

DOI:10.19651/j.cnki.emt.2108595

# 矿用输送带纵向撕裂视听融合检测方法<sup>\*</sup>

高瑜璋 乔铁柱 车 剑

(太原理工大学新型传感器与智能控制教育部重点实验室 太原 030024)

**摘要:** 输送带撕裂检测是煤矿安全生产中非常重要的部分。本文提出了一种新的输送带纵向撕裂检测方法—视听融合检测方法。视听融合方法使用可见光 CCD 和麦克风阵列采集输送带在不同运行状态下的图像和声音。通过对采集到的图像和声音进行处理和分析,分别提取出正常、撕裂和划伤的图像和声音特征。然后利用机器学习算法对提取的图像和声音特征进行融合和分类。实验结果表明,视听融合检测方法对输送带纵向撕裂的准确率为 96.23%,对于输送带划伤的准确率为 93.66%,与现有方法相比,该方法克服了传统机器视觉检测条件的局限性,对于输送带撕裂检测更加准确可靠。

**关键词:** 图像特征提取;声音特征提取;特征融合;纵向撕裂检测方法;机器学习

**中图分类号:** TP3 **文献标识码:** A **国家标准学科分类代码:** 520.2060

## Longitudinal tear detection method of conveyor belt for mine based on audio-visual fusion

Gao Yuzhang Qiao Tiezhu Che Jian

(Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education and Shanxi Province, Taiyuan University of Technology, Taiyuan 030024, China)

**Abstract:** Conveyor belt tear detection is a very important part of coal mine safety production. In this paper, a new method of detecting conveyor belt damage named audio-visual fusion (AVF) detection method is proposed. The method uses both a visible light CCD and a microphone array to collect images and sounds of the conveyor belt in different running states. By processing and analyzing the collected images and sounds, the image, and sound features of normal, tear and scratch can be extracted respectively. Then the extracted features of images and sounds are fused and classified by machine learning algorithm. The results show that the accuracy of AVF method for conveyor belt scratch is 93.66%, and the accuracy of longitudinal tear is higher than 96.23%. Compared with existing methods AVF method overcomes the limitation of visual detection condition, and is more accurate and reliable for conveyor belt tear detection.

**Keywords:** image feature extraction; audio feature extraction; feature fusion; longitudinal tear detection method; machine learning

## 0 引 言

带式输送机是煤矿生产中重要的运输设备,输送带是带式输送机中最容易发生事故的部分<sup>[1]</sup>。输送带是整个带式输送机的主要部件,在输送带运输进程中,片状煤矸石、金属螺栓、锚杆等硬杂质可能会撕裂或刮伤输送带,从而引发煤矿生产安全事故,输送带纵向撕裂检测对预防安全事故具有重要意义<sup>[2]</sup>。Pang 等<sup>[3]</sup>通过外部传感器探测预埋

在输送带里磁铁阵列磁场变化实现输送带的状态监测,此方法检测准确率高,但安装与维护成本都很高;2014 年朱大根等<sup>[4]</sup>使用超声波技术检测输送带表面撕裂,但是井下复杂的环境影响超声波回波信号,该方法在输送带纵向撕裂检测中难以推广与应用。2020 年苗长云等<sup>[5]</sup>利用高速 X 光探测器检测纵向撕裂,该方法可以有效地对输送带纵向撕裂进行检测,但由于无法完全封闭的 X 射线会对煤矿工人造成危害,此方法很难在煤矿井下应用。随着机器视觉的发展,更多的科学家把计算机视觉技术运用在撕裂检测中<sup>[6]</sup>。例如 Li 等<sup>[7]</sup>提出一种线扫描数码相机检测纵

收稿日期:2021-12-14

<sup>\*</sup> 基金项目:国家自然科学基金(NSFC-山西煤基低碳联合基金项目 U1810121)、2020 年中央引导地方科技发展资金项目(YDZX20201400001796)资助

向撕裂的 SSR 算法。Yang 等<sup>[8]</sup>提出了一种利用线阵 CCD 进行输送带纵向撕裂检测的方法。利用计算机视觉技术检测输送带纵向撕裂具有简洁方便、成本小等优点,但是由于井下黑暗、潮湿、多尘的工作环境,现有的视觉检测方法检测准确率会受到很大的影响,同时现有计算机视觉输送带纵向撕裂检测方法不能准确检测输送带划伤,而快速准确的检测输送带划伤可以防止输送带纵向撕裂的发生。相比于现有输送带纵向撕裂检测方法,视觉听觉的融合信息对目标状态的描述更为直观,且具备易于获取、无需与监测对象接触等优点。因此本文提出了一种输送带纵向撕裂视听融合检测方法,该方法克服了计算机视觉检测条件的局限性,增加了输送带撕裂检测的准确性。

## 1 检测原理

目前,随着计算机视觉和声音检测技术发展迅速,计算机视觉技术可以代替人眼进行识别、跟踪和测量,但视觉图像的质量易受障碍物遮挡和光照条件等影响。声音作为与文字、图像、影视同时存在的信息载体之一,包含了丰富的环境信息<sup>[9]</sup>,对声音所代表事件的准确识别能有效反应环境变化,因此可以用声音来辅助计算机视觉应用到输送带纵向撕裂检测中<sup>[10]</sup>。根据信息融合理论,可以在特征层、数据层和决策层 3 个层次上进行数据融合<sup>[11]</sup>。由于图像和声音数据结构不同,很难对传感器直接采集的数据在数据层进行融合。当输送带纵向撕裂时,撕裂的产生与和声音有一定的相关性,如果分别处理每个传感器数据后在决策层融合,视听信号的相关性会被分离,导致检测精度下降。特征层融合提取特征的方法是将原始数据转换为一组具有明显物理意义的特征来代表原始数据的特性,然后按照特征信息对传感器数据进行分类、识别和综合分析。特征融合的主要优点在于,它实现了不同传感器之间的融合,对于数据量庞大的原始数据实现了数据压缩,有利于实时性处理。因此,本文选择特征层进行声音和图像的数据融合,并利用机器学习模型将融合后的数据进行分类,完成输送带划伤与撕裂的检测。

## 2 视听融合检测方法

### 2.1 图像预处理

收集到的输送带纵向撕裂、划伤和正常的图像如图 1 所示,图像中输送带撕裂区域较明显。但是在图片背景中,除了受损区域,采集到的图像中会有一些划痕和污点。为了消除这些特征和较少噪声的影响,使提取的图像损伤特征更加清晰,使用中值滤波对图像进行滤波,中值滤波使用一个像素领域中灰度级的中值来代替该像素的值<sup>[12]</sup>,并对中值滤波后的图像进行灰度化处理。预处理结果如图 2 所示。

### 2.2 图像 HOG 特征提取

使用方向梯度直方图(HOG)的图像特征算法来提取

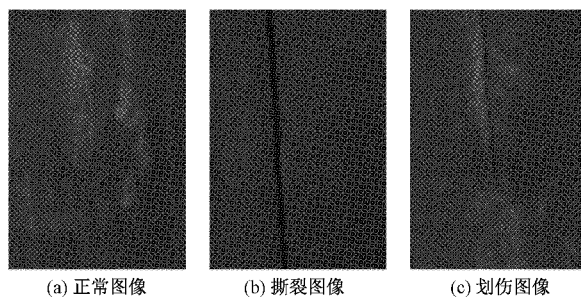


图 1 原始图像

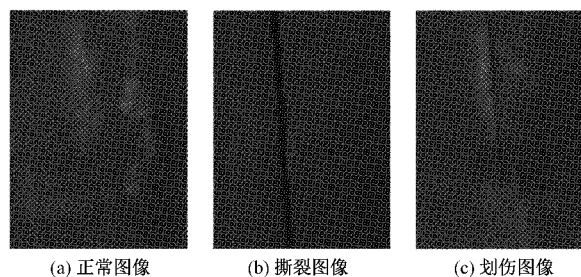


图 2 预处理后的图像

输送带纵向撕裂图像特征。HOG 算法计算并统计图像中像素的梯度和方向,计算局部图像的梯度直方图构造特征,能够很好地描述检测对象的边缘特征<sup>[13]</sup>。HOG 算法的步骤描述如下。

1)Gamma 校正。为了克服煤矿井下黑暗多尘环境造成的局部阴影和图像亮度变化的影响,突出图像受损的部分,对图像进行了伽玛(Gamma)校正,Gamma 校正公式如式(1)所示。

$$f(I) = I^\gamma \quad (1)$$

式中: $I$  为输入像素值, $\gamma$  为 Gamma 校正系数, $f(I)$  为输出像素值。本文  $\gamma = 0.8$ , 经过 Gamma 校正后的图像如图 3 所示。

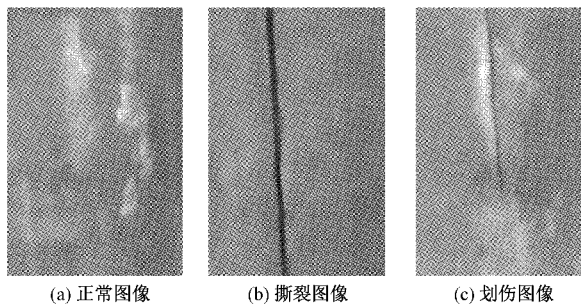


图 3 Gamma 校正后的图像

2)计算图像梯度。利用统计梯度法得到了输送机的纵向撕裂图像特征,数学上的梯度实际上是一阶导数,图像在某个像素点处的梯度可以用式(2)和(3)计算。

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \quad (2)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \quad (3)$$

式中: $G_x(x, y)$  为点  $(x, y)$  的垂直梯度, $G_y(x, y)$  为水平

梯度,  $H(x, y)$  为  $(x, y)$  点处的像素值,  $(x, y)$  点处梯度幅值和方向计算如下:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (4)$$

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \quad (5)$$

3) 统计 HOG 特征。首先, 图像被划分为  $8 \times 8$  像素的 cell, 然后  $360^\circ$  的梯度方向的 cell 可分为 9 个 bins, 9 个 bins 的直方图用于统计  $8 \times 8$  像素梯度信息, 最后将像素的梯度方向投影到相应的直方图。通过这种方式, 加权梯度大小和方向的投影单元中每个像素的直方图是相应 cell 的特征向量。使用 4 个 cell 来形成一个  $16 \times 16$  像素的 block, 它会连接形成一个  $36 \times 1$  的向量。然后选择窗口以固定的步长移动(每步 8 个像素)来统计直方图, 从而为每次移动生成标准化的  $36 \times 1$  向量。每次移动后获得的 36 个特征被连接作为图像的最终 HOG 特征。

### 2.3 声音预处理

1) 预加重。由于输送带纵向撕裂声信号采集经常受到矿井下各种噪声的影响, 高频声信号会衰减。因此, 在对声音信号进行处理之前, 应该增强声音的高频部分, 有效降低输出噪声, 获取更多的频域信息, 从而便于的声音特征提取。声音预加重传递函数通常表示为:

$$H(Z) = 1 - aZ^{-1} \quad (6)$$

式中:  $a$  为预加重系数, 本文为  $a = 0.97$ 。

2) 分帧和加窗。由于声音信号在短时间内保持不变和相对稳定, 因此可以将声音信号分成若干段进行处理, 每段称为一帧。在本文中, 帧数  $N = 1024$ , 采样频率设为 14.1 kHz, 为了获得更平滑的频谱, 使用汉明窗来分帧<sup>[14]</sup>。汉明窗计算公式如式(7)所示。

$$\omega(n) = \begin{cases} (1-a) - a \cos[2\pi n/(N-1)], & 0 \leq n \leq N \\ 0, & \text{其他} \end{cases} \quad (7)$$

式中:  $a$  为汉明窗参数, 本文取  $a = 0.46$ , 确定窗口函数后, 对预加重声音信号加进行加窗, 如式(8)所示。

$$s'(n) = s(n) \cdot \omega(n) \quad (8)$$

式中:  $s'(n)$  为声音波形时域信号,  $s(n)$  为分帧处理后的第  $n$  帧声音信号。

### 2.4 声音特征提取

1) 梅尔倒谱系数(MFCC)<sup>[15]</sup>。MFCC 被广泛的应用于语音识别, 但通过实验发现它可以用于分析输送带的损伤。人的听觉系统感知到的低频声音的频率与声音的物理频率近似为线性, 高频声音感知到的频率和声音的物理频率近似为对数, 梅尔频率  $f_{mel}$  和物理频率  $f_{hz}$  之间的关系如式(9)所示。

$$f_{mel} = 2595 \lg\left(1 + \frac{f_{hz}}{700}\right) \quad (9)$$

在对声音信号进行预处理后, 对音频信号进行时域转换, 其主要实现方法是离散傅里叶变换(DFT)。DFT 的定

义如下:

$$x_n(k) = \sum_{n=0}^{N-1} s'(n) \exp(-j2\pi kn/N), 0 \leq k \leq N \quad (10)$$

式中:  $k$  是傅里叶变换的第  $k$  个频谱,  $N$  是傅里叶变换的点数。经过离散傅里叶变换后的声音信号通过梅尔滤波器。梅尔滤波器由三角形带通滤波器组成, 可以将频谱进行平滑处理, 去除谐波的影响, 突出原始声音的共振峰, 三角形滤波器的频率响应  $H_m(k)$  为:

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1))(f(m)-f(m-1))}, & f(m-1) \leq k < f(m) \\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m))(f(m+1)-f(m))}, & f(m) \leq k \leq f(m+1) \\ 0, & k > f(m+1) \end{cases} \quad (11)$$

式中:  $H_m(k)$  满足:  $\sum_{m=0}^{M-1} H_m(k) = 1$ 。第  $m$  个梅尔滤波器输出的对数能量  $S(m)$  如式(12)所示。

$$S(m) = \ln\left(\sum_{k=0}^{N-1} |x_n(k)|^2 H_m(k)\right), 0 < m < M \quad (12)$$

式中:  $M$  是梅尔三角形滤波器的个数, 本文  $M = 26$ 。再对数能量  $S(m)$  进行离散余弦变换得到 MFCC, 计算公式如式(13)所示。

$$C(m) = \sum_{l=0}^{M-1} S(m) \cos\left(\frac{\pi l(l-0.5)}{M}\right), l = 0, 1, 2, \dots, L \quad (13)$$

式中:  $C(m)$  为梅尔倒谱参数,  $L$  为 MFCC 系数的阶数, 取值范围通常为 12~16, 本文中  $L = 13$ 。

2) 短时能量。由于声信号能量随时间的变化, 输送带撕裂、划伤和正常运行时的声能量有显著差异。因此, 对短时间能量的分析可以描述输送带在不同状态下的特性变化。短时能量  $E_n$  计算如式(14)所示。

$$E_n = \sum_{m=0}^{N-1} s_n^2(m) \quad (14)$$

式中:  $s_n(m)$  为第  $n$  帧声音信号,  $m$  为窗口位置。

3) 过零率(ZCR)是指每一帧声音信号中通过零点的次数, ZCR 主要用于背景噪声较大时的识别。计算公式如式(15)所示。

$$ZCR_n = \frac{1}{2} \sum_{m=0}^{N-1} [\text{sgn}(s_n(m)) - \text{sgn}(s_n(m+1))] \quad (15)$$

其中,

$$\text{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \quad (16)$$

4) 频谱质心。频谱质心是描述音色的重要参数之一, 它是在一定频率范围内用加权平均频率, 谱质心可描述声音的亮度, 是声音信号频率分布和能量分布的重要信息, 计算公式如式(17)所示。



$$SC_n = \frac{\sum f S_n(f)}{\sum S_n(f)} \quad (17)$$

式中： $f$  是声音信号频率， $S_n(f)$  为连续时域信号离散傅里叶变换后对应频率的谱能量。

5) 频谱滚降。频谱滚降是频谱振幅在频率低于一定设定值时的变化。根据频谱的滚转特性，可以测量频谱形的斜率。频谱滚降的计算公式如式(18)所示。

$$\sum_{k=0}^m |S_n(k)| = \theta \sum_{k=0}^{N-1} |S_n(k)| \quad (18)$$

式中： $\theta$  为阈值，其取值范围为 0.85~0.99，本文  $\theta=0.85$ 。

### 2.5 特征融合

1) 将提取到的输送带图像和声音特征进行融合。假设通过特征提取的输送带纵向撕裂的声音特征向量为  $x_i$ ，其中  $i$  表示声音特征的种类，图像的特征向量表示为  $y_j$ ， $j$  表示图像特征的种类，不同类型的声音和图像特征的权重为  $w_k$ ，则融合  $i$  类声音特征和  $j$  类图像特征后的视听特征矩阵  $T$  如所(19)所示。

$$T = (\omega_1 x_{i1}, \omega_2 x_{i2}, \omega_3 x_{i3}, \dots, \omega_i x_i, \omega_{i+1} y_{j1}, \omega_{i+2} y_{j2}, \omega_{i+3} y_{j3}, \dots, \omega_k y_j)^T \quad (19)$$

其中，特征融合的权重  $w_k$  基于该特征的准确率为权重进行融合，式中  $k = i + j$ 。

2) 将得到视听融合特征矩阵进行数据归一化处理。

$$x_j^{(i)} = \frac{a_j^{(i)} - \mu_j}{s_j} \quad (20)$$

式中： $a_j^{(i)}$  为第  $i$  个样本的  $j$  个特征值， $\mu_j$  为  $j$  个特征的均值， $s_j$  为特征的范围，即：

$$s_j = \max(a_j^{(i)}) - \min(a_j^{(i)}) \quad (21)$$

3) 主成分分析(PCA)是一种降维算法<sup>[16]</sup>，它可以将高维数据转化为低维数据，利用 PCA 算法对视听融合特征进行处理，可以提取有用信息并去除冗余信息。对数据归一化后，得到视听融合特征矩阵  $X$ ，首先计算协方差矩阵  $C$ ：

$$C = \frac{1}{m} X^T X \quad (22)$$

然后采用奇异值分解(svd)计算协方差矩阵的特征向量<sup>[17]</sup>。

$$[U, S, V] = svd(C) \quad (23)$$

得到特征向量矩阵后，可以对数据进行降维。假设降维前的值为  $x^{(i)}$ ，则降维公式如式(24)所示。

$$Z^{(i)} = U_{reduce}^T x^{(i)} \quad (24)$$

式中： $Z^{(i)}$  为降维后的值， $U_{reduce} = [u^{(1)}, u^{(2)}, \dots, u^{(k)}]$  为主成分特征矩阵。

### 2.6 机器学习模型

对收集到的输送带损伤的视听数据进行了分类。在本文中，选择了 3 个机器学习算法 K 近邻(KNN)、支持向量机(SVM)和随机森林(RF)。KNN 算法实现简单，对于数据异常点不敏感；SVM 最大的特点是可以结构化决策边界，能构造出最大距离的决策边界，具有较高的鲁棒性；

RF 算法是将多个弱分类器(决策树)组合成一个强分类器的分类算法，是对多个分类器的输出进行融合，提高了分类的准确性。

### 3 实验和分析

为了验证本文提出输送带纵向撕裂视听融合检测方法的有效性，搭建实验平台，使用可见光 CCD 和麦克风阵列来收集输送带纵向撕裂视听数据，输送带的具体参数如为 13 m×1 m×15 mm。在上下皮带之间安装金属锚杆来模拟输送带纵向撕裂，实验平台如图 4 所示。可见光 CCD 相机的分辨率为 1 280×1 080 和 60 fps 的帧速率，麦克风阵列是一个四阵列的麦克风，实验具体实施过程如下。

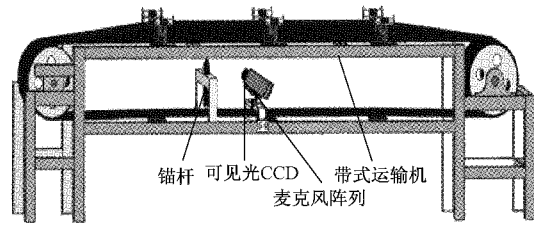


图 4 实验装置示意图

1) 在实验室现有输送机平台上搭建输送带纵向撕裂视听融合检测装置，然后调节金属锚杆的深度，模拟传送带的划伤与撕裂。根据金属锚杆插入深度的不同，分别表示传送带的正常运行，传送带的划伤和纵向撕裂，同时将在实际煤矿井下运行中的输送带收集到的噪声当作环境声音。

2) 使用 CCD 相机和麦克风阵列分别采集输送带运行图像和声音信号，标记并构建图像和声音数据样本。选取撕裂样本、划伤样本、正常样本图像和声音各 1 000 组数据循环读取完成训练，之后选取 3 组含有 200 个各类样本的测试集。在该实验中，采集不同运行状态下传送带的图像和声音，声音的采样频率设置为 44.1 kHz，每个段 1.2 s 的时间，所有的声音都为 wav 格式。使用工业计算机(I7 7700 HQ, 16 G 内存)在 Pycharm 开发平台使用 Python3.7 编译进行测试。机器学习模型的参数选取如表 1 所示。

表 1 机器学习模型参数

模型	模型主要参数
KNN	K: 1, 5, 10, 15, Distance: 'euclidean', 'cityblock'
SVM	Kernel function: rbf C: 10 <sup>-1</sup> Gamma: 2 <sup>-12</sup> , 2 <sup>-10</sup> , 2 <sup>-8</sup> , 2 <sup>-4</sup> , 2 <sup>-2</sup> , 2 <sup>0</sup> , 2 <sup>2</sup>
RF	n_estimators: 2 <sup>3</sup> , 2 <sup>4</sup> , min_sample_leaf: 0, 5, 2 <sup>5</sup> , 2 <sup>6</sup> 10, 20

3) 实时采集输送带运行图像和声音信号进行输送带撕裂测试，并计算视听融合检测模型的识别准确率。

本文中实验数据分成3份,分别是输送带正常运行时的视听数据,输送带发生纵向撕裂时的视听数据,输送带划伤时的视听数据。输送带纵向撕裂视听融合检测方法流程如图5所示,实验结果如图6所示。

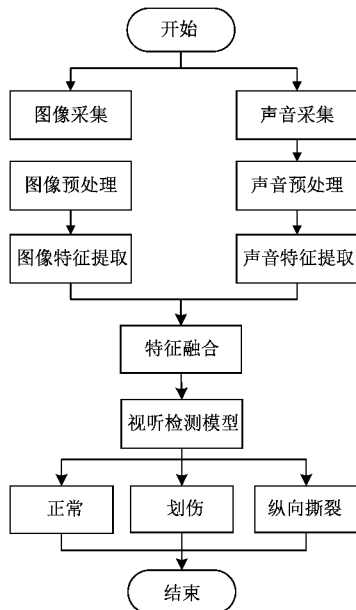


图5 输送带纵向撕裂视听融合检测方法流程

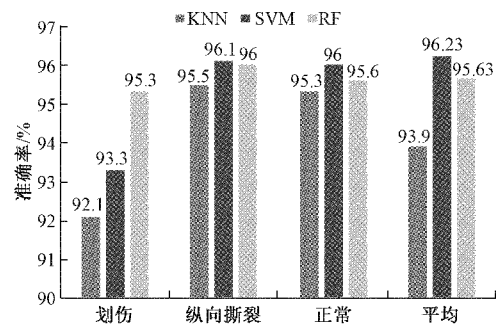


图6 实验结果

从图6可以看出使用SVM分类器的输送带纵向撕裂视听融合检测方法检测准确率较高,对于划伤的检测准确率为95.5%,对于输送带纵向撕裂检测准确率为96.1%,平均检测准确率为96.23%。为了验证视听特征与图像特征训练出来模型的差别,单独把图像HOG特征作为输入信息在基于SVM机器学习算法上建立单独图像特征的模型,模型检测结果如表2所示,从表2中能够看出SVM分类器对单独图像特征的输送带损伤检测结果使用单一图像特征撕裂检测准确率平均值为92.25%。基于SVM算法下视听融合特征的整体准确率为96.23%,优于单一图像特征。

表2 单独图像特征检测结果

样本号	撕裂样本			划伤样本			正常样本			准确率/ %
	撕裂	划伤	正常	撕裂	划伤	正常	撕裂	划伤	正常	
1	187	12	1	15	179	6	2	4	194	93.3
2	188	10	2	20	174	6	0	5	195	92.8
3	184	10	6	25	170	5	0	4	196	91.6
平均	186	11	3	17	174	6	1	4	195	92.25

## 4 结 论

为了适应煤矿复杂环境,提高输送带纵向撕裂检测精度,本文提出了一种输送带纵向撕裂视听融合检测方法,并对通过视听融合检测方法和单独图像特征检测方法检测结果进行对比,实验结果表明,输送带纵向撕裂视听融合检测方法的正确率高于单独图像特征,输送带纵向撕裂检测正确率为96.23%。综上所述,输送带纵向撕裂视听融合检测方法不仅能够适应煤矿井下复杂的工作环境,而且能够更好地实现纵向撕裂的检测。

## 参考文献

[1] HE D, PANG Y, LODEWIJKS G. Green operations of belt conveyors by means of speed control [J]. Applied Energy, 2017, DOI: 10.1016/j.apenergy.2016.12.017.

[2] 王建勋. 煤矿输送带传输故障实时监测技术[J]. 工矿

自动化, 2015, 41(1): 45-48.

- [3] PANG Y, LODEWIJKS G. A novel embedded conductive detection system for intelligent conveyor belt monitoring [C]. 2006 IEEE International Conference on Service Operations and Logistics, and Informatics, IEEE, 2006: 803-808.
- [4] 朱大根, 殷勇辉, 狄帅, 等. 基于FPGA的输送带表面检测方法研究[J]. 煤矿机电, 2014(6): 45-48.
- [5] 苗长云, 关世阳. 基于FPGA的钢丝绳芯输送带高速X光探测器[J]. 天津工业大学学报, 2020, 39(2): 49-53.
- [6] 卢金龙. 基于机器视觉的皮带撕裂检测系统设计与实现[D]. 秦皇岛: 燕山大学, 2016.
- [7] LI J, MIAO C. The conveyor belt longitudinal tear on-line detection based on improved SSR algorithm [J]. Optik, 2016, 127(19): 8002-8010.
- [8] YANG Y, MIAO C, LI X, et al. On-line conveyor

- belts inspection based on machine vision[J]. *Optik*, 2014, 125(19): 5803-5807.
- [9] D'ARCA E, ROBERTSON N M, HOPGOOD J R. Robust indoor speaker recognition in a network of audio and video sensors[J]. *Signal Processing*, 2016, 129: 137-149.
- [10] 张靓. 机械设备智能诊断中视听信息融合方法研究[D]. 天津: 河北工业大学, 2016.
- [11] DU J X, ZHAI C M, GUO Y L. Recognizing complex events in real movies by combining audio and video features[J]. *Neurocomputing*, 2014, 137(SI): 89-95.
- [12] 陈伟, 党榕. 混合噪声图像的去噪算法研究及 FPGA 实现[J]. *电子测量技术*, 2021, 44(16): 135-140.
- [13] 丰婧, 程耀瑜, 贺磊. 基于 PHOG 特征的坐姿识别方法研究[J]. *国外电子测量技术*, 2021, 40(5): 83-87.
- [14] 王志, 江青. 汉明窗在屏蔽效能计算中的应用[J]. *赤峰学院学报(自然科学版)*, 2018, 34(8): 117-119.
- [15] 张钰莎, 蒋盛益. 基于 MFCC 特征提取和改进 SVM 的语音情感数据挖掘分类识别方法研究[J]. *计算机应用与软件*, 2020, 37(8): 160-165.
- [16] 肖璞, 殷一鸣. 一种基于 PCA 降维的 SIFT 算法[J]. *现代计算机*, 2020(34): 15-18.
- [17] YUAN Y X, LIU L, ZHANG H T, et al. The solutions to the quadratic matrix equation [formula omitted][J]. *Applied Mathematics and Computation*, 2021, DOI: 10.1016/J.AMC.2021.126463.

#### 作者简介

高瑜璋, 硕士研究生, 主要研究方向为仪器与测试技术、输送带无损检测和机器视觉应用研究。  
E-mail: 15735185648@163.com