知链路,研究如何将高动态红外图像转换为固定八位输入时保留当前场景中的有效灰度层次,并设计适合数字信号处理器多核并行运行的处理方式。**方法**　将高动态红外图像转换为八位图像时,先依据当前帧实际出现的灰度内容确定需要保留的层次,再结合像素数量和灰度级占用安排输出灰度空间,通过连续曲线压缩高热区域所在的亮端,保持主体和阴影过渡的连续层次,并保留低对比目标已有的灰度差异。映射曲线形成后保存为当前帧查找表,像素转换通过查表完成。在端侧实现中,将查找表作为多个核心共同读取的只读数据,在德州仪器TMS320C6678八核数字信号处理器上按图像区域并行完成转换。**结果**　构造高热、坏点、阴影和联合干扰四类复杂场景,采用峰值信噪比、结构相似性、色调映射图像质量和目标有效灰度覆盖率评价处理结果,并与三种动态范围压缩方法比较。四种目标检测框架下,该方法在交并比阈值0.5～0.95上的平均精度均值(mean average precision,mAP@[0.5:0.95])相对于各框架中最佳对比方法提高2.9～3.7个百分点。TMS320C6678端侧实验在640×512分辨率下进行,单帧总时间由原始算法的78.33 ms降至端侧硬件加速算法的6.10 ms,帧率由12.77帧/s提高至163.93帧/s,加速比达到12.84。**结论**　本文所提出方法按照当前场景实际出现的灰度内容重新分配固定八位输出中的灰度层次,并把每帧生成的映射关系保存为查找表供多个核心复用,在保持红外图像有效层次的同时,为机载视觉处理提供了适合多核端侧实时运行的转换方式。

关键词:复杂场景;高动态红外图像;自适应处理;曲线映射;端侧硬件加速

Complex-scene high-dynamic infrared adaptive processing and edge hardware acceleration

Zhang Runhua[1], Hao Xiaohui[2], Li Bo[1], She Rui[1*]

1. School of Artificial Intelligence (Institute of Artificial Intelligence), Beihang University, Beijing 100191, China; 2. School of Computer Science and Engineering, Beihang University, Beijing 100191, China

Abstract: **Objective** Airborne infrared vision is an important sensing component for unmanned aerial vehicle inspection, night-time search and rescue, emergency monitoring, and other low-altitude applications. The image acquired by an airborne infrared sensor usually contains a wider and denser intensity range than the fixed eight-bit input interface used by subsequent airborne vision hardware. In complex scenes, large-area backgrounds, distant small targets, low-contrast targets,

收稿日期:2026-07-31;**修回日期**:2026-10-23

*通信作者:厍睿,男,准聘教授,主要研究方向为智能鲁棒感知、智能信息处理及在具身智能、多模态大模型及智能体、网联车路协同中应用。E-mail:ruishe@buaa.edu.cn

基金项目:北京市科技计划项目(批准号:Z241100004224010);国家自然科学基金(批准号:62576025)

Supported by: Beijing Municipal Science and Technology Program Project (Grant No. Z241100004224010); the National Natural Science Foundation of China (Grant No. 62576025)

© 中国图象图形学报版权所有

high-thermal regions, dead pixels, and shadows may appear simultaneously. A full-range compression strategy may allocate many output gray levels to background intervals or unoccupied intensity levels, while the high-thermal region occupies a disproportionate part of the bright end. Consequently, the gray-level differences that distinguish small targets from their surroundings, describe target boundaries, or maintain the transition between target and shadow may not be adequately represented in the fixed eight-bit image. This problem is closely related to the reliability of subsequent target localization and detection, because the detector receives only the compressed image and cannot recover gray-level distinctions that have already been merged during preprocessing. The required processing must therefore determine which intensity intervals contain effective scene information, distribute a limited number of output levels according to the current scene, and complete the resulting operation within the timing constraints of an airborne processor. The objective of this study is to develop an adaptive processing method for high-dynamic-range infrared images in complex scenes and to establish an edge-hardware implementation that preserves useful target information while providing real-time processing capability. **Method** A scene-adaptive gray-level processing method is developed from the relationship between image content distribution and the available eight-bit output levels. For each input frame, the method analyzes the number of pixels associated with different intensity intervals together with the occupancy of the corresponding gray levels. These two types of information are used to identify the intensity intervals that carry effective scene content and to exclude unoccupied intervals from consuming independent output levels. The remaining output gray levels are then assigned according to the actual distribution of the current scene. This design allows high-thermal regions to be compressed within an appropriate bright-end range, while more output levels are retained for low-contrast targets, target-background transitions, and shadow-related intensity changes. The processing curve is continuous and monotonic, so the original intensity order is preserved and abrupt gray-level discontinuities are avoided. The curve is determined from the statistics of the current frame and converted into a lookup table for pixel-level transformation. In this way, the adaptive decision is made at the frame level, whereas the pixel transformation is reduced to a uniform table-based operation. The method also introduces target effective gray-level coverage as an application-oriented image measure. This measure describes the proportion of distinct output gray levels actually used within a target region relative to the available output gray-level space. It evaluates whether the target receives sufficient representational levels after compression and therefore complements general image-quality measures that mainly describe similarity to a reference image. A higher target effective gray-level coverage indicates that more intensity distinctions within the target can reach the subsequent airborne vision algorithm, which is particularly relevant to distant small targets and low-contrast targets. For edge execution, the frame-level statistics and the corresponding adaptive curve are calculated once for the current image, after which the lookup table is stored as shared read-only data. The image is divided into independent regions and processed in parallel by the eight cores of the TMS320C6678 multi-core digital signal processor. Each core performs the same table-based conversion on its assigned region, and the converted regions form the complete eight-bit output image. This organization retains the scene-adaptive processing principle while converting the computationally intensive per-pixel curve evaluation into a regular operation suitable for multi-core execution. The implementation is evaluated using 198 paired high-dynamic-range thermal infrared images. To represent the conditions encountered in complex airborne scenes, four types of test data are constructed by introducing high-thermal interference, dead-pixel interference, shadow interference, and their combined interference. The proposed method is compared with BF&DDE, ADPHE_NCVH, and ACASP. Peak signal-to-noise ratio, structural similarity index, tone-mapped image quality index, and target effective gray-level coverage are used for image evaluation. The effect on downstream detection is assessed using YOLOv8m, ATSS, Sparse R-CNN, and Cascade R-CNN. **Result** On the complex-scene image data, the proposed method obtains a peak signal-to-noise ratio of 17.41 dB, a structural similarity index of 0.77, and a tone-mapped image quality index of 0.83. The effective gray-level coverage reaches 39.76% for person targets, 46.74% for vehicle targets, and 41.80% for all evaluated targets. These results indicate that the adaptive processing preserves more usable gray-level distinctions in target regions while maintaining the overall structure of the image. The target coverage results also show that the output gray-level allocation is related to the practical information available to the detector, especially for small and low-contrast targets. The proposed preprocessing method improves the detection results under all four detection frameworks. The mAP@[0.5;0.95] values reach 49.3% for YOLOv8m, 46.2% for ATSS, 42.5% for Sparse R-CNN, and 46.7% for Cascade R-CNN. Compared with the

© 中国图象图形学报版权所有

strongest compared compression method within each framework, the improvements are 3.1, 3.2, 2.9, and 3.7 percentage points, respectively. The corresponding mAP@0.5 values are 84.0%, 83.8%, 79.2%, and 82.0%. The improvement is also reflected in the detection of small targets, for which the mAP@small values are 36.5%, 34.4%, 32.0%, and 35.2%, respectively. These results demonstrate that the gray-level allocation mechanism improves not only general visual similarity but also the information available for downstream target detection. On the TMS320C6678 platform, the test resolution is 640 × 512 pixels, and the evaluation uses the total processing time for a single frame. The original implementation requires 78.33 ms per frame, corresponding to 12.77 frames per second. After the lookup-table-based multi-core implementation, the time decreases to 6.10 ms per frame, corresponding to 163.93 frames per second. The resulting acceleration ratio is 12.84. The comparison confirms that the shared table and multi-core regional execution provide a substantial reduction in the repeated per-pixel computation required by the adaptive curve. **Conclusion** The proposed method establishes a complete processing route for complex-scene high-dynamic-range infrared images: current-frame content determines the allocation of the fixed eight-bit output range, a continuous monotonic curve preserves the ordering and transitions of scene intensities, and a shared lookup table enables efficient execution on a multi-core airborne processor. The method directly addresses the joint requirements of high-thermal interference compression, effective gray-level allocation, target detail preservation, and edge-side processing speed. Experimental results from image quality assessment, target gray-level coverage, multi-framework detection, and TMS320C6678 implementation demonstrate that adaptive gray-level processing can improve the information representation of complex infrared scenes while meeting the processing requirements of airborne visual applications.

Key words: complex scenes; high-dynamic infrared images; adaptive processing; curve mapping; edge hardware acceleration

0 引　言

低空无人机执行电力巡检、设施巡查、夜间搜救和应急监测等任务时，机载红外传感器持续采集红外图像，呈现目标与背景的热辐射差异。传感器输出的图像通常具有较高位深，而显示、传输和视觉处理模块常使用固定低位宽的单通道图像，因此需要进行位宽转换处理。这一步处理发生在目标检测之前，直接影响图像中哪些灰度差异能够保留下来。低空飞行时，远处目标往往只占少量像素，与背景的差异也可能集中在很窄的灰度区间。如果转换时将这些灰度级合并，检测网络便无法再利用原图中的这部分差异。因此，位宽转换处理既影响图像显示效果，也影响机载目标检测可利用的灰度信息（Gade等，2014；Yu等，2020；李赓飞等，2016）。

复杂场景下，位宽转换处理需要同时兼顾差异很大的图像区域。同一帧中，火光、高温设备或强反射可能形成高热区域，海面、道路或建筑可能占据大部分画面，遮挡处会形成阴影，远距离小目标和低对比目标则往往只占少量像素。高热区域虽然像素不多，却可能拉大输入灰度范围，使其他区域能分配到的输出层次减少；大面积背景在像素统计中占优势，但像素多不代表它包含更多有助于检测目标的灰度差异；阴影使局部区域变暗，低对比目标与背景的差异则可能只有几个相邻灰度级（李俊宏等，2020；韩金辉等，2022）。因此，需要先确定当前帧实际出现了哪些灰度级，再根据这些灰度级的分布安排有限的输出灰度空间，同时保留目标与周围背景的细微差异。

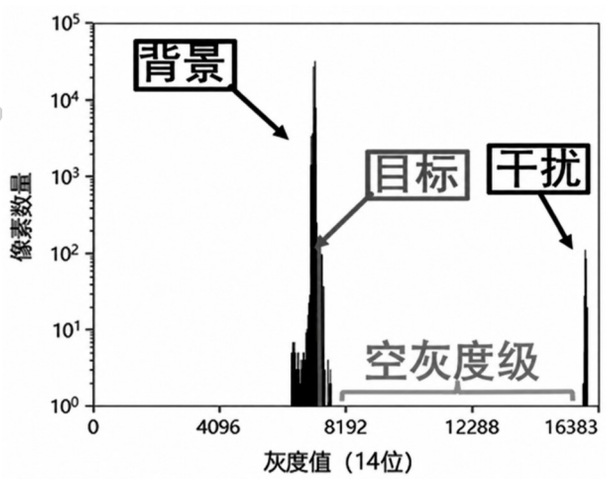


图1　复杂场景下14位高动态红外图像的灰度分布

Fig. 1　Gray-level distribution of a complex-scene 14-bit infrared image

© 中国图象图形学报版权所有

图1展示了一幅14位高动态红外图像的灰度直方图，横轴为输入灰度值，纵轴以对数刻度表示像素数量。背景像素集中在主要灰度区间，目标相关的灰度差异位于背景峰附近；高灰度端有少量干扰峰，而约8192至高灰度端之间有一段较长的空灰度区间。可见，分配输出灰度空间时，既要考虑各灰度级的像素数量，也要考虑哪些灰度级实际出现，以及它们在输入灰度轴上的位置。

已有高动态范围图像处理方法为灰度压缩提供了多种做法。针对高动态红外图像的可视化，研究者通过调整高亮区域和整体对比度改善显示层次（Branchitta等，2009）；对数映射、摄影色调映射和多尺度视网膜模型使用固定或预设的曲线压缩高对比度场景的灰度范围（Drago等，2003；Reinhard等，2002；Jobson等，1997）。自适应直方图均衡、平台直方图和双平台直方图方法根据像素分布调整灰度级（Pizer等，1987；Vickers，1996；Wang等，2006；Liang等，2012；Li等，2018；王园园等，2019；葛朋等，2020；Zuiderveld，1994）；双边滤波、梯度域重建和引导滤波方法利用局部结构或梯度关系改善细节（Durand等，2002；Fattal等，2002；He等，2013；Liu等，2014；范永杰等，2010；Zuo等，2011）。近年来，对比学习、结构保持和场景重标定也被用于色调映射与热红外图像处理（Hu等，2022；Cao等，2024；Gil等，2024；Qiu等，2024；Zamir等，2022）。对于固定低位宽的机载视觉输入，还需要明确当前帧哪些灰度级确实包含有效内容，以及如何在有限的输出层次中保留它们的灰度顺序和局部差异。

机载部署还受到处理时间、功耗和存储容量的限制。机载处理器需要同时完成图像转换、目标检测、通信和控制等任务，因此图像转换的逐像素计算应尽量简单、稳定。移动和嵌入式视觉研究表明，计算量、存储访问和并行方式都会影响系统的实时性（Sze等，2017；Chen等，2016；Sandler等，2018；Howard等，2019；Qin等，2024；Wang等，2024）。自适应映射需要依据整帧的灰度分布确定；如果转换每个像素时仍要重复访问邻域或计算映射，整帧分析的开销就会进入逐像素处理过程。因此，需要先完成当前帧的场景判断，并把结果保存为像素转换时可直接使用的映射关系。

针对这些需求，本文先找出当前帧实际出现的灰度级，确定需要保留的灰度内容，再构建连续单调的映射曲线，将其分配到固定的八位输出范围。未出现的灰度级不单独占用输出间隔；对于实际出现的灰度级，结合像素数量权重和灰度级占用权重安排输出间隔，使大面积背景和像素较少但确实存在的灰度层次都能得到考虑。高热像素集中在亮端时，通过连续改变曲线斜率压缩亮端区间；暗部补偿则作用于已有灰度差异所在的区域。曲线生成后保存为当前帧查找表，像素转换只需读取对应表项。在端侧实现中，各核心共同读取这张表，按图像区域并行完成转换，使每帧只需进行一次场景判断，并在TMS320C6678上验证处理速度。

本文的贡献主要有两个方面。首先，提出面向复杂场景的有效灰度内容判定与连续曲线重构方法，将空灰度级处理、有效灰度级选择、两类权重分配以及亮端压缩和暗部补偿纳入同一映射过程。其次，将每帧生成的映射关系保存为多核共享的查找表，使各核心按输入灰度直接查表，完成所负责图像区域的转换；通过四类复杂场景的实验、与三种动态范围压缩方法的对比，以及在四种目标检测框架上的测试，验证该方法对机载视觉输入的作用。

1 有效灰度内容驱动的复杂场景自适应处理

1.1 固定低位宽输入下的有效信息重构

设机载红外传感器输出的高动态图像为$\boldsymbol{X}(x,y)$，输出图像为$\boldsymbol{Y}(x,y)$，输入灰度位置为k。位宽转换由保持灰度顺序的单调映射T完成：

$$\boldsymbol{Y}(x,y)=T\left(\boldsymbol{X}(x,y)\right),k_1\leqslant k_2\Rightarrow T(k_1)\leqslant T(k_2)\tag{1}$$

式中，(x,y)表示像素坐标；式(1)中的映射函数将输入灰度值转换为输出灰度值，k_1与k_2表示任意两个输入灰度级。映射曲线在某一灰度位置的增量，决定相邻输入灰度级在输出端相隔多远。增量越大，原有的灰度差异越容易保留；增量越小，该区间的灰度级就被压缩得越紧。由于输出位宽固定，可用的灰度层次有限，需要根据当前帧的内容决定哪些区间分配较多层次，哪些区间适当压缩。

1.2 场景有效灰度重构

统计当前帧每个输入灰度级的像素数量，得到灰度直方图，记为$H(k)$。当$H(k)=0$时，说明该灰度级在当前帧没有对应像素，属于空灰度级。空灰

© 中国图象图形学报版权所有

度级不参与后续权重计算，也不单独分配输出间隔。查找表仍覆盖完整的输入灰度范围，以便像素按输入灰度值直接查表；当前帧的像素不会访问空灰度级对应的表项。

直方图中计数大于零的灰度级组成有效灰度级集合。后续分配输出间隔时，既考虑每个灰度级有多少像素，也考虑它是否实际出现。这样，大面积背景对应的主要灰度层能够得到保留，像素较少但确实存在的灰度级也不会被忽略。高热区域和孤立极值同样计入直方图；对于其中位于高灰度端的像素，后续通过压缩亮端区间，减少少量极高值像素对八位输出空间的占用。

1.3 灰度数量与位置占用协同的连续映射

每个有效灰度级的权重由两部分组成。像素数量权重对直方图计数进行对数压缩，使大面积背景仍能保留主体层次，同时减弱其像素数量优势；灰度级占用权重以当前灰度级实际出现为前提，并结合该级像素计数相对于有效灰度级平均计数的可信度计算，使像素较少的有效灰度级也参与输出间隔分配。两种权重按输入灰度位置融合：

$$w(k)=a(k)w_{\text{count}}(k)+\left[1-a(k)\right]w_{\text{occupancy}}(k),\tag{2}$$

式中，$w_{\text{count}}(k)$表示经过数量压缩的像素权重，$w_{\text{occupancy}}(k)$表示灰度级占用权重，$a(k)$表示随灰度位置变化的融合系数，计算公式为$a(k)=(k_{\max}-k)/(k_{\max}-k_{\min})$，$k\in\boldsymbol{\Omega}$，$k_{\max}>k_{\min}$，$0\leqslant a(k)\leqslant 1$，$k_{\min}$和$k_{\max}$分别为当前帧实际出现的最小和最大灰度级。暗部和主体区间取值较大，中高灰度区间取值逐渐减小。式(2)计算融合后的灰度级权重。

得到灰度级权重后，再根据当前帧的整体灰度位置、亮端像素的集中程度、灰度分布层次和灰度占用宽度调整映射曲线。高热像素越集中在亮端，该区间分配的输出间隔越小，从而把更多输出层次留给主体和暗部。若场景整体偏暗，且低灰度区域仍有连续的灰度差异，则适当增大暗部的输出间隔，使已有的低对比差异更容易区分。暗部补偿随输入灰度升高逐渐减弱，以保持映射曲线连续和全局灰度顺序不变。

最后，沿输入灰度轴累积各有效灰度级的权重，得到单调映射曲线：

$$T(k)=\text{round}\left(\frac{\sum_{i\in\boldsymbol{\Omega},i\leqslant k}w(i)}{\sum_{i\in\boldsymbol{\Omega}}w(i)}\times\left(2^{Q}-1\right)\right),k\in\boldsymbol{\Omega},\tag{3}$$

式中，$\boldsymbol{\Omega}$为有效灰度级集合，Q为输出位深，round(·)表示取整，求和下标表示输入灰度级。式(3)按灰度顺序累积权重，并将结果换算到输出位宽对应的灰度范围。由于累积权重不会随输入灰度升高而减小，取整后的映射值也保持非递减。空灰度级不参与求和，但查找表仍保留其输入索引；转换当前帧时，像素只会访问实际出现的灰度级对应的表项。

2 映射状态复用驱动的多核端侧加速

2.1 帧级场景判定的映射状态表达

确定映射曲线需要分析整帧的灰度分布，而转换单个像素时，只需知道该输入灰度值对应的输出值。因此，本文将处理过程分为两个阶段：先在帧级完成灰度内容判定、场景分析和曲线生成，再将映射结果保存为查找表，供像素转换直接使用。输入位深确定后，查找表的长度也随之确定，不受目标数量、目标位置或局部纹理变化的影响。

查找表记录当前帧各输入灰度值对应的输出值，保存了当前帧对高热区域、主体层次、阴影过渡和低对比目标的统一判断。建表完成后，同一帧的所有图像区域共用这张表，各处理核心只需按输入灰度读取表项，处理同一组只读映射关系即可完成转换，无需重新计算映射曲线。图2展示了当前帧建表、各核心共享查找表以及图像区域并行转换的流程。

2.2 共享映射状态支撑的多核并行转换

高动态红外图像的映射关系由整帧灰度内容决定，输出却要落实到每一个像素。端侧实现将这两个层次分开处理：先根据当前帧的灰度统计和有效灰度内容形成映射状态，再依据输入灰度完成像素转换。每帧只生成一次查找表，转换像素时根据输入灰度读取对应的输出值，从而避免在每个像素上重复进行曲线求值和场景判断。

在德州仪器 TMS320C6678 多核数字信号处理器上，当前帧映射状态写入共享存储区，作为只读信息供各核心共同使用。各核心按照图像区域分配像素任务，读取同一映射状态并写回对应输出区域；映射状态完成后即可独立推进像素转换，核心之间无需交换像素结果。因此，各核心共用同一查找表，并独立处理各自分配的图像区域。为考察共享查表对

© 中国图象图形学报版权所有

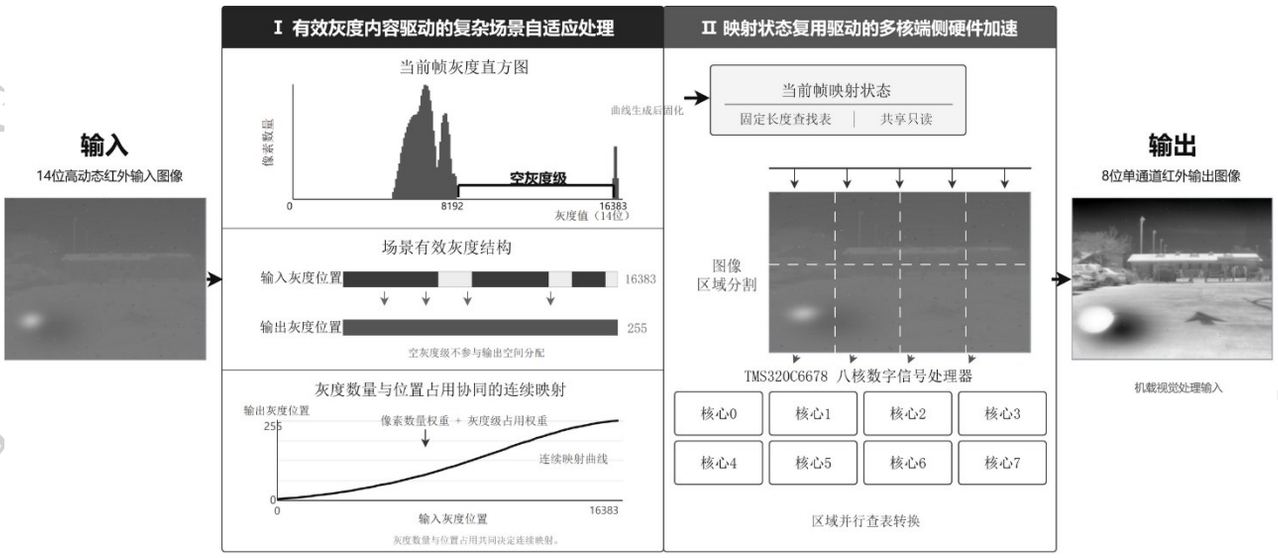


图2　复杂场景高动态红外图像自适应处理与曲线映射状态执行结构

Fig. 2　Complex-scene high-dynamic infrared adaptive processing and curve-mapping-state execution

处理速度的影响，本文比较两种多核实现。两者使用相同的输入帧、帧级统计结果和输出格式：原始实现对每个像素重复计算映射值，端侧加速实现则先生成共享查找表，再按输入灰度读取输出值。

2.3　线性复杂度与端侧实时处理能力

设图像像素数为N，输入灰度轴长度为L，局部处理窗口边长为s，梯度域重建迭代次数为K。以下用O( )表示渐近计算复杂度的量级。按各方法的直接实现计算上限，双边滤波与数字细节增强（bilateral filter and digital detail enhancement，BF&DDE）包含局部双边滤波，复杂度为$O(Ns^2)$；基于直方图归一化变异系数反馈的自适应双平台直方图均衡（adaptive double plateaus histogram equalization with feedback from the normalized coefficient of variation of the histogram，本文记为ADPHE_NCVH）主要包括直方图统计和灰度级调整，复杂度为$O(N+L)$；自适应对比度调整与结构保持（adaptive contrast adjustment and structure preservation，ACASP）需要计算局部对比度和局部方差，并进行梯度域结构重建，复杂度上限可写为$O(Ns^2+KN)$；本文方法完成一次直方图统计、一次灰度轴曲线生成和一次像素查表，单核复杂度上限为$O(N+L)$，额外存储为$O(L)$。在P个核心均衡处理像素区域时，本文方法像素阶段的并行部分约为$O(N/P)$，全局统计归约和查找表生成仍与L有关；由于14位输入下L为固定长度，本文方法的像素工作量随图像规模线性增长。

3　复杂场景处理能力与端侧实时验证

3.1　面向复杂干扰的场景构造与验证任务

定量实验采用带目标标注的高动态热红外图像，数据来源于Teledyne FLIR热红外数据集（Teledyne FLIR，2018），选取198幅具有有效目标标注的图像作为复杂场景评价样本。每幅图像保留原始目标标注，并在同一幅高动态输入上构造四类干扰场景，使不同方法在相同输入和相同目标位置下完成比较。

高热干扰在目标区域以外加入局部高热区域，用于模拟火光、高温设备和强反射引入的高灰度像素；坏点干扰加入稀疏极高值和极低值，用于模拟传感器孤立异常；阴影干扰在背景区域形成连续暗区，用于模拟建筑遮挡、阴影覆盖和局部遮挡；联合干扰同时包含高热区域、坏点和阴影，用于评估多类干扰同时存在时的处理结果。干扰区域尽量避开目标主体，以便比较各方法处理后目标区域的灰度分布。

前端动态范围压缩对比包括BF&DDE、ADPHE_NCVH、ACASP和本文方法。BF&DDE代表局部空间滤波与细节增强方法，对应范永杰等（2010）和Zuo等（2011）；ADPHE_NCVH代表自适应双平台直方图均衡方法，对应Li等（2018）；ACASP代表Qiu等（2024）提出的自适应对比度调整与结构

© 中国图象图形学报版权所有

保持方法，该方法依据局部对比度构造映射曲线，并依据局部方差调制梯度，对大梯度区域进行抑制以保持强边缘，对小梯度区域进行增强以呈现低对比细节，随后在梯度域完成结构保持重建。所有方法均由相同的高动态复杂场景输入生成低位宽输出。

3.2 有效灰度保持与目标检测评价

图像质量评价采用峰值信噪比（peak signal-to-noise ratio，PSNR）、结构相似性指数（structural similarity index measure，SSIM）、色调映射图像质量指数（tone-mapped image quality index，TMQI）和目标有效灰度覆盖率四项指标。PSNR评价输出与统一高动态参考图像之间的像素保真度，SSIM评价亮度、对比度和结构关系的保持程度，TMQI同时考虑高动态参考图像与输出图像之间的结构保真度以及输出图像的统计自然度（Yeganeh等，2013）。评价时，输入高动态图像和各方法输出均转换到统一评价域，所有方法使用相同参考和相同数据划分。

目标有效灰度覆盖率用于描述目标区域实际获得的输出灰度层次。机载红外目标通常占据较少像素，目标与背景之间可供检测利用的差异集中在有限的灰度范围内。峰值信噪比、结构相似性和色调映射图像质量主要反映整幅图像的总体保真度、结构保持程度和色调质量，对目标区域实际获得的可区分灰度层次不作直接度量。目标有效灰度覆盖率统计目标区域实际使用了多少个不同的输出灰度级。设第t个目标区域为$\boldsymbol{R}_t$，区域内实际出现的不同输出灰度级集合为$\boldsymbol{G}_t$，则目标有效灰度覆盖率定义为：

$$C_{\text{target}}=\frac{1}{M}\sum_{t=1}^{M}\frac{|\boldsymbol{G}_t|}{2^Q}\times 100\%,\tag{4}$$

式中，M为目标区域数量，$|\boldsymbol{G}_t|$为第t个目标区域实际使用的不同输出灰度级数量，Q为输出位深。该指标越高，表示目标区域使用的不同输出灰度级越多。四项图像指标按高热、坏点、阴影和联合干扰场景分别计算；目标有效灰度覆盖率按照目标区域统计。

下游检测实验覆盖单阶段、两阶段和稀疏候选框等检测路线（Redmon等，2016；Ren等，2017；Tan等，2020），采用YOLOv8中型模型（you only look once version 8 medium，YOLOv8m）、自适应训练样本选择（adaptive training sample selection，ATSS）、稀疏区域卷积神经网络（sparse region-based convolutional neural network，Sparse R-CNN）和级联区域卷积神经网络（cascade region-based convolutional neural network，Cascade R-CNN）四种目标检测框架，对应方法分别为Jocher等（2023）、Zhang等（2020）、Sun等（2021）和Cai等（2018）。各检测框架使用相同的训练集、验证集、训练策略和输入尺寸，每种动态范围压缩算法分别生成训练输入和验证输入。检测结果采用常见物体场景（common objects in context，COCO）评价协议，以平均精度均值（mean average precision，mAP）@［0.5：0.95］、mAP@0.5、mAP@small、mAP@medium和mAP@large表示。端侧实验在640×512分辨率下采用德州仪器TMS320C6678八核实现，比较原始算法与端侧硬件加速算法的单帧总时间和帧率，并据此计算加速比。

3.3 复杂干扰下的灰度空间重构效果

图3给出三组复杂干扰场景的处理结果及目标局部放大区域。ACASP和ADPHE_NCVH在部分场景中能够调整整体亮度，但目标局部的灰度过渡仍较集中；BF&DDE增强了局部边缘，同时在高热区域和细节区域产生较明显的纹理起伏。本文方法对高热区域所在亮端进行连续压缩，目标框内及局部放大图中的目标轮廓、相邻灰度过渡和车体结构保持较为完整，说明依据场景有效灰度内容安排输出灰度间隔能够同时兼顾高热干扰抑制与目标细节表达。

图4进一步给出了3组复杂场景中目标区域输出灰度的分布情况。每行对应一组测试场景，每列对应一种动态范围压缩方法，图中百分比表示目标区域实际使用的输出灰度级数量占8位输出灰度空间的比例，纵轴采用对数坐标。ACASP的目标有效灰度覆盖率分别为42.2%、29.3%和62.5%，ADPHE_NCVH分别为22.3%、14.1%和44.1%，目标灰度主要集中在较窄区间，部分输出灰度级未被使用。BF&DDE的灰度覆盖范围较宽，覆盖率分别达到66.8%、37.9%和90.2%，但直方图中同时出现多个孤立峰和稀疏尾部。本文方法的覆盖率分别为54.7%、41.0%和82.0%，3组场景均高于ACASP和ADPHE_NCVH；从直方图看，目标区域使用的灰度级主要集中在连续区间。场景2中，本文方法的覆盖率高于全部对比方法；场景1和场景3中，BF&DDE的统计覆盖率更高，其输出灰度分布也更加分散。

表1比较不同动态范围压缩方法在高热、坏点、

© 中国图象图形学报版权所有

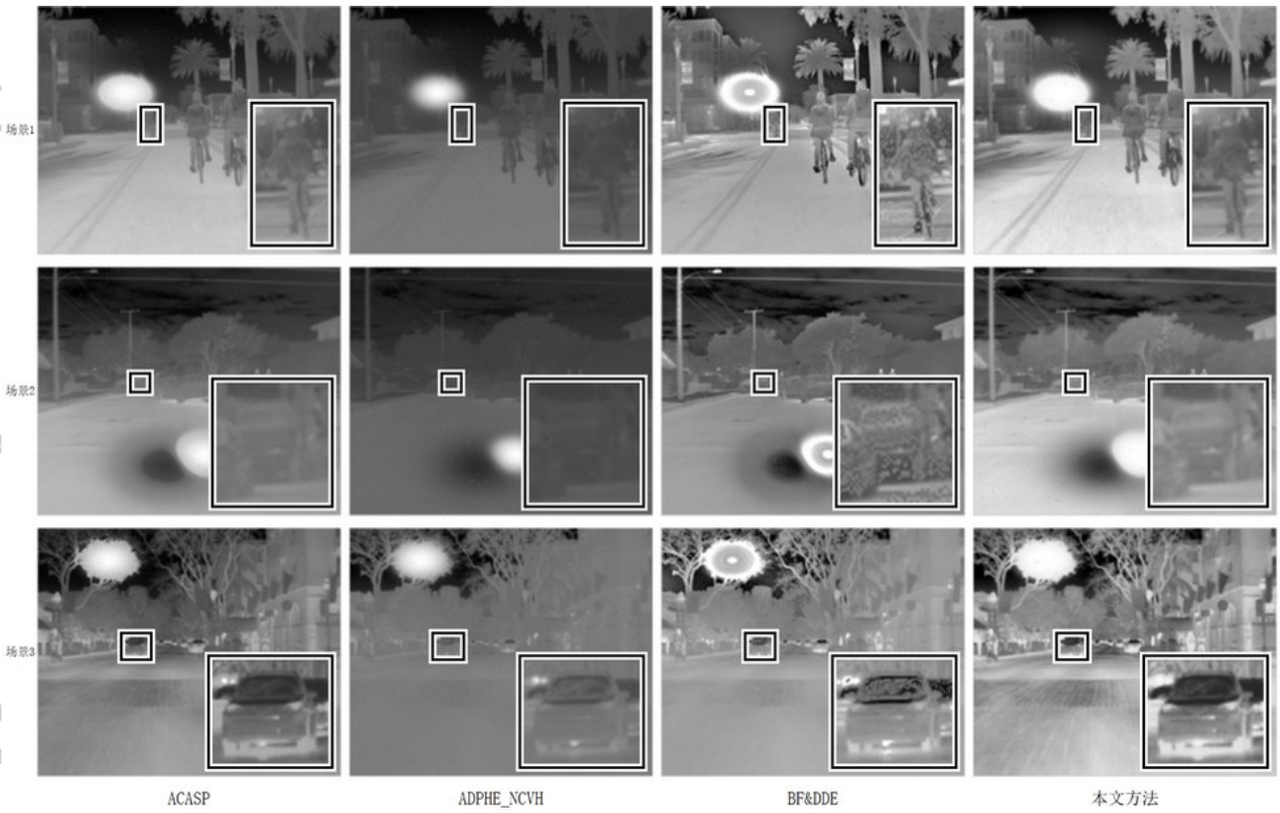


图3　复杂干扰场景下不同动态范围压缩方法的可视化比较

Fig. 3　Visual comparison of dynamic-range compression methods under complex interference

表1　复杂场景下不同动态范围压缩方法的综合评价结果

Table 1　Comprehensive evaluation of dynamic-range compression methods in complex scenes

方法	PSNR/dB	SSIM	TMQI	人目标覆盖率/%	车目标覆盖率/%	全部目标覆盖率/%
BF&DDE	13.75	0.66	0.74	32.08	41.02	36.37
ADPHE_NCVH	11.01	0.46	0.48	13.35	18.38	15.47
ACASP	15.72	0.66	0.72	23.19	31.22	27.09
本文方法	**17.41**	**0.77**	**0.83**	**39.76**	**46.74**	**41.80**

注:加粗数值表示各指标的最高值。

阴影和联合干扰四类复杂场景综合测试集上的峰值信噪比、结构相似性、色调映射图像质量,以及按人目标、车目标和全部目标统计的有效灰度覆盖率。前3项指标反映图像整体质量,3类覆盖率反映不同目标区域在固定低位宽输出中实际获得的灰度层次。四类场景采用相同目标标注和相同评价口径,表中本文方法表示场景有效灰度重构方法。本文方法的各项指标均为各方法中的最高值。结果表明,基于场景有效灰度重构的本文方法能够在复杂干扰条件下改善整体灰度表达,并为不同类型目标保留更多可供机载视觉处理利用的灰度层次。

3.4　灰度空间重构对目标检测的作用

表2比较不同动态范围压缩算法对四种检测框架的影响。每个检测框架内只改变前端动态范围压缩方法,训练数据、验证数据和检测器设置保持一致。mAP@[0.5:0.95]反映多个交并比阈值下的综合检测质量,mAP@0.5反映固定交并比阈值下的检测质量,mAP@small、mAP@medium和mAP@large分别反映小、中、大目标的检测结果。本文方法在四种检测框架的mAP@[0.5:0.95]均取得框架内最高结果,较各框架中最佳对比方法提高2.9～3.7个百分点;mAP@0.5、mAP@small、mAP@medium和

© 中国图象图形学报版权所有

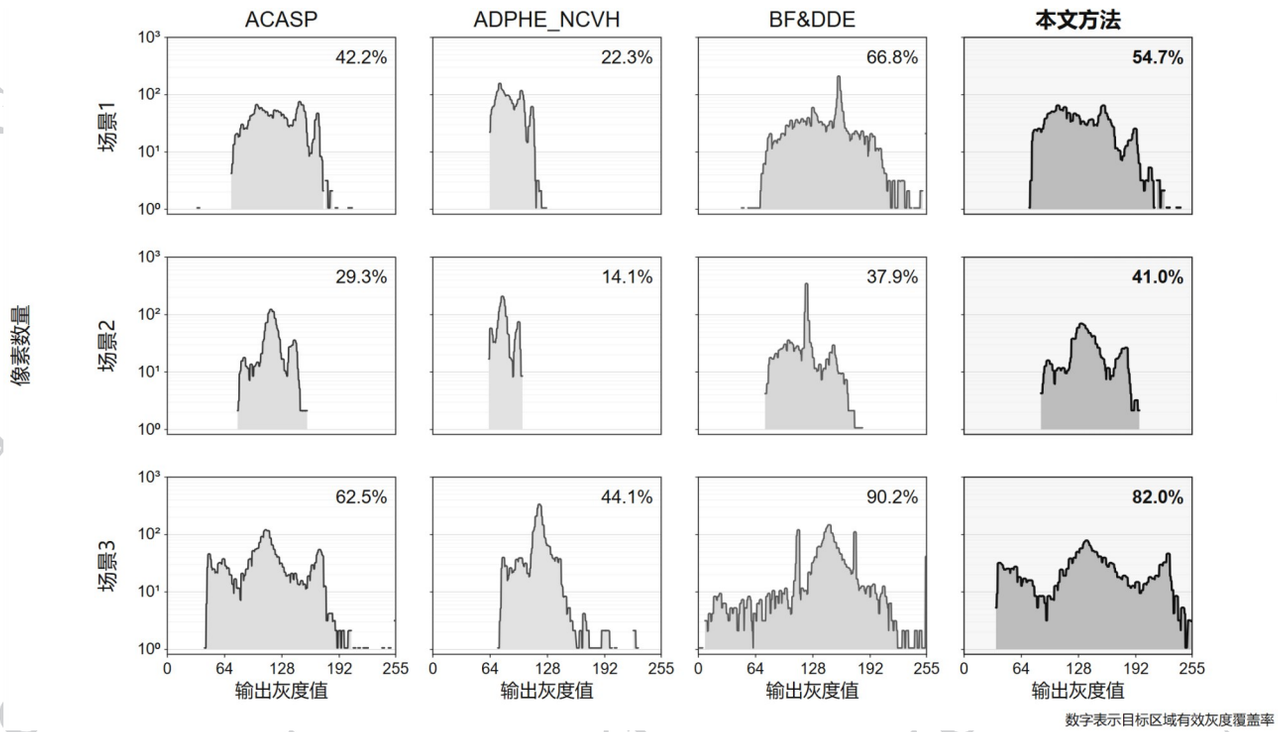


图4　复杂场景下目标区域的输出灰度直方图

Fig. 4　Output gray-level histograms of target regions under complex scenes

mAP@large同样取得框架内最高结果。不同检测器具有不同的特征提取和候选框处理机制，结果仍保持一致，说明前端曲线重构所保留的灰度层次能够转化为检测器可利用的目标信息，尤其对小目标和复杂背景下的目标检测具有直接作用。

3.5　映射状态复用的多核端侧实时性

端侧实验在德州仪器TMS320C6678八核数字信号处理器上，以640×512图像测试各方法。本文方法分别运行原始算法和端侧硬件加速算法，记录单帧总时间和帧率；单帧时间包含当前帧处理过程，预加载查找表所需时间不计入统计。

端侧性能以单帧总时间为主要指标，帧率按$R = 1000/t_{\text{frame}}$计算，其中$t_{\text{frame}}$为单帧总时间，单位为毫秒。表3列出相同处理器和640×512输入条件下各方法的结果。本文方法的端侧硬件加速算法用时6.10 ms，较原始算法的78.33 ms缩短；帧率由12.77帧/s提高至163.93帧/s，加速比为12.84。

3.6　场景有效灰度重构组成部分消融

为分析场景有效灰度重构中各处理环节的作用，以本文完整方法为基线，在与表1相同的测试图像、干扰场景和评价设置下，每次只调整一个环节，分别考察有效灰度级选择、像素数量与灰度级占用权重配置，以及亮端压缩的影响。结果见表4。

表4显示，未排除空灰度级时，目标有效灰度覆盖率由41.80%降至14.13%，PSNR、SSIM和TMQI也分别降至12.51 dB、0.63和0.35。仅使用一种权重时，完整方法的PSNR、SSIM和TMQI均更高。其中，仅用像素数量权重的目标覆盖率为42.91%，略高于完整方法的41.80%，但其余三项指标较低，说明目标覆盖率需要结合图像质量指标分析。不进行亮端压缩时，四项指标均小幅下降。这些对照结果表明，有效灰度级选择、两类权重的协同分配及亮端曲线调节共同影响完整方法的处理结果。

4　结　论

针对低空无人机红外感知中高动态图像需要转换为固定八位输入的问题，本文提出了复杂场景高动态红外图像自适应处理方法。该方法根据当前帧实际出现的灰度级及其像素数量分配输出间隔，不为未出现的灰度级单独分配间隔，并形成连续单调的映射曲线。映射曲线对高热区域所在的亮端进行渐进压缩，保持主体与阴影之间的灰度过渡，同时适度扩大低对比目标已有的灰度差异。

© 中国图象图形学报版权所有

表2　不同动态范围压缩算法在多种目标检测框架下的性能对比

Table 2　Performance comparison of dynamic-range compression methods with different object detection frameworks

目标检测算法	动态范围压缩算法	mAP@[0.5:0.95]/%	mAP@0.5/%	mAP@small/%	mAP@medium/%	mAP@large/%
YOLOv8m	BF&DDE	46.0	81.0	32.4	52.5	63.7
	ADPHE_NCVH	44.6	79.6	30.7	51.4	62.5
	ACASP	46.2	81.5	33.3	52.8	63.8
	本文方法	**49.3**	**84.0**	**36.5**	**55.9**	**68.3**
ATSS	BF&DDE	42.2	79.4	31.2	48.7	56.8
	ADPHE_NCVH	41.3	77.8	29.8	48.1	57.6
	ACASP	43.0	79.9	31.9	49.4	52.5
	本文方法	**46.2**	**83.8**	**34.4**	**52.7**	**63.7**
Sparse R-CNN	BF&DDE	38.8	75.0	28.9	44.4	55.1
	ADPHE_NCVH	35.2	71.8	25.7	42.2	49.9
	ACASP	39.6	76.4	29.6	45.1	53.8
	本文方法	**42.5**	**79.2**	**32.0**	**49.3**	**59.2**
Cascade R-CNN	BF&DDE	41.6	77.1	30.0	48.8	56.1
	ADPHE_NCVH	40.9	76.0	29.4	47.3	58.6
	ACASP	43.0	79.6	31.5	49.2	59.6
	本文方法	**46.7**	**82.0**	**35.2**	**52.3**	**63.6**

注：加粗数值表示同一目标检测算法下各指标的最高值。

表3　TMS320C6678多核端侧处理时间和帧率比较

Table 3　Multicore edge processing time and frame-rate comparison on TMS320C6678

实现方式	单帧总时间/ms	帧率/(帧/s)
ACASP	402.83	2.48
ADPHE_NCVH	101.36	9.87
BF&DDE	1929.38	0.52
本文方法(原始算法)	78.33	12.77
本文方法(端侧硬件加速算法)	**6.10**	**163.93**

注：加粗数值表示最好结果。

表4　场景有效灰度重构组成部分消融结果对比

Table 4　Ablation results for the components of scene-adaptive gray-level reconstruction

变体	PSNR/dB ↑	SSIM ↑	TMQI ↑	目标覆盖率/% ↑
未排除空灰度级	12.51	0.63	0.35	14.13
本文方法(完整)	**17.41**	**0.77**	**0.83**	**41.80**
仅用像素数量权重	16.30	0.69	0.81	**42.91**
仅用灰度级占用权重	16.94	0.73	0.79	39.22
本文方法(完整)	**17.41**	**0.77**	**0.83**	41.80
不进行亮端压缩	16.53	0.75	0.81	41.52
本文方法(完整)	**17.41**	**0.77**	**0.83**	**41.80**

注：加粗数值表示最好结果。

在端侧实现中，将当前帧的映射曲线保存为固定长度查找表，供多个核心共同读取；各核心分别处理不同图像区域，通过查表完成像素转换。针对高热、坏点、阴影及联合干扰场景，采用PSNR、SSIM、TMQI和目标有效灰度覆盖率评价图像质量与目标灰度层次。在四种检测框架下，该方法的mAP均为各框架内最高，表明在本文测试设置下，处理后的图像有利于提高目标检测精度。TMS320C6678八核实验中，采用查找表复用和多核并行后，单帧处理时间由78.33 ms降至6.10 ms，加速比为12.84。

参考文献(References)

Branchitta F, Diani M, Corsini G and Romagnoli M. 2009. New tech-

© 中国图象图形学报版权所有

nique for the visualization of high dynamic range infrared images. Optical Engineering, 48(9): 096401 [DOI: 10.1117/1.3216575].

Cai Z and Vasconcelos N. 2018. Cascade R-CNN: Delving into high quality object detection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 6154-6162 [DOI: 10.1109/CVPR.2018.00644].

Cao C, Yue H, Liu X and Yang J. 2024. Unsupervised HDR image and video tone mapping via contrastive learning. IEEE Transactions on Circuits and Systems for Video Technology, 34(2): 786-798 [DOI: 10.1109/TCSVT.2023.3290351].

Chen Y H, Emer J S and Sze V. 2016. Eyeriss: A spatial architecture for energy-efficient dataflow for convolutional neural networks// 2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture. Seoul: IEEE: 367-379 [DOI: 10.1109/ISCA.2016.40].

Drago F, Myszkowski K, Annen T and Chiba N. 2003. Adaptive logarithmic mapping for displaying high contrast scenes. Computer Graphics Forum, 22(3): 419-426 [DOI: 10.1111/1467-8659.00689].

Durand F and Dorsey J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. ACM Transactions on Graphics, 21(3): 257-266 [DOI: 10.1145/566654.566574].

Fan Y J, Jin W Q, Liu B and Liu C L. 2010. An analysis of digital detail enhancement (DDE) technology developed by FLIR. Infrared Technology, 32(3): 161-164 (范永杰, 金伟其, 刘斌, 刘崇亮. 2010. FLIR公司热成像细节增强DDE技术的分析. 红外技术, 32(3): 161-164) [DOI: 10.3969/j.issn.1001-8891.2010.03.010].

Fattal R, Lischinski D and Werman M. 2002. Gradient domain high dynamic range compression. ACM Transactions on Graphics, 21(3): 249-256 [DOI: 10.1145/566654.566573].

Gade R and Moeslund T B. 2014. Thermal cameras and applications: A survey. Machine Vision and Applications, 25(1): 245-262 [DOI: 10.1007/s00138-013-0570-5].

Ge P, Yang B, Hong W Q, Wang X D, Liu C M, Su L, et al. 2020. Dynamic range compression and detail enhancement algorithm combined with PE for high dynamic range infrared images. Infrared Technology, 42(3): 279-285 (葛朋, 杨波, 洪闻青, 王晓东, 刘传明, 苏兰, 等. 2020. 一种结合PE的高动态范围红外图像压缩及细节增强算法. 红外技术, 42(3): 279-285 [DOI: 10.3724/SP.J.7101182190]).

Gil H, Jeon M H and Kim A. 2024. Fieldscale: Locality-aware field-based adaptive rescaling for thermal infrared image. IEEE Robotics and Automation Letters, 9(7): 6424-6431 [DOI: 10.1109/LRA.2024.3406054].

Han J H, Wei Y T, Peng Z M, Zhao Q, Chen Y H, Qin Y, et al. 2022. Infrared dim and small target detection: A review. Infrared and Laser Engineering, 51(4): 20210393 (韩金辉, 魏艳涛, 彭真明, 赵骞, 陈耀弘, 覃尧, 等. 2022. 红外弱小目标检测方法综述. 红外与激光工程, 51(4): 20210393 [DOI: 10.3788/IRLA20210393].

He K M, Sun J and Tang X O. 2013. Guided image filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(6): 1397-1409 [DOI: 10.1109/TPAMI.2012.213].

Howard A, Sandler M, Chu G, Chen L C, Chen B, Tan M, et al. 2019. Searching for MobileNetV3//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE: 1314-1324 [DOI: 10.1109/ICCV.2019.00140].

Hu L, Chen H and Allebach J P. 2022. Joint multi-scale tone mapping and denoising for HDR image enhancement//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops. Waikoloa: IEEE: 729-738 [DOI: 10.1109/WACVW54805.2022.00080].

Jobson D J, Rahman Z and Woodell G A. 1997. A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Transactions on Image Processing, 6(7): 965-976 [DOI: 10.1109/83.597272].

Jocher G, Chaurasia A and Qiu J. 2023. YOLO by Ultralytics[EB/OL]. [2026-07-31]. https://github.com/ultralytics/ultralytics

Li G F, Li G J, Han G L, Liu P X and Jiang S. 2016. Real-time non-uniformity correction of infrared imaging system. Optics and Precision Engineering, 24(11): 2841-2847 (李赓飞, 李桂菊, 韩广良, 刘培勋, 江山. 2016. 红外成像系统的非均匀性实时校正. 光学精密工程, 24(11): 2841-2847) [DOI: 10.3788/OPE.20162411.2841].

Li J H, Zhang P, Wang X W and Huang S Z. 2020. Infrared small-target detection algorithms: a survey. Journal of Image and Graphics, 25(9): 1739-1753 (李俊宏, 张萍, 王晓玮, 黄世泽. 2020. 红外弱小目标检测算法综述. 中国图象图形学报, 25(9): 1739-1753) [DOI: 10.11834/jig.190574].

Li S, Jin W, Li L and Li Y Y. 2018. An improved contrast enhancement algorithm for infrared images based on adaptive double plateaus histogram equalization. Infrared Physics & Technology, 90: 164-174 [DOI: 10.1016/j.infrared.2018.03.010].

Liang K, Ma Y, Xie Y, Zhou B and Wang R. 2012. A new adaptive contrast enhancement algorithm for infrared images based on double plateaus histogram equalization. Infrared Physics & Technology, 55(4): 309-315 [DOI: 10.1016/j.infrared.2012.03.004].

Liu N and Zhao D X. 2014. Detail enhancement for high-dynamic-range infrared images based on guided image filter. Infrared Physics & Technology, 67: 138-147 [DOI: 10.1016/j.infrared.2014.07.013].

Pizer S M, Amburn E P, Austin J D, Cromartie R, Geselowitz A, Greer T, et al. 1987. Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39(3): 355-368 [DOI: 10.1016/S0734-189X(87)80186-X].

Qin D, Leichner C, Delakis M, Fornoni M, Luo S, Yang F, et al. 2024. MobileNetV4: Universal models for the mobile ecosystem// Computer Vision - ECCV 2024. Cham: Springer: 78-96 [DOI: 10.

© 中国图象图形学报版权所有

1007/978-3-031-73661-2_5].

Qiu J, Wang Z, Huang Y and Huang H. 2024. Infrared image dynamic range compression based on adaptive contrast adjustment and structure preservation. IEEE Transactions on Geoscience and Remote Sensing, 62: 1-12, Article 5006512 [DOI: 10.1109/TGRS.2024.3466388].

Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: Unified, real-time object detection//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91].

Reinhard E, Stark M, Shirley P and Ferwerda J. 2002. Photographic tone reproduction for digital images. ACM Transactions on Graphics, 21(3): 267-276 [DOI: 10.1145/566654.566575].

Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031].

Sandler M, Howard A, Zhu M, Zhmoginov A and Chen L C. 2018. MobileNetV2: Inverted residuals and linear bottlenecks//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 4510-4520 [DOI: 10.1109/CVPR.2018.00474].

Sun P, Zhang R, Jiang Y, Kong T, Xu C, Zhan W, et al. 2021. Sparse R-CNN: End-to-end object detection with learnable proposals//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE: 14449-14458 [DOI: 10.1109/CVPR46437.2021.01422].

Sze V, Chen Y H, Yang T J and Emer J S. 2017. Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE, 105 (12): 2295-2329 [DOI: 10.1109/JPROC.2017.2761740].

Tan M, Pang R and Le Q V. 2020. EfficientDet: Scalable and efficient object detection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 10778-10787 [DOI: 10.1109/CVPR42600.2020.01079].

Teledyne FLIR. 2018. FLIR thermal dataset for algorithm development [EB/OL]. [2026-10-04]. https://oem.flir.com/en-gb/solutions/automotive/adas-dataset-form/

Vickers V E. 1996. Plateau equalization algorithm for real-time display of high-quality infrared imagery. Optical Engineering, 35 (7): 1921-1926 [DOI: 10.1117/1.601006].

Wang A, Chen H, Lin Z J, Han J G and Ding G G. 2024. RepViT: Revisiting mobile CNN from ViT perspective//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 15909-15920 [DOI: 10.1109/CVPR52733.2024.01506].

Wang B J, Liu S Q, Li Q and Zhou H X. 2006. A real-time contrast enhancement algorithm for infrared images based on plateau histogram. Infrared Physics & Technology, 48 (1): 77-82 [DOI: 10.1016/j.infrared.2005.04.008].

Wang Y Y, Zhao Y H, Luo H B and Li F Z. 2019. Dynamic range compression and detail enhancement of sea-surface infrared image. Infrared and Laser Engineering, 48(1): 0126003 (王园园, 赵耀宏, 罗海波, 李方舟. 2019. 海面红外图像的动态范围压缩及细节增强. 红外与激光工程, 48(1): 0126003 [DOI: 10.3788/IRLA201948.0126003].

Yeganeh H and Wang Z. 2013. Objective quality assessment of tone-mapped images. IEEE Transactions on Image Processing, 22(2): 657-667 [DOI: 10.1109/TIP.2012.2221725].

Yu H, Li G, Zhang W, Huang Q, Du D, Tian Q, et al. 2020. The unmanned aerial vehicle benchmark: Object detection, tracking and baseline. International Journal of Computer Vision, 128(5): 1141-1159 [DOI: 10.1007/s11263-019-01266-1].

Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: Efficient transformer for high-resolution image restoration//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564].

Zhang S, Chi C, Yao Y, Lei Z and Li S Z. 2020. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 9756-9765 [DOI: 10.1109/CVPR42600.2020.00978].

Zuiderveld K. 1994. Contrast limited adaptive histogram equalization//Heckbert P S, ed. Graphics Gems IV. San Diego: Academic Press Professional, Inc.: 474-485 [DOI: 10.1016/B978-0-12-336156-1.50061-6].

Zuo C, Chen Q, Liu N, Ren J and Sui X. 2011. Display and detail enhancement for high-dynamic-range infrared images. Optical Engineering, 50(12): 127401 [DOI: 10.1117/1.3659698].

© 中国图象图形学报版权所有