基于状态识别的高炉料面视频关键帧提取方法

黄建才[1] 蒋朝辉[1,2] 桂卫华[1] 潘 冬[1] 许 川[1] 周 科[1]

摘 要 高炉料面视频关键帧是视频中的中心气流稳定、清晰、无炉料及粉尘遮挡且特征明显的图像序列, 对于及时获取炉内运行状态、指导炉顶布料操作具有重要的意义. 然而, 由于高炉内部恶劣的冶炼环境及布料的周期性和间歇性等特征, 料面视频存在信息冗余、图像质量参差不齐、状态多变等问题, 无法直接用于分析处理. 为了从大量高炉冶炼过程料面视频中自动准确筛选清晰稳定的料面图像, 提出基于状态识别的高炉料面视频关键帧提取方法. 首先, 基于高温工业内窥镜采集高炉冶炼过程中的料面视频, 并清晰完整给出料面反应新现象和形貌变化情况; 然后, 提取能够表征料面运动状态的显著性区域的特征点密集程度和像素位移特征, 并提出基于局部密度极大值高斯混合模型 (Local density maxima-based Gaussian mixture model, LDGMM) 聚类的方法识别料面状态; 最后, 基于料面状态识别结果提取每个布料周期不同状态下的关键帧. 实验结果表明, 该方法能够准确识别料面状态并剔除料面视频冗余信息, 能提取出不同状态下的料面视频关键帧, 为优化炉顶布料操作提供指导.

关键词 高炉, 料面现象, 显著性区域, 状态识别, 关键帧提取

引用格式 黄建才, 蒋朝辉, 桂卫华, 潘冬, 许川, 周科. 基于状态识别的高炉料面视频关键帧提取方法. 自动化学报, 2023, **49**(11): 2257−2271

DOI 10.16383/j.aas.c220969

Key Frame Extraction Method of Blast Furnace Burden Surface Video Based on State Recognition

HUANG Jian-Cai[1] JIANG Zhao-Hui[1,2] GUI Wei-Hua[1] PAN Dong[1] XU Chuan[1] ZHOU Ke[1]

Abstract The key frames of the blast furnace burden surface video are the clear image sequences with stable central airflow, no burden and dust occlusion, and obvious characteristics, which are of great significance for timely obtaining the running state of the blast furnace and guiding the charging operation. However, due to the harsh ironmaking environment inside the blast furnace, the periodic and intermittent characteristics of the burden distribution, the burden surface video has problems such as redundant information, uneven image quality and changeable state, which cannot be directly used for analysis and processing. To screen clear and stable burden surface images automatically and accurately from a large number of burden surface videos during the blast furnace ironmaking process, a key frame extraction method of blast furnace burden surface video based on state recognition is proposed. Firstly, the burden surface video in the blast furnace ironmaking process is collected based on the high-temperature industrial endoscope, and the new phenomenon and change of burden surface topography are given clearly and completely. Then, the feature point density and pixel displacement characteristics in the salient region that can characterize the burden surface motion state are extracted. Next, a method of local density maxima-based Gaussian mixture model (LDGMM) clustering is proposed to recognize the burden surface state. Finally, the key frames in different states of each burden distribution cycle are extracted based on the state recognition results of the burden surface. The experimental results show that this method can accurately recognize the burden surface state, eliminate the redundant information of the burden surface video, and extract the key frames of the burden surface video under different states, which provides guidance for optimizing the furnace top charging operation.

Key words Blast furnace, burden surface phenomenon, salient region, state recognition, key frame extraction

Citation Huang Jian-Cai, Jiang Zhao-Hui, Gui Wei-Hua, Pan Dong, Xu Chuan, Zhou Ke. Key frame extraction method of blast furnace burden surface video based on state recognition. *Acta Automatica Sinica*, 2023, **49**(11): 2257−2271

收稿日期 2022-12-12 录用日期 2023-02-26

Manuscript received December 12, 2022; accepted February 26, 2023

国家重大科研仪器研制项目 (61927803), 国家自然科学基金基础科学中心项目 (61988101), 湖南省科技创新计划 (2021RC4054), 湖南省研究生科研创新项目 (CX20210243), 中南大学中央高校基本科研业务费专项资金 (2021zzts0184) 资助

Supported by the National Major Scientific Research Equipment of China (61927803), National Natural Science Foundation of China Basic Science Center Project (61988101), Science and Technology Innovation Program of Hunan Province (2021RC4054), Hunan Provincial Innovation Foundation for Postgraduate (CX20210243), and Fundamental Research Funds for the Central Universities of Central South University (2021zzts0184)

本文责任编委 陈胜勇

Recommended by Associate Editor CHEN Sheng-Yong

1. 中南大学自动化学院 长沙 410000 2. 鹏城实验室 深圳 518000

1. School of Automation, Central South University, Changsha 410000 2. Peng Cheng Laboratory, Shenzhen 518000

高炉炼铁是将含铁原料 (烧结矿、球团矿等) 在焦炭、石灰石以及从炉底鼓入的热风作用下还原成铁的复杂冶炼过程. 在高炉冶炼过程中, 炉料 (烧结矿、球团矿、焦炭等) 在热风浮力和自身重力的作用下会形成料面, 炉料在发生复杂物理化学反应后会形成熔融铁水并从铁口排出, 同时炉顶会周期性撒入炉料, 并维持料面在一定的高度范围内[1−2]. 炉内料面运动状态是高炉运行过程的关键参数之一, 对于指导炉顶布料操作、防止异常工况的发生具有重要的意义. 然而, 高炉是一个大型的黑箱子, 由于炉内高温高压、密闭弱光、强粉尘等恶劣的冶炼环境, 现场难以掌握炉内料面状态.

为了监测炉内反应状况, 现场会在高炉炉顶安装视频监测设备, 如红外热成像仪[3]、炉顶可见光摄像机以及高温工业内窥镜[4]. 红外热成像仪通过红外成像原理获取炉顶料面的温度信息, 能够在一定程度上获取料面反应状态, 但是容易受到高温煤气流的影响, 且损失了料面纹理信息; 为了监测炉内料面反应程度, 一般会在高炉炉顶安装可见光摄像机, 然而由于炉内粉尘和高温煤气流的遮挡, 只能拍摄到炉内的溜槽等极其有限的信息. 高温工业内窥镜是一种新型的料面视频采集设备, 其采用内窥式可见光成像技术近距离深入炉内实时获取料面视频. 考虑到高炉布料是周期性、间歇性的, 采集到的料面视频也呈现出周期性变化的规律. 此外, 料面视频存在信息冗余、图像质量参差不齐等问题, 现场人员需要花费较多的时间去检索视频中的关键信息, 如果在后期对所有视频帧都进行分析处理会大大降低效率并影响结果. 研究高炉料面视频关键帧自动提取方法能够从大量的料面视频中筛选出清晰的、稳定的视频帧, 提高后期视频处理的精度和效率.

视频关键帧的提取是减少视频冗余、剔除无效重复信息的重要手段, 在视频检索、视频压缩存储、视频摘要、视频分类、工业视频监测等领域具有广泛的应用[5]. 现有的视频关键帧提取方法可以分为基于图像特征的方法[6]、基于镜头检测的方法[7−8]、基于聚类的方法[9]、基于运动信息的方法[10]、基于学习的方法[11−13] 等. 基于图像特征的方法通过图像的颜色、纹理、亮度等底层特征, 计算这些特征的差异并与设定的阈值做比较, 这种方法计算量比较大, 计算效率并不高; 基于镜头检测的方法先将其分成许多镜头, 并选取第一帧、最后一帧以及中间固定的几帧作为关键帧, 这种方法计算量小, 但提取的关键帧数目相对固定, 且没有考虑视频内容; 基于聚类的方法先根据聚类算法将相似的帧划分为同一类, 然后从不同类别中选取出具有代表性的视频帧, 在关键帧提取中应用比较广泛. Singh 等[14] 提出了一种对光照变化和运动鲁棒的镜头边界检测方法, 基于局部二值模式傅里叶直方图 (Local binary pattern histogram Fourier, LBP-HF) 特征和自适应阈值检测出视频关键帧. 为了提高轻轨视觉定位的精度, 一种基于像素显著性分数的关键帧提取方法被提出, 通过滑动窗口提取当前帧图像块的显著性分数, 从而筛选出关键帧, 为场景跟踪提供合适的检索窗口[15]. 文献 [16] 提出了一种基于时空分析的运动状态自适应视频关键帧提取方法, 通过分析目标运动状态并选择运动状态变化的视频帧作为关键帧. Gharbi 等[17] 提出一种基于兴趣点、可重复性网络和模块化的局部描述方法, 并引入图聚类提取关键帧, 以最小化视频的信息冗余. Lai 等[18] 提出一种时间约束聚类算法, 对内容相似的帧进行分组, 并选择了显著值最大的视频帧作为关键帧. Wu 等[19] 提出了一种基于高密度峰值搜索 (Video representation based high density peaks search, VRHDPS) 聚类算法的视频表示方法, 并将聚类中心判定为关键帧. Chu 等[20] 提出了一种共聚类方法, 通过识别视频中共享的相似镜头来同时总结相同主题的多个视频.

随着神经网络的发展, 越来越多文献研究基于深度学习的关键帧提取方法[21−22]. 为了解决视频关键帧提取存在的上下文注意力不足和分布不一致的问题, Ji 等[23] 将监督视频摘要表述为序列到序列的学习框架, 提出了一种注意力和分布一致性学习方法, 利用真实的带标注的视频训练并预测视频帧的重要性分数, 从而获得关键帧. Abed 等[24] 构建了带标签的人脸图像数据集, 采用卷积神经网络 (Convolutional neural network, CNN) 有监督预测人脸图像质量, 并将其作为关键帧提取的依据. 为了更好地帮助运动员训练以及裁判打分, 一种用于运动视频的深度关键帧提取方法被提出, 该方法以标有运动员前景和背景区域的图像为数据集, 通过全卷积网络提取运动员的感兴趣区域, 并基于卷积神经网络估计每个帧的概率, 将概率最大的帧判别为关键帧[25]. Muhammad 等[26] 提出了一种基于深度学习的工业监控场景下的视频关键帧提取方法 (Deep learning-based resource-constrained video summarization, DeepReS), 采用 ORB (Oriented FAST and rotated BRIEF) 特征和卷积神经网络获得候选关键帧, 并根据用户需求筛选出熵值高的视频帧. 针对以查询为中心的视频摘要任务, Xiao 等[27] 提出了一种由局部注意力机制和查询相关性计算模块构成的卷积分层注意力网络, 计算视频镜头与查

询之间的相似性, 从而生成视频关键帧. Zhou 等[28]将视频关键帧提取描述成顺序决策的过程, 提出了一种基于强化学习的深度摘要网络 (Deep summarization network, DSN), 设计了一个无标签的多样性–代表性奖励函数, 获得具有多样性和代表性的视频关键帧. 尽管深度学习方法在视频关键帧的提取上表现出较好的性能, 但是一般依赖于充足的训练数据来学习视频帧中的重要帧的特征, 大多数方法需要对视频进行大量的人工标注, 模型的泛化性和可解释性不强, 并且缺乏对视频的全局理解. 考虑到高炉炉内恶劣的冶炼环境, 且料面形貌会受到高炉冶炼操作、高炉尺寸、工况、布料周期等因素的影响, 料面形貌具有多样性和不确定性, 很难构建完整的料面视频数据集, 标签数据也难以获得, 因此深度学习方法并不适用于高炉料面视频关键帧的提取. 此外, 与一般场景视频关键帧提取不同的是, 从周期性变化且质量不稳定的高炉料面视频获取关键帧更加复杂, 其对反映炉内运行状态、优化高炉炉顶布料操作具有重要的指导意义. 由于高炉料面视频的特殊性, 本文定义的高炉料面关键帧为视频中不同状态下的中心气流稳定、清晰、无炉料及粉尘遮挡且特征明显的图像序列.

为从大量视频中自动准确提取清晰稳定的高炉料面图像, 本文提出基于状态识别的高炉料面视频关键帧提取方法. 首先基于高温工业内窥镜采集高炉冶炼过程的料面视频, 并对料面视频进行分析, 将料面分为静止、缓慢下沉、快速下沉三个状态. 为剔除视频的冗余信息, 计算料面图像显著性区域的关键特征, 并提出基于局部密度极大值高斯混合模型 (Local density maxima-based Gaussian mixture model, LDGMM) 聚类的方法识别料面所处的状态, 根据料面所处的状态实现每个布料周期不同状态下中心气流稳定、清晰、无炉料及粉尘遮挡且图像特征明显的关键帧的自动判别.

1　高炉料面视频采集及分析

1.1　料面视频采集系统

在高炉冶炼过程中, 通常炉顶的温度会达到 200 ~ 400 ℃, 中心局部温度可能会达到 600 ~ 800 ℃. 由于炉料下降与煤气流的相对运动, 会产生大量的粉尘和噪声, 而高炉冶炼过程中产生的高温煤气流有毒有害, 不能随意排出, 因此高炉需要严格密闭, 炉内无法提供额外的光源条件. 高炉炉顶高温高压、密闭弱光、强粉尘多噪声的恶劣环境特点, 导致清晰的料面图像难以获取. 为了克服炉内恶劣的冶炼环境并采集到高炉冶炼过程中的料面视频, 采用了某炼铁厂安装的一套实时高炉料面视频采集系统[29]. 该系统包括用于采集炉内料面视频的高温工业内窥镜、保证设备长期稳定运行的水–气双冷却管、用于信号传输的光纤及防爆控制箱、位于中控室的上位机等, 如图 1 所示. 得益于采集系统的冷却装置及具有极强的微光成像能力的光学性能[30], 料面采集系统能够在高温高压密闭弱光的炉内环境中安全稳定运行并实时采集高炉冶炼过程中的料面视频, 设备安装的实物示意图如图 2 所示.

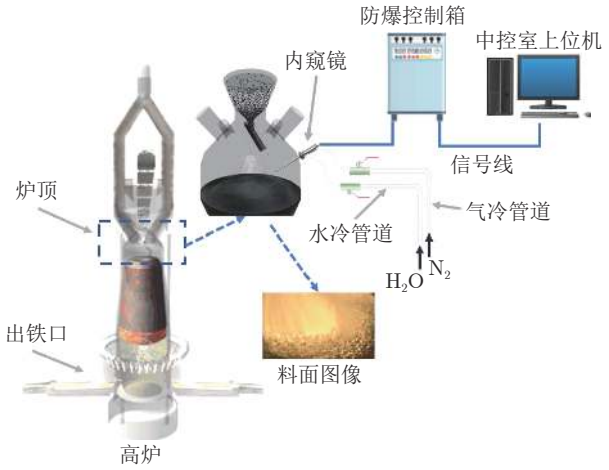


图 1　料面视频采集系统

Fig. 1　Burden surface video acquisition system

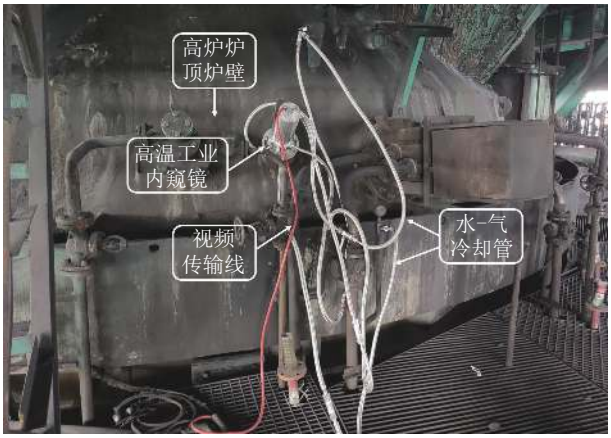


图 2　设备安装图

Fig. 2　Equipment installation diagram

通过高炉料面视频采集系统采集到了大量高炉冶炼过程中的料面视频, 为现场操作人员认识炉内料面反应状况、指导炉顶布料操作提供了直观的视频图像信息.

1.2　料面反应新现象

高炉炼铁是一个周期性加料、间歇性出铁的复

杂生产过程, 可以将高炉一个布料周期分为布料过程和布料间歇, 其中布料过程包括环形布料和中心加焦, 而高炉运行过程中炉料会在焦炭等助燃剂的作用下及炉底鼓入的热风环境中发生复杂的物理化学反应, 并生成铁水从铁口排出. 通过安装的高炉料面视频采集系统采集到的料面视频清晰地监测到炉料周期性的反应过程, 并揭示了料面变化新现象. 图 3 给出了采集到的某炼铁厂 2#高炉一个布料周期内的料面反应现象.

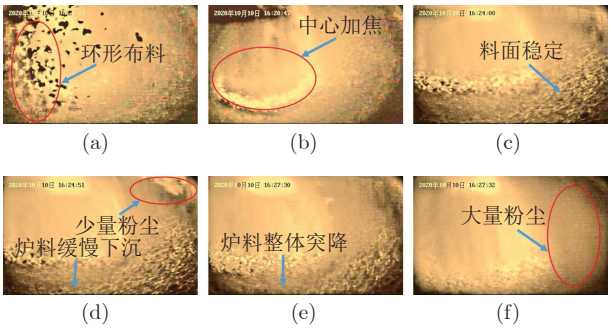


图 3 一个布料周期内的料面反应现象

Fig. 3 Burden surface reaction phenomenon in a burden cycle

在高炉布料过程中, 当料面料位到达预设的最低位置, 通过设定的布料矩阵将炉料环形撒入炉内, 此时炉料会遮挡料面有效信息, 如图 3(a) 所示. 在加料之后会将焦炭以另一个布料矩阵撒入高炉中心, 中心气流逐渐变窄, 并最终被焦炭完全遮挡, 如图 3(b) 所示. 随之, 中心气流宽度会逐渐变大, 反映了中心加焦能够改善料面中心的透气性. 此时料面会存在一段时间的布料间歇, 料面视频较清晰, 炉料颗粒清晰可见. 从采集的料面视频可以发现, 在正常工况下, 料面存在静止、缓慢下沉、快速下沉三种不同的状态, 一般来说, 静止料面图像清晰度最高, 如图 3(c) 所示. 炉料缓慢下沉会产生少量的粉尘, 而快速下沉会由于运动导致图像出现一定的模糊, 在快速下沉之后, 会出现大量的粉尘遮挡住料面的有效信息, 如图 3(d) ~ 3(f). 此外, 从图 3 中可以发现, 布料过程的料面有效信息被遮挡较多, 而布料间歇的料面图像清晰度要高于布料过程. 采集到的料面视频监测了高炉冶炼过程中料面变化的新现象, 为现场操作人员认清炉内反应状况提供了重要的信息.

考虑到高炉炼铁是一个大时滞、慢变的过程, 高炉布料具有周期性和间歇性, 并且料面形貌与布料制度、炉况等因素有关. 为了更全面地分析料面视频的特征, 采用抽帧方式 (每间隔 600 帧抽取一帧, 即采样频率为 3 帧/min) 展示了高炉一个布料周期约为 6 min 的料面视频帧, 如图 4 所示. 图 4(a) ~ 4(h) 是高炉布料过程的料面图像, 图 4(i) ~ 4(r) 是高炉布料间歇的料面图像. 在高炉布料过程, 料面变化较大, 视频中的炉料由少到多, 并逐渐遮挡整幅图像, 采用环形布料一定圈数之后, 料面粉尘量增大; 而在中心加焦之后, 料面气流宽度变窄, 而炉内无额外光照, 料面图像中充满噪点, 缺乏有效信息. 相比于布料过程的图像, 布料间歇的图像更加清晰、直观, 包含的料面有效信息更加丰富, 这是为布料操作提供信息反馈及后续图像分析的主要数据来源.

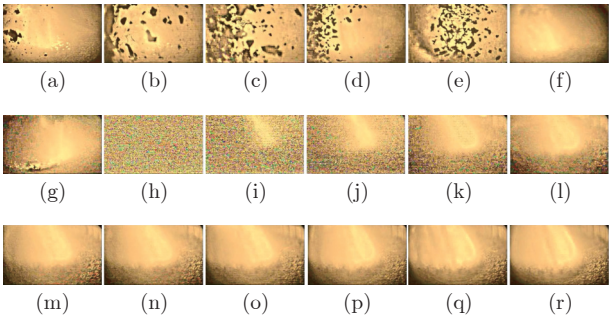


图 4 一个布料周期内料面视频帧

Fig. 4 Burden surface video frames in a burden cycle

与一般场景视频不同, 料面视频具有状态多变 (静止、缓慢下沉、快速下沉)、形貌多样 (中心平坦型、中心井口型、中心喷溅型等)、图像质量不一 (周期性变化, 布料过程图像清晰度低, 有效信息较少; 而布料间歇图像清晰度较高, 能够监测到炉内反应状况)、信息重复冗余 (料面是慢变的, 视频中存在很多重复、冗余的帧) 等特点. 从现场采集的料面视频无法直接用于分析理解, 且如果对所有视频帧都进行处理会大大降低效率并影响处理的结果, 为此有必要研究一种能够自动从料面视频中提取关键帧的方法. 考虑到料面所处的状态与高炉炉况及图像清晰度息息相关, 将料面所处的状态作为关键帧提取的依据更加符合视频的特征及现场的需求.

2 高炉料面视频关键帧提取

2.1 料面显著性区域关键特征提取

为了从周期性变化、质量不一、状态多变的高炉料面视频中提取出关键帧, 本文将不同状态下中心气流稳定、清晰、无炉料及粉尘遮挡且特征明显的图像定义为视频关键帧. 料面图像的特征点密集程度及像素位移大小与高炉料面的运动状态具有很强的相关性. 一般来说, 料面图像特征点越密集, 像

素位移越小, 料面下降越慢, 料面越稳定; 而图像特征点越稀疏, 像素位移越大, 料面下降越快, 料面越不稳定. 为此本文选择料面图像特征点密集程度及运动的像素位移作为料面视频关键帧提取的两个关键特征, 其中特征点的密集程度采用特征点数目表示, 图像运动的像素位移用平均光流矢量表示.

基于高温工业内窥镜拍摄的料面图像如图 5(a) 所示, 高炉炉内恶劣的冶炼环境导致料面图像与一般场景图像不同. 料面图像包含煤气流区域、炉壁区域等非料面区域和料面区域, 并且不易分割, 其中料面区域是本文的感兴趣区域. 图像中煤气流区域过曝, 而远离煤气流区域偏暗. 在计算料面关键特征时, 非料面区域会对特征提取的精度造成影响, 且高炉炉内煤气流是变化的, 基于经典的光流法计算出的料面图像光流矢量具有一定的误差. 受启发于人类的视觉注意力机制, 为了提高关键特征提取的精度和效率, 首先对图像进行显著性检测, 选择图像中具有重要信息的区域[31]. 本文定义图像中亮度均匀、纹理明显的料面区域为显著性区域, 并基于图像像素间的颜色特征、纹理特征、亮度特征以及空间特征的差异度作为显著值计算的依据. 像素 p_i, p_j 间的总体差异度 $d(p_i, p_j)$ 定义如下

$$d(p_i, p_j) = \frac{\sum_{t=1,2,3} d_t(p_i, p_j)}{1 + \alpha d_s(p_i, p_j)} \tag{1}$$

其中, $d_1(p_i, p_j)$, $d_2(p_i, p_j)$, $d_3(p_i, p_j)$ 分别表示料面图像像素 p_i, p_j 之间的颜色特征、亮度特征和纹理特征的欧氏距离, $d_s(p_i, p_j)$ 表示像素 p_i, p_j 之间空间特征的欧氏距离, α 表示一个常数, 定义为 3. 式 (1) 可以理解为显著性区域是颜色特征、亮度特征、纹理特征差异较大并且位置相对集中的区域.

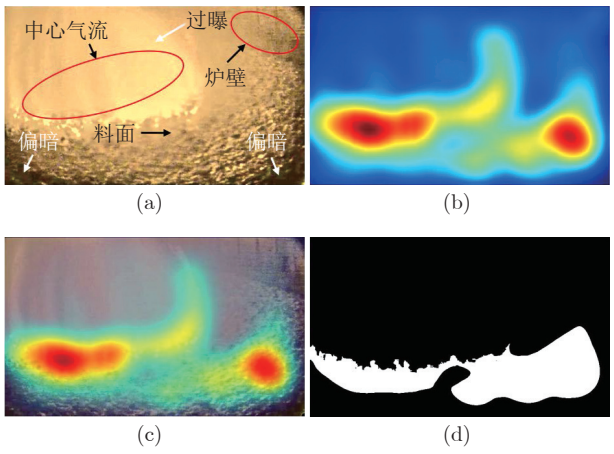


图 5　料面显著性区域提取

Fig. 5　Salient region extraction of burden surface

在得到像素间的差异度之后, 像素 p_i 的显著值 S_i 可以表示为

$$S_i = 1 - \exp\left(-\frac{1}{L}\sum_{l=1}^{L} d(p_i, p_l)\right) \tag{2}$$

其中, p_l 表示以像素 p_i 为中心的 8×8 邻域内的像素点, $d(p_i, p_l)$ 越大, S_i 越大.

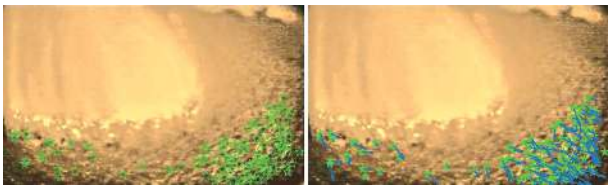
料面显著值计算结果如图 5(b) 和 5(c) 所示, 其中红色越深, 表示显著值越大, 蓝色越深表示显著值越小. 为了进一步减小煤气流区域对精度的影响, 采用 OTSU 阈值分割方法对料面图像进行二值化, 将料面图像分为煤气流区域和非煤气流区域, 并作为掩膜剔除掉显著性区域的煤气流区域, 定位的显著性区域的位置如图 5(d) 中的白色区域.

在定位料面显著性区域之后, 考虑料面运动越快, 光流矢量越大, 而由于运动, 图像清晰度会有所下降, 从而检测到的特征点数目减少, 因此计算特征点密集程度和特征点光流矢量作为料面运动状态识别的关键特征. 特征点集合及特征点光流矢量集合如式 (3) 和式 (4) 所示, 提取结果如图 6 所示

$$K = \{k_1, k_2, k_3, \cdots, k_m\} \tag{3}$$

$$F = \{f_{k_1}, f_{k_2}, f_{k_3}, \cdots, f_{k_m}\} \tag{4}$$

其中, K 表示料面图像特征点集合, F 表示图像的特征点光流矢量的集合.



(a) 提取的特征点
(a) Extracted feature points

(b) 特征点光流矢量
(b) Feature point optical flow vectors

图 6　提取的特征点及其光流矢量

Fig. 6　Extracted feature points and their corresponding optical flow vectors

2.2　基于局部密度极大值 GMM 的料面状态识别

在获取料面图像显著性区域的特征点密集程度及特征点运动的像素位移之后, 需要基于关键特征从视频中判别出料面所处的运动状态. 考虑到高斯混合模型 (Gaussian mixture model, GMM) 可以拟合任意数据的分布, 而传统的 GMM 对于初始值的选取十分敏感, 容易因初始值的选取不佳而陷入局部最优, 无法获得准确的结果. 为了准确识别料面所处的状态, 本文提出了 LDGMM 料面状态识

别方法, 可以在无参数的条件下基于数据分布特征, 找到局部密度最大的点, 并作为 LDGMM输入的初始值, 以加快 LDGMM 迭代的速度, 优化 LDGMM 拟合的结果, 具体步骤如下:

1) 考虑到正常工况下, 料面是相对稳定的, 料面发生整体运动, 为了简化运算, 以特征点平均光流矢量作为某个时刻料面的运动像素位移

$$f = \frac{1}{m}\sum_{a=1}^{m}\sqrt{u_a^2 + v_a^2} \tag{5}$$

其中, f 表示特征点平均光流矢量, m 表示图像特征点的数目, u_a, v_a 分别表示图像特征点 a 的像素位移水平分量和垂直分量.

2) 由于料面视频中包含不稳定的状态, 在布料、加焦和炉料快速下沉后会激起大量的粉尘. 料面的有效信息被炉料或者粉尘遮挡, 此时料面图像的特征极其不明显, 给定较小的表征特征点数目的阈值剔除这个时刻的数据 (通过大量的实验, 本文的阈值选择了 10, 即 $T_0 = 10$), 然后采用等置信概率剔除其他时刻的一些异常值, 并对特征点数目及像素位移矢量进行归一化处理

$$\begin{cases} |d_0 - m(d)| > w \times s(d) \\ d^* = \dfrac{d_i - d_{\min}}{d_{\max} - d_{\min}} \end{cases} \tag{6}$$

其中, w 表示肖维勒系数, d_0, $m(d)$, $s(d)$ 分别表示数据离群点、均值、标准差, d^*, d_i, $d_{\min}$, $d_{\max}$ 分别表示归一化后的数据、归一化前的数据、数据最小值和最大值.

3) 在数据预处理之后, 总共获取了不同布料周期不同状态下的料面图像数据 1 500 个, 采用核密度估计函数得到数据分布的概率密度图, 核密度估计函数如下

$$f_h(d) = \frac{1}{nh}\sum_{j=1}^{n} K_h\left(\frac{d - d_j}{h}\right) \tag{7}$$

其中, $f_h(\cdot)$ 表示核密度估计函数, $K_h(\cdot)$ 表示核函数, h 表示平滑参数, 满足 $h > 0$, d 表示观测值, d_j 表示第 j 个样本, n 表示样本的总数.

特征点光流概率密度分布如图 7 所示, 其中不同的颜色表征数据的密度大小. 可以发现, 所提取的料面关键特征 (特征点密集程度和平均光流矢量) 基本满足对角分布规律, 即特征点越密集, 平均光流矢量越小; 特征点越稀疏, 平均光流矢量越大, 获取的料面图像关键特征能够被比较明显地分为三个密集分布. 根据光流矢量的分布可以得到概率密度分布直方图, 如图 8 所示, 可以得到三个概率密

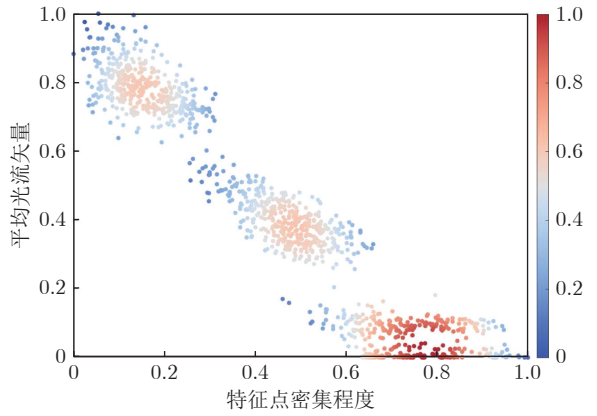


图 7 特征点密集程度和平均光流矢量的概率密度分布图

Fig. 7 Probability density distribution map of feature point density and average optical flow vectors

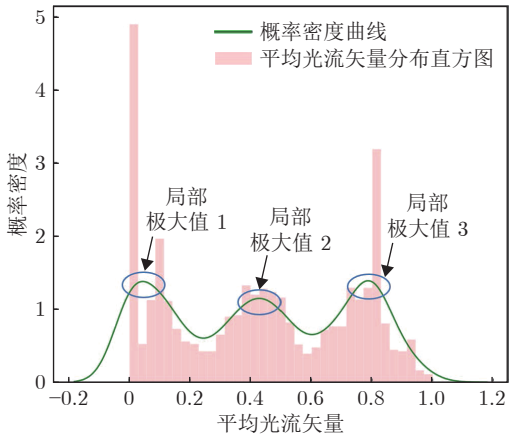


图 8 特征点平均光流矢量概率分布直方图

Fig. 8 Probability distribution histogram of average feature point optical flow vectors

度局部极大值, 作为 GMM 的初始聚类中心. 值得注意的是, 这是一种非参数估计方法, 可以在不知道具体类别数目的前提下仅根据关键特征的数据分布来估计出数据的概率密度分布, 并从中获得高斯分布数量和局部密度极大值

$$C = \{c_1(n_1, f_1), c_2(n_2, f_2), c_3(n_3, f_3)\} \tag{8}$$

其中, C 表示局部密度极大值点集合, c_1, c_2, c_3 表示三个局部密度极大值点, (n_1, f_1), (n_2, f_2), (n_3, f_3) 表示三个局部密度极大值对应的特征点密集程度及平均光流矢量.

4) 将上一步获取的局部密度最大的三个点作为初始的聚类中心, 根据局部密度极大值设定 GMM 参数的初始值, GMM 的概率密度函数如下

$$P(x) = \sum_{\gamma=1}^{K}\alpha_\gamma p\left(\frac{x}{\mu_\gamma}, \sum_\gamma\right), \sum_{\gamma=1}^{K}\alpha_\gamma = 1 \tag{9}$$

其中, $P(x)$ 表示 GMM 的概率密度函数, $p(\cdot)$ 表示单个高斯模型的概率密度函数. α_γ, μ_γ, $\sum_\gamma$ 是 GMM 的三个参数, 分别表示当前帧属于类别 γ 的概率、第 γ 个高斯分布的均值和协方差矩阵.

采用期望最大化 (Expectation-maximization, EM) 算法对 GMM 参数进行更新, 直至似然函数达到最大, 模型收敛

$$f\left(\alpha_\gamma, \mu_\gamma, \sum\nolimits_\gamma\right) = \arg \max_{\alpha_\gamma, \mu_\gamma, \sum_\gamma} L\left(\alpha_\gamma, \mu_\gamma, \sum\nolimits_\gamma\right) = \arg\max \sum_{\lambda=1}^{n} \ln P(x_\lambda) \tag{10}$$

其中, $f(\alpha_\gamma, \mu_\gamma, \sum_\gamma)$ 表示模型优化的目标函数, $L(\alpha_\gamma, \mu_\gamma, \sum_\gamma)$ 表示对数似然函数, x_λ 表示第 λ 个数据.

图 9 给出了 5 次采用经典 GMM 和所提方法训练过程, 可以发现 GMM 平均需要 41.4 次更新才能达到最优, 并且容易陷入局部最优, 而所提聚类模型只需要迭代 13 次就可以迅速达到最优, 一定程度上验证了所提模型是高效准确的.

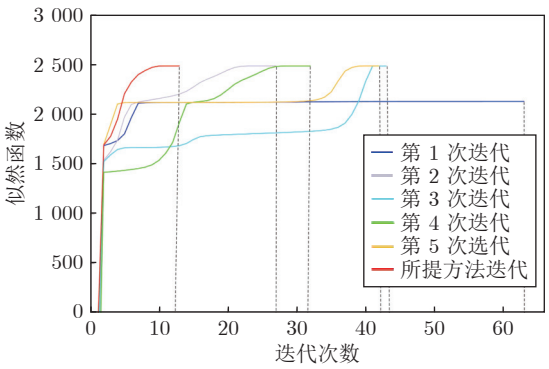


图 9　不同方法不同训练过程似然函数变化情况

Fig. 9　Change of likelihood function of different methods in multiple training processes

2.3　基于状态识别的料面关键帧提取

料面所处的状态是提取关键帧的主要依据, 在识别料面所处的状态后, 可以根据后期不同的处理需求, 得到满足不同条件的关键帧集合, 本文提取了不同状态下中心气流稳定、清晰、无炉料及粉尘遮挡且特征明显的关键帧.

考虑到高炉炉顶布料操作是周期性、间歇性的, 且操作制度在不同周期会有所改变, 采集到的高炉料面视频在不同的布料周期的形貌、图像清晰度等具有较大的区分度, 不能对整个料面视频直接进行关键帧提取. 为此, 本文根据布料周期对料面视频进行镜头分割. 值得注意的是, 高炉布料周期 T 并不是固定的, 且 T 未知, 周期不能简单地根据时间进行划分. 从采集的料面视频可以发现, 高炉由布料间歇到布料的切换过程中, 图像特征变化明显, 图像有效信息降低, 清晰度明显下降. 本文基于复合帧间差异度函数获取料面视频中的突变帧进而实现料面视频不同周期的识别, 其中复合帧间差异度函数由视频帧的熵、平均梯度、特征点密集程度的差异组成.

$$\begin{cases} E = -\sum\limits_{i=1}^{n_k} p(x_i)\log_2 p(x_i) \\ G = \dfrac{\sqrt{2}}{2(w-1)(h-1)} \sum\limits_{x=1}^{w-1}\sum\limits_{y=1}^{h-1} \sqrt{\left(\dfrac{\partial I}{\partial x}\right)^2 + \left(\dfrac{\partial I}{\partial y}\right)^2} \\ N = \dfrac{m - m_{\min}}{m_{\max} - m_{\min}} \end{cases} \tag{11}$$

其中, E, G, N 分别表示视频帧的熵、平均梯度和特征点密集程度; $p(x_i)$ 表示灰度值出现概率; w, h 表示图像的宽和高; I 表示图像灰度值; $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial y}$ 分别表示图像水平和垂直方向的梯度; m, $m_{\min}$, $m_{\max}$ 分别表示当前帧图像特征点数目, 最小特征点数目和最大特征点数目.

为了减小视频帧关键特征计算的随机误差, 取每秒所有图像熵的平均值、平均梯度的平均值和特征点密集程度的平均值作为当前时刻的特征值, 并分别归一化到 (0, 1], 然后进行加权得到计算的结果, 计算相邻时刻图像的帧间差异度

$$S_t = \alpha_1 E_t + \alpha_2 G_t + \alpha_3 N_t \tag{12}$$

$$D_t = S_{t+1} - S_t \tag{13}$$

其中, E_t, G_t, N_t 分别表示视频帧归一化后的熵、平均梯度和特征点密集程度. S_t 表示图像特征值, D_t 表示视频帧间差异度, α_1, α_2, α_3 表示权重, 本文取 $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$.

视频帧间差异度曲线如图 10 所示, 图中给出几个不同时刻下的料面图像. 从图 10 可以发现, 在布料或者加焦过程中, 视频帧间差异度较大, 图像有效信息被粉尘或者炉料遮挡; 而在布料间歇料面逐渐稳定, 视频帧间差异度较小, 图像更加清晰. 基于视频帧间差异度可以检测到视频的边界帧, 进而对料面视频进行周期划分, 边界帧需要满足以下条件

$$\begin{cases} D_c \geq T_1, \quad D_{c+1} \geq T_1 \\ t_{c+1} - t_c > T_t \end{cases} \tag{14}$$

其中, D_{c+1}, D_c 表示相邻的两个边界帧, 边界帧的差异度需要大于阈值 T_1, 并且需要至少间隔 T_t 以避

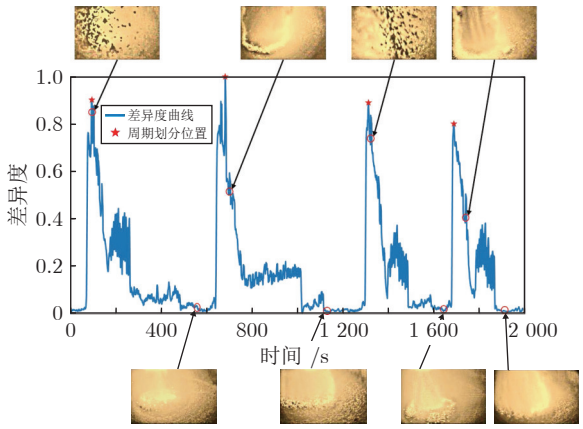


图 10 料面视频差异度曲线

Fig. 10 Difference curve of burden surface video

免同一个布料周期多个边界帧被检测到. T_1, T_t 是固定的阈值, 通过实验统计得到.

在高炉炉顶布料过程中, 会产生大量的粉尘、噪声, 由于炉料、粉尘、噪声等对料面有效信息的遮挡, 图像清晰度下降, 能够检测的特征点将大大减少, 这一类图像并不适合作为后期处理的数据来源. 为了保证所提取的关键帧特征明显、图像清晰, 在关键帧提取前剔除有效信息缺失的料面图像. 为了简化运算并提高效率, 统计当前帧的特征点数目 n_t, 并与阈值 T_0 比较, 若 $n_t \le T_0$, 则不属于关键帧, 直接剔除; 若 $n_t > T_0$, 则当前帧进入下一步判断, 从而剔除特征不明显、有效信息缺失较多的视频帧

$$\begin{cases} n_t > T_0,\ n_t \in V \\ n_t \le T_0,\ n_t \notin V \end{cases} \tag{15}$$

其中, n_t 表示当前帧特征点数目, V 表示初步筛选后留下的图像集合.

计算 V 中视频帧与不同聚类中心的欧氏距离, 获取当前帧所处的状态, 若为静止状态, 则被判定为周期 T 的关键帧集合 A 中的候选关键帧; 若为缓慢下沉状态, 则被判定为周期 T 的关键帧集合 B 中的候选关键帧; 若处于快速下沉状态, 则初步判定为周期 T 的关键帧集合 C 中的候选关键帧

$$\begin{cases} D_i(n_{t_f},\ f_{t_f}) = \sqrt{(n_{t_f} - n_{c_i})^2 + (f_{t_f} - f_{c_i})^2} \\ \min D_i(n_{t_f},\ f_{t_f}) = \begin{cases} D_1(n_{t_f},\ f_{t_f}),\ t_f \in A \\ D_2(n_{t_f},\ f_{t_f}),\ t_f \in B \\ D_3(n_{t_f},\ f_{t_f}),\ t_f \in C \end{cases} \end{cases} \tag{16}$$

其中, $(n_{c_i},\ f_{c_i})$, $i = 1, 2, 3$ 分别表示静止、缓慢下沉、快速下沉三种运动状态的聚类中心, t_f 表示当前时刻帧, $D_i(n_{t_f},\ f_{t_f})$ 表示当前时刻视频帧与三个聚类中心的欧氏距离.

从三个候选关键帧集合中分别将聚类中心对应的视频帧、按照固定采样频率得到的图像作为关键帧, 得到一个周期的关键帧集合

$$\begin{cases} A_t = \{a_{k_1},\ a_{k_2},\ a_{k_3},\ \cdots,\ a_{k_k}\} \\ B_t = \{b_{k_1},\ b_{k_2},\ b_{k_3},\ \cdots,\ b_{k_m}\} \\ C_t = \{c_{k_1},\ c_{k_2},\ a_{k_3},\ \cdots,\ a_{k_n}\} \end{cases} \tag{17}$$

其中, A_t, B_t, C_t 分别表示一个布料周期内的静止、缓慢下沉、快速下沉三种运动状态的关键帧集合.

一个布料周期的料面视频关键帧提取方法如算法 1 所示. 在提取了一个布料周期料面视频关键帧之后, 采用相同方法获取不同布料周期的料面视频关键帧, 得到最终的关键帧集合.

算法 1. 一个布料周期的料面视频关键帧提取

输入. n_e, t_f, N_m, T_h //帧数, 当前帧, 总帧数, 当前周期

输出. A_t, B_t, C_t //不同状态下的关键帧集合

1) **for** $n_e = 1, 2, 3, \cdots, N_m$ **do**

2) **while** $n_t > T_0$ **do**

3) Count $D_t \leftarrow S_{t+1} - S_t$ //计算帧间差异度

4) **if** $D_c < T_1$ 或 $D_{c+1} < T_1$ 或 $t_{c+1} - t_c \le T_t$ **then**

5) Set $T_h \leftarrow T_h$

6) Set $\min D_i(n_{t_f},\ f_{t_f}) = \begin{cases} D_1(n_{t_f},\ f_{t_f}),\ t_f \in A \\ D_2(n_{t_f},\ f_{t_f}),\ t_f \in B \\ D_3(n_{t_f},\ f_{t_f}),\ t_f \in C \end{cases}$

//判断当前帧所处状态

7) **else**

8) Set $T_h \leftarrow T_h + 1$

9) Sample from A, B, C //获得不同状态下的关键帧集合

$$A_t \leftarrow \{a_{k_1},\ a_{k_2},\ a_{k_3},\ \cdots,\ a_{k_k}\}$$

$$B_t \leftarrow \{b_{k_1},\ b_{k_2},\ b_{k_3},\ \cdots,\ b_{k_m}\}$$

$$C_t \leftarrow \{c_{k_1},\ c_{k_2},\ a_{k_3},\ \cdots,\ a_{k_n}\}$$

10) **end if**

11) **end while**

12) **end for**

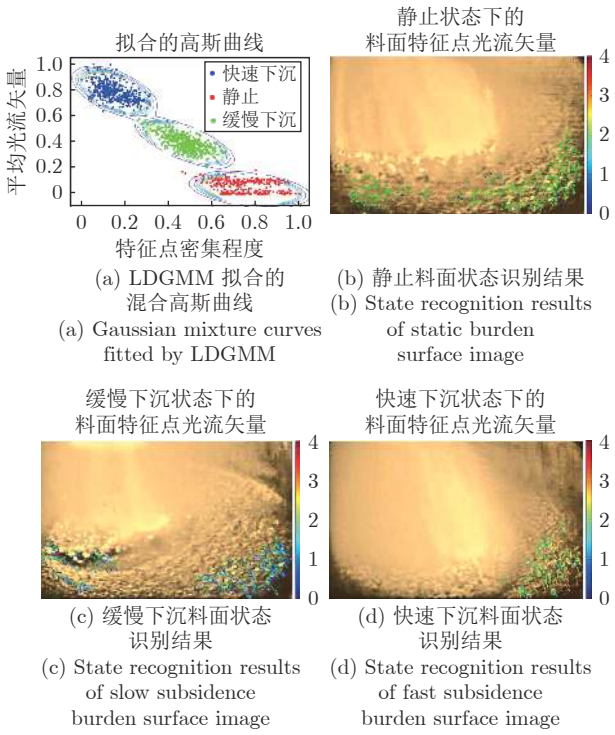
13) **return** A_t, B_t, C_t

3 实验验证与结果分析

本文实验的软件环境: Windows10 操作系统下的 MATLAB R2020a 及 PyCharm 2020; 硬件环境: 处理器为 Intel Core i7-11 700 2.50 GHz, 内存为 32 GB. 测试的视频来源于高温工业内窥镜采集到的某钢铁厂 2#高炉运行过程的料面视频, 采样频率为 30 帧/秒, 图像分辨率大小为 828 × 494.

3.1　状态识别的性能分析

在采集高炉运行过程不同布料周期的料面视频之后, 提取图像的特征点并计算特征点光流矢量, 基于所提的 LDGMM 识别当前帧的运动状态, 结果如图 11 所示, 其中图 11(a) 表示 LDGMM 识别结果, 将高斯分布曲线内或者接近曲线的数据归为一个状态, 从图中可以得出, 三种不同的状态可以比较明显地被区分开. 图 11(b) ~ 11(d) 分别表示三个不同布料周期中静止、缓慢下沉、快速下沉三种状态的可视化识别结果. 为了能够更加直观地区分不同的状态, 采用不同颜色、不同长度的箭头表示料面特征点运动的大小, 其中红色越深, 箭头越长, 表示特征点运动的光流矢量越大; 蓝色越深, 箭头越短, 表示特征点运动的光流矢量越小. 从图 11 可以获得料面上不同位置的特征点像素位移大小, 此外, 静止料面能够提取更多的特征点, 其对应的光流矢量较小; 相比于静止料面, 缓慢下沉的料面提取的特征点数量变少, 但光流矢量变大; 而快速下沉的料面光流矢量最大, 提取的特征点数量最少, 直观上说明所提方法能够明显区分出料面的不同状态.



(a) LDGMM 拟合的混合高斯曲线
(a) Gaussian mixture curves fitted by LDGMM

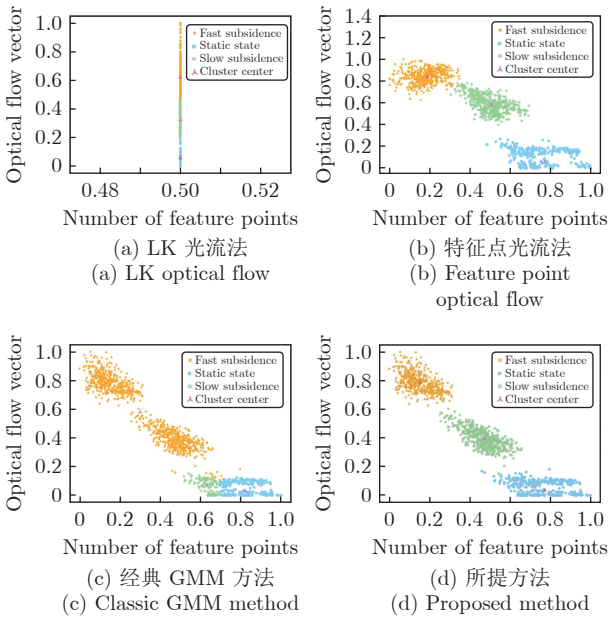
(b) 静止料面状态识别结果
(b) State recognition results of static burden surface image

(c) 缓慢下沉料面状态识别结果
(c) State recognition results of slow subsidence burden surface image

(d) 快速下沉料面状态识别结果
(d) State recognition results of fast subsidence burden surface image

图 11　不同布料周期不同状态的识别结果
Fig. 11　Various state recognition results in different burden cycles

为了验证本文所提状态识别方法的有效性, 进行消融实验, 分别比较 LK (Lucas-Kanade) 光流法、无显著性区域定位的特征点光流法、经典 GMM 算法以及本文方法得到的聚类结果, 如图 12 所示. 可以发现, LK 光流法只依据图像的像素位移进行状态的划分, 数据的区分度并不大, 如图 12(a) 所示; 在不引入显著性区域定位时, 由于图像包含炉壁和煤气流区域, 提取的部分特征点并不在料面区域, 计算的像素位移大小可能会比实际值大, 并且快速下沉和缓慢下沉两个状态分离度不高, 如图 12(b) 所示; 由于经典 GMM 算法初始的聚类中心是随机的, 若初始聚类中心选择不当, 会使得模型陷入局部最优, 如图 12(c) 所示; 采用本文所提方法得到的结果中三种状态的分离较明显, 能够快速达到最优的聚类效果, 不同状态区分度较大, 如图 12(d) 所示.



(a) LK 光流法
(a) LK optical flow

(b) 特征点光流法
(b) Feature point optical flow

(c) 经典 GMM 方法
(c) Classic GMM method

(d) 所提方法
(d) Proposed method

图 12　相同料面图像不同方法的聚类结果
Fig. 12　Clustering results of different methods for the same burden surface image

为了评估不同方法的聚类结果, 选取戴维森堡丁 (Davies-Bouldin, DB) 指数、CH (Calinski-Harabasz) 指数、轮廓系数 (Silhouette coefficient, SC)、分离度 (Separation, SP) 作为评价指标, 比较基于不同方法得到的不同状态的类间距离和类内距离, 其中 DB 值越小, CH、SC 及 SP 值越大, 说明聚类效果越好, 评价结果如表 1 所示. 从表 1 可以看出, 所提方法的 DB 指标值最小, CH、SC 和 SP 指标值最大, 综合分析可得所提方法的类间数据最分散, 类内数据最集中, 状态聚类的效果在这几种方法里是最好的.

进一步地, 采用 Cai 等[32] 提出的动态光照下光流计算方法 (Dynamic illumination optical flow computing, DIFlow)、Liu 等[33] 提出的从无标注数

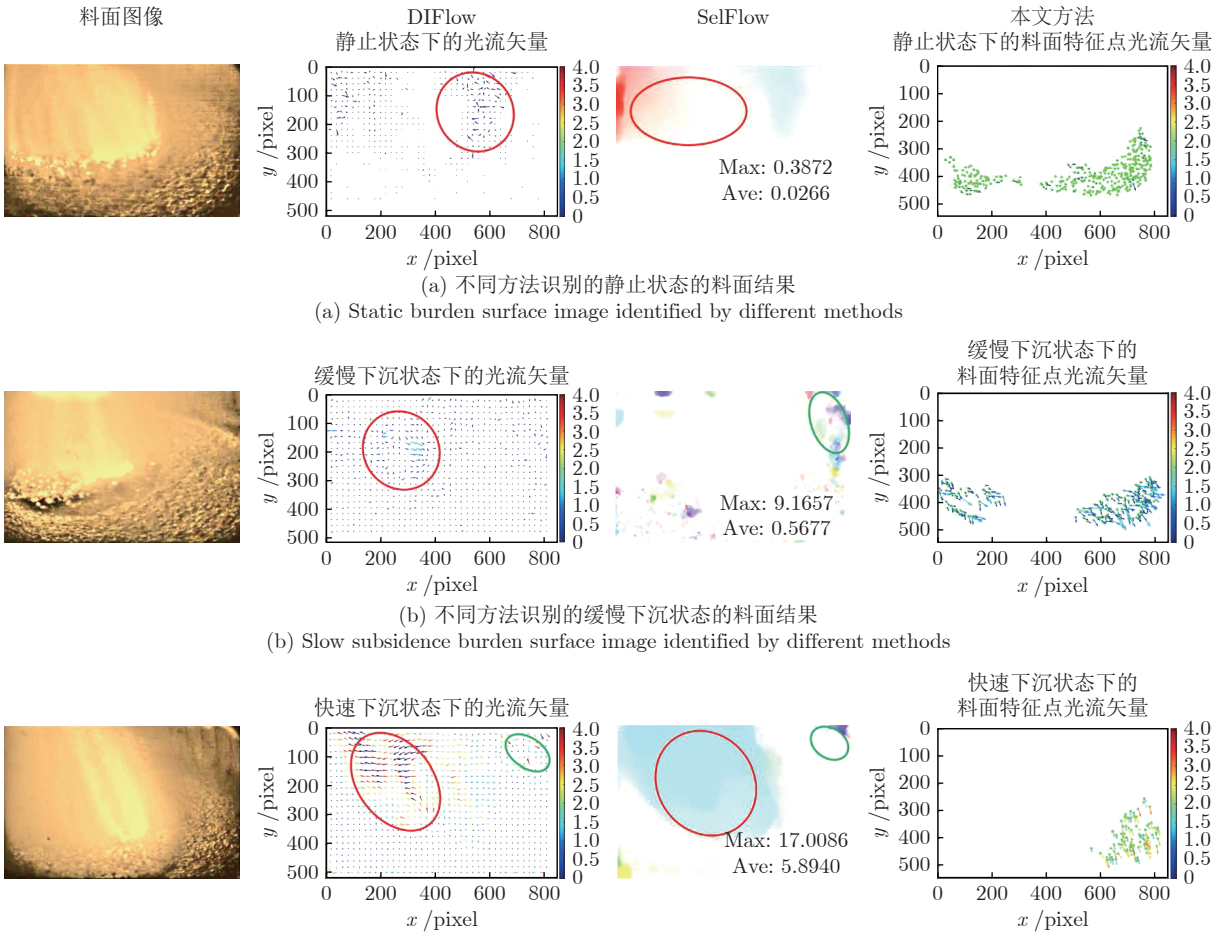
表 1 不同方法的聚类效果比较

Table 1 Comparison of clustering performance of different methods

	指标			
	DB	CH	SC	SP
LK 光流	0.2603	474.41	0.9826	0.6013
特征点光流	0.2867	5392.80	0.9949	0.7129
GMM	0.1376	1347.30	0.9816	0.8018
本文方法	0.0010	7762.36	0.9989	0.9537

据中学习光流自监督网络 (Self-supervised learning of optical flow, SelFlow) 及本文所提方法识别相同料面视频的不同状态, 不同方法识别不同料面状态的结果如图 13 所示. DIFlow 考虑了光照对光流估计的影响, 将料面图像均匀地分成图像块, 并计算了料面图像块的像素位移; SelFlow 采用无监督深度学习方法计算了整幅料面图像的稠密光流. 在图 13 中, 给出了 SelFlow 计算的最大光流矢量值和平均光流矢量值, 并采用不同颜色的椭圆标记出存在较大误差的位置, 其中红色椭圆部分表示计算的中心气流的光流矢量, 绿色椭圆部分表示计算的炉壁部分的光流矢量. 从图 13 可以看出, DIFlow 及 SelFlow 方法会受到中心气流和炉壁等非料面区域的影响, 计算的光流结果偏大, 料面状态识别的结果并不直观, 而本文方法能够极大减少非料面区域对识别结果的影响, 可以明显区分出不同的料面状态.

在高炉冶炼现场, 为了能够检测料位, 一般会在高炉炉顶安装 2 ~ 3 个机械探尺. 机械探尺通过提尺、放尺、浮尺三个阶段按照一定的采样频率 (约 10 s 采样一次) 来测量料面某个位置上的料位, 其中浮尺阶段的数据被认为是真实有效的料面料位值. 根据机械探尺的测量值能够得到布料间歇料面的高度, 结合相邻采样间隔的机械探尺的测量数据及人工经验获取了料面的状态. 采用不同方法的识



(a) 不同方法识别的静止状态的料面结果

(a) Static burden surface image identified by different methods

(b) 不同方法识别的缓慢下沉状态的料面结果

(b) Slow subsidence burden surface image identified by different methods

(c) 不同方法识别的快速下沉状态的料面结果

(c) Fast subsidence burden surface image identified by different methods

图 13 不同方法识别不同料面状态的对比结果

Fig. 13 Comparison results of different burden surface states identified by different methods

别结果如图 14 所示, 从图中可以发现所提方法的识别状态与真实状态吻合较好, 而 DIFlow 和 SelFlow 识别的料面状态与真实状态有较大的偏差, 一定程度上验证了所提方法是可靠的, 并且优于 DIFlow 和 SelFlow 方法.

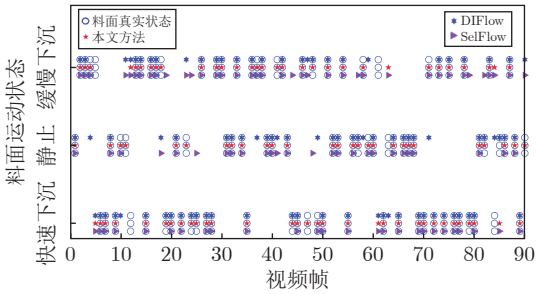


图 14　不同方法识别的不同状态

Fig. 14　Different states identified by different methods

为了定量验证所提状态识别方法的精度, 分别选取调整兰德系数 (Adjusted Rand index, ARI)、标准化互信息 (Normalized mutual information, NMI)、熵 (Entropy, E)、纯度 (Purity, P) 评价不同方法识别的精度, 结果如表 2 所示, 其中 ARI、NMI、P 的值越大, E 的值越小, 表明与真实结果更接近. 从表 2 可得, 对于给定评价指标, 所提方法均表现出最优的性能, 体现了所提的料面状态识别方法准确有效, 识别的料面状态的精度高于其他两种方法.

表 2　不同方法的识别精度比较

Table 2　Accuracy comparison of recognition results of different methods

	指标			
	ARI	NMI	E	P
DIFlow	0.4731	0.5105	1.0125	0.7666
SelFlow	0.4133	0.4276	1.0629	0.7344
本文方法	0.7669	0.7602	0.5212	0.9083

料面的不同运动状态能够反映高炉的炉况和运行效率, 若高炉料面基本保持在缓慢下沉状态, 则此时调控较好, 炉况比较稳定, 运行效率较高; 若高炉料面长期处于静止状态, 则容易发生悬料; 若高炉料面长期处于快速下沉状态, 则容易发生塌料. 为了得到不同状态的识别精度, 给出所提运动状态识别方法的混淆矩阵, 如图 15 所示. 从图 15 可以发现, 快速下沉、静止、缓慢下沉三种状态的识别精度分别为 87.88%、93.94%、91.18%. 其中 9.09% 的快速下沉状态会被识别为缓慢下沉状态, 3.03% 的静止状态会被识别为缓慢下沉状态, 而 5.88% 的缓慢下沉状态会被识别为快速下沉状态, 精度能够满足现场的需求.

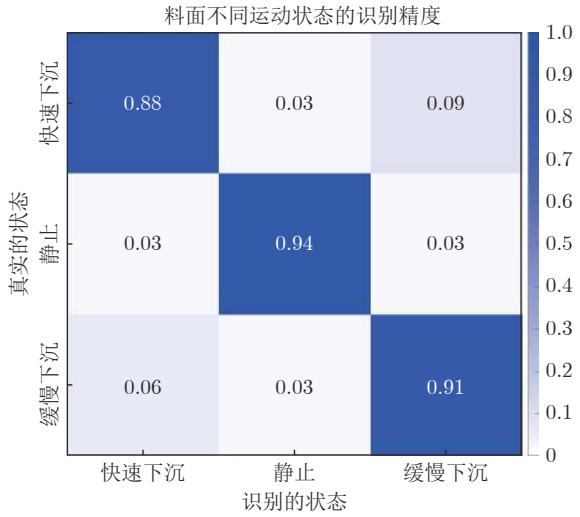


图 15　所提方法的混淆矩阵

Fig. 15　Confusion matrix of the proposed method

3.2　关键帧提取的性能分析

在获取料面状态后, 可以根据料面的不同状态筛选出关键帧. 为了验证所提关键帧提取方法的有效性和先进性, 随机选取一个布料周期为 451 s 的料面视频 (共 13 544 帧), 采用双阶段镜头边界检测方法 (Dual-stage-based shot boundary detection, DSBD)[14]、DeepReS[26]、DSN[28] 及本文方法提取料面视频的关键帧, 其中 DSBD 基于提取的 LBP-HF 特征和自适应阈值检测关键帧, DeepReS 和 DSN 采用深度学习方法识别视频关键帧, 图 16 给出了不同方法提取的部分关键帧结果. 从图 16 可以看出, 由于高炉布料过程料面图像变化较大, 而在布料间歇料面图像变化不易察觉, DSBD 将料面视频中图像视觉特征变化较大的视频帧作为关键帧, 获取了高炉布料过程的料面图像, 提取的关键帧图像不清晰, 有效信息遮挡较多, 难以反映当前时刻的料面反应状态, 且由于高炉是环形周期性布料, 在布料阶段图像变化较大, 采用 DSBD 方法提取的关键帧仍然有大量的视频冗余. DeepReS 方法能够去除布料过程部分冗余, 但是仍然存在布料过程不清晰的图像, 无法获取不同状态下稳定的料面图像. DSN 将视频帧的多样性和代表性作为奖励函数, 相比前两种方法获取的视频帧更加清晰, 但是缺乏对料面视频内容的分析, 且仍然存在部分布料过程不清晰的图像. 本文方法能够剔除细节模糊、不稳定的图像, 同时精简了视频, 去除了大量冗余, 增加了对料面视频运动状态的分析, 提取出清晰的不同状态下的料面图像, 对高炉炉顶布料具有一定的指导意义.

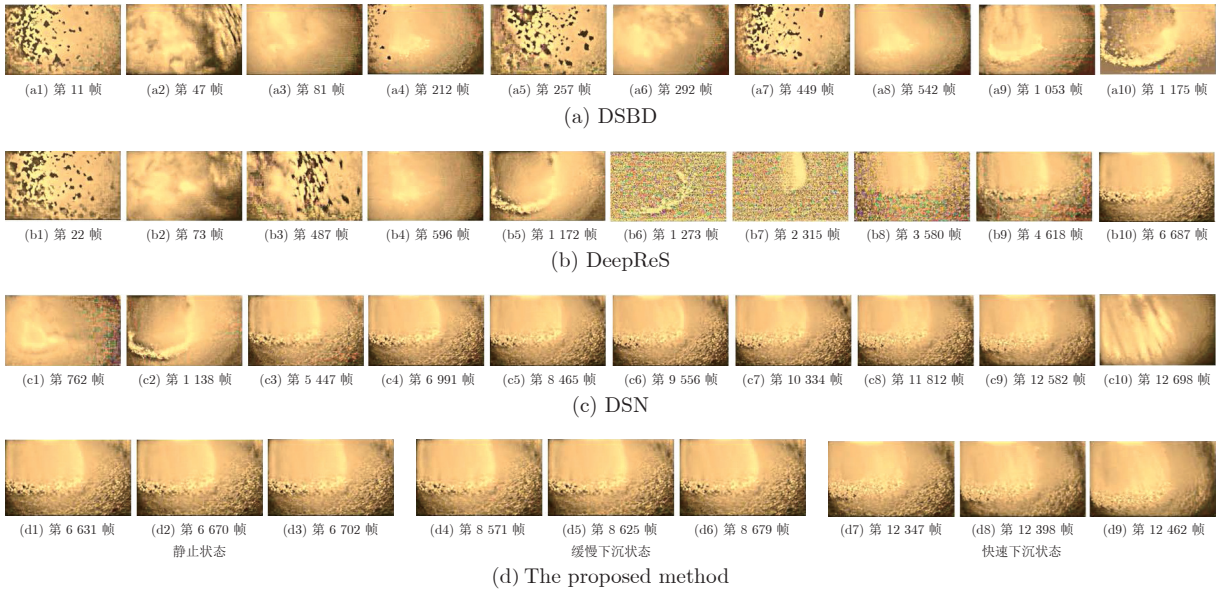


图 16 采用不同方法提取的同一布料周期的部分关键帧

Fig. 16 Extracted partial key frames of the same burden cycle using different methods

为了定量评价关键帧提取方法的性能, 以人工提取的关键帧为参考帧, 统计提取的关键帧集合的帧数、正确帧数、错误帧数、漏检帧数, 并计算查全率 R、准确率 P、$F1$ 值. 表 3 给出了采用不同方法提取的一段 2 323 s 料面视频 (共 66 332 帧) 的结果.

$$R = \frac{n_r}{n_r + n_f} \tag{18}$$

$$P = \frac{n_r}{n_r + n_w} \tag{19}$$

$$F1 = \frac{2PR}{P + R} \tag{20}$$

其中, n_r, n_f, n_w 分别表示正确提取的关键帧数量、漏检的关键帧数量和错误提取的关键帧数量.

表 3 不同方法提取的关键帧精度比较

Table 3 Accuracy comparison of key frames extracted by different methods

方法	关键帧	查全率	准确率	$F1$ 值
DSBD	679	60.3%	28.9%	0.3904
DeepReS	451	76.0%	54.8%	0.6366
DSN	394	85.2%	70.3%	0.7705
人工经验	325	—	—	—
本文方法	338	92.0%	88.5%	0.9020

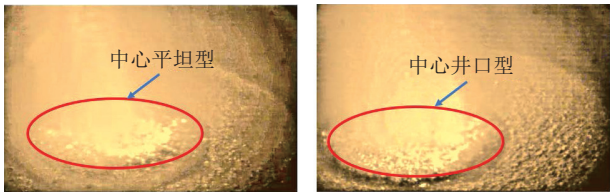
从表 3 可以看出, 受到高炉布料和炉内粉尘的影响, DSBD、DeepReS 和 DSN 提取的料面视频关键帧具有较多的冗余, 存在一定的漏检, 查全率和准确率并不高, 而所提方法所提取的料面视频冗余帧和漏检帧较少, 查全率、准确率及 $F1$ 值均高于其他三种方法, 定量验证了所提方法的准确性和可靠性.

在不同的布料周期, 料面形貌会有所变化, 采用本文所提的关键帧提取方法提取了不同布料周期的料面视频关键帧, 剔除了布料、加焦等不稳定时刻的料面图像, 图 17(a) ~ 17(d) 分别为 4 个不同布料周期提取的平坦型、井口型、喷溅型、气流发展型的料面图像关键帧. 所提取的关键帧能够清晰直观展示高炉不同布料周期的形貌, 为高炉冶炼现场提供极具价值的炉内反应状况信息及布料反馈信息.

考虑到高炉冶炼过程中不同布料周期的料面形貌可能会不同, 采用本文方法提取了 4 个不同布料周期不同典型形貌的料面视频的关键帧, 并分别计算了准确率、查全率和 $F1$ 值, 结果如图 18 所示. 从图中可得, 虽然料面形貌有所差异, 但是关键帧提取的准确率、查全率及 $F1$ 值均较高, 且变化并不明显, 说明所提方法在不同布料周期能够准确地提取料面视频关键帧, 且精度受布料周期和料面形貌不同的影响并不大.

4 结论

针对高炉料面视频存在的冗余信息多、图像质量不稳定、状态多变等问题, 本文提出了基于状态识别的高炉料面视频关键帧提取方法. 基于显著性区域特征点光流聚类方法识别了料面所处的状态, 并依据料面状态从冗余、质量不一的料面视频中筛选出不同周期不同状态下形貌变化较大且清晰稳定



(a) 中心平坦型
(a) Central flat type

(b) 中心井口型
(b) Central wellhead type

(c) 中心喷溅型
(c) Central splash type

(d) 中心气流发展型
(d) Central gas flow development type

图 17　提取的不同布料周期的关键帧

Fig. 17　Extracted key frames of different burden cycles

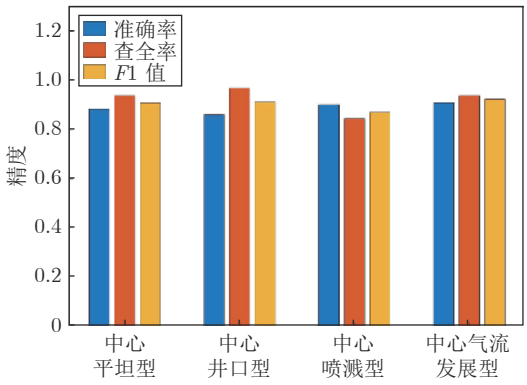


图 18　不同形貌料面视频关键帧提取精度

Fig. 18　Extraction accuracy of video key frames with different shapes and surfaces

的关键帧, 方便后期图像处理, 同时也有利于从大量的料面视频中快速准确地捕获到有效信息. 通过大量的对比实验表明本文所提方法能够准确识别料面所处的状态, 同时能够从视频中剔除细节模糊、不稳定的视频帧, 得到不同状态下中心气流稳定、清晰、无炉料及粉尘遮挡且特征明显的关键帧, 为现场操作人员认清炉内反应状况、指导炉顶布料操作提供重要反馈信息. 与其他现有的状态识别和视频关键帧提取方法相比, 所提方法结合了料面视频变化特征, 性能优越, 算法复杂度不高, 实用价值高, 能够为现场操作人员提供实时清晰的炉内反应过程图像信息, 并为炉顶布料提供反馈信息. 此外, 本文所提方法提供了一种复杂多变且周期性变化的工业视频关键帧提取的有效途径, 为其他复杂恶劣环境中视频图像处理提供了新思路.

References

1 Zhou Ping, Liu Ji-Ping, Liang Meng-Yuan, Zhang Rui-Yao. KPLS robust reconstruction error based monitoring and anomaly identification of fuel ratio in blast furnace ironmaking. *Acta Automatica Sinica*, 2021, **47**(7): 1661−1671
(周平, 刘记平, 梁梦圆, 张瑞垚. 基于 KPLS 鲁棒重构误差的高炉燃料比监测与异常识别. 自动化学报, 2021, **47**(7): 1661−1671)

2 Jiang Zhao-Hui, Xu Chuang, Gui Wei-Hua, Jiang Ke. Prediction method of hot metal silicon content in blast furnace based on optimal smelting condition migration. *Acta Automatica Sinica*, 2021, **48**(1): 207−219
(蒋朝辉, 许川, 桂卫华, 蒋珂. 基于最优工况迁移的高炉铁水硅含量预测方法. 自动化学报, 2021, **48**(1): 207−219)

3 Shi L, Wen Y B, Zhao G S, Yu T. Recognition of blast furnace gas flow center distribution based on infrared image processing. *Journal of Iron and Steel Research International*, 2016, **23**(3): 203−209

4 Chen Z P, Jiang Z H, Gui W H, Yang C H. A novel device for optical imaging of blast furnace burden surface: Parallel low-light-loss backlight high-temperature industrial endoscope. *IEEE Sensors Journal*, 2016, **16**(17): 6703−6717

5 Zhang Xiao-Yu, Zhang Yun-Hua. Video keyframe extraction method based on fusion feature. *Computer Systems & Applications*, 2019, **28**(11): 176−181
(张晓宇, 张云华. 基于融合特征的视频关键帧提取方法. 计算机系统应用, 2019, **28**(11): 176−181)

6 Xu T X, Chen Z P, Jiang Z H, Huang J C, Gui W H. A real-time 3D measurement system for the blast furnace burden surface using high-temperature industrial endoscope. *Sensors*, 2020, **20**(3): Article No. 869

7 Nandini H M, Chethan H K, Rashmi B S. Shot based keyframe extraction using edge-LBP approach. *Journal of King Saud University-Computer and Information Sciences*, 2022, **34**(7): 4537−4545

8 Zhi Min, Cai An-Ni. Shot boundary detection with main color. *Acta Automatica Sinica*, 2007, **33**(6): 655−657
(智敏, 蔡安妮. 基于基色调的镜头边界检测方法. 自动化学报, 2007, **33**(6): 655−657)

9 Tang H, Liu H, Xiao W, Sebe N. Fast and robust dynamic hand gesture recognition via key frames extraction and feature fusion. *Neurocomputing*, 2019, **331**: 424−433

10 Yuan Y, Lu Z, Yang Z, Jian M, Wu L F, Li Z Y, et al. Key frame extraction based on global motion statistics for team-sport videos. *Multimedia Systems*, 2022, **28**(2): 387−401

11 Li Z N, Li Y J, Tan B Y, Ding S X, Xie S L. Structured sparse coding with the group log-regularizer for key frame extraction. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(10): 1818−1830

12 Li X L, Zhao B, Lu X Q. Key frame extraction in the summary space. *IEEE Transactions on Cybernetics*, 2018, **48**(6): 1923−1934

13 Zhao B, Gong M G, Li X L. Hierarchical multimodal transformer to summarize videos. *Neurocomputing*, 2022, **468**: 360−369

14 Singh A, Thounaojam D M, Chakraborty S. A novel automatic shot boundary detection algorithm: Robust to illumination and motion effect. *Signal, Image and Video Processing*, 2020, **14**(4): 645−653

15 Wang Ting-Xian, Jia Ke-Bin, Yao Meng. Real-time visual localization method for light-rail with high accuracy. *Acta Automatica Sinica*, 2021, **47**(9): 2194−2204

(王婷娴, 贾克斌, 姚萌. 面向轻轨的高精度实时视觉定位方法. 自动化学报, 2021, **47**(9): 2194−2204)

16 Zhang Y Z, Tao R, Wang Y. Motion-state-adaptive video summarization via spatiotemporal analysis. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, **27**(6): 1340−1352

17 Gharbi H, Bahroun S, Zagrouba E. Key frame extraction for video summarization using local description and repeatability graph clustering. *Signal, Image and Video Processing*, 2019, **13**(3): 507−515

18 Lai J L, Yi Y. Key frame extraction based on visual attention model. *Journal of Visual Communication and Image Representation*, 2012, **23**(1): 114−125

19 Wu J X, Zhong S H, Jiang J M, Yang Y Y. A novel clustering method for static video summarization. *Multimedia Tools and Applications*, 2017, **76**(7): 9625−9641

20 Chu W S, Song Y, Jaimes A. Video co-summarization: Video summarization by visual co-occurrence. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015. 3584−3592

21 Elahi G M M E, Yang Y H. Online learnable keyframe extraction in videos and its application with semantic word vector in action recognition. *Pattern Recognition*, 2022, **122**: Article No. 108273

22 Wu G D, Lin J Z, Silva C T. IntentVizor: Towards generic query guided interactive video summarization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022. 10493−10502

23 Ji Z, Zhao Y X, Pang Y W, Li X, Han J G. Deep attentive video summarization with distribution consistency learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(4): 1765−1775

24 Abed R, Bahroun S, Zagrouba E. Keyframe extraction based on face quality measurement and convolutional neural network for efficient face recognition in videos. *Multimedia Tools and Applications*, 2021, **80**(15): 23157−23179

25 Jian M, Zhang S, Wu L F, Zhang S J, Wang X D, He Y H. Deep key frame extraction for sport training. *Neurocomputing*, 2019, **328**: 147−156

26 Muhammad K, Hussain T, Ser J D, Palade V, de Albuquerque V H C. DeepReS: A deep learning-based video summarization strategy for resource-constrained industrial surveillance scenarios. *IEEE Transactions on Industrial Informatics*, 2020, **16**(9): 5938−5947

27 Xiao S W, Zhao Z, Zhang Z J, Yan X H, Yang M. Convolutional hierarchical attention network for query-focused video summarization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, **34**(7): 12426−12433

28 Zhou K Y, Qiao Y, Xiang T. Deep reinforcement learning for unsupervised video summarization with diversity-representativeness reward. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018, **32**(1): 7582−7589

29 Huang J C, Jiang Z H, Gui W H, Yi Z H, Pan D, Zhou K, et al. Depth estimation from a single image of blast furnace burden surface based on edge defocus tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, **32**(9): 6044−6057

30 Yi Z H, Chen Z P, Jiang Z H, Gui W H. A novel 3-D high-temperature industrial endoscope with large field depth and wide field. *IEEE Transactions on Instrumentation and Measurement*, 2020, **69**(9): 6530−6543

31 Li Dong-Min, Li Jing, Liang Da-Chuan, Wang Chao. Multiple salient objects detection using multi-scale prior and deep features. *Acta Automatica Sinica*, 2019, **45**(11): 2058−2070
(李东民, 李静, 梁大川, 王超. 基于多尺度先验深度特征的多目标显著性检测方法. 自动化学报, 2019, **45**(11): 2058−2070)

32 Cai S Z, Huang Y B, Ye B, Xu C. Dynamic illumination optical flow computing for sensing multiple mobile robots from a drone. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, **48**(8): 1370−1382

33 Liu P P, Lyu M, King I, Xu J. SelFlow: Self-supervised learning of optical flow. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019. 4566−4575

黄建才 中南大学自动化学院博士研究生. 2018 年获重庆大学学士学位. 主要研究方向为图像处理, 计算机视觉和复杂工业过程参数检测.
E-mail: huangjiancai@csu.edu.cn
(**HUANG Jian-Cai** Ph.D. candidate at the School of Automation, Central South University. He received his bachelor degree from Chongqing University in 2018. His research interest covers image processing, computer vision, and parameter detection of complex industrial processes.)

蒋朝辉 中南大学自动化学院教授. 2011 年获中南大学博士学位. 主要研究方向为智能传感与检测技术, 图像处理与智能识别和人工智能与机器学习. 本文通信作者.
E-mail: jzh0903@csu.edu.cn
(**JIANG Zhao-Hui** Professor at the School of Automation, Central South University. He received his Ph.D. degree from Central South University in 2011. His research interest covers intelligent sensing and detection technology, image processing and intelligent recognition, artificial intelligence and machine learning. Corresponding author of this paper.)

桂卫华 中南大学自动化学院教授. 1981 年获得中南矿冶学院硕士学位. 主要研究方向为复杂工业过程建模, 优化与控制应用和故障诊断与分布式鲁棒控制. E-mail: gwh@csu.edu.cn
(**GUI Wei-Hua** Professor at the School of Automation, Central

South University. He received his master degree from Central South Institute of Mining and Metallurgy in 1981. His research interest covers complex industrial process modeling, optimization and control applications, fault diagnosis and distributed robust control.)

潘 冬 中南大学自动化学院讲师. 分别于 2015 年和 2021 年获中南大学学士学位和博士学位. 主要研究方向为红外热成像, 视觉检测, 图像处理和深度学习.

E-mail: pandong@csu.edu.cn

(**PAN Dong** Lecturer at the School of Automation, Central South University. He received his bachelor degree and Ph.D. degree from Central South University in 2015 and 2021, respectively. His research interest covers infrared thermography, vision-based measurement, image processing, and deep learning.)

许 川 中南大学自动化学院博士研究生. 2021 年获中南大学硕士学位. 主要研究方向为数据分析, 机器学习和复杂工业过程建模.

E-mail: csuxuchuan@csu.edu.cn

(**XU Chuan** Ph.D. candidate at the School of Automation, Central South University. He received his master degree from Central South University in 2021. His research interest covers data analysis, machine learning, and complex industrial process modeling.)

周 科 中南大学自动化学院博士研究生. 2018 年获重庆大学学士学位. 主要研究方向为复杂工业过程的数据挖掘、建模与优化控制.

E-mail: zhouke95@csu.edu.cn

(**ZHOU Ke** Ph.D. candidate at the School of Automation, Central South University. He received his bachelor degree from Chongqing University in 2018. His research interest covers data mining, modeling and optimal control of complex industrial processes.)