

## 基于脑电的虚拟现实诱发下情绪状态分类

张进<sup>1</sup>, 许子明<sup>2</sup>, 周月莹<sup>2</sup>, 王澎湃<sup>2</sup>, 张道强<sup>2</sup>

(1. 南京航空航天大学艺术学院, 南京 211106; 2. 南京航空航天大学计算机科学与技术学院, 南京 211106)

**摘要:** 情绪在歌唱活动中发挥着重要作用, 但现阶段高校声乐课程缺乏对情绪调动的有效训练。由于虚拟现实技术具有真实性、沉浸性的特征, 本文提出将其应用于声乐教学过程中的情绪诱发阶段。为了验证虚拟现实技术对情绪诱发的有效性, 本文提取了不同类型的脑电特征, 对情绪自我想象和虚拟现实情绪诱发两种场景下的情绪状态进行分类, 并对比了情绪分类准确率、情绪自评分数和声乐自评分数, 从主观和客观两方面探究虚拟现实技术对参与者情绪调动的影响。实验结果表明, 相比于传统自我想象, 虚拟现实技术可以极大地诱发参与者情绪, 提升演唱效果, 从而为声乐演唱教学提供一种新的辅助手段。

**关键词:** 脑电; 虚拟现实; 自我想象; 情绪分类; 声乐教学

**中图分类号:** TP391      **文献标志码:** A

## Emotion Classification Induced by Virtual Reality Based on EEG

ZHANG Jin<sup>1</sup>, XU Ziming<sup>2</sup>, ZHOU Yueying<sup>2</sup>, WANG Pengpai<sup>2</sup>, ZHANG Daoqiang<sup>2</sup>

(1. College of Art, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China; 2. College of Computer Science and Technology, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

**Abstract:** Emotion plays an important role in singing. At present, vocal music courses in universities lack effective training for emotion mobilization. Due to the authenticity and immersive of virtual reality (VR) technology, this article applies it to the emotional induction stage in vocal music teaching. To verify the effectiveness of VR technology for emotion induction, different types of electroencephalogram (EEG) features are extracted, and then emotions are classified into two scenarios: emotional self-imagination and VR-induced. The accuracy of emotion classification, self-evaluation score of emotion and self-evaluation score of vocal music are compared to explore the influence of VR on participants' emotion mobilization from both subjective and objective aspects. Experimental results show that compared with traditional self-imagination, VR technology can greatly induce the emotions of participants and enhance the singing performance, thus providing a new auxiliary method for vocal singing teaching.

**Key words:** electroencephalogram (EEG); virtual reality (VR); self-imagination; emotion classification; vocal teaching

## 引言

情绪是人对客观事物的态度体验及相应的行为反应,在人类的行为和心理健康方面有着重要的影响<sup>[1]</sup>。神经科学和认知科学的研究结果表明,情绪的产生与生理活动,特别是与大脑活动密切相关<sup>[2]</sup>,因此可以通过分析大脑活动来识别情绪状态。由于脑电具有高时间分辨率、采集方便和价格低廉等优点<sup>[3]</sup>,基于脑电的情绪识别得到了广泛关注<sup>[4]</sup>。基于脑电的情绪识别一般收集情绪视频、情绪图片等刺激下的脑电数据,提取相关脑电特征,探究脑电特征和不同情绪类别的相关性<sup>[5]</sup>;或使用机器学习算法,基于脑电特征进行情绪分类<sup>[6-7]</sup>。目前,针对脑电的情绪识别技术已应用于对意识障碍患者康复治疗<sup>[8]</sup>、士兵精神状态评估<sup>[9]</sup>、驾驶员驾驶状态监测<sup>[10]</sup>等场景,但尚未涉及对歌唱者情绪诱发状态的研究。

歌唱活动是歌唱者将内在情感呈现出来以供观众感受和认知的形式<sup>[11]</sup>。目前高校的声乐演唱学习多以行业训练和声乐理论知识传授为主,缺乏对声乐学习者进行情感体验和想象力训练的手段,导致声乐演唱训练环节枯燥、固化。因此声乐学习者的感知能力和情感思维能力不足,在歌曲演绎过程中情感表达不够理想,歌唱情绪状态调动不够积极,很难到达声情并茂的歌唱状态要求。

近年来,随着虚拟现实(Virtual reality, VR)技术的发展,它已经被应用于情感诱发,并在大多数研究中成功诱发了情绪<sup>[12]</sup>。此外,虚拟现实结合脑电技术已被应用于情感计算<sup>[13]</sup>、教育教学<sup>[14]</sup>等领域,尤其是沉浸式VR和脑电技术被不断应用于教育教学过程和相关研究,正逐步成为一种发展趋势<sup>[15]</sup>。因此,本文将虚拟现实技术和脑电技术相结合,应用于声乐教学过程中的歌唱情绪诱发和情绪诱发效果评价。

通过对自评量表进行显著性统计分析,对脑电数据进行多种类别的特征提取和情绪分类,从主观和客观两方面探究VR技术对参与者情绪和演唱表现的影响。本文的完整流程如图1所示,采用基于

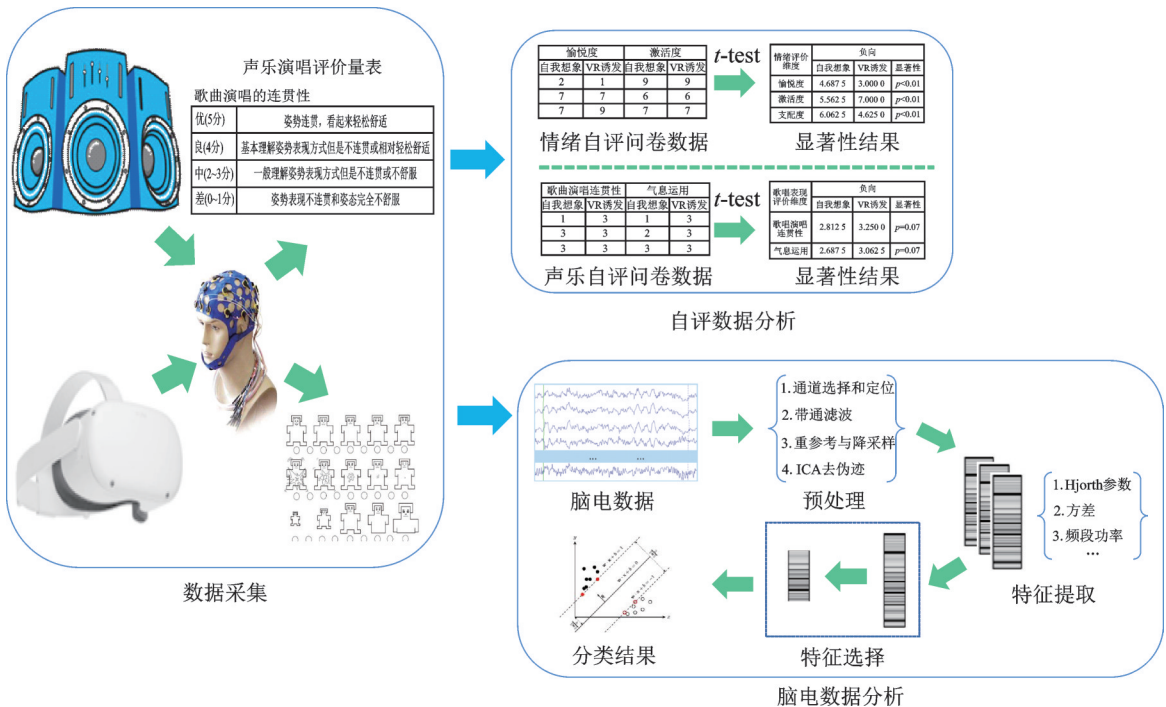


图1 实验流程图

Fig.1 Flow chart of experiment

脑电的情绪分类、情绪自评分数统计分析和歌唱自评分数统计分析3种方式,对16名参与者在自我想象式的情绪调动和VR情绪诱发两种场景下的情绪进行研究。研究表明,相比传统情绪自我调动,沉浸式VR可以更好地诱发参与者的情绪,提升歌唱表现,可为声乐演唱者的学习和教学提供一种新的辅助方式和手段。

## 1 实验

### 1.1 参与者

实验参与者为来自南京航空航天大学艺术学院声乐表演专业的16名大学生(8名女生和8名男生,平均年龄为 $19.5 \pm 1.54$ 岁)。所有参与者听力均正常,视力正常或矫正正常,没有任何脑部或精神疾病,所有参与者均为右利手。实验前一天晚上要求参与者保持充足的睡眠时间。所有参与者均签署一份书面知情同意书。

### 1.2 实验设计

选取3种情绪类别(正向、中性、负向)的6首歌曲作为情绪刺激材料,每种情绪两首歌曲,分别为正性:《我和我的祖国》《在希望的田野上》;负性:《烛光里的妈妈》《时间都去哪儿了》;中性:《牧歌》《贝加尔湖畔》。每首歌曲被剪辑到3 min左右,去除每首歌的歌词演唱,只保留背景音乐。

在实验中,每个被试需要进行两个阶段的实验,情绪自我想象和VR诱发情绪,完整的实验范式如图2所示。情绪自我想象阶段模拟声乐教学场景,播放剪辑的歌曲,要求被试根据听到的歌曲自我想象所对应的场景和情绪。为了更好地模拟真实声乐教学场景,在情绪自我想象阶段的歌曲背景音乐播放之前加入一段情绪引导解说,介绍歌曲的相关背景以及所需要的情绪。在VR诱发情绪方面,针对不同歌曲的具体内容使用对应的素材构建沉浸式虚拟现实全景视频,使得制作出的视频与歌曲的情感和内容吻合,从而达到更好的VR情绪诱发效果。如图3视频截图所示,《我和我的祖国》使用故宫等全国各地景色;《在希望的田野上》使用田野场景;《烛光里的妈妈》使用怀孕母亲的剪影等其他与母亲相关的素材;《时间都去哪了》使用钟表特写等;《贝加尔湖畔》使用贝加尔湖风景;《牧歌》使用草原放牧景色。这些场景的建立是通过搜集与歌曲内容相关的全景视频或图片,并将其剪辑成一段完整的视频作为情绪刺激的沉浸式VR全景视频,视频的背景音乐使用去除歌词后的对应歌曲。在VR诱发情绪阶段,要求被试佩戴VR显示眼镜(Quest2, Oculus, 美国),认真观看上述不同歌曲对应的沉浸式虚拟现

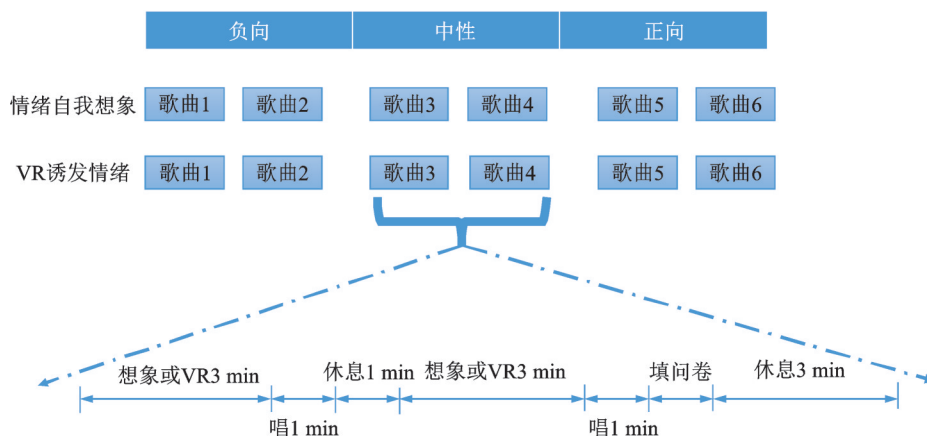


图2 实验范式

Fig.2 Experimental paradigm

实全景视频。

根据情绪类别,将实验的每个阶段划分为3组(负向、中性、正向),每组包括同种情绪类别两首歌曲的实验。每首歌曲的实验流程包括情绪自我想象或观看VR视频3 min,以实现被试情绪的诱发;唱歌1 min;使被试静息休息1 min,休息完成后开始下一首歌曲的实验。每组两首歌曲的实验完成后,要求被试填写两个自评量表:情绪自评量表(Self-assessment manikin, SAM)量表<sup>[16]</sup>和声乐自评量表。在实验中,按照负向、中性、正向的顺序依次播放每组的两首歌曲,每两组之间休息3 min。每个阶段的3组实验结束后休息10 min。在实验中,被试被要求尽可能保持静止,减少肌电等伪迹对脑电的影响。

自评量表中的SAM量表通过1~9分自我评分的形式衡量情绪的愉悦度、激活度和支配度,分数越高表示情绪越强烈(高愉悦、高唤醒、高支配)。声乐自评量表从歌曲演唱连贯性、气息运用、共鸣运用、音准节奏、语言咬字、乐感和情感表达7个方面对歌唱表现进行自评打分,分为优、良、中和差4个等级(对应分数分别为4、3、2和1),分数越高,歌唱表现越好。



图3 6首歌曲虚拟现实全景视频画面

Fig.3 Virtual reality panoramic video pictures of six songs

### 1.3 脑电图记录和预处理

使用64通道便携式无线脑电图系统(NeuSen.W64,博睿康,中国)进行数据采集,采样率设置为1 000 Hz。根据国际标准10-20系统,记录59个电极的脑电数据,以CPz为参考电极,AFz为接地电极。在整个实验过程中,所有电极阻抗都保持在5 k $\Omega$ 以下。

使用EEGLAB<sup>[17]</sup>工具箱对脑电进行预处理。在通道定位后,原始脑电数据经过1~45 Hz的带通滤波,进行全脑平均重参考,降采样到250 Hz。然后,使用独立成分分析(Independent component analysis, ICA)将脑电信号分解为独立成分,使用ICLabel<sup>[18]</sup>插件去除眼动、肌肉运动等伪迹的影响。在预处理之后,将每首歌曲最后150 s的脑电信号划分成1 s时间段的脑电信号样本,用来获得最大的情绪反应<sup>[19]</sup>。

## 2 脑电信号处理方法

### 2.1 脑电特征提取

参考相关研究<sup>[20-21]</sup>,本文主要提取两类特征,包括线性特征和非线性特征。计算出的特征的具体类型和数量如表1所示。其中,每个被试的脑电数据表示为一个矩阵 $X(t)$ ,大小为 $c \times N$ , $N$ 表示时间样本数, $c$ 表示脑电通道数。为方便起见,在下文中,将使用 $x(t)$ 来表示单个通道的时间序列。

表1 本文提取的特征  
Table 1 Features extracted in this research

特征	线性特征	非线性特征
时域特征	Hjorth参数(活动性、移动性、复杂性)	
	均方值	近似熵
	平均绝对振幅	样本熵
	方差	
频域特征	总功率	
	最大功率谱密度	
	最大功率谱密度对应频率	C0复杂度
	5个频段的平均功率(delta、theta、alpha、beta、gamma)	
	半球不对称性	

### 2.1.1 脑电时频域线性特征

#### (1) 时域特征

Hjorth参数<sup>[21-22]</sup>是脑电信号最重要的时域特征之一。Hjorth参数包括3类参数:活动性(Activity)、移动性(Mobility)和复杂性(Complexity),分别代表信号的平均功率、平均频率和频率的变化<sup>[21]</sup>;其定义分别为

$$\text{Activity}(x(t)) = \text{var}(x(t)) \quad (1)$$

$$\text{Mobility}(x(t)) = \sqrt{\frac{\text{Activity}(x'(t))}{\text{Activity}(x(t))}} \quad (2)$$

$$\text{Complexity}(x(t)) = \frac{\text{Mobility}(x'(t))}{\text{Mobility}(x(t))} \quad (3)$$

式中 $x'(t)$ 表示 $x(t)$ 的一阶导数。

对脑电信号计算了均方值(Mean square)、平均绝对振幅(Mean absolute amplitude)和方差(Variance)3个统计特征,其定义分别为

$$\text{Mean square} = E(x^2(t)) = \frac{1}{N} \sum_{i=1}^N x_i^2 \quad (4)$$

$$\text{Mean absolute amplitude} = E(|x(t)|) = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (5)$$

$$\text{Variance} = E\{(x(t) - E(x^2(t)))\} \quad (6)$$

#### (2) 频域特征

利用Welch法和汉宁窗计算信号 $x(t)$ 的频谱,并提取9个线性特征:总功率、最大功率谱密度、最大功率谱密度对应频率、5个频段的平均功率( $\delta$ :1~4 Hz; $\theta$ :4~8 Hz; $\alpha$ :8~13 Hz; $\beta$ :13~30 Hz; $\gamma$ :30~45 Hz)、和半球不对称性。其中,半球不对称性<sup>[21,23-24]</sup>通常通过右半球功率的自然对数减左半球功率的自然对数( $\ln(\text{右功率}) - \ln(\text{左功率})$ )来计算。在59个通道中,最终得到26个半球不对称性特征。

### 2.1.2 脑电非线性特征

研究表明脑电信号有很强的非线性结构<sup>[25]</sup>,因此非线性特征也常用于情绪识别。

近似熵(Approximate entropy, ApEn)<sup>[26-27]</sup>是一种测量时间序列复杂性和规律性的分析方法,可以表示当维数变化时时间序列中产生新模式的概率的大小。为了提取近似熵,首先将时间序列 $\{x(t), t =$

$1, 2, \dots, N\}$ 转换为  $m$  维向量  $X_i$

$$X_i = \{x(t), x(t+1), \dots, x(t+m-1)\} \tag{7}$$

然后计算任意向量  $X_i$  和  $X_j$  之间的距离 ( $j = 1, 2, \dots, N - m + 1$ ), 即

$$d_{ij} = \max |x(t+k) - x(j+k)| \quad k = 0, 1, \dots, m-1 \tag{8}$$

对于给定的阈值  $r$ , 计算每个向量的距离  $d_{ij} < R (R = r \times \text{SD})$  的数目 (SD 为序列的标准差), 并且该数目与总的向量个数  $N - m + 1$  的比值记为  $C_i^m(r)$ , 可以得到

$$\varphi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{9}$$

令  $m = m + 1$ , 重复上述过程, 得到  $C_i^{m+1}(r)$  和  $\varphi^{m+1}(r)$ 。近似熵可计算为

$$\text{ApEn} = \varphi^m(r) - \varphi^{m+1}(r) \tag{10}$$

样本熵 (Sample entropy, SampEn)<sup>[28]</sup> 是对近似熵的改进。与近似熵相比, 样本熵的计算不依赖数据长度, 具有更好的一致性。样本熵在计算每个向量的距离  $d_{ij} < R (R = r \times \text{SD})$  的数目时不包含向量到自身的距离, 并且该数目与不包含自身的总向量个数  $N - m$  的比值记为  $C_i^m(r)$ , 可以得到

$$\varphi^m(r) = \frac{1}{N - m} \sum_{i=1}^{N-m} C_i^m(r) \tag{11}$$

令  $m = m + 1$ , 重复上述过程, 样本熵可计算为

$$\text{SampEn} = \ln \varphi^m(r) - \ln \varphi^{m+1}(r) \tag{12}$$

C0 复杂度<sup>[29]</sup> 描述序列的非规则程度, 可以反映时间序列的随机性。对于时间序列  $x(t), t = 1, 2, 3, \dots, N$ , 相应的傅里叶变换表示为

$$F_N(j) = \frac{1}{N} \sum_{t=1}^N x(t) W_N^{-jt} \quad j = 1, 2, \dots, N \tag{13}$$

$$W_N = e^{-\frac{2\pi i}{N}} \tag{14}$$

则  $F_N(j) (j = 1, 2, \dots, N)$  的均值  $G_N$  计算为

$$G_N = \frac{1}{N} \sum_{j=1}^N |F_N(j)|^2 \tag{15}$$

$$\bar{F}_N(j) = \begin{cases} F_N(j) & |F_N(j)|^2 > G_N \\ 0 & |F_N(j)|^2 \leq G_N \end{cases} \tag{16}$$

$\bar{F}_N(j) (j = 1, 2, \dots, N)$  的傅里叶逆变换为  $\bar{x}(t)$ 。则 C0 定义为

$$\text{C0} = \frac{\sum_{t=1}^N |x(t) - \bar{x}(t)|^2}{\sum_{t=1}^N |x(t)|^2} \tag{17}$$

## 2.2 特征选择和情绪分类

根据歌曲的情绪种类将情绪分为负向、中性和正向 3 种类别。对于每个被试的样本分别提取线性和非线性特征, 使用 Z-score 对提取后的特征进行标准化处理, 并根据不同情绪种类的歌曲标记不同的标签。

在本研究中, 使用最大相关最小冗余 (Max-relevance, min-redundancy, MRMR) 算法<sup>[30]</sup> 进行特征选择。该算法使用互信息来衡量特征与类别、特征与特征之间的关系, 核心思想是最大化特征与分类变量之间的相关性, 而最小化特征与特征之间的冗余性, 即从特征空间中寻找与目标类别有最大相关

性且相互之间具有最少冗余性的  $m$  个特征。

本文使用径向基核函数(Radial basis function, RBF)的支持向量机(Support vector machine, SVM)用于情绪识别。SVM 作为一种广泛使用的分类器,被证明是一种实用且鲁棒的脑解码方法<sup>[31]</sup>。本文采用十折交叉验证的方式计算分类的准确性,并通过遍历特征选择后的特征维数  $k$ (10到50,步长为10),SVM的核函数参数  $g$ (0.5到4,步长为0.5)和惩罚系数  $c$ ( $10^{-2}$ 到 $10^2$ ,步长为10)以得到最优参数。为了评估情绪自我想象和VR诱发情绪两种场景下分类准确性的统计显著性,在最优参数下,分别对两种场景下的数据进行了100次十折交叉验证,并对两种场景下100次的平均分类准确率进行显著性检验。

### 3 实验结果与分析

#### 3.1 自评量表结果

##### 3.1.1 SAM量表

从16名被试中收集SAM量表数据(情绪的愉悦度、激活度和控制度,分数范围1~9),并对不同情绪类别下16名被试在自我想象情绪调动和VR情绪诱发两种场景下的情绪自评分数进行了  $t$ -test 检验,量表平均分数和显著性检验结果如表2所示。在负向情绪下,情绪的愉悦度、激活度和控制度3个方面的自评分数在两种场景下都存在显著性差异(所有  $p < 0.01$ )。在中性情绪下,情绪的愉悦度、激活度和控制度3个方面的自评分数在两种场景下都没有显著性差异(所有  $p > 0.05$ )。在正向情绪下,情绪的愉悦度和激活度自评分数在两种场景下存在显著性差异( $p < 0.01$ ),情绪的控制度自评分数在两种场景下没有显著性差异( $p = 0.33$ )。在两个场景下的情绪自评平均分数符合预期,即负向情绪的VR刺激使愉悦度降低,正向情绪的VR刺激使愉悦度升高,两种情绪的VR刺激都使得激活度升高。由于中性情绪的效价和唤醒度一般都为中,所以在两种场景下的愉悦度、唤醒度和支配度都不存在显著性差异。

表2 SAM量表平均分数和显著性检验结果

Table 2 Average scores and significance test results of SAM

情绪评价 维度	负向			中性			正向		
	自我想象	VR诱发	显著性 $p$	自我想象	VR诱发	显著性 $p$	自我想象	VR诱发	显著性 $p$
愉悦度	4.687 5	3.000 0	<0.01	5.937 5	6.187 5	0.43	7.000 0	8.062 5	<0.01
激活度	5.562 5	7.000 0	<0.01	5.687 5	6.125 0	0.34	6.312 5	7.562 5	<0.01
支配度	6.062 5	4.625 0	<0.01	6.500 0	6.687 5	0.59	6.437 5	6.062 5	0.33

##### 3.1.2 声乐自评量表

从16名被试中收集声乐自评量表数据(分数范围1~4),并对不同情绪类别下16名被试在7个评价维度的声乐自评分数进行了  $t$ -test 检验,量表平均分数和显著性检验结果如表3所示。在负向情绪中,VR情绪诱发下的平均分在所有歌唱表现评价维度都大于自我想象的平均分,在共鸣运用、情感表达方面有显著性差异。

在中性情绪的自评量表中,VR诱发下的平均分在所有歌唱表现评价维度都大于自我想象的平均分,在共鸣运用、情感表达方面有显著性差异。在正向情绪的自评量表中,除了乐感方面,VR诱发下的平均分在所有其他歌唱表现评价维度都大于自我想象的平均分,在语言咬字和情感表达方面有显著性差异。自我想象和VR情绪诱发下的歌唱情感表达自评分数在3种情绪下都存在显著性差异,且VR情绪诱发的平均分数显著高于自我想象,说明VR可以很好地诱发出歌唱情绪。

表3 声乐自评量表平均分数和显著性检验结果

Table 3 Average score and significance test result of vocal self-rating scale

歌唱表现评价维度	负向			中性			正向		
	自我想象	VR诱发	显著性 $p$	自我想象	VR诱发	显著性 $p$	自我想象	VR诱发	显著性 $p$
连贯性	2.812 5	3.250 0	0.07	3.250 0	3.375 0	0.22	3.312 5	3.437 5	0.17
气息运用	2.687 5	3.062 5	0.07	3.187 5	3.375 0	0.09	3.125 0	3.312 5	0.09
共鸣运用	2.500 0	3.062 5	<0.01	2.875 0	3.250 0	<0.05	3.062 5	3.187 5	0.25
音准节奏	2.937 5	3.062 5	0.27	3.125 0	3.312 5	0.14	2.687 5	3.000 0	0.07
语言咬字	2.937 5	3.062 5	0.22	3.062 5	3.375 0	0.06	3.000 0	3.312 5	<0.05
乐感	3.125 0	3.437 5	0.06	3.437 5	3.562 5	0.17	3.375 0	3.312 5	0.38
情感表达	2.750 0	3.312 5	<0.01	3.000 0	3.437 5	<0.01	3.250 0	3.625 0	<0.05

### 3.2 分类性能

表4给出了自我想象和VR情绪诱发两种场景下,基于MRMR算法,并使用最优特征维数( $k=30$ )和最优参数下RBF核SVM进行情绪识别的三分类准确率(1 000个分类准确率的平均值)。VR诱发场景下的情绪分类准确率比自我想象高2.72%,且两种场景1 000次分类准确率存在显著性差异( $p<0.01$ )。图4(a)和图4(b)分别给出了自我想象和VR情绪诱发两种场景的分类混淆矩阵,VR诱发场景下的负向情绪和正向情绪分类准确率比自我想象场景有明显提升(5.93%和3.17%),中性情绪的分类准确率变化较小(0.84%)。结果说明,VR对于负向情绪和正向情绪诱发更为显著。

表4 情绪识别分类准确率

Table 4 Accuracy of emotion recognition classification %

场景	自我想象	VR诱发
准确率	73.614 5	76.332 7

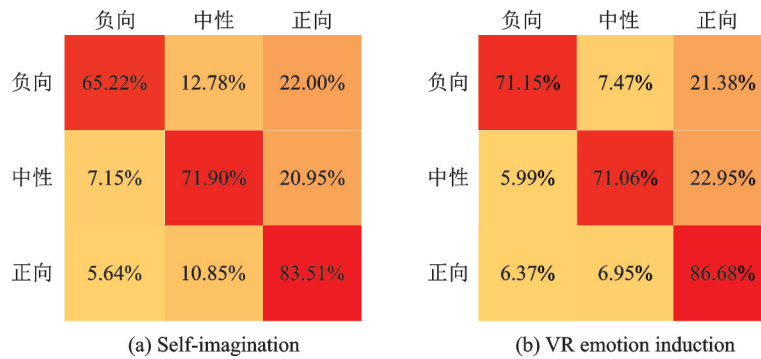


图4 情绪分类混淆矩阵

Fig.4 Emotion classification confusion matrix

### 3.3 判别性特征

表5显示了在自我想象和VR情绪诱发场景下,在1 000次特征选择中出现次数大于950次的特征。

在自我想象场景中,共有17个特征,包括16个线性特征和1个非线性特征。线性特征包括6个时域特征(3个为均值、3个为方差)和10个频域特征(1个为最大功率谱密度、1个为最大功率谱密度对应频率、1个为总功率、2个为频带平均功率、5个为半球不对称性特征)。1个非线性特征为AF4电极的



表5 特征选择中出现次数大于950次的特征

Table 5 Features with more than 950 occurrences in feature selection

场景	电极位置与特征
自我想象	AF4:均方值,Fz:均方值,F6:均方值,AF7:方差,F1:方差,PO3:方差,PO5:最大功率谱密度,C2:最大功率谱密度对应频率,PO4:总功率,P7:theta频带平均功率,CP5:alpha频带平均功率,AF4:C0复杂度,FP1-FP2、AF7-AF8、C5-C6、CP1-CP2、PO7-PO8:半球不对称性
VR情绪诱发	Fz:均方值,PO3:平均绝对振幅,Oz:Hjorth活动性,P5:alpha频带平均功率,Fp1:beta频带平均功率,AF3:近似熵,PO6:C0复杂度,AF3-AF4、F7-F8、FC3-FC4、T7-T8、CP3-CP4、TP7-TP8、PO7-PO8、O1-O2:半球不对称性

C0复杂度。

在VR情绪诱发场景中,共有15个特征,包括13个线性特征和2个非线性特征。线性特征包括3个时域特征(1个为均方值、1个为平均绝对振幅、1个为Hjorth活动性)和10个频域特征(2个为频带平均功率、8个为半球不对称性)。2个非线性特征为AF3电极的近似熵和PO6电极的C0复杂度。

在自我想象场景中选择出的时域特征主要集中在均方值和方差两种类型,而VR情绪诱发场景的时域特征类型比较分散。在两种场景选择出的特征中,频域特征占大多数(10/16和10/15),自我想象场景下的频域特征类别较为分散(共6种),VR情绪诱发场景下的频域特征集中在频带平均功率和半球不对称性两类。在两种场景的频域特征中,半球不对称性特征更为重要(占比5/10和8/10)。这可能是因为情绪具有半球不对称性<sup>[23-24]</sup>,且以往的研究也表明虚拟现实环境下的情绪诱发和游戏操作也会导致半球不对称性<sup>[32-34]</sup>。

### 3.4 问题与展望

虽然本文提出了较好的对比结果,说明了VR应用于辅助声乐教学的可行性,但仍然存在被试数量较少、分布较窄等问题。在进一步的研究中,将邀请更多声乐表演专业的学生增加数据量,从而得到更有普遍性的结果。此外,也可以进一步扩展研究背景,邀请非声乐表演专业的学生,研究虚拟现实技术对非专业演唱者的情绪诱发和演唱表现的影响。并在此基础上,对比不同情绪诱发方式对专业演唱者和非专业演唱者在情绪诱发效果上的差异,以及其对应大脑机制上的异同。

## 4 结束语

本文通过主观量表和客观脑电信号两种数据,对比了自我想象和VR情绪诱发两个场景下被试的情绪状态,并探究了两种场景下与情绪相关的脑电特征。量表的统计性结果和脑电情绪分类的结果表明,相比传统自我想象方式,VR情绪诱发可以激活演唱者的情绪状态,从而提升演唱表现。本文第一次将VR和脑电结合运用在声乐教学和表演中,不仅为基于脑电的情绪识别应用于歌唱表演场景下开辟一个新的研究领域,同时也为VR辅助音乐教育的教学模式提供了新的研究路径。

### 参考文献:

- [1] 张冠华,余旻婧,陈果,等.面向情绪识别的脑电特征研究综述[J].中国科学:信息科学,2019,49(9):1097-1118.  
ZHANG Guanhua, YU Minjing, CHEN Guo, et al. A review of EEG features for emotion recognition[J]. Science in China(Information Sciences), 2019, 49(9): 1097-1118.
- [2] LEDOUX J E. Emotion circuits in the brain[J]. Annual Review of Neuroscience, 2000, 23(1): 155-184.
- [3] BAIG M Z, KAVAKLI M. Analyzing novice and expert user's cognitive load in using a multi-modal interface system[C]// Proceedings of 2018 26th International Conference on Systems Engineering (ICSEng). Sydney, Australia: IEEE, 2018: 1-7.

- [4] ALARCAO S M, FONSECA M J. Emotions recognition using EEG signals: A survey[J]. *IEEE Transactions on Affective Computing*, 2017, 10(3): 374-393.
- [5] HAJCAK G, MACNAMARA A, OLVET D M. Event-related potentials, emotion, and emotion regulation: An integrative review[J]. *Developmental Neuropsychology*, 2010, 35(2): 129-155.
- [6] WANG Xiaowei, NIE Dan, LU Baoliang. Emotional state classification from EEG data using machine learning approach[J]. *Neurocomputing*, 2014, 129: 94-106.
- [7] LI Peiyang, LIU Huan, SI Yajing, et al. EEG based emotion recognition by combining functional connectivity network and local activations[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(10): 2869-2881.
- [8] HUANG Haiyun, XIE Qiyou, PAN Jiahui, et al. An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness[J]. *IEEE Transactions on Affective Computing*, 2019. DOI: 10.1109/TAFFC.2019.2901456.
- [9] LIN T, GILAM G, RAZ G, et al. Accessible neurobehavioral anger-related markers for vulnerability to post-traumatic stress symptoms in a population of male soldiers[J]. *Frontiers in Behavioral Neuroscience*, 2017, 11: 38.
- [10] HALIM Z, REHAN M. On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning[J]. *Information Fusion*, 2020, 53: 66-79.
- [11] 石惟正. 声乐学基础[M]. 北京: 人民音乐出版社, 2002.  
SHI Weizheng. *Fundamentals of vocal music*[M]. Beijing: People's Music Publishing House, 2002.
- [12] BERNARDO P D, BAINS A, WESTWOOD S, et al. Mood induction using virtual reality: A systematic review of recent findings[J]. *Journal of Technology in Behavioral Science*, 2021, 6(1): 3-24.
- [13] MARÍN-MORALES J, HIGUERA-TRUJILLO J L, GRECO A, et al. Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors[J]. *Scientific Reports*, 2018, 8(1): 1-15.
- [14] SOOD S K, SINGH K D. An optical-fog assisted EEG-based virtual reality framework for enhancing E-learning through educational games[J]. *Computer Applications in Engineering Education*, 2018, 26(5): 1565-1576.
- [15] 杨晓哲, 任友群. 虚拟现实与脑电联动系统的开发及其教育研究功能探索[J]. *远程教育杂志*, 2019, 37(1): 47-54.  
YANG Xiaozhe, REN Youqun. Development of virtual reality and EEG linkage system and exploration of its educational research function[J]. *Journal of Distance Education*, 2019, 37(1): 47-54.
- [16] BRADLEY M M, LANG P J. Measuring emotion: The self-assessment manikin and the semantic differential[J]. *Journal of Behavior Therapy and Experimental Psychiatry*, 1994, 25(1): 49-59.
- [17] DELORME A, MAKEIG S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis[J]. *Journal of Neuroscience Methods*, 2004, 134(1): 9-21.
- [18] PION-TONACHINI L, KREUTZ-DELGADO K, MAKEIG S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website[J]. *Neuroimage*, 2019, 198: 181-197.
- [19] LI Wenyu, HU Xin, LONG Xuefei, et al. EEG responses to emotional videos can quantitatively predict big-five personality traits[J]. *Neurocomputing*, 2020, 415: 368-381.
- [20] JENKE R, PEER A, BUSS M. Feature extraction and selection for emotion recognition from EEG[J]. *IEEE Transactions on Affective Computing*, 2014, 5(3): 327-339.
- [21] LI Ziyu, WU Xia, XU Xueyuan, et al. The recognition of multiple anxiety levels based on electroencephalograph[J]. *IEEE Transactions on Affective Computing*, 2019. DOI: 10.1109/TAFFC.2019.2936198.
- [22] HJORTH B. EEG analysis based on time domain properties[J]. *Electroencephalography and Clinical Neurophysiology*, 1970, 29(3): 306-310.
- [23] LI Yang, ZHENG Wenming, CUI Zhen, et al. A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition[C]//*Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI)*. Stockholm, Sweden: Morgan Kaufmann, 2018: 1561-1567.
- [24] ZHAO Guozhen, ZHANG Yulin, GE Yan. Frontal EEG asymmetry and middle line power difference in discrete emotions[J]. *Frontiers in Behavioral Neuroscience*, 2018, 12: 225.
- [25] WANG Fei, WU Shichao, ZHANG Weiwei, et al. Multiple nonlinear features fusion based driving fatigue detection[J].

- Biomedical Signal Processing and Control, 2020, 62: 102075.
- [26] LI Xiang, SONG Dawei, ZHANG Peng, et al. Exploring EEG features in cross-subject emotion recognition[J]. *Frontiers in Neuroscience*, 2018, 12: 162.
- [27] 李昕, 蔡二娟, 田彦秀, 等. 一种改进脑电特征提取算法及其在情感识别中的应用[J]. *生物医学工程学杂志*, 2017, 34(4): 510-517, 528.
- LI Xin, CAI Erjuan, TIAN Yanxiu, et al. An improved electroencephalogram feature extraction algorithm and its application in emotion recognition[J]. *Journal of Biomedical Engineering*, 2017, 34(4): 510-517, 528.
- [28] JIE Xiang, CAO Rui, LI Li. Emotion recognition based on the sample entropy of EEG[J]. *Bio-medical Materials and Engineering*, 2014, 24(1): 1185-1192.
- [29] 卞宁艳, 曹洋, 王斌, 等. 基于二阶C0复杂度的癫痫发作预测[J]. *生物物理学报*, 2007(1): 67-74.
- BIAN Ningyan, CAO Yang, WANG Bin, et al. Prediction of epileptic seizures based on second-order C0 complexity[J]. *Acta Biophysica Sinica*, 2007(1): 67-74.
- [30] PENG Hanchuan, LONG Fuhui, DING C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(8): 1226-1238.
- [31] BASHIVAN P, YEASIN M, BIDELEMAN G M. Single trial prediction of normal and excessive cognitive load through EEG feature fusion[C]//*Proceedings of 2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. Philadelphia, USA: IEEE, 2015: 1-5.
- [32] QIN Yi, MAO Mei, ZHANG Huayu, et al. Features of brainwave induced by 3D auditory stimulation[C]//*Proceedings of 2020 International Conference on Culture-Oriented Science & Technology (ICCST)*. Pattaya, Thailand: IEEE, 2020: 250-254.
- [33] TARRANT J, COPE H. Combining frontal Gamma asymmetry neurofeedback with virtual reality: A proof of concept case study[J]. *NeuroRegulation*, 2018, 5(2): 57-66.
- [34] OLIVEIRA S M S, MEDEIROS C S P, PACHECO T B F, et al. Electroencephalographic changes using virtual reality program[J]. *Neurological Research*, 2018, 40(3): 160-165.

#### 作者简介:



张进(1974-),女,副教授,硕士生导师,研究方向:声乐演唱与计算机可视化研究, E-mail: 550311098@qq.com。



许子明(1997-),男,硕士研究生,研究方向:机器学习与脑电信号分析, E-mail: xuziming@nuaa.edu.cn。



周月莹(1992-),女,博士研究生,研究方向:脑机接口、脑电分析和机器学习。



王澎湃(1992-),男,博士研究生,研究方向:脑机接口和深度学习。



张道强(1978-),通信作者,男,教授,博士生导师,研究方向:机器学习、模式识别以及数据挖掘和医学影像分析, E-mail: dqzhang@nuaa.edu.cn。

(编辑:王静)