

DOI: 10.19650/j.cnki.cjsi.J2007210

基于语音频谱图像特征的人体疲劳检测方法^{*}

李响¹, 李国正², 邓明君¹, 万平¹, 严利鑫¹

(1. 华东交通大学交通运输与物流学院 南昌 330013; 2. 北京交通大学机械与电子控制工程学院 北京 100044)

摘要:为了将语谱图的可视化图像分析手段有效应用于人体疲劳检测,提出一种基于语音频谱图像特征的人体疲劳检测方法。首先,在研究分析人体疲劳对语谱图影响机理的基础上,对语谱图进行基于听觉感知理论的 Mel 频率拉伸变换,以突出易受疲劳影响的感兴趣区域。其次,将 Mel 频率拉伸后的语谱图分割为 24 个相互交叠的临界频带子图,并从各子图在 4 个方向上的灰度共生矩阵中分别提取了 15 种纹理特征参数用于语音疲劳信息的定量表征。最后,建立多子带疲劳信息融合的人体疲劳检测模型,针对各临界频带子图特征分别设计特征层分类器进行分布检测,并通过决策层的多分类器融合判决得到最终的疲劳检测结果。实验结果表明,该方法所提取的语音频谱图像特征具有比传统声学特征更好的疲劳表征能力,同时该方法的人体疲劳检测效果也优于现有的语谱图特征识别方法。

关键词: 语音; 人体疲劳检测; 频谱图; 灰度共生矩阵; 融合判决

中图分类号: TN912.3 TH701 **文献标识码:** A **国家标准学科分类代码:** 510.40

A human fatigue detection method based on speech spectrogram features

Li Xiang¹, Li Guozheng², Deng Mingjun¹, Wan Ping¹, Yan Lixin¹

(1. School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China;

2. School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: To apply the visual image analysis of speech spectrogram to human fatigue detection effectively, a human fatigue detection method based on speech spectrogram features is proposed. Firstly, the influence mechanism analysis of human fatigue on speech spectrogram is analyzed. The Mel frequency stretching transform of speech spectrogram based on the auditory perception theory is used to highlight the region of interest which is susceptible to fatigue. Secondly, the Mel frequency stretched spectrogram is divided into 24 overlapping critical frequency band sub-images, and 15 texture features are extracted from the gray level co-occurrence matrixes of each sub-image in 4 directions to quantitatively describe the fatigue information. Finally, a human fatigue detection model based on multi sub-bands fatigue information fusion is formulated by designing the feature-layer classifier for distribution detecting the features of each critical frequency band. In this way, the fatigue detection result can be achieved, which is based on the decision-level multi-classifiers fusion decision. Experimental results show that the extracted speech spectrogram features have stronger fatigue classification ability than traditional acoustic features. The fatigue detection effectiveness of this method is also better than the existing spectrogram feature recognition methods.

Keywords: speech; human fatigue detection; spectrogram; gray level co-occurrence matrix; fusion decision

0 引言

在交通运输领域,人体疲劳对驾驶人员的作业稳定性、可靠性和安全性均有较大影响,极易引发交通事故,

因此人体疲劳的检测技术一直是业界的热点^[1]。鉴于航空及铁路等高安全等级的交通运输行业,均要求其驾驶人员在值乘过程中全程采用标准作业用语^[2];同时在公路交通方面,车载语音交互系统也逐渐成为车辆标配,驾驶人员可通过语音来实现对车辆的辅助监控^[3],这

收稿日期:2020-12-08 Received Date: 2020-12-08

^{*} 基金项目:国家自然科学基金(51965021,52062014)、江西省自然科学基金(20202BABL202017)项目资助

就为应用语音信号检测人体疲劳状态提供了广阔的应用前景。并且人体疲劳的语音检测方法还可以弥补现有面部图像检测方法^[4]和生理指标检测方法^[5]在实时性、非接触性、环境适应性和设备复杂昂贵等方面的不足^[6-7]。因此,人体疲劳状态的语音检测方法成为了国内外学者的重点研究方向^[6-8]。

在人体疲劳状态的语音检测方法中,疲劳相关语音特征的提取是最为关键的一环,将直接决定检测效果^[6]。现有研究常用的疲劳相关语音特征主要集中在传统声学特征范畴,包括音质、韵律、频域和非线性特征等几类^[7]。但这些特征往往局限于单一的时域或频域分析,忽略了人体疲劳对语音信号在时、频域共同影响的相关性,以致分析角度不够全面和细致,检测效果有待提升^[8]。近些年逐渐兴起的针对语音频谱的二维可视化图像分析手段^[9-12]为人体疲劳检测提供了新思路。语音频谱图像,又简称为语谱图,能够将语音信号的时、频域联合分布进行可视化表达,从中分析并提取相关的图像特征可以突破传统声学特征的单一性,从时、频域联合分析的角度获取更多传统声学特征无法表征的语音信息,在诸多语音识别领域也取得了一些十分有意义的成果^[9-12]。

然而当前语谱图像特征的实际应用还不够成熟,所采用的分析方法大多仍沿用传统空间图像的处理方法,存在着较多的问题和挑战^[9]。语谱图的横、竖坐标轴及各像素的取值,均有其自身显著的声学特点,在物理意义上与传统空间图像是完全不同的^[10]。那么采用传统空间图像的处理方法来提取语谱图特征是否合适,是值得进一步探讨的。尤其是传统空间图像处理方法的侧重点大多集中在抑制图像旋转变换、空间尺度变化和光照(阴影)变化所带来的干扰问题^[11-12]上,为了提高空间图像识别的准确性和鲁棒性,所提出的一系列图像预处理、图像分割、底层描述子和特征提取计算等经典处理步骤,往往会损失上述图像变化信息。而语谱图中的这些图像变化信息大多又是与人体疲劳高度相关的语音信息,例如语谱图的横向尺度变化代表语速变化,纵向尺度变化代表频率变化,像素点的取值则代表语音的能量或幅度。如果在语谱图特征提取过程中引入空间尺度或旋转不变性等常规图像处理时,这些重要的语音信息就会有所损失,便难以充分表达语音中所包含的人体疲劳信息。因此将语谱图像特征应用于人体疲劳信息的提取与检测是十分有潜力的,但是在具体实现上还有待更为深入的研究。

为此,本文在研究分析人体疲劳对语谱图影响机理的基础上,遵循语谱图的物理意义和声学特点,提出一种基于语音频谱图像特征的人体疲劳检测方法。首先对语谱图进行基于听觉感知理论的 Mel 频率拉伸变换,以获得疲劳声纹变化更为清晰细致的可视化 Mel 频谱图像表

达。随后分割并提取其 24 个临界频率子带图像在 4 个方向上的各 15 种灰度共生矩阵特征,用以表征语谱图所包含的疲劳信息。最后通过构建多子带疲劳信息融合的检测模型,获得人体疲劳状态的检测结果。实验阶段验证了该方法的可行性和有效性。

1 语谱图的疲劳影响机理分析与预处理

1.1 语音频谱图像的生成

对时域连续采样的语音信号 $s(t)$ 进行分帧、加窗和短时离散傅里叶变换处理^[10],有:

$$X(n, f) = \frac{1}{N} \sum_{t=0}^{N-1} h(t) s(t, n) e^{-j \frac{2\pi}{N} f t} \quad (1)$$

式中: t 为时域采样点数; n 为语音帧数; f 为频率点数; $s(t, n)$ 代表第 n 帧语音信号; $h(t)$ 为窗函数; N 为窗口长度。随后便可计算语音信号的对数能量谱图为:

$$F(n, f) = 20 \lg[|X(n, f)|] \quad (2)$$

由此获得的语谱图 $F(n, f)$ 是一种能够直观表示语音频谱随时间变化的二维图像,其横轴 n 对应时间,纵轴 f 对应频率,各坐标点的取值则对应某时刻某频率成分的能量大小。据此,声音随时间的谐振频率变化在语谱图中就会呈现出不同的可视化图像纹理变化^[12]。

1.2 人体疲劳对语谱图的影响机理分析

在语音识别领域,区别于语音情感识别的“人体短期状态检测”(情感状态主要受心理因素影响,在短时就可能发生剧烈变化)以及性别、年龄、身份等生物特征识别的“人体长期状态检测”(此类状态在长时期内不会有较大改变),人体疲劳检测则属于“人体中期状态检测”(还包括醉酒态和疾病态等),此类中期状态往往会对人体及其发声系统产生持续数十分钟乃至数天的负面影响,且自身难以或无法主观控制状态的改变^[6]。那么人体疲劳对发声器官及语音信号的影响^[6-8]主要体现在:人体疲劳时,呼吸平缓使得声压能量降低;声带松弛导致基音频率下降;脑活力及语音规划能力下降,导致语音清晰度降低、韵律异常和语速变慢;声道、喉部和面部肌肉的松弛会使得频谱中共振峰频率整体下降,并伴随着带宽变宽;同时,声道壁温度和粘弹性的变化,会加大声道壁与语音气流的摩擦,从而进一步降低共振峰的位置并相应增加带宽,该现象在发浊音段的 200~3 000 Hz 低频部分尤为明显^[8]。

这种疲劳影响可以在语谱图中直观地表现出来,例如同一说话人在正常状态与疲劳状态下,说出同一句语音“司机明白”的语谱图(语音采样率为 22.05 kHz,帧长取 1 024,帧间覆盖率取 50%,窗函数为汉宁窗)如图 1 所示。对比图 1(c)和(d)的不同状态语谱图可以看出语音纹理在图中的上述疲劳影响变化。但是纵观整个语谱图,语音信号在浊音段的 200~3 000 Hz 低频部分所占的

比例较低,在图像中体现为横向声纹区域较小且不够清晰。根据上述疲劳影响机理分析可知,该区域也是最易

受疲劳影响的感兴趣区域,其尺寸过小和分辨率较低,对后续疲劳信息的提取与检测是不利的。

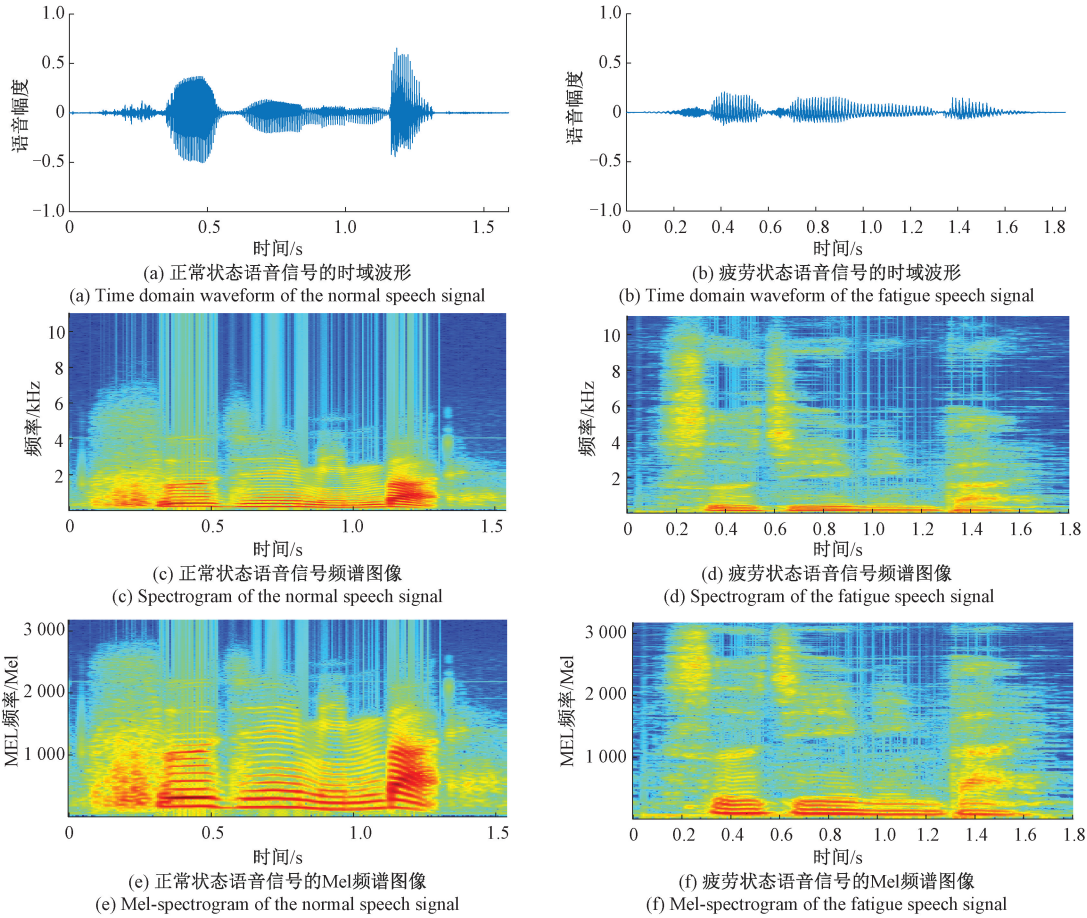


图 1 不同状态下语音信号的频谱图像及其 Mel 频率变换结果

Fig. 1 Spectrograms in different states and their Mel-frequency transformation results

1.3 基于听觉感知特性的语谱图预处理过程

引入基于听觉感知理论的 Mel 频率变换^[13]对语谱图进行预处理。Mel 频率与线性频率之间的变换方法为

$$Mel(f) = 2595 \lg(1 + f/700) \quad (3)$$

该式所描述的线性频率与 Mel 频率之间的映射关系如图 2 所示。对比横、竖坐标可见,在原有线性频率尺度下,对疲劳敏感度较高的 200~3 000 Hz 低频部分,仅占所有线性频谱范围的 25.5%;当映射至 Mel 频率尺度后,这部分感兴趣区域的频谱占比约为整个 Mel 频谱范围的 50.2%,具有明显的放大效果。

图 1(c) 和(d) 的 2 幅语谱图经 Mel 频率变换后的 Mel 频谱图,如图 1(e) 和(f) 所示。可以直观看出,将语谱图由线性频率尺度变换至 Mel 频率尺度后,语音信号对疲劳敏感度较高的浊音段低频部分被给予了高分辨率的非线性拉伸,在图像中具有感兴趣区域语音纹理明显放大的效果,从而能够更为细致地

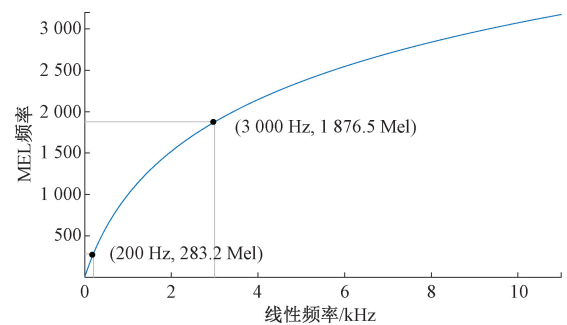


图 2 线性频率与 Mel 频率之间的映射关系

Fig. 2 Mapping relationship between linear and Mel frequency

反映人体疲劳对语音信号的影响,为后续图像纹理特征的提取提供良好的疲劳信息表达。此外,这一遵循听觉感知理论的预处理过程,还可为后续图像分割、灰度级量化等特征提取参数的确定,提供理论借鉴和科学依据。

2 基于 GLCM 的语谱图纹理特征提取

2.1 语谱图纹理特征的选择

现有研究在提取语音频谱图像特征时,通常采用 Gabor 滤波^[11-12]结合局部二值模式(local binary pattern, LBP)特征^[11]或图像几何不变 Hu 矩^[12]等图像纹理特征来量化描述语谱图中的声纹信息。但这些纹理特征在用于疲劳检测时,会存在一定的不足:一方面,Gabor 滤波和 LBP 特征的特点在于舍弃图像整体的低频灰度信息,以突出邻近像素点间的高频纹理变化或符号信息,进而获得对环境光照变化的良好适应性,但是对于语谱图来说,整体的幅度信息表现的是语音声压能量的变化,这是与疲劳高度相关的信息而不宜舍弃;另一方面,LBP 特征和 Hu 矩特征还具有图像旋转不变性的特点,由于语谱图在空间尺度上并不会旋转,其中部分纹理方向的改变,描述得往往是语音音节过渡时音调和韵律的联合变化,旋转不变性反而会导致这些语音变化信息丢失,从而会影响其疲劳表征能力。

为此,本文采用基于灰度共生矩阵(gray-level co-occurrence matrix, GLCM)的图像纹理特征^[14]来描述语谱图所包含的疲劳信息。相较于现有常用的图像纹理特征而言,GLCM 是对图像中保持一定距离的两像素点分别具有特定灰度的情况进行统计得到的,其中不仅包含了多个方向上像素间的灰度变化信息,还在一定程度上保留了像素点本身的灰度分布信息,因此通过 GLCM 所构建的一系列图像纹理特征将更为适用于语谱图中人体疲劳信息的提取与分类。

2.2 灰度共生矩阵的生成

对于尺寸为 $m \times n$ 的图像 $I(x, y)$,该图像的 GLCM 是 $I(x, y)$ 中所有像素点在给定偏移量处灰度阶跃变化的统计矩阵。设其每个像素点有 L 个灰度级,那么该图像 GLCM 的尺寸为 $L \times L$,其中第 i 行第 j 列元素的取值为

$$g(i, j) = \sum_{x=1}^m \sum_{y=1}^n \begin{cases} 1, & I(x, y) = i \text{ 且 } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{其他} \end{cases} \quad (4)$$

式中: $I(x, y)$ 为图像中横、竖坐标为 x 和 y 处像素的灰度值; i 为原始像素点的灰度值, j 为偏移像素点的灰度值, i 和 j 的取值均为 $0, 1, 2, \dots, L-1$; Δx 和 Δy 分别为坐标点的横、竖坐标偏移量。坐标偏移量通常采用方向 θ 和步距 d 这 2 个参数来表达,它们与 Δx 、 Δy 之间的转换可以表示为 $\theta = \arctan(\Delta y / \Delta x)$ 和 $d = \sqrt{\Delta x^2 + \Delta y^2}$ 。

根据式(4)的定义,GLCM 统计了图像中各像素点及其偏移点的灰度值分别为 i 和 j 的像素对数量。可见 GLCM 中各个元素值的大小对图像中像素灰度值的周期结构非常敏感,因此由其构建的一系列特征参数常被用于图像纹理的定量描述^[14]。

2.3 语音频谱图像 GLCM 特征的提取

1) GLCM 特征提取窗口的确定

为了区分不同区域图像纹理特征的分类能力,同时减少全图扫描的庞大计算量,在特征提取之前需要对图像进行分割,以确定特征提取窗口^[15]。传统空间图像的特征提取窗口主要是基于图像纹理的空间尺度来确定,通常以固定尺寸的滑动窗口来生成 GLCM。但对于语谱图的分割,应符合其物理意义和声学特点。

对预处理后的语音 Mel 频谱图,首先在横向时域尺度上采用语段的动态划分方法进行分割,具体算法采用 VAD(voice activity detection) 端点检测法^[16]将语音信号在时间上划分为各个语音段。这种在时域上进行的动态语段划分,不仅符合语谱图的声学特点,还在一定程度上保留了疲劳相关的语速(语段时长)信息。

对于纵向 Mel 频域尺度上的图像分割,则根据听觉感知理论,借鉴 MFCC 特征在提取时,设计三角滤波器组的临界频带分割思想^[13],将 Mel 尺度均匀划分为 24 个相互交叠的临界频率子带。由式(3)计算可得各临界频率子带的上、下界频率,如表 1 所示。表中原始数值为 Mel 频率,括号内为对应的线性频率。据此,便可将语音 Mel 频谱图像动态分割为每语段各 24 幅相互交叠的临界频带子图,作为其 GLCM 特征的提取窗口。

表 1 各临界频率子带的上、下界频率范围

Table 1 The up and low range of each critical frequency band

序号	下界 Mel 频率 (线性频率/Hz)	上界 Mel 频率 (线性频率/Hz)
1	0.0 (0.0)	254.1 (177.0)
2	127.1 (83.5)	381.2 (281.7)
3	254.1 (177.0)	508.2 (398.9)
4	381.2 (281.7)	635.3 (530.0)
5	508.2 (398.9)	762.3 (676.8)
6	635.3 (530.0)	889.4 (841.1)
7	762.3 (676.8)	1 016.4 (1 025)
8	889.4 (841.1)	1 143.5 (1 230.8)
9	1 016.4 (1 025.0)	1 270.5 (1 461.2)
10	1 143.5 (1 230.8)	1 397.6 (1 719.1)
11	1 270.5 (1 461.2)	1 524.6 (2 007.8)
12	1 397.6 (1 719.1)	1 651.7 (2 331.0)
13	1 524.6 (2 007.8)	1 778.7 (2 692.7)
14	1 651.7 (2 331.0)	1 905.8 (3 097.5)
15	1 778.7 (2 692.7)	2 032.8 (3 550.7)
16	1 905.8 (3 097.5)	2 159.9 (4 058.0)
17	2 032.8 (3 550.7)	2 286.9 (4 625.8)
18	2 159.9 (4 058.0)	2 414.0 (5 261.3)
19	2 286.9 (4 625.8)	2 541.0 (5 972.7)
20	2 414.0 (5 261.3)	2 668.1 (6 769.0)
21	2 541.0 (5 972.7)	2 795.1 (7 660.4)
22	2 668.1 (6 769.0)	2 922.2 (8 658.0)
23	2 795.1 (7 660.4)	3 049.2 (9 774.8)
24	2 922.2 (8 658.0)	3 176.3 (11 025.0)

2) 语音频谱图像灰度级 L 的确定

GLCM 的尺寸与灰度级 L 有关,为了降低 GLCM 特征提取的计算量,传统空间图像通常要将 256 个灰度级均匀压缩成较低的 16 个灰度级,进而降低 GLCM 的阶数,提高计算效率^[15]。但语谱图中像素点的取值大小代表的是语音的幅度或能量,其单位为 dB,灰度级也并非 256 级,如果按传统方法来均匀压缩灰度级是不够合理的。根据语音的声学特点和听觉感知理论,语音的能量幅度约在 30~75 dB 范围内,人耳的响度分辨率可设为 3dB^[17]。据此对语谱图的灰度级进行分段压缩,将低于 30 dB 的像素点取值设为最低级 0 级,30~75 dB 之间按 3 dB 一个梯度划分为 1~15 级,高于 75 dB 的设为最高 15 级。此时,语谱图的灰度级仍为 16 级,所生成的 GLCM 尺寸仍为 16×16,但灰度级的压缩过程将语音幅度之外的像素点取值分别压缩至 0 级和 15 级,在后续提取图像纹理特征时不再考虑这些像素的灰度变化,相较于传统图像的灰度级压缩方法更为科学合理。

3) GLCM 生成方向 θ 和步距 d 的确定

GLCM 的生成还与像素对之间的方向 θ 和步距 d 这两个坐标偏移量有关。对于传统空间图像,方向 θ 一般取 0°、45°、90° 和 135° 来分别生成 4 个方向上的 GLCM^[14-15]。传统方法为使其特征值获得旋转不变性,还需对这 4 个方向上的 GLCM 求取平均值,以得到最终的纹理特征^[15]。但旋转不变性并不适用于语谱图像的纹理特征提取,会导致语音纹理的方向信息丢失,为此本文保留这 4 个方向上的 GLCM 特征值共同用于疲劳检测,其有效性在实验部分进行分析讨论。而对于步距 d ,一般根据图像纹理的特点取 1~10 之间的特定值^[14],同样在实验部分对其进行参数寻优确定。

4) GLCM 特征参数的提取计算

GLCM 一般不直接用于图像纹理的描述,而是由其构造出一系列统计参数作为纹理特征用于图像的分类识别^[14]。本文选取了 15 种 GLCM 特征参数^[14-15]用于语音疲劳分类,其计算方法和描述说明,如表 2 所示。

表 2 GLCM 特征的计算公式及描述说明

Table 2 The calculation formulas and descriptions of GLCM features

GLCM 特征	计算公式	描述说明
角二阶矩 Angular Second Moment (ASM)	$ASM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} g^2(i, j)$	角二阶矩又称为能量,主要描述图像灰度分布的均匀程度和纹理的粗细程度。
对比度 Contrast (CON)	$CON = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [(i-j)^2 \cdot g(i, j)]$	对比度的计算权重随 GLCM 各元素与对角线的距离以指数方式增长,用以反映图像纹理的清晰度。
相异性 Dissimilarity (DIS)	$DIS = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [i-j \cdot g(i, j)]$	相异性的计算权重随 GLCM 各元素与对角线的距离以线性方式增长,以描述图像纹理沟纹的深浅。
逆差矩 Inverse Differential Moment (IDM)	$IDM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{g(i, j)}{1 + (i-j)^2}$	逆差矩又称为同质性,主要描述图像局部纹理的粗糙度,粗纹理的逆差矩较大,细纹理的逆差矩较小。
熵 Entropy (ENT)	$ENT = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} g(i, j) \lg[g(i, j)]$	熵是图像包含信息量的随机性度量,其大小反映了图像纹理的复杂程度。
和熵 Sum Entropy (SEN)	$SEN = - \sum_{k=2}^{2(L-1)} g_{i+j}(k) \lg[g_{i+j}(k)]$	和熵主要描述图像中像素对灰度值之和的重复性,其大小反映了图像中灰度变化的平缓或突变程度。
差熵 Difference Entropy (DEN)	$DEN = - \sum_{t=0}^{L-1} g_{i-j}(t) \lg[g_{i-j}(t)]$	差熵主要描述图像中邻域灰度值差异的随机性或变异性。
集群阴影 Cluster Shade (CSHD)	$CSHD = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [(i+j-\mu_i-\mu_j)^3 g(i, j)]$	集群阴影用于测量 GLCM 的偏度和均匀性,其大小可反映纹理的褶皱程度,褶皱越多,CSHD 值越大。
集群突出 Cluster Prominence (CPRM)	$CPRM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [(i+j-\mu_i-\mu_j)^4 g(i, j)]$	集群突出反映图像纹理中的突兀情况,突出物和规则纹理的反差越大,CPRM 值越大。
最大概率 Maximum Probability (MPRB)	$MPRB = \max_{i,j} \{g(i, j)\}$	最大概率表示图像中出现次数最多的纹理特征。
相关性 Correlation (COR)	$COR = \frac{1}{\sigma_i \sigma_j} \left[\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i j g(i, j) - \mu_i \mu_j \right]$	相关性描述 GLCM 元素在行或列方向上的相似程度,COR 的大小反映了图像局部纹理的一致性。
自相关 Autocorrelation (ACOR)	$ACOR = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} [i \cdot j \cdot g(i, j)]$	自相关在计算过程中没有考虑均值和方差(相较于相关性),主要反映图像纹理在整体上的相似程度。
均值和 Sum average (SAVG)	$SAVG = \sum_{k=0}^{2L-2} [k \cdot g_{i+j}(k)]$	均值和主要用于图像纹理明暗变化的统计学表征。
方差和 Sum variance (SVAR)	$SVAR = \sum_{k=0}^{2L-2} [(k-ASM)^2 \cdot g_{i+j}(k)]$	方差和主要用于图像纹理周期大小的统计学表征。
差异方差 Difference Variance (DIV)	$DIV = \sum_{t=0}^{L-1} \left\{ \left[t - \sum_{i=0}^{L-1} t \cdot g_{i-j}(t) \right]^2 \cdot g_{i-j}(t) \right\}$	差异方差是相邻像素对之间灰度值差异的方差,主要用于纹理变化情况的统计学表征。

其中, μ_i 和 μ_j 分别为 GLCM 各行与各列元素的均值, σ_i 和 σ_j 分别为 GLCM 各行与各列元素的均方差; $g_{i+j}(k)$ 与 $g_{i-j}(t)$ 的计算方法为:

$$\begin{cases} g_{i+j}(k) = \sum_{i=0}^{L-1} \sum_{\substack{j=0 \\ |i+j|=k}}^{L-1} g(i, j), k = 0, 1, \dots, 2L-2 \\ g_{i-j}(k) = \sum_{i=0}^{L-1} \sum_{\substack{j=0 \\ |i-j|=k}}^{L-1} g(i, j), t = 0, 1, \dots, L-1 \end{cases} \quad (5)$$

从表 2 所示的 GLCM 特征计算方法及其描述说明可见, 这 15 种 GLCM 特征可分别从不同角度充分描述语谱图的声纹纹理信息。

3 多子带疲劳信息融合的人体疲劳检测模型

在分类检测模型的搭建方面, 首先需要考虑的是特征参数的分类能力和空间维度。一方面, 根据人体疲劳

对语音信号的影响机理分析可知, 语音信号在不同频段内对人体疲劳的敏感程度不同, 也即图像分割后各临界频带子图 GLCM 特征对人体疲劳的分类表征能力是不同的; 另一方面, 各段语音信号可通过 24 幅临界频带子图, 在 4 个方向上, 各获得 15 个 GLCM 特征, 共计 $24 \times 4 \times 15 = 1440$ 维, 特征维数显然过高。如果将这 1440 维分类能力不同的特征参数串接成 1 个特征向量, 仅设计 1 个分类器来进行疲劳分类的话, 势必其疲劳检测效果不够理想, 同时过高的特征维度还会引起维度灾难问题^[7], 导致分类器无法训练或性能退化。

为此, 提出一种多子带疲劳信息融合的人体疲劳检测模型。针对 24 个临界频带子图的 60 维 GLCM 特征 (4 个方向各 15 个特征), 分别设计 24 个分类器进行特征层的分布检测; 然后通过决策层的二次建模, 实现多分类器分布检测结果的融合判决, 以得到最终的人体疲劳检测结果。该检测模型的多分类器级联结构与检测流程, 如图 3 所示。

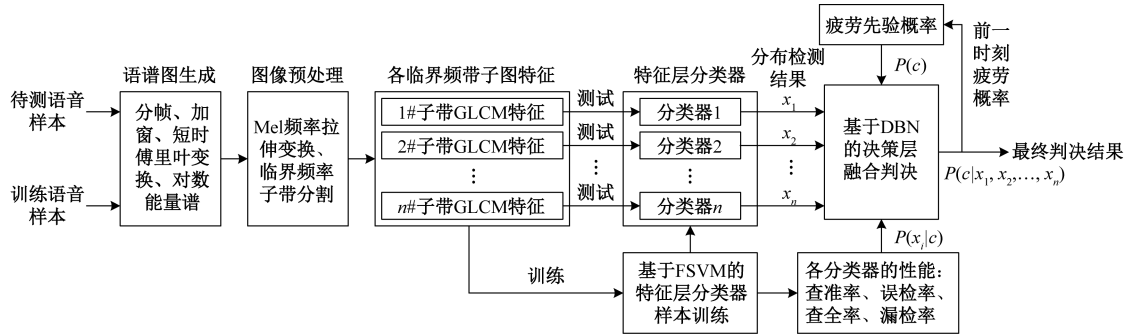


图 3 人体疲劳检测模型的结构与检测流程

Fig. 3 Structure and detection flow of the human fatigue detection model

在 24 个特征层分类器的设计上, 鉴于支持向量机算法在语谱图特征分类上的优势和广泛应用^[11-13], 本文采用模糊化改进后的支持向量机 (fuzzy support vector machine, FSVM) 算法^[8]来进一步抑制中间态噪声和孤立点对分类器性能的不利影响, 以解决语音样本在疲劳分类时的不确定性问题。由于各临界频带子图的 GLCM 特征对疲劳的表征能力不同, 那么各特征层 FSVM 分类器的检测性能也不一定相同。对这 24 个特征层分类器的检测结果进行决策层二次融合判决时, 还需考虑分类器检测性能差异的影响。

在决策层融合判决模型的设计上, 采用动态贝叶斯网络 (dynamic Bayesian network, DBN) 算法^[8]实现。设语音样本的类别变量为 c , 令 $c=1$ 为疲劳类, $c=0$ 为正常类。各特征层分类器的检测结果为 $x_1, x_2, \dots, x_n$, 同样取值为 1 代表检测结果为疲劳状态, 取值为 0 代表正常状态。各分类器的检测性能便可根据极大似然法以条件概率分布 $P(x_i|c)$ 表示^[8], 例如第 i 个分类器在样本训练时统计得到的查准率可作为条件概率 $P(x_i=0|c=0)$ 的

无偏估计, 同理误检率为 $P(x_i=1|c=0)$ 的估计值, 漏检率为 $P(x_i=0|c=1)$ 的估计值, 查全率为 $P(x_i=1|c=1)$ 的估计值。那么当所有特征层分类器的检测结果 $x_1, x_2, \dots, x_n$ 及其检测性能的条件概率分布 $P(x_i|c)$ 均为已知的情况下, 根据贝叶斯理论, 语音样本类别的后验概率可以表示为:

$$\begin{aligned} P(c|x_1, x_2, \dots, x_n) &= \frac{P(c, x_1, x_2, \dots, x_n)}{P(x_1, x_2, \dots, x_n)} = \\ &= \frac{P(c) \prod_{i=1}^n P(x_i|c)}{\sum_{c=0}^1 \left[\prod_{i=1}^n P(x_i|c) P(c) \right]} \\ &\propto \alpha P(c) \prod_{i=1}^n P(x_i|c) \end{aligned} \quad (6)$$

式中: α 为规范化常数。由此便可计算得出待测语音样本类别为正常类的后验概率 $P(c=0|x_1, x_2, \dots, x_n)$ 以及为疲劳类的后验概率 $P(c=1|x_1, x_2, \dots, x_n)$ 。取其中概率最大者, 即为最终的决策层融合判决结果 C , 有:

$$C = \arg \max_{c(x_1, x_2, \dots, x_n)} \left\{ P(c) \prod_{i=1}^n P(x_i | c) \right\} \quad (7)$$

其中,先验概率分布 $P(c)$ 的确定,通常会关系到贝叶斯推理结果的可信度,尤其是当多个特征层分类器的检测结果不一致而存在判决矛盾时,先验概率 $P(c)$ 的大小将对最终的判决结果产生影响。鉴于语音信号的连续性以及人体疲劳的时变规律,也即人体疲劳状态的产生与恢复在时间上应是连续相关的,短时不会发生剧烈变化,那么其概率值 $P(c)$ 在相邻时序的语音信号间也同样不会剧烈变化。因此可根据隐马尔可夫理论,引入时间变量 t 来建立动态贝叶斯网络模型^[8],将前一“ $t-1$ ”时刻的后验概率作为当前“ t ”时刻的先验概率估计并实时动态更新,有:

$$P_t(c) = P_{t-1}(c | x_1, x_2, \dots, x_n) \quad (8)$$

此外,对于初始“ $t=0$ ”时刻的先验概率 $P_{t=0}(c)$,或是在时间上不连续的语音样本的先验概率,为了避免主观因素对贝叶斯概率推理的影响,可参考文献^[18]的疲劳风险动态量化评价方法,根据被测者在不同作业条件下的疲劳风险予以客观确定。

4 实验

为符合实际应用场景,人体疲劳实验安排在室内列车模拟驾驶操作台^[7]上进行。被试者共计 60 人,男性 32 人,女性 28 人,年龄 20~45 岁,身体状况良好,饮食情况正常,实验前均清楚实验目的并进行了模拟驾驶操作台的适应性训练。实验真实模拟铁路机车驾驶人的实际作业过程,均为单次持续 8 小时左右的列车驾驶任务(昼夜均有值乘任务)且部分被试者在实验前还进行了不同程度的睡眠剥夺。实验期间,被试者需按要求说出“操作、联控、播报、呼唤、应答、通信”等多类覆盖驾驶作业全过程的标准作业用语,并以此作为实验语音样本数据,语音采集设备参照铁路驾驶舱语音记录仪设置,22.05 kHz 采样率,16 位量化精度。

对实验语音样本的类别标记,采用半小时填写 1 次的疲劳症状自测量表^[5](10 分制,分值越小代表疲劳程度越高)与实时心率信号监测相结合的人体疲劳状态判别方法^[7],即疲劳量表低于 5 分且短时心率低于平均心率时段内的语音样本标记为疲劳状态样本,其余标记为正常状态语音样本。所有实验共采集语音样本 5 427 份,其中正常状态语音样本 4 456 份,疲劳状态语音样本 971 份。实验采用十折交叉验证法将所有样本随机均分为 10 份,其中 9 份为训练样本,1 份作为测试样本,取 10 次测试结果的均值用于检测效果的对比分析。

4.1 特征提取参数对比实验

为了验证不同特征提取参数对检测结果的影响,进行 GLCM 生成方向 θ 和步距 d 的参数对比实验。在 GLCM 生成方向 θ 上,本文算法与传统方法的不同之处在于保留了 4 个方向的 GLCM 特征,而没有对其取均值来获得 GLCM 特征的旋转不变性,并以此将语音纹理的方向信息也用于疲劳检测。为验证其有效性进行了参数对比实验,检测结果如图 4 所示。同时,实验还对比了步距 d 在不同取值下的检测结果,步距 d 是影响 GLCM 的重要参数,其大小不仅直接决定了 GLCM 在生成时像素对之间的距离(对方向上纹理粗细的表达较为敏感),还决定了方向上图像边缘有多少像素能够参与到 GLCM 的统计中(即图像纹理信息的丰富程度)。由于步距 d 较大时,图像边缘将有较多的像素信息丢失,因此步距 d 仅取 1~10 进行对比, d 在不同取值下所有语音样本人体疲劳检测的平均正确率曲线也如图 4 所示。

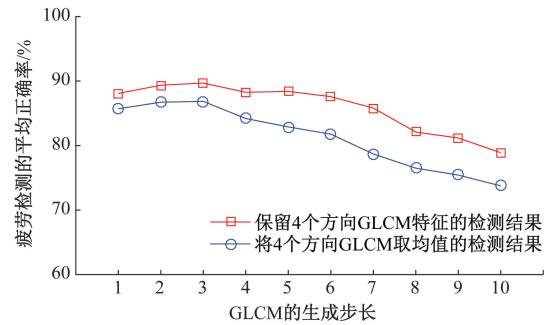


图4 不同特征提取参数的人体疲劳检测结果

Fig. 4 Human fatigue detection results with different feature extraction parameters

首先对比图 4 的两条平均正确率曲线可见,4 个方向取均值的 GLCM 特征在应用于人体疲劳检测时,其平均正确率显著低于保留 4 个方向的 GLCM 特征,这在一定程度上验证了图像特征的旋转不变性会使得语音谱图中语音声纹的方向信息丢失,从而影响了 GLCM 特征对音节过渡段音调和韵律变化的准确描述,最终导致疲劳分类能力的降低。同时,考察不同步距 d 的取值对检测结果的影响。如图 4 的上部曲线所示,当步距取值较小($d=1$ 或 2)时,其检测效果并不是最佳,这是由于当纹理较粗时,邻近像素间的变化较为缓慢,GLCM 的元素大多集中在对角线附近而难以得到图像纹理的良好表达;而当步距取值较大时,图像边缘信息的大幅丢失又会对疲劳检测带来不利影响,以致检测效果较差。最终在步距取 $d=3$ 时人体疲劳检测的平均正确率最高,达到 89.7%,由此可确定最佳的步距取值为 $d=3$ 。

值得指出的是,保留 4 个方向的 GLCM 特征(各分类

器的特征维数由 15 维升至 $4 \times 15 = 60$ 维) 以及步距 d 的具体取值, 均会对特征提取的计算量和疲劳检测的实时性造成影响。当确定上述特征提取参数后, 本文方法对语段时长在 0.5~20 s 范围内的实验语音样本, 在 PC 平台(4 核 3.6 GHz CPU, 8G 内存) 上的疲劳检测耗时约为 0.87~2.56 s, 相较于人体疲劳状态的变化速度以及实际语音的持续、间隔时间而言, 基本能够满足人体疲劳检测的实时性要求。

4.2 传统声学特征对比实验

目前用于人体疲劳检测的传统声学特征主要有

2011 年 INTERSPEECH 人体中期状态(疲劳、醉酒及病态等) 语音识别挑战赛所采用的声学特征集^[6], 以及在此基础上优化并增加浊音段非线性动力学特征的组合特征集^[7]。本文方法与这 2 种传统声学特征检测方法的底层描述子、语段统计参数、特征维数和实验语音样本的疲劳检测结果, 如表 3 所示。其中, 前 2 种传统声学特征检测方法的特征选择、特征提取、特征降维和分类检测算法分别参照文献[6]和[7]确定, 所有方法的检测结果也以查准率(与错检率互补)、查全率(与漏检率互补)和平均正确率这 3 项常用指标^[8]表示。

表 3 与传统声学特征的检测结果对比

Table 3 Comparison with the detection results of traditional acoustic features

检测方法	底层描述子(包括帧间变化)	语段统计参数	维数	实验语音样本的疲劳检测结果/%		
				查准率	查全率	平均正确率
INTERSPEECH 2011 挑战赛方法	4 种韵律特征+5 种音质特征+ 50 维频域特征	33 个语段统计值+ 6 个帧数统计值	3 996	82.1	83.0	82.3
文献[7]方法	3 种韵律特征+4 种音质特征+ 12 阶 MFCC+3 种非线性特征	11 个语段统计值+ 6 个帧数统计值	864	86.5	88.3	86.8
本文方法	24 幅 Mel 频谱临界频带子图× 4 个方向上的 GLCM	15 个 GLCM 纹理 特征参数	1 440	89.3	91.4	89.7

从表 3 所示的实验语音样本疲劳检测结果可以看出, 本文方法的查准率、查全率和平均正确率分别为 89.3%、91.4%、和 89.7%, 均高于前 2 种采用传统声学特征的检测方法。说明本文基于可视化时、频域联合分析所提取的语音频谱图像 GLCM 特征, 包含了更多现有声学特征无法表征的语音疲劳信息, 从而在应用于人体疲劳检测时, 拥有较好于传统声学特征的疲劳分类能力。此外, 对比疲劳检测的实时性, 在相同的 PC 处理平台下, 2011 年 INTERSPEECH 挑战赛方法对实验语音样本的检测耗时约为 1.06~4.54 s, 文献[7]方法的检测耗时约为 0.48~1.36 s, 本文方法为 0.87~2.56 s, 耗时介于两种方法之间, 计算量较传统特征检测方法而言并无显著提升, 但均能满足疲劳检测的实时性要求。由此可见, 较佳的检测效果, 不弱的检测效率, 表明了本文方法在实际应用时的可行性和有效性。

4.3 常用语音频谱图像特征对比实验

为了进一步验证本文方法所提取的 GLCM 特征, 所提出的图像预处理与分割方法, 以及所建立的分类检测模型是否更适用于人体疲劳检测, 将本文方法与目前常用的语谱图纹理特征识别方法^[11-12]分别进行了不同的语谱图特征、预处理方法、特征提取窗口和分类检测模型条件下, 实验语音样本的人体疲劳检测实验, 检测结果如表 4 所示。其中, 方法 1 和方法 2 的语谱图纹理特征均

采用局部 LBP 特征, 区别在于方法 1 参照文献[11]采用常规的 Gabor 滤波预处理方法, 特征提取窗口也按参数寻优方法^[11]确定为 32×32 的固定尺寸滑动窗口, 分类器选为 SVM; 而方法 2 的预处理方法和特征提取窗口, 均采用本文提出的 Mel 频率变换方法和临界频带子图分割方法, 分类检测模型也采用本文提出的 FSVM 结合 DBN 的多分类器级联结构。同样, 方法 3 和方法 4 的语谱图纹理特征均采用局部 hu 矩特征, 区别在于方法 3 参照文献[12]采用 Gabor 滤波预处理方法和固定尺寸的滑动特征提取窗口, 分类器选为 SVM; 方法 4 则采用本文提出的图像预处理步骤、临界频带子图的特征提取窗口和多分类级联的检测模型结构。

对比表 4 中方法 1 和方法 2 的检测结果, 以及方法 3 和方法 4 的检测结果可以看出, 不论选择何种图像纹理特征, 当采用本文的 Mel 频率变换预处理方法并据此进行临界频带子图分割来确定特征提取窗口时, 对实验语音样本进行人体疲劳检测的效果均优于采用常规 Gabor 滤波预处理和固定尺寸滑动提取窗口的检测方法。这说明传统空间图像的预处理方法和特征提取窗口并不完全适用于语谱图的疲劳分析与识别, 传统空间图像处理办法为了突出图像纹理而采用的 Gabor 滤波往往会损失语谱图的低频能量变化信息, 而不利于疲劳信息的提取和识别。实际应用时, 应遵循语谱图自身的声学特点及

表4 常用语谱图纹理特征的检测结果对比

Table 4 Comparison with the detection results of common speech spectrogram texture features

检测方法	语谱图特征	预处理方法	特征提取窗口	分类器模型	实验语音样本的疲劳检测结果/%		
					查准率	查全率	平均正确率
方法1	LBP特征	Gabor滤波	固定尺寸滑动窗口	SVM	84.6	86.3	84.9
方法2	LBP特征	Mel频率变换	临界频带子图分割	FSVM+DBN	86.8	88.5	87.1
方法3	Hu矩特征	Gabor滤波	固定尺寸滑动窗口	SVM	82.7	84.4	83.0
方法4	Hu矩特征	Mel频率变换	临界频带子图分割	FSVM+DBN	84.9	86.7	85.2
本文方法	GLCM特征	Mel频率变换	临界频带子图分割	FSVM+DBN	89.3	91.4	89.7

其语音纹理的疲劳表征机理,来突出感兴趣区域才能有利于后续的疲劳特征提取。同时从分类检测模型的区别上也可以看出,如果将语谱图中对疲劳表征能力不同的各频段局部纹理特征串接在一起,仅用一个分类器来进行疲劳识别,其检测能力势必不佳。这也在一定程度上说明本文所提出的多分类器级联结构是更为合理有效的人体疲劳检测模型。

同时,将本文方法与方法2及方法4的检测结果进行对比,以验证所选用的GLCM特征相较于现有常用的LBP特征和Hu矩特征的有效性。从表4可以看出,本文方法的检测结果要优于方法2和方法4,其对实验语音样本进行疲劳检测的查准率、查全率和平均正确率相较于方法2,分别高出2.5%、2.9%和2.6%,相较于方法4则分别高出4.4%、4.7%和4.5%。这说明GLCM特征相较于现有常用的LBP特征和Hu矩特征而言,不仅能够良好描述语谱图中疲劳相关的声纹变化信息,还在一定程度上保留了像素的原始灰度(幅度、能量)和纹理方向(音调、韵律)变化信息,从而获得更好的疲劳分类检测效果。此外,鉴于LBP特征和Hu矩特征也兼有旋转不变性的特点,方法2和方法4的检测效果不佳,也能在一定程度上再次说明此类传统空间图像特征的旋转不变性并不适用于语音频谱图像的人体疲劳检测。

最终通过一系列对比实验可以得出,所提出的基于语音频谱图像特征的人体疲劳检测方法,在图像预处理、图像分割、特征选用、特征提取和分类检测模型搭建等具体处理步骤上均遵循语谱图的物理意义、声学特点和疲劳表征机理,从而获得较现有方法更好的疲劳检测效果。所提方法对所有实验语音样本进行人体疲劳检测的查准率、查全率和平均正确率均达到最高的89.3%、91.4%和89.7%。由于实际应用时,作业人员的疲劳极易引发事故,而本文方法较好的查全率指标(也即较低的漏检率),则表明绝大多数的人体疲劳状态均可被该方法检出,具有一定的实际应用价值。

5 结 论

为了将语音频谱的二维可视化图像分析手段良好应

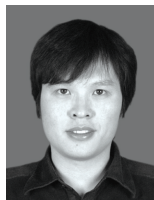
用于人体疲劳检测,提出一种基于语音频谱图像特征的人体疲劳检测方法。在深入分析人体疲劳对语音信号影响机理的基础上,对语谱图进行Mel频率变换以突显易受疲劳影响的感兴趣区域,所得的语音Mel频谱图像能够更为清晰细致地表达疲劳相关的声纹变化。同时,根据听觉感知理论进一步对语音Mel频谱图进行的临界频带分割,可以更为科学合理地确定纹理特征提取窗口。在此基础上提取的24个临界频带子图在4个方向上的15种GLCM纹理特征,不仅包含了语谱图纹理变化信息,还保留了语音幅度和声纹方向相关信息,具有较好的疲劳表征能力。最终建立的多子带疲劳信息融合检测模型,能够在保证检测实时性的同时,获得较为准确可信的疲劳检测结果。实验结果显示,该方法应用于人体疲劳检测的查准率、查全率和平均正确率均高于传统声学特征检测方法和现有语谱图像特征检测方法,具有较好的科学意义和实际的应用价值。

参考文献

- [1] 谢平, 齐孟松, 张园园, 等. 基于多生理信息及迁移学习的驾驶疲劳评估[J]. 仪器仪表学报, 2018, 39(10): 223-231.
XIE P, QI M S, ZHANG Y Y, et al. Driver fatigue assessment based on multi-physiological signals and transfer learning [J]. Chinese Journal of Scientific Instrument, 2018, 39(10): 223-231.
- [2] 谢湘, 唐刚, 肖泽革, 等. 飞行驾驶员的应答方式识别[J]. 北京理工大学学报, 2017, 37(7): 744-747.
XIE X, TANG G, XIAO Z P, et al. Speaking style recognition of pilots in flight [J]. Transactions of Beijing Institute of Technology, 2017, 37(7): 744-747.
- [3] ALICIA F R, KLAS I, MARC W, et al. Towards affect-aware vehicles for increasing safety and comfort: Recognising driver emotions from audio recordings in a realistic driving study [J]. IET Intelligent Transport Systems, 2020, 14(10): 1265-1277.
- [4] LI K, GONG Y, REN Z. A fatigue driving detection algorithm based on facial multi-feature fusion [J]. IEEE Access, 2020, 8(6): 101244-101259.

- [5] 张朕, 姜劲, 傅嘉豪, 等. 基于功能近红外光谱的多生理脑力疲劳检测 [J]. 仪器仪表学报, 2017, 38(6): 1345-1352.
ZHANG ZH, JIANG J, FU J H, et al. Multi-physiological mental-fatigue detection based on the functional near infrared spectroscopy [J]. Chinese Journal of Scientific Instrument, 2017, 38 (6): 1345-1352.
- [6] SCHULLER B, STEIDL S, BATLINER A, et al. Medium-term speaker states-A review on intoxication, sleepiness and the first challenge [J]. Computer Speech and Language, 2014, 28(2): 346-374.
- [7] 李响, 李国正, 彭理群, 等. 基于语音特征迁移学习的驾驶疲劳检测 [J]. 铁道学报, 2020, 42(4): 74-81.
LI X, LI G ZH, PENG L Q, et al. Driver fatigue detection based on speech feature transfer learning [J]. Journal of China Railway Society, 2020, 42(4): 74-81.
- [8] 李响, 李国正, 石俊刚, 等. 基于语音心理声学分析的驾驶疲劳检测 [J]. 仪器仪表学报, 2018, 39(10): 166-175.
LI X, LI G ZH, SHI J G, et al. Fatigue driving detection based on speech psychoacoustic analysis [J]. Chinese Journal of Scientific Instrument, 2018, 39 (10): 166-175.
- [9] HUZAIFAH B M S, WYSE L. Applying visual domain style transfer and texture synthesis techniques to audio: Insights and challenges [J]. Neural Computing and Applications, 2020, 32(4): 1051-1065.
- [10] ZHANG Y, DAI S, SONG W, et al. Exposing speech resampling manipulation by local texture analysis on spectrogram images [J]. Electronics, 2020, 9 (1): 23/1-16.
- [11] ABIDIN S, TOGNERI R, SOHEL F. Spectrotemporal analysis using local binary pattern variants for acoustic scene classification [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018, 26(11): 2112-2121.
- [12] TAO H, LIANG R, ZHA C, et al. Spectral features based on local Hu moments of Gabor spectrograms for speech emotion recognition [J]. IEICE Transactions on Information & Systems, 2016, 99(8): 2186-2189.
- [13] LI Q, YANG Y, LAN T, et al. MSP-MFCC: Energy-efficient MFCC feature extraction method with mixed-signal processing architecture for wearable speech recognition applications [J]. IEEE Access, 2020, 8(3): 48720-48730.
- [14] 焦敬品, 李思源, 常予, 等. 集箱管接头内焊缝表面缺陷识别方法研究 [J]. 仪器仪表学报, 2017, 38(12): 3044-3052.
JIAO J P, LI S Y, CHANG Y, et al. Defect classification of weld surface in header pipe joint [J]. Chinese Journal of Scientific Instrument, 2017, 38(12): 3044-3052.
- [15] XING Z, JIA H. Multilevel color image segmentation based on GLCM and improved salp swarm algorithm [J]. IEEE Access, 2019, 7(3): 37672-37690.
- [16] HENNI R, DJEBARI M, DJENDI M. A new joint noise reduction and echo suppression system based on FBSS and automatic voice activity detector [J]. Applied Acoustics, 2020, 168: 1-12.
- [17] SUM K S, PAN J. Subjective evaluation of reverberation times of sound fields with non-exponential decays [J]. Acta Acustica United with Acustica, 2006, 92 (4): 583-592.
- [18] 李响, 徐玉萍, 章海亮. 机车司机驾驶疲劳风险动态量化评价研究 [J]. 中国安全科学学报, 2017, 27(2): 18-23.
LI X, XU Y P, ZHANG H L. A study on dynamic quantitative evaluation of train driver's fatigue risk [J]. China Safety Science Journal, 2017, 27(2): 18-23.

作者简介



李响(通信作者), 2004 年于重庆大学获得学士学位, 分别在 2007 年和 2016 年于北京交通大学获得硕士和博士学位, 现为华东交通大学副教授, 主要研究方向为安全工程、信号处理和人工智能。

E-mail: lixiang@ecjtu.edu.cn

Li Xiang (Corresponding author) received his B.Sc. degree from Chongqing University in 2004, and received his M.Sc. degree and Ph.D. degree both from Beijing Jiaotong University in 2007 and 2016, respectively. He is currently an associate professor at East China Jiaotong University. His main research interests include safety engineering, signal processing and artificial intelligence.



李国正, 分别在 2005 年和 2014 年于北京交通大学先后获得学士和博士学位, 现为北京交通大学副教授, 主要研究方向为信号检测和模式识别。

E-mail: liguozheng@bjtu.edu.cn

Li Guozheng received his B.Sc. degree and Ph.D. degree both from Beijing Jiaotong University in 2005 and 2014, respectively. He is currently an associate professor at Beijing Jiaotong University. His research directions are signal detection and pattern recognition.