

面向红外目标识别的图像复杂度量化模型

在无人机监督的场景下，目标场景所包含的空中与地面背景的红外辐射特性变化极大、分布也可能极不均匀，易在图像上“淹没”目标并使目标与背景无法区分；而受环境电磁干扰、探测成像系统内部非均匀性及噪声，也会使成像出现模糊、失真、噪声增大等劣化情况并增加目标识别的难度。红外图像复杂度量方法可以用于描述目标识别面临的复杂场景变化，在红外成像目标识别算法性能评估、改进目标识别性能模型等方面具有重要作用。

无人机场景中需检测目标主要为小目标，对红外目标识别图像复杂度的客观评价主要基于目标局部复杂度和图像全局复杂度的特性分布差异，我们结合目标局部复杂度与全局复杂度的综合考虑，提出一种多尺度融合的面向红外目标识别的图像复杂度量化算法。

一、 面向红外目标识别的图像复杂度量化算法

红外目标识别图像复杂度量化中，除了干扰红外成像质量的因素，主要考虑红外图像中局部目标显著性和全局的目标显著性对目标识别性能和图像复杂度量度的影响。局部目标显著性是对图像中目标区域与限定邻近区域的相对度量，表征了目标与邻近区域中其他物体相区别的特征，体现了从背景中提取目标的能力。全局的目标显著性是用来表征整个图像与目标相似的程度，用来评估由于红外图像中背景存在与目标相似的特征而导致目标识别系统产生虚警的可能。

在无人机跟踪运动目标时，随着目标的不断移动，同一目标在不同场景中提取的难易程度可能相差较大。因此我们提出面向目标识别的红外图像复杂度量化，将红外图像中目标识别提取的难易程度列入复杂度量化的指标，综合考虑待识别目标的局部边缘特性、目标灰度相对其限定邻近区域的灰度强弱等因素。结合局部目标显著性和全局目标显著性特征度量，进一步改进为多尺度融合的图像复杂度量化方法，如图 1 所示：

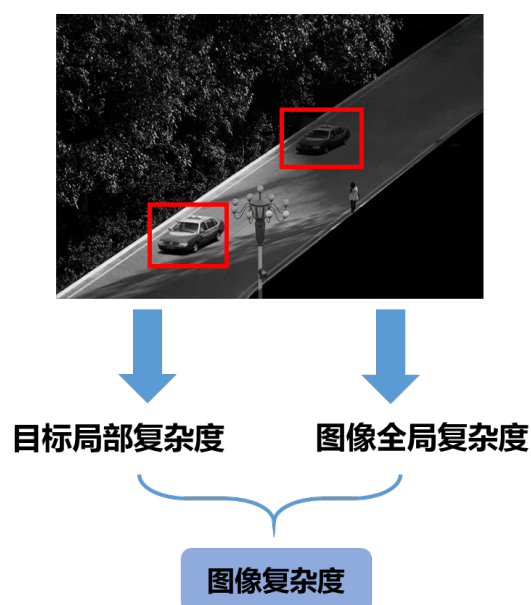


图 1 红外目标识别的图像复杂度量化

首先对红外灰度图像中目标的边缘、灰度分布等局部特征进行提取分析，并计算识别目标与邻近区域背景之间的特征差异，评估目标提取、识别难易度作为目标的局部复杂度；然后进一步计算图像的全局灰度分布、特征等因素，评估全局复杂度。最后结合两者的输出结果综合评估红外图像的复杂度，给出多尺度的图像全局复杂度量化值。

二、 算法实现

本方法针对无人机拍摄的红外图像进行复杂度量化，以图像的结构特征、噪声等因素作为全局复杂度的量化指标，以标记目标和其限定邻近区域的各项差异特征作为目标局部复杂度量化指标。考虑到无人机拍摄的图像存在目标较小的特点，主要从图像灰度分布上对图像进行处理和复杂度评估。结合局部目标和全局目标显著性特征度量，给出一种多尺度融合的图像复杂度量化方法。

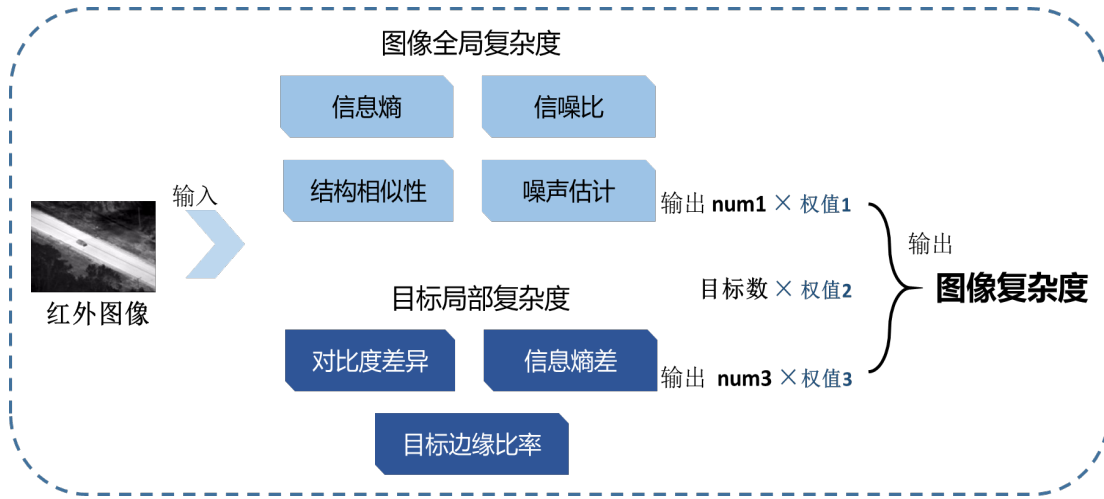


图 2 多尺度融合的红外图像复杂度量化方法

复杂度量化的基本流程如图 2 所示，输入一张红外灰度图像，从全局和局部两个方面进行复杂度量化，用实现的各个量化算法分别对图像进行量化输出，然后对各算法输出的数值进行归一化并赋权值，结合图像中的待识别目标数目，最后输出图像复杂度的值。算法的权值是依据人的视觉感官原理和目标显著性的情况来调整。下面详细介绍多尺度融合图像复杂度量化方法。

$$C = \alpha \cdot C_{global} + \beta \cdot C_{local} + \gamma \cdot N$$

我们设 C 为图像的复杂度量化值，其中 C_{global} 与 C_{local} 分别为算法对图像全局复杂度和目标局部复杂度的输出结果。 α 与 β 为对全局和目标局部量化结果所赋的权值，比例为 1:1。 N 为人工标记的图像客观目标数， γ 为目标数的权值。

（一）图像全局复杂度量化

针对图像的全局和目标局部特征的不同，我们融合了多种尺度对复杂度进行量化，图像的全局复杂度量化方法表达式如下：

$$C_{global} = \alpha_1 \cdot H + \alpha_2 \cdot PSNR + \alpha_3 \cdot SSIM + \alpha_4 \cdot \sigma_n$$

全局量化方法表达式中， H 代表图像的信息熵，图像的信息熵是一种特征的统计形式，它反映了图像中平均信息量的多少，我们使用统计方法通过对图像灰度级的出现概率进行计算，得到图像灰度的信息熵值，用以反映图像灰度域的内在复杂度情况。图像的一维信息熵可以表示图像灰度分布的聚集特征，却不能反映图像灰度分布的空间特征，为了表征这种空间特征，我们在一维熵的基础上引入能够反映灰度分布空间特征的特征量来组成图像的二维熵。选择图像的邻域灰

度均值作为灰度分布的空间特征量，与图像的像素灰度组成特征二元组，记为 (i, j) ，其中 i 表示目标像素的灰度值 $(0 \leq i \leq 255)$ ， j 表示领域像素的灰度均值 $(0 \leq j \leq 255)$ ，这里首先用 P_{ij} 表示某个灰度在该图像中出现的概率：

$$P_{ij} = \frac{f(i, j)}{N^2}$$

其中 $f(i, j)$ 为特征二元组 (i, j) 出现的频数， N 为图像的尺度。在此基础上可定义图像的二维信息熵为：

$$H = - \sum_0^{255} P_{ij} \cdot \lg \left(\frac{P_{ij}}{N} \right)$$

即可反映某像素位置上的灰度值与其周围像素的灰度分布的综合特征。

全局量化方法表达式中，PSNR（Peak Signal to Noise Ratio） 峰值信噪比是最普遍和使用最为广泛的一种图像客观评价指标，它是基于图像与变换图像之间对应像素点间的误差，即基于误差敏感的图像质量评价。这里我们将原图像的目标作为前景提取出来，将目标的背景改为体现目标全局显著性的单一颜色背景。将原图作为图像 X ，提取目标变换后的图像作为 Y 。如下图 3 所示：

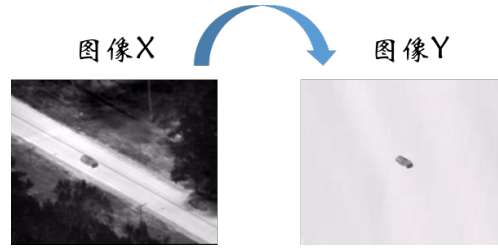


图 3 针对全局显著性的目标提取方法

为了计算 PSNR，首先求出两图像的均方误差 MSE：

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2$$

其中，MSE 表示当前图像 X 和变换图像 Y 的均方误差（Mean Square Error）， H 、 W 分别为图像的高度和宽度。然后进一步峰值信噪比 PSNR：

$$PSNR = 10 \log_{10} \frac{(2n-1)^2}{MSE}$$

其中， n 为每像素的比特数，一般取 8，即像素的灰阶数为 $2^8 = 256$ 。

全局量化方法表达式中，SSIM（structural similarity）结构相似性，同 PSNR 一样，是一种全参考的图像质量评价指标。由于人类视觉很容易从图像中抽取出结构信息，因此计算两幅图像结构信息的相似性就可以用来作为一种检测图像质量

的好坏的方法。图像质量的好坏也受到亮度信息和对比度信息的制约，因此在计算图像质量好坏时,在考虑结构信息的同时也需要考虑这两者的影响。SSIM 中使用的 X, Y 图像同上述 PSNR 的变换方式一致。SSIM 使用的计算表达式如下：

$$l(X,Y)=\frac{2\mu_X\mu_Y+C1}{\mu_X^2+\mu_Y^2+C1} \quad c(X,Y)=\frac{2\sigma_X\sigma_Y+C2}{\sigma_X^2+\sigma_Y^2+C2} \quad s(X,Y)=\frac{\sigma_{XY}+C3}{\sigma_X+\sigma_Y+C3}$$

$$SSIM(X,Y)=l(X,Y) \cdot c(X,Y) \cdot s(X,Y)$$

μ_X 、 μ_Y 分别表示图像 X 和 Y 的均值， σ_X 、 σ_Y 分别表示图像 X 和 Y 的方差， σ_{XY} 表示图像 X 和 Y 的协方差。其中，l, c, s 分别表示两图像在亮度、对比度和结构信息上的相似性。C1、C2、C3 为常数，用来增加计算结果的稳定性。

全局量化方法表达式中， σ_n 代表图像的噪声估计，采用基于算术平均滤波的噪声估计模型（Filter-Based Approach Using Arithmetic Averaging）。该类型算法是基于因为图像边缘结构具有很强的二阶差分特性，所以图像是对 Laplacian Mask 的噪声统计是敏感的，算法通过两个 Laplacian Mask 组成的 kernel 来进行卷积操作。

$$L1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad L2 = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad N = 2L2 - L1 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

然后直接通过对图像的一次卷积即可：

$$\sigma_n^2 = \frac{1}{36(W-2)(H-2)} \sum_{Image} (I(X,Y) * N)^2$$

但该算法对每个像素进行了一次乘方，进行改进后，通过

$$\sigma_n = \sqrt{\frac{\pi}{2}} \frac{1}{6(W-2)(H-2)} \sum_{Image} |I(X,Y) * N|$$

即可一次算出来，该算法整体表现良好。

（二）目标局部复杂度量化

目标局部复杂度量化是对图像中目标区域与其限定邻近区域的相对度量，表征了目标识别系统初始探测目标并标记为感兴趣区域，以及从背景中提取目标的能力。在此类方法中，我们将目标的限定邻近区域定义为以目标为中心、长宽尺寸为目标的 2 倍的矩形区域。如图 4 所示：

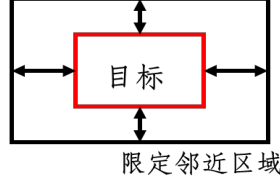


图 4 目标/局部背景

通过计算比较标记目标与其临近区域的相对差异，多尺度融合的图像的目标局部复杂度量化方式如下：

$$C_{local} = \beta_1 \cdot RSS + \beta_2 \cdot H_{local} + \beta_3 \cdot edge$$

面向目标的局部复杂度量化表达式中，RSS 为平方和根（root sum of squares, RSS）度量方法，该方法由目标与其邻近区域的对比度差异的计算方法改进而来。首先计算目标与其邻近区域的对比度差：

$$\Delta \mu = |\mu_T - \mu_B|$$

其中， μ_T 为目标的灰度均值； μ_B 为背景的灰度均值。而度量尺度没有考虑目标和背景的内部结构差异，目标和背景可能会有同样的灰度均值，但是也可能通过目标的内部结构探测到目标。为了纠正这个问题，有学者提出了 RSS 度量的改进方法。即：

$$RSS = [(\mu_T - \mu_B)^2 + \sigma_T^2]^{\frac{1}{2}}$$

其中， σ_T 代表目标的灰度标准差。当 $\sigma_T = 0$ 时，RSS 度量与灰度均值差 $\Delta \mu$ 相等。

H_{local} 为目标区域和其邻近区域的信息熵差值，信息熵的计算方式与全局复杂度量化中的信息熵计算方式一致。信息熵差 H_{local} 表征了目标与局部背景区域灰度分布的一致性，信息熵差越小，表明目标越难以从背景中分离出来，表达式如下：

$$H_{local} = |E_T - E_B|$$

其中， E_T 和 E_B 分别为目标和背景区域的熵。

edge 为边缘比率，边缘是目标的显著特征，同时也是人视觉感官系统较为敏感的特征。这里采用了 canny 作为边缘检测算法，对标记目标边缘进行提取，然后计算目标边缘灰度占限定邻近区域的比例，并将此结果设为边缘比率。如下图 5 所示：

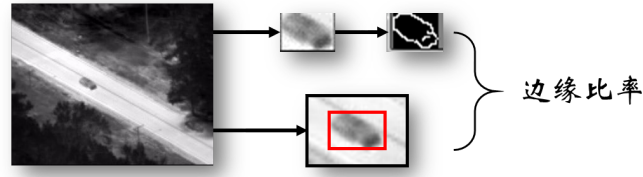


图 5 目标边缘比率计算示意图

边缘比率的计算公式为：

$$\text{edge} = \frac{P_{\text{edge}}}{M \times N}$$

其中， P_{edge} 为目标边缘的像素灰度值之和， M 与 N 为目标限定邻近区域的宽高尺寸。通过计算边缘比率，能够量化目标在其邻近区域内的显著性。

量化方法表达式中， $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_1, \beta_2, \beta_3$ 分别为各个量化尺度设置的权值。由于不同尺度输出的数值区间不同，本方法首先对各尺度输出的量化值进行归一化处理，然后结合人的视觉感官对图像的复杂度量化结果进行权值调整，修改各个权值的比例，以得到合适的结果。经过多次的修改与验证，目前 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的比值为 5:2:2:1， $\beta_1, \beta_2, \beta_3$ 比值为 1:1:1。由于图像复杂度受输入图像的影响，针对不同的图像场景与情况，所赋权值可灵活修改。

三、 图像复杂度量化结果与分析

为了更好地量化并分析红外图像的复杂度，我们将图像复杂度分为三个等级，结合量化方法的输出结果，我们认定复杂度小于 0.3 的红外图像为低等复杂场景，复杂度在 0.7 以上的红外图像为高等复杂场景，而复杂度处于 0.3—0.7 之间的红外图像为中等复杂场景。图 6 所示为测试图像的复杂度量化结果。

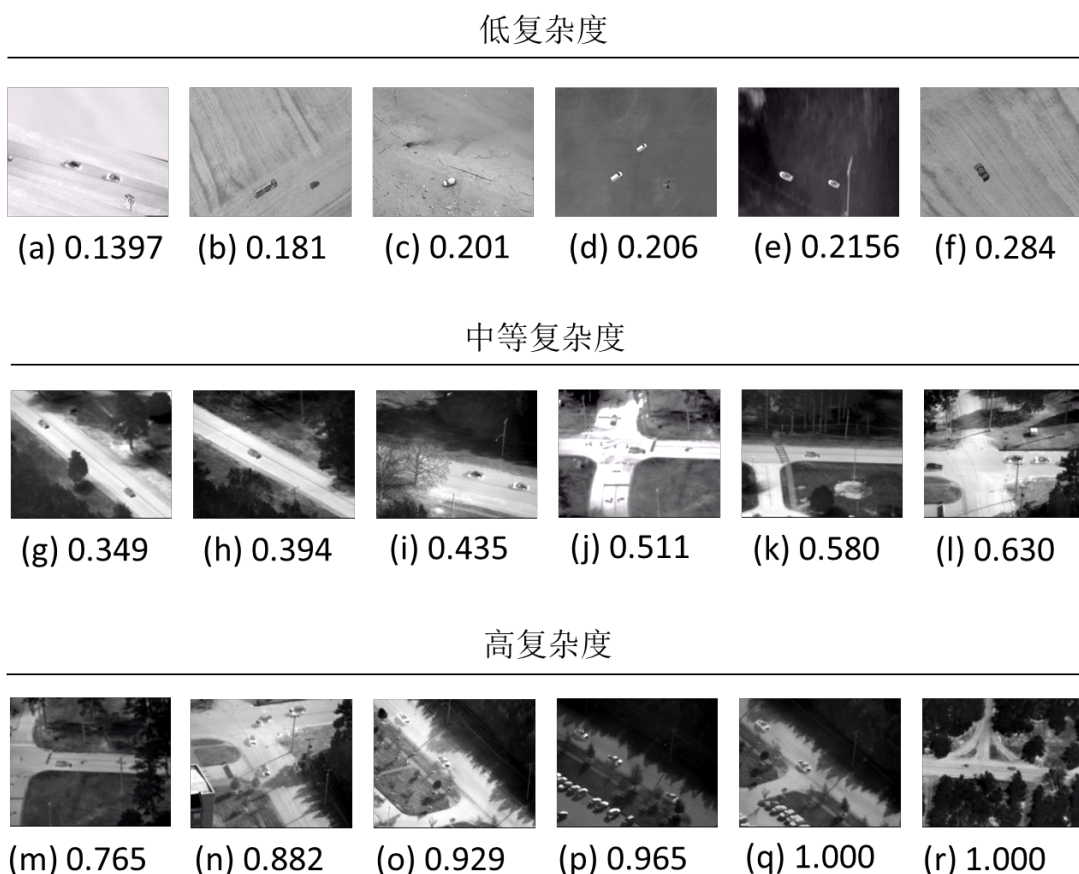


图6 图像复杂度量化结果

第一行为图像库中量化测试结果表现为低等复杂的红外目标识别图像。不难看出，对于背景单一、目标明显的红外目标识别图像，本方法量化复杂度的结果数值较低。

第二行为图像库中量化测试结果表现为中等复杂的红外目标识别图像。不难看出，在局部目标不明显、场景相对复杂的场景下，红外目标识别图像的复杂度量化数值有所提高。在无人机监督的场景下，我们主要对本方法量化结果表现为中等复杂度的红外图像进行目标进行检测识别，并在量化结果的基础上进一步比较各目标检测识别算法的优劣，以期达到良好的目标检测识别效果。

第三行为量化结果为高等复杂的红外目标识别图像，观察各图可以看出，相对中等复杂图像，高等复杂图像中明显存在影响目标识别检测结果的遮挡和噪声，且图像整体分布复杂，符合人眼视觉感官对复杂场景中识别目标难度提升因而造成识别检测难度增加的判断，且目标数较多且所占图像比例偏小也增大了目标识别检测的难度。

四、 复杂度量化方法的讨论与评价

当前我们提出的多尺度融合的红外图像复杂度量化方法仍存在不足之处，这是由于红外图像的复杂度，不仅与输入的场景有关，而且与提取信息的类型和提取方法有关。面对不同的情况，各个复杂度的量化方法侧重也有所不同。因此在选择红外图像复杂度度量方法时，应当考虑多方面因素：

- 1) 目标与背景的基本特性。红外图像复杂度度量方法均是针对图像中目标和背景提供的信息而构建，在选择方法时也要考虑特定任务的目标和背景特性。例如对于红外图像中的小目标，仅能提供像素灰度值，缺乏形状和纹理等信息，而大目标图像中包含的目标信息较为丰富，如灰度、边缘、形状和纹理等特征。
- 2) 一般要符合人眼的视觉特性。通常情况下，人眼判定两幅红外图像的复杂度不同时，同一度量尺度的计算结果也应当不同，但有时也存在与人眼视觉不同的情况。
- 3) 与目标识别算法性能的相关性要强。在度量尺度构建的过程中，应当使其与目标识别算法性能建立良好的相关性。红外图像的复杂度升高，相应目标识别算法的检测概率应当降低，虚警率升高。
- 4) 能够描述场景参数的变化。
- 5) 避免复杂的算法。
- 6) 易于软件或硬件实现。
- 7) 具有足够的鲁棒性和可靠性，对于场景变化的适应性要强，能够为一定数量规模、背景动态变化范围较大的图像提供精确和稳定的复杂度特征描述。
- 8) 具有充分的敏感性和动态范围，可以识别相似的场景内容，区分不同的场景的变化。

为了验证多尺度融合的图像复杂度量化方法符合人眼视觉感官的判断，我们也开展了相关的调研验证活动，找到多名视觉感知系统正常的用户，对图像数据进行复杂度分类判断，并在限定范围内对图像复杂度高低进行打分。考虑到不同用户之间的感知差异，对各用户所得的分数结果进行求平均值，将平均值与本方法的量化结果进行对比，当差值小于允许范围内时，我们认为本方法对该图像的复杂度量输出判定符合人眼的视觉感知判定。

图像数据集来源：

<http://vision.cse.psu.edu/data/vividEval/datasets/datasets.html>

<http://projects.asl.ethz.ch/datasets/doku.php?id=ir:iricra2014>