

结合视觉感知的调焦窗口构建

刘雪超^{1,2} 吴志勇¹ 黄德天³ 梁敏华¹

(¹ 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033
² 中国科学院大学, 北京 100049; ³ 华侨大学工学院, 福建 泉州 362021)

摘要 为使靶场光测设备在进行基于图像的自动调焦过程中建立适当的调焦窗口以降低计算量并克服背景干扰, 提出了结合视觉感知机制的调焦窗口构建方法。对图像进行金字塔分层获得多层子图; 采用 Gabor 滤波器和基于高斯差分(DoG)的模型模拟人类大脑视皮层细胞功能, 对子图层进行信息提取和特征融合, 获得基于视觉感知的特征图; 对该特征图进行阈值去噪, 以减少干扰, 并以边界扩展法建立适合于后续处理的矩形调焦窗口。实验结果证明, 对于各种调焦状态下的图像, 尤其是深度离焦状态, 都能快速、准确地针对目标构建适当的调焦窗口, 处理 720 pixel×576 pixel 的目标图像耗时小于 130 ms。该方法能够满足靶场的实际要求, 为整个调焦过程实时建立稳定、准确的调焦窗口, 具有广泛的工程应用价值。

关键词 图像处理; 调焦; 视觉感知; 调焦窗口; 特征提取

中图分类号 TN919.8 **文献标识码** A **doi**: 10.3788/CJL201441.0109001

Designing of Focus Window Combined with Visual Perception

Liu Xuechao^{1,2} Wu Zhiyong¹ Huang Detian³ Liang Minhua¹

¹ Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin 130033, China
² University of Chinese Academy of Sciences, Beijing 100049, China
³ College of Engineering, Huaqiao University, Quanzhou, Fujian 362021, China

Abstract In order to build right focus window for optical measure equipment in shooting range, a novel method combined with visual perception is proposed. It can cut down compute time and overcome disturbance of background during auto-focus procedure based on image. The image is separated into several sub-images by using pyramid structure. The function and principle of cells in human visual cortex are simulated by Gabor filter and difference of Gaussian (DoG) model. Feature maps from the sub-images are extracted and combined together to get final feature map based on visual perception. The threshold is used to denoise disturbance in the feature map and build regular rectangle focus window area by spreading boundary. After plenty of experiments, the results show that this method can build focus window rapidly with exact size and right position for the target in image. And it is available for every kinds of focus states, especially for deep-defocus state. To deal with the image of 720 pixel×576 pixel, the compute time is less than 130 ms. The proposed method can construct steady and exact focus window during the whole focus procedure, which satisfies the practical demand of shooting range and is of great and wide value for engineering application.

Key words image processing; focus; visual perception; focus window; feature extraction

OCIS codes 100.2000; 100.3008

收稿日期: 2013-07-15; **收到修改稿日期**: 2013-08-23

基金项目: 国家 863 计划(2008AA121803)、国家自然科学基金(61108066)

作者简介: 刘雪超(1985—), 女, 博士研究生, 主要从事图像处理与机器视觉等方面的研究。

E-mail: chaochao27819@126.com

导师简介: 吴志勇(1965—), 男, 研究员, 博士生导师, 主要从事光电测控设备总体技术、视频调光调焦技术和光纤通信技术等方面的研究。E-mail: wuzy@ciomp.ac.cn

1 引 言

光测设备作为现代靶场的重要装备之一,能够为战区指挥控制中心提供大量的数据信息,对指挥决策的做出具有重要的参考价值。快速、精准的调焦系统对光测设备的目标跟踪、识别和提取等起着至关重要的作用。基于图像处理的调焦系统以光测设备捕获的图像质量为依据进行调焦控制,该过程包含多个步骤:调焦窗口的建立、图像质量的评价和搜索策略的执行。调焦窗口的建立作为整个调焦过程的第一步,应针对观测场景中的目标划分出大小合适、位置准确的有效区域作为后续调焦过程的范围,以避免背景和噪声干扰,并降低运算量。目前广泛使用的对焦深度法(DFF)聚焦时通常需要采集几十帧,甚至上百帧的图像进行质量评价才能确定正焦位置。因此,构建合理的调焦窗口对提高整个光测设备的性能具有重要意义。

传统的调焦窗口通常采用固定尺寸和固定位置,如中央区域选择法、多区域选择法和黄金分割法等^[1]。但是,随着智能化技术的发展和现代靶场的客观要求,亟需一种能够根据目标位置和大小自动建立调焦窗口的方法^[2-3]。为此,文献[1,4-6]进行了相关研究,并给出实验结果。文献[1]在目标已知的情况下采集多帧图像,通过脱靶量推断目标位置建立调焦窗口;文献[4]以图像细节丰富程度建立调焦窗口,处理外界干扰小的图像效果较好;文献[5]从图像配准理论出发,以提取出的图像轮廓信息建立调焦窗口;文献[6]在人工参与下采用交互方式和自动方式建立调焦窗口。

靶场光测设备在调焦过程中,捕获的可能是正焦、轻度离焦,甚至是深度离焦的视频图像,图像的特点通常是背景简单,但存在一定程度的自然景物干扰。因此,在调焦过程中单纯地从数字图像处理的角度出发很难提取出有效的调焦窗口。本文受生物视觉系统成像过程的启发,提出一种结合视觉感知的调焦窗口构建方法。以生物视觉特性理论为基础,以机器视觉为载体,模拟人眼的观察方式,对图像进行处理,在不需对比库的情况下,无论图像正焦与否,仅需1帧就能准确划分出调焦窗口,且不易受靶场环境影响,具有很强的稳健性。

2 生物视觉感知机制

人类的视觉系统可视为一套强大的数据系统,使人在面对复杂场景时也能快速发现目标,并对其

特征做出判断^[7]。视觉机制起源于20世纪50年代末 Hubel 和 Wiesel 对视皮层细胞的研究。依据视皮层细胞的复杂程度,将视觉信息处理过程建立成从初级到高级的层次性结构。随后的感受野假说也进一步完善了该理论。

为更好地应用这种高效的生物机制,众多学者对其进行了建模,典型代表是 Serre 模型和 Itti 模型^[8-10]。Serre 模型从视觉机制的生物层面出发,为信号在细胞间的传递过程建立了详尽的模型结构。相比之下,Itti 将显著特性引入生物视觉,使建立的模型更加简单明了。本文在 Itti 模型和显著特性的基础上针对调焦系统建立了符合调焦环境和调焦要求的视觉感知模型,以获得与目标相对应的调焦窗口。

3 针对调焦窗口的视觉感知模型

靶场环境的光测设备捕获的图像通常为灰度图像,特点是背景简单、伴有自然景物、目标种类和位置不确定、距离较远,而调焦之前捕获的图像又会因各种离焦状态导致模糊。调焦窗口的建立就是要在这种条件下实时地构建一个包含目标主体的有限范围,保障后续调焦步骤的顺利进行。采用视觉感知机制建立调焦窗口能有效避免因离焦造成的影响。因此,根据生物视觉感知机制的原理,提出适用于靶场环境下的调焦窗口构建模型,如图1所示。

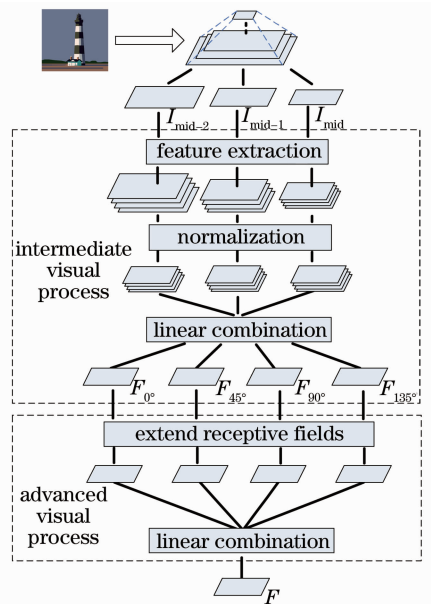


图1 结合视觉感知的调焦窗口构建模型
Fig.1 Model of focus window based on visual perception

将采集的 1 帧图像进行金字塔化分层,模拟初级视觉过程;对分层后得到的子图层进行多尺度特征提取,模拟中级视觉过程;以更大的感受野作用范围处理特征信息,模拟高级视觉过程。通过该模型能获得原图像的特征图,为调焦窗口的最终建立提供有效依据。

3.1 金字塔化图像

通过金字塔化可将对高分辨率原始图像的处理转变为对低分辨率子图的处理,从而降低算法运算量。常用的变换模型为高斯金字塔,变换过程如下:

$$I_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 \omega(m, n) I_{l-1}(2i+m, 2j+n) \quad (1)$$

$$(1 \leq l \leq P, 0 \leq i \leq R_l, 0 \leq j \leq C_l), \quad (1)$$

$$P = \min(\lceil \lg M \rceil, \lceil \lg N \rceil), \quad (2)$$

式中 I 为原始图像; R_l 、 C_l 为当前第 l 层的行、列数; ω 为高斯窗口函数,大小为 5×5 ^[11]; P 为可金字塔化的总层数,图像大小为 $M \times N$ 。金字塔化后得到低分辨率的子图层序列 $\{I_1, I_2, \dots, I_m, \dots, I_{P-1}, I_P\}$ 作为中级视觉过程的输入。其中, I_m 为分层图像序列中的中间层图像,当 P 为偶数时, $I_m = I_{P/2}$ 。

3.2 提取图像特征信息

为实现视觉机制的中级过程,对获得的低分辨率子图层进行特征提取,并将提取结果作为高级视觉过程的输入。Hubel 等^[11]于 1962 年发现人类的视皮层细胞具有方向敏感性,而 Gabor 滤波器能将图像中的方向、空间位置和频率特性同时提取出,该特点较其他特征提取方法更符合人类的视觉特性,并被广泛应用于图像处理领域^[10,12-17],其数学模型如下式所示:

$$g(x, y) = h(x', y') \cdot \exp(i2\pi f x'), \quad (3)$$

$$\begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases}, \quad (4)$$

$$h(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right], \quad (5)$$

式中 σ_x 和 σ_y 分别为 x 和 y 方向上的标准差,也是滤波器的尺度参数; θ 是方向,取 0° 、 45° 、 90° 和 135° ; f 是中心频率。

采用多尺度的 Gabor 滤波器对不同分辨率的子图层进行滤波可保证特征提取的完整性。用于特征提取的图层分辨率过高会增加计算量,分辨率过低会使特征缺失,从 3.1 节获得的图层序列中仅选取 $\{I_{m-2}, I_{m-1}, I_m\}$ 三层作为中级视觉过程的输入进行 Gabor 滤波。经过不同尺度、不同方向下的特征

提取,使 3 个不同分辨率下的子图均获得 4 个方向上的特征图。因输入的图层有限,为充分提取图像的特征信息,采用最近邻差值法按 I_m 的分辨率进行规范,再沿方向 θ 融合,得到 4 个不同方向上的特征图 F_θ 。若输入的原始图像分辨率为 $720 \text{ pixel} \times 576 \text{ pixel}$,则中级视觉过程的输出结果仅有 4 幅图,且分辨率都仅为 $45 \text{ pixel} \times 35 \text{ pixel}$,与原始 Itti 模型需获得 42 幅特征图相比,运算量极大降低。实验发现,Gabor 滤波器的尺寸过大或过小会使最终得到的特征图出现噪声过大或特征缺失的现象。因为,中级视觉过程的输入为低分辨率子图像,根据文献^[10,18]中对生物数据的分析,为图层 $\{I_{m-2}, I_{m-1}, I_m\}$ 分配尺寸为 $21 \text{ pixel} \times 21 \text{ pixel}$ 、 $15 \text{ pixel} \times 15 \text{ pixel}$ 和 $7 \text{ pixel} \times 7 \text{ pixel}$ 的滤波器,并以文献^[18]中的经验公式确定对应尺寸下的 σ_x 、 σ_y 和 f 值。通过第 5 节的实验证明,该滤波器设置比较合理。

3.3 扩大感受野

感受野的扩大可以提高对信息综合理解和处理的能力,该特点符合高级视觉过程的要求。1965 年 Hubel 通过分析猴子的大脑皮层提出了感受野假说,并得到众多学者验证^[13,19-20]。随后, Rodieck 等^[20]建立了经典的高斯差分(DoG)模型来模拟这一机制。本文采用 Itti 等^[12]改进的 DoG 模型模拟感受野扩大的过程,提高对特征图 F_θ 的理解能力,表达式为

$$F'_\theta = |F_\theta + F_\theta F_{\text{DoG}} - G_{\text{in}}|, \quad (6)$$

式中 G_{in} 是抑制因子,用于限定特征图数值范围; DoG 模型如下式所示:

$$F_{\text{DoG}}(x, y) = \frac{c_{\text{ex}}^2}{2\pi\sigma_{\text{ex}}^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_{\text{ex}}^2}\right) - \frac{c_{\text{in}}^2}{2\pi\sigma_{\text{in}}^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_{\text{in}}^2}\right). \quad (7)$$

根据本文内容和实验环境,以及已有的理论分析和验证^[8,13,19],这里将 c_{ex} 和 c_{in} 定义为增强和抑制系数,取值为 0.5 和 1.5; σ_{ex} 和 σ_{in} 定义为增强和抑制标准差,取值为 14 和 100。(6)式中的参数设置为 $G_{\text{in}} = 5$ 。

将各方向上扩大感受野后所得的特征图进行线性融合,得到最终的特征图 F ,为调焦窗口的建立提供重要的目标区域信息,可表示为

$$F = \sum F'_\theta, \quad \{\theta \in (0^\circ, 45^\circ, 90^\circ, 135^\circ)\}. \quad (8)$$

4 建立调焦窗口

4.1 特征图去噪

经过上述视觉感知模型建立的特征图能将目标区域显示出来,但仍存在一些干扰,影响机器视觉的判断。通过设定阈值去除干扰,可进一步突出目标区域。

对特征图 F 进行 $3 \text{ pixel} \times 3 \text{ pixel}$ 范围内的局部灰度累加和非 0 像素个数统计,以获得局部灰度累加均值 \overline{S}_k 作为去噪阈值,如下式所示:

$$S_k = \sum_{i=1}^H \sum_{j=1}^W \sum_{n=-1}^1 \sum_{m=-1}^1 F_k(i+n, j+m), \quad (9)$$

$$F_k(i+n, j+m) > 0, \quad (9)$$

$$\overline{S}_k = S_k / N_k, \quad (10)$$

式中,特征图 F 尺寸为 $H \times W$; k 为迭代次数; N_k 为非 0 像素个数; S_k 为统计累加后的局部灰度值;经试验证明 k 取值为 2 时效果较好。去噪过程为

$$F_{k+1}(i, j) = \begin{cases} F_k, & F_k(i, j) \geq \overline{S}_k \\ 0, & \text{else} \end{cases} \quad (11)$$

为使目标区域在特征图中更直观地体现出,也为方便后续计算,对去噪后的特征图按下式进行二值化处理,得二值化特征图 F' :

$$F'(i, j) = \begin{cases} 1, & F_{k+1}(i, j) > 0 \\ 0, & F_{k+1}(i, j) \leq 0 \end{cases} \quad (12)$$

4.2 边界扩展法建立调焦窗口

由二值化特征图 F' 中目标区域直接建立的调焦窗口边界不规则,不利于后续调焦处理。本文提出边界扩展法,依据 F' 中的目标区域建立最终规则的矩形调焦窗口。

将二值化特征图 F' 中灰度值非 0 的像素视为有效像素,有效像素的集合构成目标区域;反之,灰度值为 0 的像素就为无效像素,无效像素的集合构成非目标区域;与非目标区域相邻的目标区域像素集合构成目标区域的边界。目标区域不规则正是因为存在 45° 、 135° 、 225° 和 315° 的边界走势。为规范调焦窗口,应将具有上述走势的边界进行扩展,有效化邻近的无效像素,使扩展后的区域成为仅具有水平和垂直方向边界走势的矩形窗口,且恰好包含全部目标区域。因不规则边界的角度特性,采用“十”字模板进行边界扩展并建立调焦窗口。

该过程应考虑以下三方面因素:

1) 当无效像素的“十”字范围存在的有效像素多于 1 个时,该无效像素应有效化成为扩展点,实现边界扩展;

2) 扩展点可能会改变相邻无效像素的性质;

3) 当存在多个矩形窗口时,选择特征权重最大的窗口作为调焦窗口。

用“十”字模板对 F' 进行卷积,根据因素 1) 进行扩展判断和操作。考虑因素 2) 的情况,如果某一无效点因周围出现了一个扩展点使其满足因素 1),那么该点也应当有效化。同时,为获得规范的矩形调焦窗口,需充分考虑因素 2) 的可能性,进行多次迭代处理。每次进行扩展时记录扩展点“十”字范围内的无效像素点 P_n ,并生成预处理序列 $L = \{P_1, P_2, \dots, P_n, \dots\}$,对该序列进行因素 1) 的再判断和扩展,并刷新。该过程应迭代直至 L 为空,得到规则的矩形窗口。

当采集到的图像因深度离焦或存在过多背景导致扩展出多个规则矩形窗口时,以因素 3) 确定唯一的调焦窗口,特征权重 W_I 及其最大值为

$$W_I = \sum_{i \in I, j \in I} F(i, j) / N_I, \quad (13)$$

$$W = \max_{I=1,2,\dots} \{W_I\}, \quad (14)$$

式中 N_I 为目标区域 I 内的有效像素个数, W 为通过特征图确立的调焦窗口。将 W 按原始图像与特征图分辨率的比例进行扩展,获得原始图像的最终调焦窗口。

5 实验结果与分析

为验证本方法在靶场环境中应用的有效性和可靠性,进行了大量实验验证,实验平台为靶场环境下的某型号经纬仪,程序运行环境为 VC6.0,处理的数据源为 $720 \text{ pixel} \times 576 \text{ pixel}$ 大小的正焦、轻度离焦和深度离焦状态下具有靶场环境特点的图像,并与原始 Itti 模型相比较。因篇幅有限,仅列出四组典型实验及深度离焦下的图像对比结果如图 2~5 所示。

第一组实验如图 2 所示,靶标图像为典型的靶场图像,具有单一小目标、噪声小和存在少量背景干扰等特点。图 2(a)~(d) 显示,离焦导致图像模糊,深度离焦更导致背景与目标区别不明显,影响调焦窗口的建立;图 2(c) 中,原始 Itti 模型建立的调焦窗口边界不规则;图 2(d) 显示,本文方法对深度离焦图像建立的调焦窗口与正焦和轻度离焦状态下的相比略有不同,但位置和尺寸仍合理,且边界规则,有利于后续调焦步骤进行;图 2(e)~(g) 显示,深度离焦使图像特征不明显,导致其特征图存在干扰,经阈值去噪和边界扩展滤除干扰,得到规则的矩形调焦窗口。

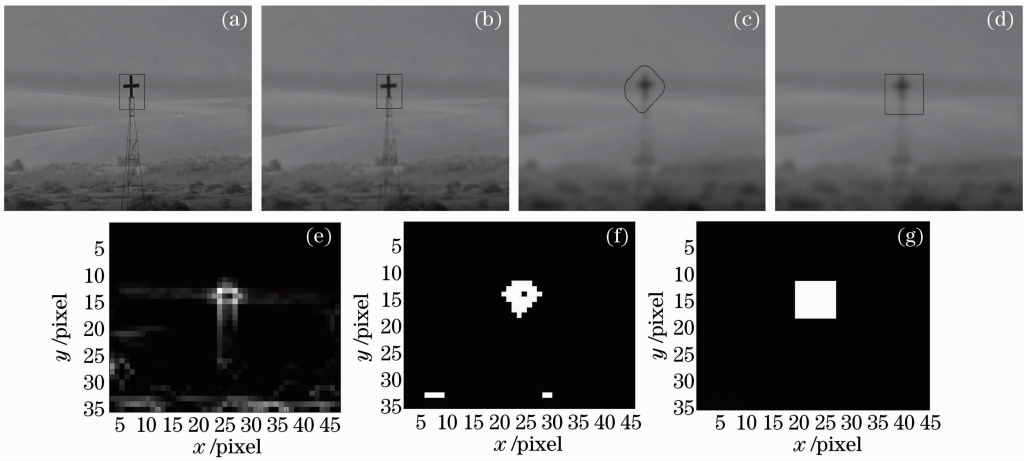


图 2 靶标图像构建调焦窗口。(a)、(b)、(d)本文方法对正焦、轻度离焦和深度离焦图像建立调焦窗口；(c)原始 Itti 模型对重度离焦图像建立调焦窗口；(e)~(g)本文方法在深度离焦下提取的特整图、去噪后二值化特征图和边界扩展后获得的调焦窗口

Fig. 2 Focus window for target picture. (a)、(b)、(d) Focus windows creation under focus state, light-defocus state and deep-defocus state by proposed method; (c) focus window creation by original Itti model for deep-defocus state; (e)~(g) feature map, binary feature image with denoising and rectangle region after being extended for deep-defocus state by proposed method

第二组实验如图 3 所示,建筑物图像是靶场环境中的常见类型之一,具有目标较大、背景简单和存在其他小目标干扰等特点。图 3(c)显示,原始 Itti 模型在处理深度离焦下的较大目标时,仅能针对目

标中特征最明显的部分建立窗口,导致信息缺失;而图 3(d)中的调焦窗口基本包含了整个目标主体,说明本文方法所建立的区域连续性好,能对大目标建立调焦窗口。

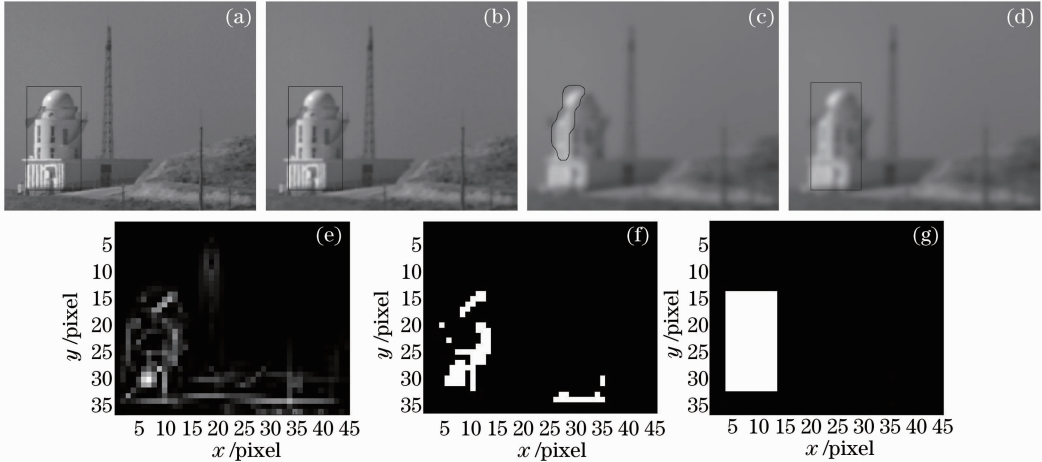


图 3 建筑图像构建调焦窗口。(a)、(b)、(d) 本文方法对正焦、轻度离焦和深度离焦图像建立调焦窗口；(c)原始 Itti 模型对重度离焦图像建立调焦窗口；(e)~(g)本文方法在深度离焦下提取的特整图、去噪后二值化特征图和边界扩展后获得的调焦窗口

Fig. 3 Focus window for building picture. (a)、(b)、(d) Focus windows creation under focus state, light-defocus state and deep-defocus state by proposed method; (c) focus window creation by original Itti model for deep-defocus state; (e)~(g) feature map, binary feature image with denoising and rectangle region after being extended for deep-defocus state by proposed method

第三组实验如图 4 所示,飞机图像的特点是目标较小且单一、背景简单,但存在大量噪声干扰,这也是光测设备使用中的常见问题。图 4(a)~(d)显示,正焦和轻度离焦图像中的目标相对明显,但存在

大量噪声;深度离焦下背景噪声减弱,但也影响了图像质量。该组实验说明,本方法不仅能克服调焦窗口建立过程中离焦造成的影响,也能克服大量噪声的干扰。

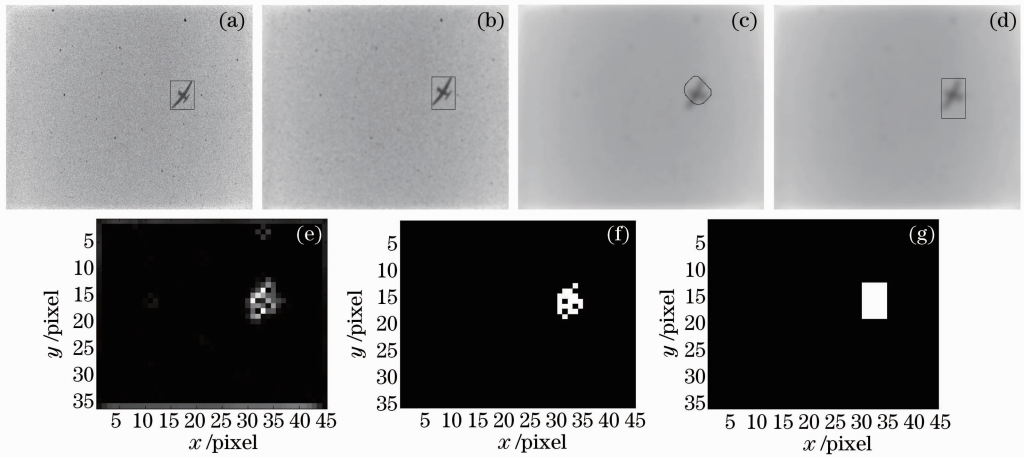


图 4 飞机图像构建调焦窗口。(a)、(b)、(d)本文方法对正焦、轻度离焦和深度离焦图像建立调焦窗口；(c)原始 Itti 模型对重度离焦图像建立调焦窗口；(e)~(g)本文方法在深度离焦下提取的特整图、去噪后二值化特征图和边界扩展后获得的调焦窗口

Fig.4 Focus window for plane picture. (a)、(b)、(d) Focus windows creation under focus state, light-defocus state and deep-defocus state by proposed method; (c) focus window creation by original Itti model for deep-defocus state; (e)~(g) feature map, binary feature image with denoising and rectangle region after being extended for deep-defocus state by proposed method

第四组实验如图 5 所示,路灯图像的特点是背景复杂,目标大小适中、细节丰富。图 5(e)~(g)显示,复杂纹理和丰富细节会增加特征图中的干扰,经去噪和边界扩展后选取特征权重最大的区域为调焦窗口,保证唯一性;图 5(c)显示,原始 Itti 模型选取

的目标内容严重缺失;图 5(a)~(d)中各调焦状态下建立的调焦窗口基本一致,且均包含目标主体。该组实验虽然不是典型的靶场实验,但说明本方法能克服复杂背景对调焦窗口建立的影响,稳健性强,也进一步保证了其在靶场环境下使用的有效性。

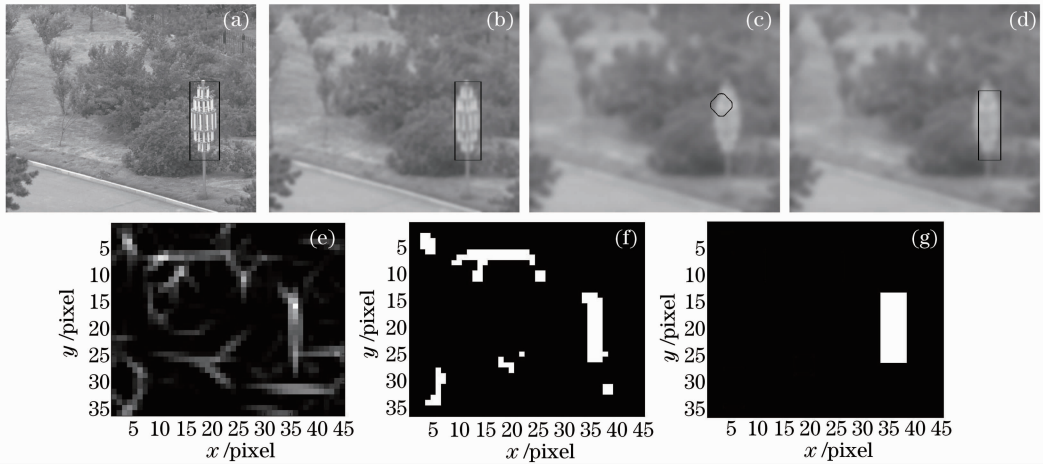


图 5 路灯图像构建调焦窗口。(a)、(b)、(d)本文方法对正焦、轻度离焦和深度离焦图像建立调焦窗口；(c)原始 Itti 模型对重度离焦图像建立调焦窗口；(e)~(g)本文方法在深度离焦下提取的特整图、去噪后二值化特征图和边界扩展后获得的调焦窗口

Fig.5 Focus window for lamp picture. (a)、(b)、(d) Focus windows creation under focus state, light-defocus state and deep-defocus state by proposed method; (c) focus window creation by original Itti model for deep-defocus state; (e)~(g) feature map, binary feature image with denoising and rectangle region after being extended for deep-defocus state by proposed method

图 2~5 中的目标均不在中央区域且位置随机、含有背景和噪声,如果采用传统的固定调焦窗口,不仅不能准确找到目标位置,更易受到各种干扰,影响调焦精度,甚至导致调焦失败;同时,实验也显示出

原始 Itti 模型具有局限性。

为建立统计性数据,将内容相同、调焦状态不同的 3 幅图像归为一组,选取实验中 80 组(共 240 幅)图像的处理结果进行统计。其中,包括 40 组靶场环

境图像和 40 组普通环境图像；每种环境图像又由 20 组观测目标较大图像和 20 组观测目标较小图像构成。按观测目标大小和图像环境分类得评价结果分别如表 1 和表 2 所示，整体评价结果如表 3 所示。窗口建立的准确性由人眼观察决定，在位置准确的前提下，当窗口覆盖范围不小于目标主体 3/4、且不

大于 2 倍时，认为该窗口准确建立。从表 1 和表 2 可见，原始 Itti 模型与本文算法在处理正焦状态下的小目标图像时效果均较好；随着目标变大、图像离焦，本文算法的准确率远远大于 Itti 模型。表 3 从整体统计的角度显示出本文方法对各调焦状态图像的处理结果优于 Itti 模型。

表 1 观测目标大小不同时本文方法与原始 Itti 模型比较

Table 1 Comparison of proposed method and original Itti model for different sizes of targets

Images		Original Itti model		Proposed method	
		Exact window	Ratio /%	Exact window	Ratio /%
Big target (40 groups)	Focus	12	30	37	92.5
	Light-defocus	10	25	37	92.5
	Deep-defocus	8	20	36	90
Small target (40 groups)	Focus	39	97.5	39	97.5
	Light-defocus	34	85	38	95
	Deep-defocus	29	72.5	35	87.5

表 2 图像环境不同时本文方法与原始 Itti 模型比较

Table 2 Comparison of proposed method and original Itti model for different scenes

Images		Original Itti model		Proposed method	
		Exact window	Ratio /%	Exact window	Ratio /%
Target environment (40 groups)	Focus	26	65	39	97.5
	Light-defocus	23	57.5	39	97.5
	Deep-defocus	20	50	37	92.5
Common environment (40 groups)	Focus	25	62.5	37	92.5
	Light-defocus	21	52.5	36	90
	Deep-defocus	17	42.5	34	85

表 3 本文方法与原始 Itti 模型处理结果整体对比

Table 3 Integrated comparison of proposed method and original Itti model

Images	Original Itti model		Proposed method	
	Exact window	Ratio /%	Exact window	Ratio /%
Focus	51	63.7	76	95
Light-defocus	44	55	75	93.7
Deep-defocus (80 groups)	37	46.2	71	88.8

实验中两种算法的时间对比统计如图 6 所示，可见本方法的计算耗时少于 130 ms，远低于原始

Itti 模型。

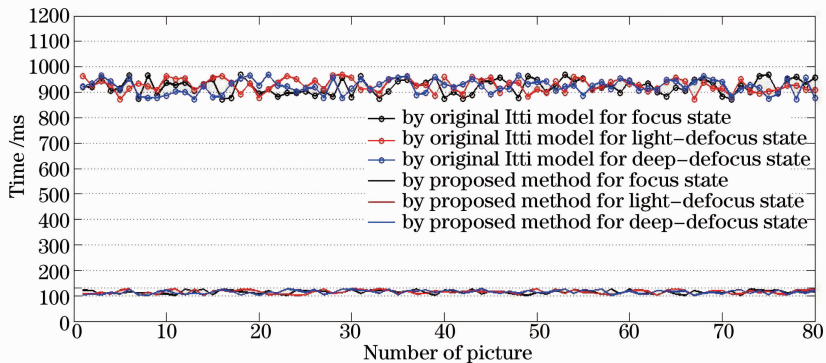


图 6 本方法与原始 Itti 模型耗时对比

Fig. 6 Cost by proposed method and original Itti model

上述实验分析和统计结果显示,本方法能有效克服噪声和背景的影响,对于大目标和小目标都能在正焦、轻度离焦和深度离焦状态下建立位置准确、大小合适的调焦窗口,且耗时少,满足靶场环境要求。

6 结 论

针对靶场环境下基于图像的自动调焦,本文提出了一种新的调焦窗口构建方法。该方法引入了视觉感知机制,从人眼成像过程出发,以 Itti 模型为基础,通过金字塔分层、多尺度 Gabor 滤波器和基于 DoG 的感受野模型模拟视觉感知的过程,对图像进行特征提取,并设定阈值对特征图去噪,以滤除背景、噪声和模糊等因素产生的干扰,最后以边界扩展获得规则的矩形调焦窗口。通过在某型号经纬仪上的大量实验证明,该方法能有效克服靶场环境对调焦窗口的影响,对于各种调焦状态,尤其是深度离焦状态下的图像,都能构建出位置准确、大小合适的调焦窗口,且对 $720 \text{ pixel} \times 576 \text{ pixel}$ 的图像处理时间少于 130 ms ,满足实时性要求,有利于调焦过程的顺利进行,具有重要的工程应用价值。但是,实验中也发现,本方法在处理距离过近、运动速度过快的目标时会出现调焦窗口滞后的现象,因此,在该方面本算法有待进一步研究。

参 考 文 献

- Liang Minhua, Wu Zhiyong, Chen Tao. Auto-focusing adjustment of theodolites by largest the gradient method[J]. Optics and Precision Engineering, 2009, 17(12): 3016-3021.
梁敏华,吴志勇,陈涛.采用最大灰度梯度法实现经纬仪自动调焦控制[J].光学精密工程,2009,17(12):3016-3021.
- Zuo Baojun, Liang Yanfeng, Fan Zhigang. Detection system of optical athermal design effect based on image processing[J]. Infrared and Laser Engineering, 2012, 41(6): 1600-1604.
左保军,梁炎锋,范志刚.基于图像处理的光学无热化设计效果检测系统[J].红外与激光工程,2012,41(6):1600-1604.
- Chen She, Meng Yonghong, Jing Gang. Study on the autofocus evaluation function in ellipsometric imaging system[J]. Acta Optica Sinica, 2005, 25(7): 923-928.
陈涉,孟永宏,靳刚.椭圆偏光显微成像系统聚焦评价函数的研究[J].光学学报,2005,25(7):923-928.
- Hu Tao, Chen Shizhe, Liu Guodong, et al.. Algorithm of selecting the optimal focusing region[J]. Optical Technique, 2005, 32(6): 851-854.
胡涛,陈世哲,刘国栋,等.图像法自动调焦的最佳调焦区域选取算法[J].光学技术,2005,32(6):851-854.
- Lü Hengyi, Liu Yang, Xue Xucheng, et al.. Intelligent focusing method for the remote sensing cameras[J]. Infrared and Laser Engineering, 2012, 41(5): 1262-1265.
吕恒毅,刘杨,薛旭成,等.遥感相机的智能调焦方法[J].红外与激光工程,2012,41(5):1262-1265.
- Yuan Yukai, Jia Wei, Pei Yuntian. Automatic selection of

- evaluation patch in autofocusing[J]. Science Technology and Engineering, 2006, 6(23): 4763-4765.
- 袁育凯,贾伟,裴云天.自动调焦中的评价区域自动选取[J].科学技术与工程,2006,6(23):4763-4765.
- Zhang Xuewu, Ding Yanqiong, Yan Ping. Vision inspection of metal surface defects based on infrared imaging[J]. Acta Optica Sinica, 2011, 31(3): 0312004.
张学武,丁燕琼,闫萍.一种基于红外成像的强反射金属表面缺陷视觉检测方法[J].光学学报,2011,31(3):0312004.
- Jiang Mei. Target Recognition Based on Biological Visual Perception[D]. Nanjing: Nanjing Technological University, 2011.
江梅.基于生物视觉感知模型的目标识别[D].南京:南京理工大学,2011.
- Hu Dekun. The Research on Technology of Image Understanding Based on Biological Visual Perception[D]. Chengdu: University of Electronic Science and Technology of China, 2012.
胡德昆.基于生物视觉感知机制的图像理解技术研究[D].成都:电子科技大学,2012.
- Li Zuojin. Visual Cognition Based Recognition of Objects in Natural Images[D]. Chongqing: Chongqing University, 2010.
李作进.基于视觉认知的自然图像目标识别研究[D].重庆:重庆大学,2010.
- D H Hubel, T N Wiesel. Receptive fields and functional architecture of the monkey striate cortex[J]. J Physiol, 1965, 195(1): 215-243.
- J G Dangman. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters[J]. J Opt Soc Am A, 1985, 2(7): 1160-1169.
- L Itti, C Koch. Feature combination strategies for saliency-based visual attention systems[J]. J Electron Imaging, 2001, 10(1): 161-169.
- Zhang Libao, Wang Pengfei. Fast detection of regions of interest in high resolution remote sensing image[J]. Chinese J Lasers, 2012, 39(7): 0714001.
张立保,王鹏飞.高分辨率遥感影像感兴趣区域快速检测[J].中国激光,2012,39(7):0714001.
- Zhang Libao. Fast detection of visual saliency regions in remote sensing image based on region growing[J]. Chinese J Lasers, 2012, 39(11): 1114001.
张立保.基于区域增长的遥感影像视觉显著目标快速检测[J].中国激光,2012,39(11):1114001.
- Yuan Weiqi, Feng Suyue. Simulation system of improved non-contact on-line palmprint recognition[J]. Acta Optica Sinica, 2011, 31(7): 0712003.
苑玮琦,冯素月.改进的非接触式在线掌纹识别模拟系统[J].光学学报,2011,31(7):0712003.
- Wang Xiang, Ding Yong. Full reference image quality assessment based on Gabor filter[J]. Journal of Zhejiang University, 2013, 47(3): 422-430.
王翔,丁勇.基于Gabor滤波器的全参考图像质量评价方法[J].浙江大学学报,2013,47(3):422-430.
- T Serre, M Riesenhuber. Realistic Modeling of Simple and Complex Cell Tuning in the HMAX Model and Implications for Invariant Object Recognition in Cortex[D]. MIT: Cambridge, CSAIL, 2004.
- Orin S Packer, Dennis M Dacey. Synergistic center-surround receptive field model of monkey H1 horizontal cells[J]. J Vision, 2005, 5(11): 1038-1054.
- R W Rodieck, J Stone. Analysis of receptive fields of cat retinal ganglion cells[J]. J Neurophysiology, 1965, 28(5): 833-849.

栏目编辑:何卓铭