

文章编号: 2095-2163(2021)06-0036-08

中图分类号: TP751

文献标志码: A

基于微振动增强的伪装人体检测算法

黄飞龙^{1,2}, 杨学志^{2,3}, 吴克伟^{1,2}, 沈晶^{4,5}, 陈鲸¹

(1 合肥工业大学 计算机与信息学院, 合肥 230009; 2 工业安全与应急技术安徽省重点实验室, 合肥 230009;

3 合肥工业大学 软件学院, 合肥 230009; 4 合肥师范学院 电子信息与电气工程学院, 合肥 230061;

5 安徽微波与通信工程技术研究中心, 合肥 230061)

摘要: 伪装人体检测在视频监控领域具有重要的研究和应用价值。本文针对目前基于图像表层特征提取的伪装人体检测方法无法有效检测出没有明显移动的伪装人体目标, 充分利用了人体目标具有一定范围呼吸率的独特特性, 提取并增强伪装人体的微振动特征, 利用振动特征实现了伪装人体目标进行检测。首先, 提出了自己的伪装人体视频数据集和基于微振动特征的伪装人体检测模型; 其次, 利用训练集估计伪装人体目标的最优呼吸率区间, 利用估计的最优呼吸率区间增强了视频中伪装人体的微振动; 最后, 根据增强了微振动的视频和提出的检测模型, 实现伪装人体目标的定位检测, 并通过图像形态学去噪后处理降低检测结果的噪声。本文提取微振动特征的检测算法通过充分的消融实验和对比实验进行了验证, 在提出的视频数据集上, 检测效果 IOU 达到了 0.526, Precision 达到了 0.738, 优于其它图像特征最先进的检测方法。

关键词: 伪装人体; 振动特征; 呼吸率区间; 微振动

Camouflaged human detection algorithm based on micro-vibration enhancement

HUANG Feilong^{1,2}, YANG Xuezhizhi^{2,3}, WU Kewei^{1,2}, SHEN Jing^{4,5}, CHEN Jing¹

(1 School of Computer and Information, Hefei University of Technology, Hefei 230009, China;

2 Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei 230009, China;

3 School of Software, Hefei University of Technology, Hefei 230009, China;

4 School of Electronic Information and Electrical Engineering, Hefei Normal University, Hefei 230061, China;

5 Anhui Microwave and Communication Engineering Technology Research Center, Hefei 230061, China)

[Abstract] Camouflaged human detection has important research and application value in the field of video surveillance. In view of the fact that the current detection methods for camouflaged human based on image surface feature extraction cannot effectively detect the target without obvious movement, this paper makes full use of the human's unique characteristics with a certain range of respiration rates. This paper extracts and enhances the micro-vibration features of the camouflaged human to realize the detection of the target. First, a camouflaged human video data set and a camouflaged human detection model based on micro-vibration features are proposed. And then the training set is used to estimate the optimal respiration rate interval of camouflaged human. Next, the estimated optimal respiration rate interval is used to enhance the micro-vibration of the human. Finally, according to the enhanced micro-vibration video and the proposed detection model, the camouflaged human in video gets detected and segmented, and the post-processing of image morphology denoising is used to reduce the noise of the result. The detection algorithm for extracting micro-vibration features in this paper is verified by sufficient ablation experiments and comparative experiments. On the proposed video data set, the detection IOU reaches 0.526, and the precision reaches 0.738, which is better than other advanced detection methods.

[Key words] camouflaged human; vibration features; respiration rate interval; micro-vibration

0 引言

人体的检测在现实生活中有着许多重要的应用, 例如交通监控、自动驾驶、智能监控、行人监控等, 这些场景都需要实现准确的人体检测。人体的

检测一直是计算机视觉领域的研究热点, 许多基于人脸和人体姿态估计的人体检测算法已经被提出来并得到了很好的完善。过去几年, 人们的重点关注在于行人的检测, 在公开可用的研究数据集中, 最受欢迎的是 INRIA 数据集。除此以外, 显著性目标的

作者简介: 黄飞龙(1994-), 男, 硕士研究生, 主要研究方向: 计算机视觉; 杨学志(1970-), 男, 博士, 教授, 主要研究方向: 图像处理、模式识别、SAR 图像处理与解译; 吴克伟(1984-), 男, 博士, 副教授, 主要研究方向: 计算机视觉; 沈晶(1984-), 女, 博士, 副教授, 主要研究方向: 计算电磁学; 陈鲸(1940-), 男, 学士, 教授, 主要研究方向: 通信与信息系统。

通讯作者: 杨学志 Email: xzyang@hfut.edu.cn

收稿日期: 2021-03-16

检测也得到了很充分的研究,为行人的识别研究提供了很好的参考。然而,在一些场景下,当人体的外观特征并不明显的时候,仅仅依靠图像的表层特征是无法检测出来的。例如,在复杂环境中经过了一定程度伪装的人体目标。伪装人体就是经过迷彩服装、颜料等手段,使自己的外观颜色、纹理特征和背景高度融合,以至于人很难通过眼睛发现的人体目标。伪装人体的检测算法可以很好检测出图像特征不明显的人体目标,对于边防监控、入侵检测、灾害救援活人检测、军事目标发现等方面具有重要的作用。

伪装人体检测是目前计算机视觉的一个具有挑战性的任务。使用计算机视觉的检测方法处理的数据通常是单张图像或者是具有时间信息的视频。近年来,已有一些伪装人体检测相关的研究,这些研究根据提取的视觉特征不同,可以分成4类:亮度或颜色特征、纹理特征、梯度特征、深度特征。亮度和颜色特征是最基本的特征,利用亮度或者颜色特征来检测伪装人体目标,设计了一种加权区域合并的迭代方法,来检测图像序列中颜色非常接近的伪装目标的方法^[1]。这种方法的效果主要依赖于图像之间的差异,如果检测的目标不移动或缓慢移动,这种方法就不能检测出伪装人体目标;利用纹理特征进行伪装检测,提出了一种方法来检测目标的伪装部分,并将伪装部分从给定图像的环境中提取出来^[2]。采用基于灰度共生矩阵的纹理特征和树状图来检测伪装目标,但是这种方法非常耗时,因为其需要将给定的图像分割成多个块或更小的区域,在包含阴影效果和背景与目标包含相似纹理的图像中,这种方法无法有效检测出伪装的目标。使用梯度特征进行伪装人体检测,提出了一种 D_{arg} 算子来增强对应于三维目标凸面的阴影区域,从而将该区域从具有颜色和纹理等相似特征的平坦背景中分离出来^[3]。 D_{arg} 算子直接应用图像的灰度函数,对目标上光滑的三维凸面做出响应,不会受到任何特定光源及其反射函数的限制,但是使用 D_{arg} 算子检测的结果高度依赖于阈值,选择一个合适的阈值本身也是一个重要的问题。此外,这种方法不适用于包含凹面特性的背景和深色的目标。深度特征检测伪装人体目标,提出了强语义扩张网络来检测伪装的人^[4]。充分利用了卷积神经网络的语义信息,并利用了空洞卷积来扩大接受野,以发现伪装的人体目标。在一个标准的图像数据集上给出了令人印象深刻的结果,但是如果有一些数据集中未知场景的伪装人体不一定能检测出来。

针对以上问题,本文提出了一种基于微振动增强的伪装人体检测算法。首先,提出了自己的伪装人体视频数据集和基于微振动特征的伪装人体检测模型,利用训练集估计伪装人体目标的最优的呼吸率区间;其次,利用估计的最优呼吸率区间增强了视频中伪装人体的微振动;最后,根据增强了微振动的视频和提出的检测模型实现伪装人体目标的定位检测,并通过图像形态学去噪后处理降低检测结果的噪声。本文的振动特征用于伪装人体检测取得了比其它特征更好的效果,如图1所示。

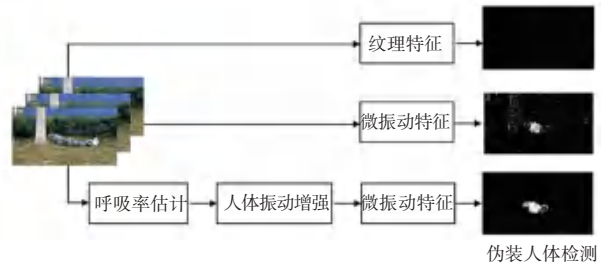


图1 不同特征的检测效果

Fig. 1 Detection effect of different features

1 本文算法

1.1 提出的数据集

由于现有的伪装数据集收集的伪装图像无法用于基于视频的伪装人体检测。因此,使用带有三脚架的佳能数码相机收集了10个被伪装人员的300个稳定的视频。视频微小振动检测面临的挑战是长距离的目标,因此这项工作进一步在不同的拍摄距离上收集了不同比例的伪装人类。拍摄距离为10 m、20 m、30 m的视频数分别为100个,每个视频的持续时间为15 s,分辨率为720×360,帧速率为50 fps。在一天中的不同时间,在足够的光照条件下采集视频。视频数据集分为训练集和测试集,其中每种拍摄距离的50个视频用于训练,其余视频用于测试。测试集中不存在训练集中的伪装人员,可以验证呼吸率区间估计的鲁棒性。

这些视频数据使用标注工具 labelme 对第一帧图像的伪装人体目标进行像素级别的标注和外接包围框,并带有拍摄条件,包括:环境、振动类型、遮挡类型、高光和阴影类型、颜色深浅类型和拍摄距离。数据集部分视频及其标注,如图2所示。

环境:视频包括各种自然环境,例如:草(图2(a)),灌木丛(图2(b)),树木(图2(c)),树叶(图2(d))。

振动类型:视频具有2种人体振动类型,包括腹

部向上(图2(e)和图2(g))和腹部向下(图2(f)和图2(h))。

遮挡类型:伪装的人体没有背景遮挡(图2(e)和图2(f))或有背景遮挡(图2(g)和图2(h))。

高光和阴影类型:在视频中存在高光和阴影(图2(b))或没有高光和没有阴影(图2(a))。

颜色深浅类型:背景和伪装的人体具有深色外观(图2(b))或浅色的外观(图2(i))。

拍摄距离:伪装的人体距离为10 m(图2(i)), 20 m(图2(j)), 30 m(图2(k))。

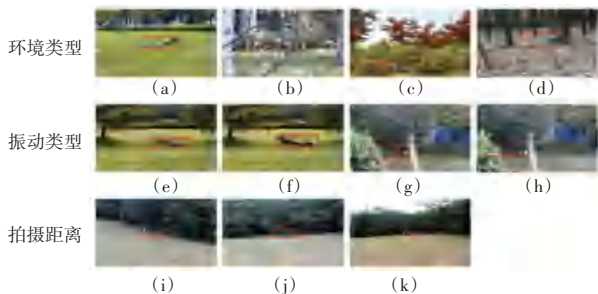


图2 视频数据集部分数据

Fig. 2 Part of the video data set

1.2 提出的方法

本文提出了基于微振动特征的伪装人体检测模型,根据视频训练集估计了更准确的伪装人体呼吸率区间,根据估计的呼吸率区间增强视频中伪装人体振动信号,利用检测模型和形态学去噪方法得到检测结果。检测框架如图3所示。

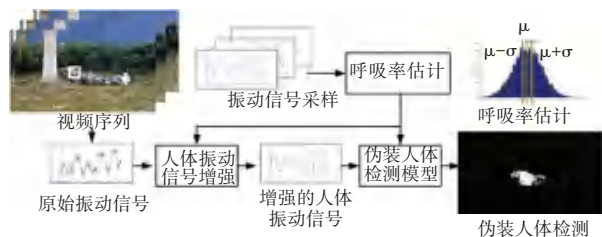


图3 伪装人体检测框架

Fig. 3 Camouflage human detection framework

1.2.1 伪装人体微振动检测模型

人体的呼吸频率是人体胸部和背部微振动的频率。根据医学统计,人体的呼吸率数值区间为0.20~0.33 Hz。因此,人体胸部和背部的振动频率和人体的呼吸率是一样的。给定包含伪装人体的视频,并将该视频转换到YIQ色彩空间,其中通道Y是亮度通道。像素 x 在各个时间 t 的亮度可以记为序列 $f(x,t)$,则 $f(x,t)$ 是振动信号;使用一维快速傅立叶变换(FFT)将时域信号 $f(x,t)$ 转换为频域信号: $f_{\omega}(x,\omega) = \text{fft}(f(x,t))$,频域信号 $f_{\omega}(x,\omega)$ 的模的最大值是其幅值,如公式(1)所示。

$$A(x) = \max(|f_{\omega}(x,\omega)|). \quad (1)$$

频域信号的振动频率与其幅值相对应,如公式(2)所示。

$$H(x) = \arg \max_{\omega}(f_{\omega}(x,\omega)). \quad (2)$$

假设人体的呼吸频率区间为 $H(a,b)$,则基于微振动特征的伪装人体检测模型如公式(3)所示。

$$R(x) = \begin{cases} 1, & H_a < H(x) < H_b, A(x) > \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

$R(x)$ 表示伪装人体检测结果。

1.2.2 伪装人体呼吸率估计

由于医学统计的人体的呼吸频率区间是在人体正常情况下测量的,而在本文的数据集中,伪装人体的呼吸频率高于正常人体的呼吸率区间,这可能是因为人体的伪装姿势会影响到呼吸。因此,需要在视频数据集人体伪装姿势情况下估计呼吸率区间。本文使用视频训练集估计伪装人体的呼吸率的区间,训练集包含每个拍摄距离的50个视频。呼吸率区间估计过程如图4所示。在每个训练的视频真值分割图中采样200个像素,并形成振动信号集 $F_p(x,t) = \{f_p(x_i,t)\}$,其中 p 表示伪装人体的真值位置, i 是训练视频的编号。呼吸率区间估计变成了发现信号集振动频率和频谱幅值的数学分布,振动信号集的振动频率和频谱幅值都服从正态分布,如公式(4)、(5)所示。

$$H_p(x) = \{H_p(x_i)\} \sim N(\mu_H, \sigma_H^2), \quad (4)$$

$$A_p(x) = \{A_p(x_i)\} \sim N(\mu_A, \sigma_A^2). \quad (5)$$

在本文的训练集中,振动频率的均值 μ_H 为0.447、标准差 σ_H 为0.133,频谱幅值的均值 μ_A 为22.830,标准差 σ_A 为15.590。数据集中伪装人体目标的呼吸率区间和频谱幅值的阈值分别取为等式(6)、(7)。

$$(H_a, H_b) = (\mu_H - k\sigma_H, \mu_H + k\sigma_H). \quad (6)$$

其中, k 是区间估计因子,当 $k=1$ 时,估计的区间包含95%的真值图像的像素。

$$\tau = \mu_A - \frac{1}{2}\sigma_A. \quad (7)$$

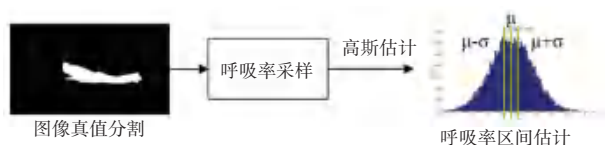


图4 呼吸率区间估计示意图

Fig. 4 An overview of respiration rate interval estimation

1.2.3 伪装人体微振动增强

伪装人体目标在视频中表现出来的振动是微小

的,很难直接分析到这种信号。伪装人体的振动还会受到环境的干扰,例如背景中的草、树叶等植物在风的作用下的晃动,需要抑制这些环境中植物的振动。微振动增强就好比一个显微镜,可以将特定频率范围的振动信号进行增强,抑制其它不感兴趣频率的振动,可以更好的捕捉到需要的信号。在这里不能使用标准的傅立叶变换方法,因为其仅适用于图像中的整体相位的变化。

使用了基于相位变化的微振动增强方法来处理局部相位的变化^[5]。本文基于其源码修改,以实现伪装人体的振动信号增强。为了测量图像中的局部

相位振动,使用了复数数域的可控金字塔,类似于将图像划分为不同空间结构的局部傅里叶变换。复数数域的可控金字塔将图像分解为对应于不同图像尺度和方向的不同子带,每个子带的每个像素位置都用局部幅度和局部相位来表示。用时间带通滤波器进一步处理随时间变化的相位信号,以提取感兴趣振动的像素位置,这些经过时域带通滤波的相位差乘以一个放大因子 α 。在图像的局部相位放大处理之后,将所有的子带求和以重建视频,得到伪装人体微振动增强的视频。伪装人体微振动增强的整个过程如图 5 所示。

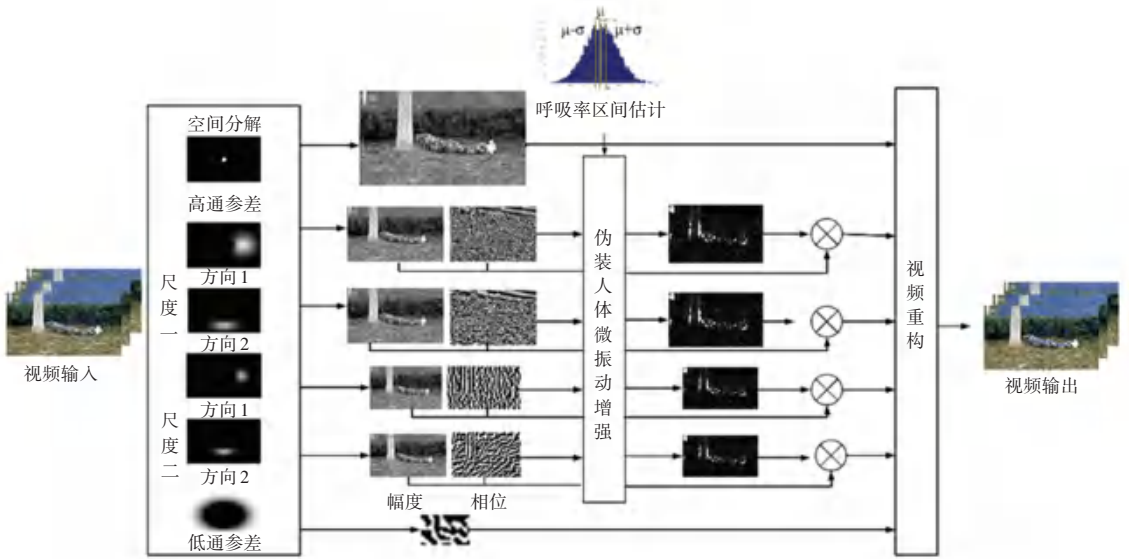


图 5 伪装人体微振动增强示意图

Fig. 5 An overview of micro-vibration enhancement of camouflaged human

假设图像尺度 j 的视频信号为 $f_j(x + \delta(t))$,其中 f 表示尺度 j 的图像的亮度平面, $\delta(t)$ 表示伪装人体的微振动位移。根据傅里叶级数展开式, $f_j(x + \delta(t))$ 可以写成一系列复数正弦信号 $S_{j\omega}(x, t) = A_{j\omega} e^{i\omega(x + \delta(t))}$ 的和,而且这些复数正弦信号分别对应一个频率 ω ,如公式(8)所示。

$$f_j(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} S_{j\omega}(s, t) = \sum_{\omega=-\infty}^{\infty} A_{j\omega} e^{i\omega(x + \delta(t))}. \quad (8)$$

其中, $S_{j\omega}(x, t)$ 是子带信号 $f_j(x + \delta(t))$ 的频率为 ω 的基信号。因为 $S_{j\omega}(x, t)$ 是正弦信号,其相位信息 $\omega(x + \delta(t))$ 包含了伪装人体的微小振动的相位信息和直流分量 ωx 。在使用时域带通滤波器去除直流分量 ωx 后,伪装人体的微振动相位信息是 $B_{\omega}(x, t) = \omega\delta(t)$,这在每一个视频尺度和每一个子带下都是一样的,然后使用放大因子 α 对表示伪装人体微振动相位信息进行放大处理,得到放大了相位信息的基信号,如公式(9)所示。

$$\hat{S}_{j\omega}(x, t) = S_{j\omega}(x, t) e^{i\alpha B_{\omega}(x, t)} = \sum_{\omega=-\infty}^{\infty} A_{j\omega} e^{i\alpha\omega t + i\omega(x + \delta(t))} = \sum_{\omega=-\infty}^{\infty} A_{j\omega} e^{i\omega(x + (1 + \alpha)\delta(t))}. \quad (9)$$

其中, $\hat{S}_{j\omega}(x, t)$ 表示放大了相位信息的基信号。把所有放大了相位信息的基信号加起来,得到一个增强的子带信号,如公式(10)所示。

$$\hat{f}_j(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} \hat{S}_{j\omega}(s, t) = \sum_{\omega=-\infty}^{\infty} A_{j\omega} e^{i\omega(x + \delta(t))}. \quad (10)$$

其中, $\hat{f}_j(x + \delta(t))$ 表示尺度 j 下增强的子带信号。最后,将所有尺度的子带信号加起来进行视频重构,得到了一个增强了伪装人体微振动信号的视频,如公式(11)所示。

$$\hat{f}(x + \delta(t)) = \sum_{j=1}^n \hat{f}_j(x + \delta(t)). \quad (11)$$

其中, $\hat{f}(x + \delta(t))$ 表示增强了伪装人体振动信

号的视频信号。

1.2.4 伪装人体检测后处理

经过振动增强模块,原视频中伪装人体的微小振动信号得到增强,对进一步进行伪装人体检测很有帮助。对增强信号后的视频进行处理,计算出增强视频的整个空间的频率平面 $\hat{H}(x)$ 和幅值平面 $\hat{A}(x)$, 利用公式(3)表示的检测模型,根据频率平面 $\hat{H}(x)$ 、幅值平面 $\hat{A}(x)$ 、公式(6)、公式(7),得到

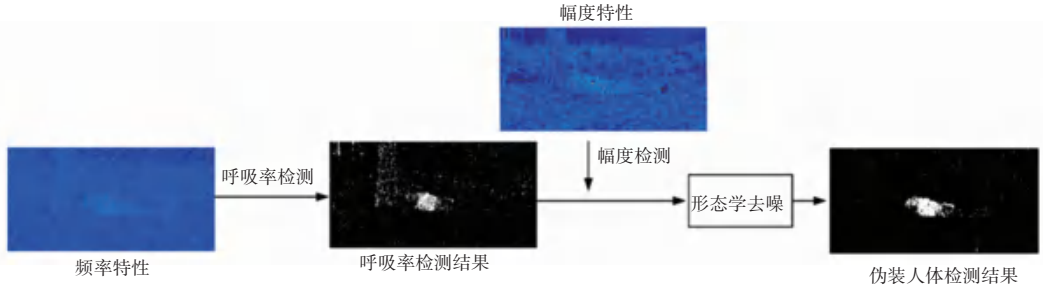


图6 伪装人体检测后处理过程示意图

Fig. 6 An overview of the post-processing process of camouflaged human detection

基于微振动增强的伪装人体检测的框架可以使用算法1进行描述。

算法1 基于微振动增强的伪装人体检测

输入:原始视频信号 $f(x, t)$, 估计的呼吸率区间 $\omega = (H_a, H_b)$, 估计的幅值阈值 τ

输出:伪装人体图像分割结果 $R(x)$

(1) for all j do;

(2) 提取尺度为 j 的视频的相位信息: $f_j(x +$

$$\delta(t)) = \sum_{\omega=-\pi}^{\pi} A_{j\omega} e^{i\omega(x+\delta(t))};$$

(3) 使用估计的呼吸率区间 ω 放大基信号 $S_{j\omega}(x,$

$$t) \text{ 的相位信息: } \hat{S}_{j\omega}(x, t) = \sum_{\omega=-\pi}^{\pi} A_{j\omega} e^{i\omega(x+(1+\alpha)\delta(t))};$$

(4) 重建尺度为 j 的视频信号 $f_j(x + \delta(t))$: $\hat{f}_j(x$

$$+ \delta(t)) = \sum_{\omega=-\pi}^{\pi} A_{j\omega} e^{i\omega(x+\delta(t))};$$

(5) end for;

(6) 用所有尺度的视频重建微振动增强的视

$$\hat{f}(x + t) = \sum_{j=1}^n \hat{f}_j(x + \delta(t));$$

(7) 将增强了伪装人体微振动的视频从空间域

$$\text{转到频域: } f_{\omega}(x, \omega) = \text{fft}(\hat{f}(x + t));$$

(8) 对转到频域的视频计算其频率平面 $\hat{H}(x)$

和幅值平面 $\hat{A}(x)$:

$$\hat{A}(x) = \max(|\hat{f}_{\omega}(x, \omega)|), \hat{H}(x) = \arg \max_{\omega}(\hat{f}_{\omega}(x,$$

了伪装人体目标检测的结果 $R(x)$ 。

为了降低检测结果的噪声、填补目标的空洞,最后加入了一个图像形态学去噪模块对检测结果进行后处理操作。使用大小为 3×3 的结构元素,对检测结果 $R(x)$ 进行一次图像开运算,以降低检测结果的噪声。一次图像开运算是先对图像进行一次膨胀操作,再对图像进行了一次腐蚀操作,前者可以填补目标的空洞,后者可以去除检测结果的噪声。这个过程如图6所示。

$\omega)$);

(9) 根据 $\hat{H}(x)$ 、 $\hat{A}(x)$ 、 τ 检测伪装人体目标:

$$R(x) = \begin{cases} 1, & H_a < \hat{H}(x) < H_b, \hat{A}(x) > \tau, \\ 0, & \text{otherwise.} \end{cases}$$

(10) 使用一个 3×3 的结构元素对 $R(x)$ 进行一次图像开运算的形态学去噪。

2 实验结果与分析

2.1 评价指标

使用重叠度 (IOU) 来评价本文检测算法的检测能力。重叠度指标经常用在物体检测任务上,用来评价检测的物体区间和物体真实区间之间的匹配程度。本文实验里,重叠度是检测区间面积与物体真实区间面积的比例。如果检测到的目标的区间面积、目标真实区间的面积分别是 S_p 、 S_g , 那么重叠度 IOU 的计算方式是公式(12)。

$$IOU = \frac{S_p \cap S_g}{S_p \cup S_g}. \quad (12)$$

也使用了精确度 ($Precision$) 来评价本文的检测模型。精确度指标是在分类模型里评价模型检测出正确数据点的能力,定义为公式(13)。

$$Precision = \frac{TP}{TP + FP}. \quad (13)$$

IOU 和 $Precision$ 是机器学习目标检测中广泛使用的评价标准,两个评价标准可以很好评价目标

检测的准确性,本文选择了这两个评价标准来进行性能度量。

2.2 算法时间效率

使用 Matlab 2019 在处理器为 i7Core、3.60 Ghz CPU 和 8 G 内存的 PC 上实现的。为了减少运行时间,将视频原始图像降采样为 360×180 的分辨率。输入的振动信号是 360×180×750,其中 750 是帧率为 50 fps 的 15 s 的视频帧数。一个视频的平均处理时间为 237 s,其中快速傅立叶变换、形态学去噪处理模块、微振动增强模块的运行时间分别为 50 s, 14 s 和 173 s。

2.3 消融实验

2.3.1 呼吸率区间的影响

在本文的检测框架里,主要估计了数据集的伪装人体的呼吸率区间。为了证明使用估计的呼吸率比使用经验值效果好,本文改变呼吸率的值做了 3 组实验,见表 1。当 $k = 1$ 时,使用估计的呼吸率比经验值有更好的检测效果,这是因为估计的呼吸率比经验值更好的符合数据集中伪装人体的呼吸率。相比之下, IOU 提高了 0.143, $Precision$ 提高了 0.141。当 k 等于其它值,例如 $k = 2$,检测效果反而下降。通过实验,认定 $k = 1$ 时估计的呼吸率最好。其它的实验也使用了 $k = 1$ 的呼吸率区间。

表 1 不同呼吸率区间的影响

Tab. 1 The effect of different breathing rate respiration rate interval

呼吸率区间/Hz	IOU	$Precision$
0.20-0.33(经验值)	0.383	0.597
0.31-0.59 ($k = 1$)	0.526	0.738
0.18-0.71 ($k = 2$)	0.274	0.461

2.3.2 视频微振动增强的影响

为了分析微振动增强对不同距离的人体检测的影响,本文分别在微振动不增强视频和微振动增强视频的两种情况下,验证了 10 m、20 m、30 m 的检测结果,见表 2,部分视觉检测效果如图 7 所示。可以发现,在没有增强的视频中,对于 20 m 和 30 m 的较远的目标,检测的 IOU 和 $Precision$ 都比 10 m 的要低;在增强的视频中,对于 20 m、30 m 的目标,检测准确性大大提高,整体上 IOU 提高了 0.24, $Precision$ 提高了 0.278。可以得出一个结论,微振动增强有助于检测出 20 m 远的小尺度目标。

表 2 微振动增强对检测结果的影响

Tab. 2 The effect of micro-vibration enhancement

距离	IOU				$Precision$			
	10 m	20 m	30 m	all	10 m	20 m	30 m	all
不增强	0.376	0.293	0.189	0.286	0.586	0.467	0.327	0.460
增强	0.585	0.526	0.467	0.526	0.796	0.763	0.655	0.738

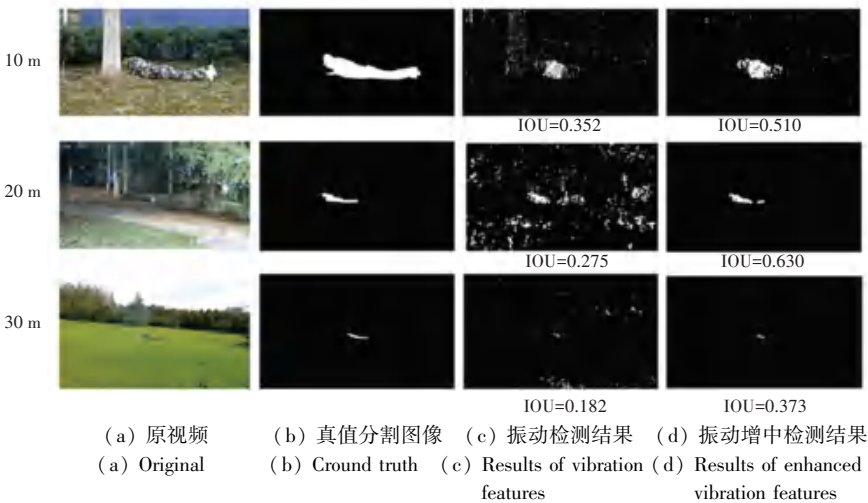


图 7 微振动增强对检测结果的影响

Fig. 7 The effect of micro-vibration enhancement

2.3.3 不同模块的影响

为了验证本文的检测框架中微振动增强、幅值估计模块、形态学去噪 3 个模块对伪装人体检测准确性的效益,本文设计了模块的消融实验,实验的结果见表 3。从本文方法 1 与基线方法比较可知,视频放大模块对检测结果提升明显, IOU 提升了

0.102, $Precision$ 提升了 0.120;本文方法 2 与基线方法比较可知,视频放大模块结合幅值估计模块可以大幅度提升检测的效果, IOU 提升了 0.183, $Precision$ 提升了 0.194。因为对放大的视频使用快速傅里叶变换求各像素的振动频率时,在非振动区域中得到的频谱去除了一些没有明显波峰的数据,

使得估计的像素振动频率更准确;本文方法 3 与本文方法 2 比较,形态学去噪模块对于提升检测效果也有一定的作用, IOU 提升了 0.057, $Precision$ 提升了 0.084。形态学去噪模块之所以有作用,是因为其对初步的检测结果消除了零散的噪声,并填补了结果的空洞。

表 3 不同模块的影响

Tab. 3 The effect of different modules

方法	微振动增强模块	幅值估计	形态学去噪	IOU	$Precision$
基线方法				0.286	0.460
本文方法 1	✓			0.388	0.580
本文方法 2	✓	✓		0.469	0.654
本文方法 3	✓	✓	✓	0.526	0.738

2.4 对比实验

为了证明本文的算法有效,还实现了基于梯度特征的检测方法、基于纹理特征的检测方法、基于深度特征的检测方法,并与其检测效果进行了比较,这几个算法分别是梯度特征、纹理特征、深度特征等不同特征的先进方法。由于这几个研究或是没有公开其数据集,或是其数据集不适用于本文的算法,为了比较更加公平,把这些算法用在本文的数据集上进行测试。这几种检测方法的输入数据是图像,因此把所有验证视频的第一帧作为其输入数据。检测结果见表 4,本文算法的检测准确性是最高的, IOU 达

到 0.526, $Precision$ 达到了 0.738。

表 4 对比实验结果

Tab. 4 Comparison of the detection results

方法	IOU	$Precision$
梯度特征法	0.09	0.11
纹理特征法	0.103	0.155
深度特征法	0.330	0.556
本文方法	0.526	0.738

部分检测效果如图 8 所示。可以看出,本文的检测算法效果比较清晰,可以较好的保持伪装人体目标的外形和边缘轮廓,具有较高辨识度。这是因为本文增强并提取的是伪装人体目标的微小振动特征,由人体目标的呼吸而产生,与人的呼吸一样具有内在稳定性,不受背景中纹理、阴影、颜色深浅的影响,较梯度特征法、纹理特征法具有更好的抗噪声能力。本文使用视频训练集估计出伪装人体的振动频率区间,得到的结果可视为先验知识,在未知场景中同样适用,较深度特征法是更普遍通用的伪装人体检测框架。梯度特征法只是定位出一个目标点,检测准确性比较低,会将图像中纹理的边缘错检测成伪装人体。纹理特征法的检测准确率也很低,并且检测出来比较暗淡模糊,无法辨识伪装目标。深度特征法虽然善于检测正常的图像,但是不擅长检测有遮挡的伪装人体,通常将背景的物体错误检测成伪装目标。

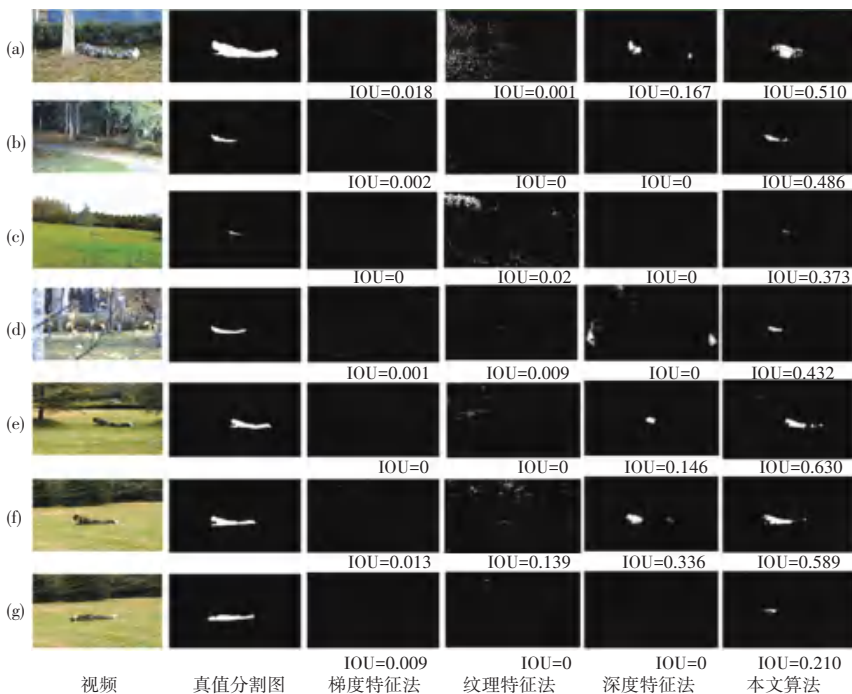


图 8 对比实验检测效果

Fig. 8 Comparison of detection results