

# 红外热成像信号处理技术的研究进展\*

曹彦鹏<sup>1,2</sup>, 许宝杯<sup>1,2</sup>, 何泽威<sup>1,2</sup>, 官大衍<sup>1,2</sup>

(1. 浙江大学流体动力与机电系统国家重点实验室 杭州, 310027)

(2. 浙江大学浙江省先进制造技术重点研究实验室 杭州, 310027)

**摘要** 红外热成像技术通过二维图像的形式准确反映物体表面温度分布情况,广泛应用于医疗诊断、故障检测、能耗监控及智能驾驶等众多工业领域。针对制约红外热成像技术发展的 3 个主要瓶颈问题,分别阐述相关研究领域内最新的科研进展。首先,介绍红外信号固定模式噪声建模及降噪方法,解决信号采集系统中关键部件焦平面阵列传感器存在的严重噪声干扰;其次,介绍红外信号特征重建理论的研究,克服红外焦平面阵列探测器制作工艺复杂、成品率低及价格昂贵等制约了红外热成像技术广泛工业应用的问题;最后,探索多源信息融合技术,减少二维图像采集过程中降维映射导致的信息损失,提高基于多光谱信息进行目标检测系统的性能。红外热成像信号处理技术的最新科研进展,为新一代高成像精度、高分辨率、低制作成本的三维红外成像仪器的研发提供新理论与技术。

**关键词** 红外热成像;噪声建模;特征重建;信息融合

**中图分类号** TP752.1;TH17

## 1 问题的引出

自然界中任何物体的温度都高于绝对零度,会不断向周围空间发射红外电磁能量<sup>[1]</sup>。波长在 0.75~1 000  $\mu\text{m}$  范围内的红外光位于电磁波谱中微波光和可见光之间。如图 1 所示,这一光谱范围可进一步细分为短波红外(0.76~1.5  $\mu\text{m}$ )、中波红外(1.5~5.6  $\mu\text{m}$ )和长波红外(7.8~13.5  $\mu\text{m}$ )。常用的可见光成像系统无法检测红外波长范围内的能量信号,因此无法直接测量物体的温度信息。如图 2 所示,红外热像系统采集物体表面的红外辐射能量,经光学系统会聚,把接收的红外光波映射至二维的红外焦平面探测器(focal plane arrays,简称 FPA)上。焦平面探测器中的单个像元中的吸收层接收到红外辐射能量后产生温度变化,进而导致非晶硅热敏电阻的阻值变化,最后,互补式金属氧化物半导体(complementary metal-oxide-semiconductor,简称 CMOS)电路将热敏电阻阻值的变化转变为差分电流并进行积分放大,得到红外热焦平面阵列中单个像元的输出值。通过以上红外热成像技术,可以探测物体本身温度导致的红外热辐射,产生与景物热辐射分布相对应的红外热成像图像。

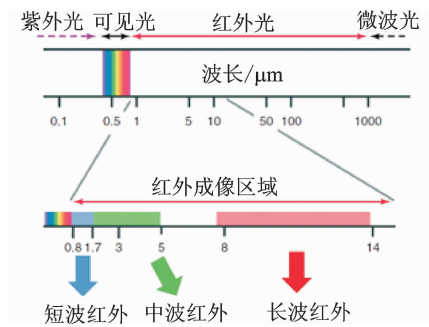


图 1 红外电磁波频率分布

Fig. 1 Overview of thermal infrared electromagnetic waves

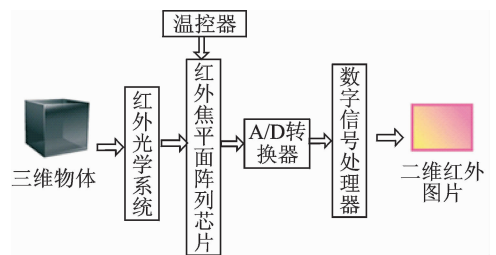


图 2 红外信号采集系统

Fig. 2 Overview of a thermal infrared imaging system

红外焦平面探测器是红外热成像系统的核心部件,根据其成像原理和工作温度分为制冷和非制冷

\* 国家自然科学基金资助项目(51605428,51575486,U1664264)

收稿日期:2017-12-01; 修回日期:2017-12-30

两类。制冷红外焦平面探测器基于光子探测原理进行红外辐射成像,具有极高的测温灵敏性,能准确地测量物体的绝对温度,且红外探测距离远。然而,制冷红外探测器的工作性能受环境温度变化影响严重。环境温度升高会导致探测器材料固有的热辐射能耗迅速增强,暗电流和噪声的增大将严重降低探测器的性能,甚至无法正常工作成像。因此制冷探测器的正常工作温度在 200 K 温度以下,通常需要用液氮制冷,导致其制作和维护成本昂贵,目前主要应用于高端军事装备。非制冷红外焦平面探测器能够在室温状态下工作,具有启动快、功耗低、体积小、重量轻、寿命长及成本低等诸多优点<sup>[2]</sup>。目前,现有的非制冷红外焦平面探测器制作工艺可以精确捕捉微小的物体表面温度变化,例如基于氧化钒(vanadium oxide,简称 VOx)热敏材料的非制冷红外焦平面探测器的单个成像元在 25 °C 室温环境下可以测量到±0.05 °C 的温度变化。虽然非制冷红外焦平面探测器在测温灵敏度上与制冷器件尚有一定差距,但针对许多工业检测应用已经足够,且性价比高,因此具有更加广阔的市场应用前景。

红外热成像技术可以将光谱中不可见的红外辐射信号转化成为可见的二维图像,实现无光环境下的优秀成像,也可对物体表面温度进行非接触式的准确测量。红外热成像技术被广泛应用于军事和民用领域,如生物医疗、故障诊断、防火消防、智能驾驶及夜视安防等<sup>[3]</sup>。然而红外热成像技术的发展也面临一些主要瓶颈制约问题,包括:a. 信号采集系统中焦平面阵列传感器存在固定模式噪声干扰,严重影响红外图像信号的成像质量;b. 红外探测器焦平面传感器制作工业复杂、成品率低、价格昂贵,制约了红外热成像技术的工业广泛应用;c. 二维图像采集过程中降维映射导致信息损失,影响了基于红外温度信息进行设备状态监测系统的性能。针对以上 3 个主要问题,文中分别阐述在红外信号固定模式噪声建模和除噪、红外信号特征重建、多源信息融合 3 个红外信号处理研究领域的最新科研进展。

## 2 红外热成像技术基本原理

在热辐射理论,黑体(black body)吸收所有入射放射能量,并根据普朗克定律产生如下辐射能量

$$M_{\lambda} = \frac{c_1}{\lambda^5} \frac{1}{\exp\left(\frac{c_2}{\lambda T}\right) - 1} \quad (1)$$

其中: $\lambda$  为辐射的波长( $\mu\text{m}$ ); $M_{\lambda}$  为绝对黑体的光谱

辐射出射度( $\text{W} \cdot \text{cm}^2 / \mu\text{m}$ ); $T$  为绝对温度(K); $c_1$  和  $c_2$  分别为第一和第二辐射常量。

在全部波长范围内对普朗克公式积分,得到从黑体单位面积辐射至半球空间的总辐射功率,即总辐射出射度的表达式,通常称为斯蒂芬-玻尔兹曼定律

$$M = \int_0^{\infty} M_{\lambda} d\lambda = \sigma T^4 \quad (2)$$

其中: $M$  表示黑体的总辐射出射度( $\text{W} \cdot \text{cm}^2$ ); $\sigma$  为斯蒂芬-玻尔兹曼常数( $5.676 \times 10^{-8} \text{W} \cdot \text{m}^2 \cdot \text{K}^{-4}$ )。

斯蒂芬-玻尔兹曼定律表明了黑体辐射功率和绝对温度之间的线性耦合关系,它是基于物体红外辐射信号采集实现物体温度测量的主要理论依据。根据维恩位移定律,黑体辐射光谱的极大值对应的波长为

$$\lambda_{\max} = b/T \quad (3)$$

其中: $b = 2\,897.8 \mu\text{m} \cdot \text{K}$ 。

维恩位移定律表明,光谱辐射出射度的峰值波长与绝对温度成反比,物体温度愈高,对应辐射的峰值波长越短。根据式(3),在 200~3 800 K 度范围内物体目标的光谱出射度峰值波长均落在红外波谱中。因此,可以利用红外热成像技术实现大温度范围内物体表面温度的非接触式测量。

## 3 红外热成像系统噪声形态建模及补偿方法

红外焦平面探测器作为红外热成像系统的核心部件,需要有极高的温度响应率和灵敏度,因此要求探测器像元微桥具有良好的热绝缘性。同时为保证红外热成像的高频信号输出,需使像元的热容尽量小以保证足够小的热时间常数。图 3 为焦平面中单个探测元的结构示意图<sup>[4]</sup>。最新红外像元制作工艺利用细长的微悬臂梁支撑热敏材料以提高绝热性能。制作轻、薄的桥面结构以减小热质量,并在衬底制作反射层,与桥面之间形成谐振腔,提高红外吸收效率。利用悬臂梁的两端将像元微桥与衬底内的 CMOS 读出电路连接。CMOS 电路将热敏电阻阻值变化转变为差分电流并进行积分放大,采样后得到红外热图像中单个像元的输出值。

红外焦平面探测器的制作工艺复杂,难以保证每个单独探测元的温度响应特性一致。此外,非制冷红外焦平面探测器单个探测元的温度-输出响应曲线会随着相机工作温度的变化发生迁移。如图 4 所示,采集的红外原始图像(raw image)中包含明显的固定模式噪声(fixed-pattern noise,简称 FPN),需要对其进行非均匀性矫正(non-uniformity cor-

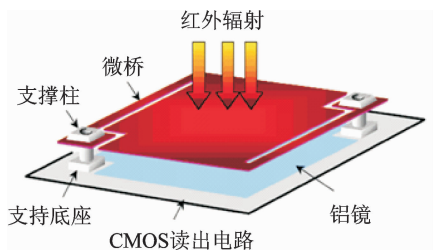


图3 红外焦平面探测器像元结构示意图

Fig. 3 Diagram of infrared FPA detector pixel structure

rection),降低噪声信号的干扰,获得真实反映物体温度场信息的红外图像。

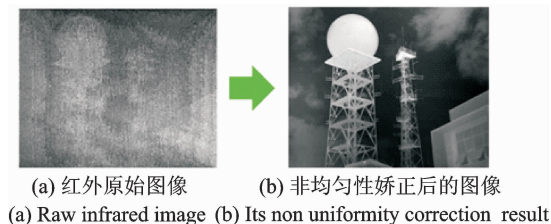


图4 红外图像除噪

Fig. 4 Denoising of infrared image

利用红外辐射源对焦平面探测元阵列进行温度响应曲线标定是最常用的非均匀性矫正方法之一。在文献[5]中,通过采集标定源的红外图像,利用“两点矫正法”或“相机挡板矫正法”计算单个成像元的非均匀性误差的校正参数(如增益参数和偏移参数),并通过线性矫正模型应用于原始图像中的对应成像元消除阵列空间非均匀性导致的信号误差。然而,基于辐射标定的非均匀性矫正方法需要使用外置热源,严重影响了红外热成像系统的尺寸、重量、功率及制造成本。此外,在进行定期校准标定的过程中,红外信息的采集会被迫中断数秒钟,降低了红外热成像仪器的动态实时性,无法捕捉场景中快速变化的目标信息。

为了克服上述缺陷,在红外信号研究领域已经提出了多个基于红外视频场景分析技术的无基准红外像元非均匀性矫正方法。Torres等<sup>[6]</sup>提出了一种基于场景的红外热成像阵列非均匀性自适应校正方法。利用离散时间域内的高斯-马尔可夫(Gauss - Marko)建模方法,准确获得红外热成像阵列传感器随环境温度变化导致固定模式噪声中低频和高频漂移信号的不同形态规律。通过卡尔曼滤波器(Kalman filter)对阵列像元的输出信号进行时间域滤波分析,计算出探测器像元非均匀性矫正所需的增益和偏差参数。Vera等<sup>[7]</sup>提出了一种基于场景辐照度变化估值最小化的红外阵列非均匀性矫

正方法。利用交替最小化的策略对定义的各向同性总变差损失函数进行优化,以更快的收敛速度准确计算焦平面阵列中各个探测器的非均匀性矫正参数。然而,基于场景分析的非均匀性矫正方法通常需要缓存并处理一定数目的红外图像帧,才能完成非均匀性矫正参数的稳定计算,因此难以实现基于现场可编程门阵列(field programable gate array,简称FPGA)等信号处理硬件的实时计算。此外,基于场景分析的矫正方法需要其处理的图像序列中包含足够的场景信息变化,否则前帧图像中的物体会作为伪影信息叠加至当前帧。如图5所示,此类伪影信息的产生将严重影响红外图像信息的真实性<sup>[8]</sup>。

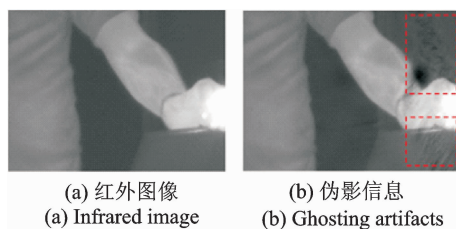


图5 基于场景分析的红外图像非均匀性矫正过程中引入的伪影信息

Fig. 5 Ghosting artifacts generated during scene-based non uniformity correction

由于以上基于热源标定和场景分析的非均匀性矫正方法的性能缺陷,红外信号处理领域的研究焦点已逐渐转移至基于单帧图像处理的无基准红外焦平面阵列矫正方法的研发。Tendero等<sup>[9]</sup>提出了一种基于中位直方图均衡处理技术的单帧红外图像非均匀性矫正方法。如图6所示,通过对像元列中不同输出信号的出现频率进行概率统计,可以在保持红外细节信息的同时有效消除红外图像中的条状固定模式噪声。如图7所示,笔者设计并搭建固定模式噪声信号模拟实验平台,分析不同辐射强度输入在红外焦平面阵列上导致固定模式噪声的形态特征<sup>[10]</sup>,发现红外热成像系统中的条状固定模式噪声与辐射输入呈局部线性耦合关系。

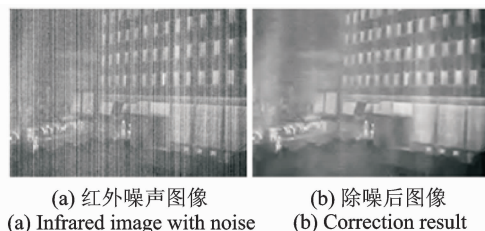


图6 进行条状固定模式噪声矫正前后的红外图

Fig. 6 Infrared images before and after non uniformity correction



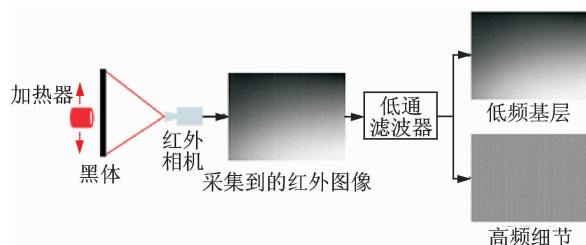


图 7 固定模式噪声信号模拟实验平台示意图

Fig. 7 Diagram of experimental setup for investigating characteristics of FPN

在对干扰噪声进行形态建模的基础上,笔者提出基于一维引导滤波器的单帧红外图像非均匀性矫正方法,将高频的纹理信息和干扰噪声进行准确分离,并可以通过 FPGA 可编程信号处理器进行实时图像矫正处理<sup>[11]</sup>。最近,科研人员开始尝试将深度学习领域中的科研成果应用于解决红外领域中的除噪问题。Kuang 等<sup>[12]</sup>在不同场景采集无噪声红外图像,并通过条状噪声形态模型,模拟产生有固定模式噪声的输入图像,通过搭建双层卷积神经网络,以监督学习的方式计算最优的除噪模型参数,并通过图形处理器(graphics processing unit,简称 GPU)并行编程实现视频流的实时处理。然而,基于深度网络学习的红外图像除噪研究尚处在起步阶段。文献<sup>[12]</sup>中提出的条状除噪方法仅通过简单的线性噪声模型产生训练样本,因此训练的深度网络模型只能在模拟产生的含噪图像上取得了较好的除噪效果,无法应用于包含复杂噪声形态变化的真实红外图像上。如何实现训练深度网络模型所需大规模实验数据的快速采集,是将深度学习技术应用于红外信号除噪领域的关键科学问题之一。

#### 4 红外信号特征提取及重建方法

高像素红外焦平面阵列传感器的制作工艺复杂、成品率低,导致其制作成本昂贵,严重制约了红外热成像技术的广泛工业应用。如何有效对低分辨率红外图片中的细节信息进行提取、分析和拟合,通过特征重建的方式掌握低像素红外图像与高像素红外图像的复杂映射关系,是目前红外信号处理领域的热点研究方向。这种对低分辨率图像进行增强从而获得其高分辨率版本的处理技术,在图像处理领域中被称作超分辨率。目前,大部分的超分辨率方案都是针对可见光频谱中的彩色或者灰度图像研发设计的,而适用于红外图像的超分辨率技术的研究尚处于探索阶段。

超分辨率技术通常可以分为基于数据差插值、基于特征组合和基于样本学习 3 大类。传统的双三次插值方法计算简单,运行速度较快,但是得到的重建图像比较模糊,并且会引入伪影。基于特征组合的超像素方法通过引入先验知识(如梯度一致、非局部均值等)来约束超分辨率重构病态问题的解集空间,获得较好的重建效果。近些年来,基于大数据学习的超像素方法发展迅速,重建性能远超基于数据差插值和基于特征组合的超像素方法。其基本原理是对大量匹配的低分辨率和高分辨率图像的对应特征变化规律进行学习拟合,从而得到低分辨率至高分辨率图像的映射函数。

稀疏编码是一种具有代表性的机器学习方法<sup>[13]</sup>。其基本假设是低分辨率和高分辨率图像具有相似的稀疏系数。稀疏编码法通过对图像进行编码分析,分别得到低分辨率和高分辨率图像对应的特征字典。当输入一幅低分辨率图像时,计算其基于低分辨率字典中的稀疏表示系数。然后,使用相同的系数与高分辨率字典重建得到其对应的高分辨率图像。类似的机器学习算法还包括领域嵌入法<sup>[14]</sup>,该方法认为低分辨率和高分辨图像在低维非线性流场上具有相似的几何结构。

利用深度学习的方法,Dong 等<sup>[15]</sup>提出了基于卷积神经网络的图像超分辨算法(super-resolution convolutional neural network,简称 SRCNN),从大规模样本数据集里学习低分辨率至高分辨率图像的最优映射函数。如图 8 所示,虽然 SRCNN 的网络结构相对简单(仅包含 3 层卷积层),但是可以取得超过稀疏编码方法<sup>[13]</sup>和领域嵌入方法<sup>[14]</sup>的图像重建效果。此外, SRCNN 精简的网络结构也满足实时视频流处理的要求。为了进一步提升深度网络模型的运行速度, Shi 等<sup>[16]</sup>引入基于亚像素层机构的高效亚像素卷积神经网络(efficient sub-pixel convolutional neural network,简称 ESPCN)网络模型(如图 9 所示)。这种网络模型直接将低分辨率图像作为训练输入,所有的卷积操作也都在低分辨率图上进行,大大减少了每幅图片样本训练时所需的卷积运算次数。

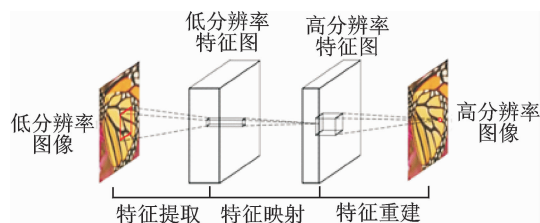


图 8 SRCNN 网络结构

Fig. 8 Network architecture of SRCNN

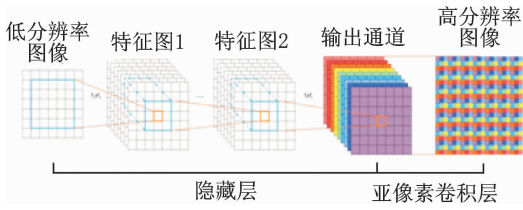


图 9 ESPCN 网络结构

Fig. 9 Network architecture of ESPCN

为了进一步提升深度网络模型的重建效果, Kim 等<sup>[17]</sup>提出了深层超像素网络模型(very deep super resolution, 简称 VDSR), 如图 10 所示。通过将 SRCNN 进行网络深度扩展(从 3 层增加至 20 层), 并且引入了跳跃连接, 学习低分辨率与高分辨率图像之间残差信息的映射关系。通过对网络层数扩展, 其包含的参数数目大大增加, 从而可以更加精确地拟合低分辨率和高分辨率图片之间的复杂映射关系。然而, 随着网络深度的增加, 回传误差极大化或极小化的问题也变得愈发严重, 导致模型训练难度大大增加。文献<sup>[17]</sup>通过使用较大的学习率加快了训练收敛速度, 同时采用了可调节的回传梯度限制, 缓解了回传误差梯度极大化或极小化问题。

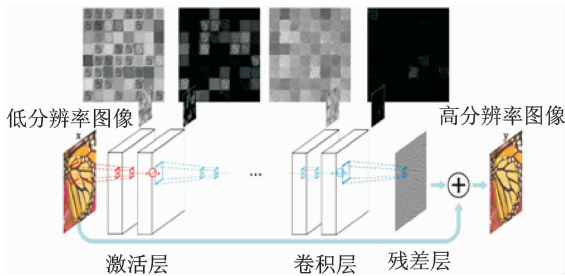


图 10 VDSR 网络结构

Fig. 10 Network architecture of VDSR

由于复杂的网络模型会导致庞大模型参数占用大量存储空间的问题, Kim 等<sup>[18]</sup>提出了基于深度监督和参数共享技术的深层递归卷积网络(deep-recursive convolutional network, 简称 DRCN)。多个模块之间通过参数共享, 大大减少模型的参数总和。并且每个模块都通过跳跃连接将特征层输出到重建层进行重建, 每个重建后的图像都进行互相监督, 并回传误差梯度。DRCN 网络模型在大大减少模型参数的情况下, 取得了类似于 VDSR 模型的特征重建效果。如图 11 所示, Tai 等<sup>[19]</sup>通过叠加了多个残差网络模块(ResNet), 构建了一个包含 52 层网络的深层递归残差网络(deep recursive residual network, 简称 DRRN)结构, 取得了远超 VDSR 方法的重建效果。为了减少模型参数数目, 每个残差模

块之间共享参数。需要指出, 虽然以上的参数共享的技术可以有效节省网络模型参数的存储空间, 但在运行模型时仍存在运算速度过慢的问题, 因此仅适用于对运行时间要求不高的离线图像增强应用中。

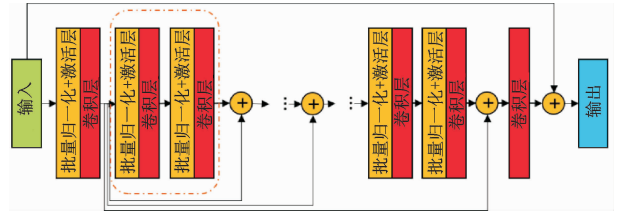


图 11 DRRN 网络结构

Fig. 11 Network architecture of DRRN

鉴于深度学习应用于可见光图像超像素取得的优良性能, Choi 等<sup>[20]</sup>提出了 4 层红外增强卷积神经网络(thermal image enhancement network, 简称 TEN), 将深度学习应用于红外图像的超像素处理。如图 12 所示, 与 SRCNN 网络相似, TEN 的网络结构也比较简单, 方便对网络参数进行训练但难以达到较高的特征重建效果。此外, 由于高分辨率的红外图像数据集较难获得, 且可见光图像的纹理变化比红外图像更加丰富, Choi 等<sup>[20]</sup>认为可以利用可见光图像进行训练, 并将训练得到的网络模型直接应用于红外图像, 同样能取得较好的效果。然后, 由于没有分析红外和可见光图像特征的本质区别, TEN 的增强效果十分有限。因此, 研究基于深度学习方法的红外图像超像素技术, 首先需要完成大规模红外图像特征库采集, 并针对红外图像的纹理信息和结构信息的不同特征, 提出相应的学习方法分别掌握, 在提高拟合精度的同时简化网络结构。此外, 基于节省成本的要求, 低分辨率红外传感器的像元数量通常极低(比如 60 像素×80 像素), 因此要求的超分辨率方法的放大倍数也相对较大, 这也给红外超分辨率技术的发展提出了更高的要求。研发一种针对红外图像的高倍超分辨率算法也是广大科研人员需要解决的问题之一。

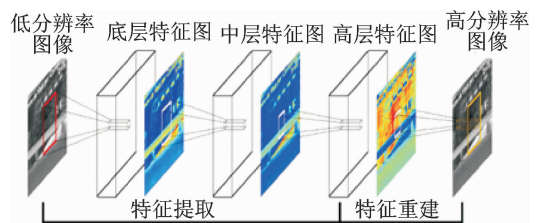


图 12 TEN 网络结构

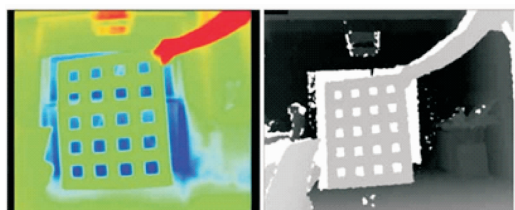
Fig. 12 Network architecture of TEN



## 5 多源信息采集、融合及互补技术

使用二维红外图像表述三维物体的表面温度分布具有严重局限性,包括只能获取单个视角下物体的温度信息、缺少温度异常区域的三维信息、无法精确定位热点位置等。因此,基于二维红外图像和三维深度信息融合互补技术的三维温度场重建技术是红外热成像领域的热点研究方向。现有的三维温度场重建系统通常包含以下主要步骤:a.对红外相机与深度传感器进行外参标定和视频同步,实现单视角采集条件下的多维信息融合;b.将深度传感器采集的物体单点三维信息进行坐标系统一匹配,获取物体的完整三维信息;c.将采集的红外温度信息根据物体三维几何信息进行校正补偿,并完成三维物体模型至二维红外图像的投影映射。

Skala 等<sup>[21]</sup>使用结构光深度相机和红外相机搭建三维温度场重建系统。如图 13 所示,他们制作特殊标定板,使用立体标定法计算出深度相机与红外相机的外参矩阵,利用相机外参矩阵直接将温度信息映射到三维模型上。Vidas 等<sup>[22]</sup>采集同时具备形状和温度明显变化的物体(如工作状态下的显示器)的红外图像和深度图像,通过对显著边缘信息的提取和匹配,计算红外相机与深度传感器之间的相对位姿,并完成多源传感器的同步标定。然而,由于标定板边缘处的深度信息不准确和难以对低像素红外图像中的角点准确定位等干扰因素,导致以上外参矩阵的计算方法存在误差。



(a) 红外图像  
(a) Thermal camera  
(b) 深度图像  
(b) Depth camera

图 13 特殊标定板对应的图像

Fig. 13 Designed calibration board seen by the thermal and depth camera

为了获取物体的三维信息,Ju 等<sup>[23]</sup>搭建一套基于双目立体视觉的三维重建系统。该系统使用两个高分辨率的彩色相机,可以在光照条件良好的环境里,精确获得物体的三维信息。然而该方案的测量精度易受环境光照、场景阴影等因素的影响,且配置多个高分辨率彩色相机价格昂贵,双目成像系统体积大,无法胜任快速、灵活的三维扫描任务。Tanno 等<sup>[24]</sup>提出使用两个红外相机实现物体三维

温度场重建。但由于红外图片低分辨率、低纹理信息、缺乏特征点的成像特性,该方案获取的物体三维信息精度较差。Ham 等<sup>[25]</sup>提出基于动态、生成结构(structure from motion,简称 SFM)技术的三维温度场重建系统。然而,SFM 算法只能生成稀疏点云且三维模型无尺度信息。激光雷达可获得精确的三维点云并且受环境因素影响较小,Alba 等<sup>[26]</sup>使用激光雷达和红外相机搭建三维温度场重建系统。然而,基于激光雷达的三维信息采集方案的价格昂贵并且体积重量大等。近些年,基于主动式结构光投射原理的三维测量技术获得了迅速发展。开发的深度传感器(如 Microsoft Kinect 相机)具有价格低、重量轻、体积小和可夜间成像等诸多优势。然而,基于结构光的深度相机普遍存在着噪声干扰严重,测量精度不高等问题。Izadi 等<sup>[27]</sup>提出的基于 GPU 并行编程技术的 Kinect Fusion 算法,通过采集同一场景在临近视角下的多帧深度信息,利用获得的冗余三维信息,对传感器噪声进行有效抑制,提高三维信息测量精度,可以实现±2 cm 精度的实时三维场景重建。

随着多视角三维重建技术的发展,Vidas 等<sup>[22]</sup>提出基于结构光深度相机的多视角全景三维温度场重建系统。如图 14 所示,该系统利用迭代最近点(iterative closest point,简称 ICP)算法来追踪相机位姿得到深度相机位姿,并进行多传感器时间轴同步,估算出最近时刻红外相机的位姿,最后将该视角下的三维信息与温度信息融合到统一的全局坐标系中,得到大场景的三维温度场重建结果。然而,该系统使用的 ICP 匹配算法仅仅通过寻找几何邻近点的简单方式建立三维匹配点,未能考虑其他有效的场景信息。因此,ICP 算法在相机位姿变化剧烈的情况下容易失效,并且在全景三维重建的过程中无法有效处理累积误差,导致重建的三维模型精度难以得到保障。

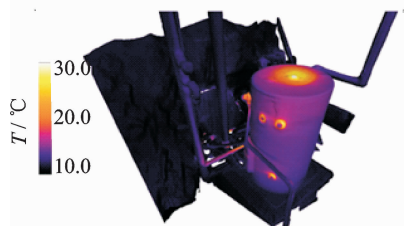


图 14 基于 Kinect 深度相机的三维温度场重建结果

Fig. 14 Multi-views 3D thermal mapping reconstruction based on Kinect depth camera

如图 15 所示,文献<sup>[28]</sup>将基于视频的位姿估计(video-based pose estimation,简称 VBPE)和 ICP

算法结合,提高追踪相机位姿的稳定性。为了保证三维温度场重建系统的实时性,该系统只能提取简单的图像特征点(如角点)对视频图像进行匹配,因此仍然无法处理前后两帧之间相机位姿变化大和物体缺少纹理信息的 ICP 失效问题。

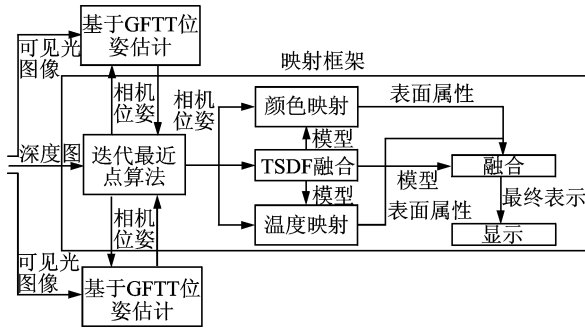


图 15 VBPE 和 ICP 结合的三维温度场重建系统框图  
Fig. 15 3D thermal mapping reconstruction system diagram combined VBPE and ICP

在获得准确匹配的二维温度场和三维深度信息后,可结合结构变化对红外辐射发射率的影响函数,对二维温度信息进行三维视角正规化补偿。为了获得更高精度的三维温度场,Muller 等<sup>[29]</sup>研究了一系列影响三维温度场重建精度的因素,如深度相机分辨率、相机轨迹追踪算法的精度以及鲁棒性、物体离红外相机的距离、红外线在大气中传播的能量损失、红外线发射角度以及反光的影响等。如何有效解决大位移条件下多维传感器系统运动轨迹的稳定计算,以及如何建立结构变化与红外辐射发射率之间的耦合函数,对二维温度信息进行三维视角正规化补偿,是三维温度场重建技术未来发展的重要研究方向。

## 6 多频谱红外和可见光信息互补技术

目标物体在光照条件较好的白天环境中,可见光图像能够提供更加丰富的颜色纹理信息,而在光照条件较差的夜间环境中,红外图像具有比可见光图像更加清晰的物体温度信息。因此利用不同传感器的互补性,获得目标更全面、更丰富、更多细节的多源信息,能有效提升红外热成像系统目标检测准确度。以智能驾驶应用中的行人检测为例,Hwang 等<sup>[30]</sup>搭建车载多频谱信号采集系统(如图 16 所示),对城市中不同光照条件下的行人目标进行图像采集和样本标注。Hwang 等<sup>[30]</sup>通过无视差的可见光与红外图像采集系统得到了完全配准好的多光谱图像数据集,采用集合频道特征(aggregate channel features,简称 ACF)特征提取算子和 Adaboost 分类器训练得到的可见光与红外目标检测器性能在白天和夜晚测试集上均超过了单一光谱(可见光或红

外)目标检测器性能。Choi 等<sup>[31]</sup>采用深度神经网络(deep convolutional neural networks,简称 DCNNs)进行目标特征的提取,使用支持向量回归(support vector regression,简称 SVR)取得了比传统特征方法更好的效果。Liu 等<sup>[32]</sup>将深度神经网络用于分类器的训练中,进一步地提升了检测器的性能。König 等<sup>[33]</sup>针对快速区域卷积神经网络特征(fast regions with CNN features,简称 Fast-RCNN)对尺度较小目标分类性能较差的特点提出了采用 AdaBoost 对区域建议网络(region proposal network,简称 RPN)特征进行精分类,提升了目标检测性能。目前对可见光与红外行人目标检测模型的优化主要还是在特征提取方法和分类器模型的选择上进行调整。笔者将同步检测分割区域建议网络(simultaneous detection segmentation & region proposal network,简称 SDS-RPN)<sup>[35]</sup>可见光行人目标检测模型用于红外与可见光行人目标检测器的训练与测试中,取得了目前最低的漏检率,SDS-RPN 同时作为特征提取器和分类器,在检测速度上更有优势。基于可见光与红外多频谱信息融合的行人目标检测模型的性能对比如表 1 所示。目前研究的红外与可见光目标检测器均采用相同的算法对红外与可见光图像进行特征提取。针对红外成像特性研究适用于红外图像的特征提取方法,是今后提升基于红外与可见光信息融合目标检测性能的一个重要方向。

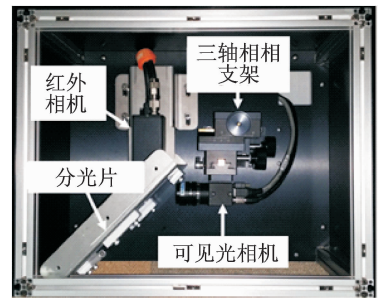


图 16 可见光与红外图像采集系统  
Fig. 16 Multispectral images acquisition system

表 1 基于可见光与红外多频谱行人目标检测模型对比  
Tab. 1 Comparison of multispectral pedestrian detection

检测模型	特征提取方法	分类器	漏检率/%
ACF	传统方法	传统方法	54.74
Adaboost	(ACF)	(Adaboost)	47.31
DCNNs+	深度神经网络	传统方法	36.99
SVR	(DCNNs)	(SVR)	29.83
Halfway	深度神经网络	深度神经网络	28.46
Fusion	(RPN)	(Fast RCNN)	
RPN+	深度神经网络	传统方法	
Adaboost	(RPN)	(Adaboost)	
SDS-		深度神经网络	
RPN		(SDS-RPN)	

## 7 结束语

红外热成像技术通过二维图像的形式准确反映物体表面温度分布情况,实现无光环境下的优秀成像,也可对物体表面温度进行非接触式的准确测量,广泛应用于医疗诊断、故障诊断、能耗监控、智能驾驶等众多工业领域。在对红外热成像技术进行的过程中,笔者发现红外热成像技术的发展也面临一些主要瓶颈制约问题。由于红外传感器制作工艺的特殊性,存在严重的噪声干扰,影响红外图像信号的成像质量。此外,红外探测器焦平面传感器价格昂贵,难以进行广泛的工业应用推广。最后,红外图像只包含物体的温度信息,无法单独为准确诊断提供依据,然而多源信息的采集融合工作目前还处于起步阶段。

针对以上3个主要问题,笔者详细介绍在红外信号固定模式建模和除噪、图像信息特征重建、多源信息融合三个红外信号处理研究方向的最新科研进展。未来拟开展的科研方向包括:a. 红外成像信号的噪声构成分析及形态建模方法;b. 可见光和低像素红外为信息源的红外图像重建理论;c. 红外图像、可见光图像、三维深度信息融合互补融合方法。研究成果将为新一代高成像精度、高分辨率、低制作成本的三维红外成像仪器的研发具有重要的参考价值。

### 参 考 文 献

- [1] Vollmer M, Mollmann K P. Infrared thermal imaging [J]. *European Journal of Physics*, 2010, 32(5): 8-25.
- [2] 何雯瑾, 太云见, 李玉英, 等. 热释电非制冷红外焦平面探测器热绝缘结构[J]. *红外与激光工程*, 2011, 40(8):1407-1411.  
He Wenjin, Tai Yunjian, Li Yuying, et al. Thermal isolation structure for pyroelectric uncooled infrared focal plane array detector [J]. *Infrared and Laser Engineering*, 2011, 40(8):1407-1411. (in Chinese)
- [3] 杰哈. 红外技术应用[M]. 北京: 化学工业出版社, 2004:1-28.
- [4] 李淑玉. 非制冷红外焦平面阵列封装及其关键工艺研究[D]. 北京: 北京大学, 2014.
- [5] Dereniak E L. Linear theory of nonuniformity correction in infrared staring sensors[J]. *Optical Engineering*, 1993, 32(8):1854-1859.
- [6] Torres S N, Hayat M M. Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays. [J]. *Journal of the Optical Society of America a Optics Image Science & Vision*, 2003, 20(3): 470-480.
- [7] Vera E, Meza P, Torres S. Total variation approach for adaptive nonuniformity correction in focal-plane arrays. [J]. *Optics Letters*, 2011, 36(2):172-174.
- [8] Rossi A, Diani M. Bilateral filter-based adaptive non-uniformity correction for infrared focal-plane array systems[J]. *Optical Engineering*, 2010, 49(5):057003.
- [9] Tendero Y, Gilles J. ADMIRE: a locally adaptive single-image, non-uniformity correction and denoising algorithm: application to uncooled IR camera[C]// *Infrared Technology and Applications*. Baltimore, Maryland, United States: SPIE, 2012:124-134.
- [10] Cao Yanpeng, Li Yiqun. Strip non-uniformity correction in uncooled long-wave infrared focal plane array based on noise source characterization [J]. *Optics Communications*, 2015, 339(15):236-242.
- [11] Cao Yanpeng, Yang M Y, Tisse C L. Effective strip noise removal for low-textured infrared images based on 1-D guided filtering[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2016, 26(12):2176-2188.
- [12] Kuang Xiaodong, Sui Xiubao, Chen Qian, et al. Single infrared image stripe noise removal using deep convolutional networks[J]. *IEEE Photonics Journal*, 2017, 9(4):1-13.
- [13] Yang Jianchao, Wright J, Huang T S, et al. Image super-resolution via sparse representation [J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
- [14] Chang Hong, Yeung D Y, Xiong Yimin. Super-resolution through neighbor embedding[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New Jersey, United States: IEEE, 2004:275-282.
- [15] Dong Chao, Chen C L, Tang Xiaou. Accelerating the super-resolution convolutional neural network [C]// *Proceedings of European Conference on Computer Vision*, Amsterdam. The Netherlands: Springer, Cham, 2016:391-407.
- [16] Shi Wenzhe, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New Jersey, United States: IEEE, 2016:1874-1883.
- [17] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]// *Proceedings of IEEE Conference on Computer Vision*



- and Pattern Recognition, New Jersey, United States; IEEE, 2016:1646-1654.
- [18] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New Jersey, United States; IEEE, 2016:1637-1645.
- [19] Tai Ying, Yang Jian, Liu Xiaoming. Image super-resolution via deep recursive residual network[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New Jersey, United States; IEEE, 2017:2790-2798.
- [20] Choi Y, Kim N, Hwang S, et al. Thermal image enhancement using convolutional neural network[C]// Proceedings of IEEE International Conference on Intelligent Robots. New Jersey, United States; IEEE, 2016:16503487.
- [21] Skala K, Lipič T, Sovič I, et al. 4D thermal imaging system for medical applications[J]. Periodicum Biologorum, 2011, 113(4):407-416.
- [22] Vidas S, Moghadam P, Bosse M. 3D thermal mapping of building interiors using an RGB-D and thermal camera[C]//Proceedings of IEEE International Conference on Robotics and Automation. New Jersey, United States; IEEE, 2013:2311-2318.
- [23] Ju X, Nebel J C, Siebert J P. 3D thermography imaging standardization technique for inflammation diagnosis[J]. Proceedings of SPIE - The International Society for Optical Engineering, 2004;56(40):266-274.
- [24] Tanno K, Kurose R, Makino H, et al. 3D Mapping of surface temperature using thermal stereo[C]// Proceedings of IEEE International Conference on Control Automation Robotics and Vision. New Jersey, United States; IEEE, 2007:1-4.
- [25] Ham Y, Golparvar-Fard M. An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery[J]. Advanced Engineering Informatics, 2013, 27(3):395-409.
- [26] Alba M I, Barazzetti L, Scaioni M, et al. Mapping infrared data on terrestrial laser scanning 3D models of buildings[J]. Remote Sensing, 2011, 3(9):1847-1870.
- [27] Izadi S, Kim D, Hilliges O, et al. KinectFusion:real-time 3D reconstruction and interaction using a moving depth camera[C]// Proceedings of ACM Symposium on User Interface Software and Technology. Santa Barbara, USA;CA, 2011:559-568.
- [28] Vidas S, Moghadam P, Sridharan S. Real-time mobile 3D Temperature mapping[J]. IEEE Sensors Journal, 2015, 15(2):1145-1152.
- [29] Muller A O, Kroll A. Generating high fidelity 3D thermograms with a handheld Real-time thermal imaging system[J]. IEEE Sensors Journal, 2017, 17(3):774-783.
- [30] Hwang S, Park J, Kim N, et al. Multispectral pedestrian detection: Benchmark dataset and baseline[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New Jersey, United States; IEEE, 2015:1037-1045.
- [31] Choi H, Kim S, Park K, et al. Multi-spectral pedestrian detection based on accumulated object proposal with fully convolutional networks[C]// Proceedings of IEEE Conference on Computer Vision and Pattern recognition. New Jersey, United States; IEEE, 2017:621-626.
- [32] Liu Jingjing, Zhang Shaoting, Wang Shu, et al. Multispectral deep neural networks for pedestrian detection [C]// Proceedings of the British Machine Vision Conference. York, UK; BMVA Press, 2016:1-73.
- [33] König D, Adam M, Jarvers C, et al. Fully convolutional region proposal networks for multispectral person detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern recognition. Workshops, New Jersey, United States; IEEE, 2017:243-250.
- [34] Girshick R. Fast RCNN[C]//Proceedings of the IEEE International Conference on Computer Vision. New Jersey, United States; IEEE, 2015:1440-1448.
- [35] Brazil G, Yin X, Liu X. Illuminating Pedestrians via Simultaneous Detection & Segmentation[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New Jersey, United States; IEEE, 2017:4950-4959.



**第一作者简介:**曹彦鹏,男,1981年7月生,博士、研究员、博士生导师,浙江大学“百人计划”。现为全国高校机械工程测试技术研究会常务理事、副秘书长,中国振动工程学会动态测试专业委员会常务委员、副秘书长。主要研究方向为红外、可见光、深度信息的采集处理和融合互补技术等。目前获得海外(欧洲)发明专利1项,已发表(SCI/EI收录)论文20余篇。

E-mail:caoy@zju.edu.cn