

# 基于粒子群算法优化的音频特征应用研究

王志强 郭 宁 傅向华

(深圳大学计算机与软件学院 深圳 518060)

**摘 要** 在深入研究音频特征的基础上,提取响度特征和音调特征,并利用粒子群算法优化特征权重。提出一种对歌唱片段进行自动评价的方法,用于视频点歌系统的实时评分模块。实验结果表明,该系统能够反映演唱者歌声和歌曲原唱两者内容的相似程度,从而给出了有效的评分标准。

**关键词** 音频特征,PSO,视频点歌系统

**中图分类号** TP301 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.10.010

## Application Research of Audio Feature Based on Particle Swarm Optimization Algorithm

WANG Zhi-qiang GUO Ning FU Xiang-hua

(College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China)

**Abstract** Based on the research of audio feature, this paper extracted the features of loudness and pitch, and selected their feature weights by PSO. We proposed an automatic evaluation method of singing segment which was already applied to video song-on-demand marking system. According to the results of the experiments carried out, the systems characterizes similar degree of the singer singing and singing original sound in real time so that the marking standards are efficient.

**Keywords** Audio feature, PSO, Video song-on-demand systems

## 1 引言

随着互联网的发展和普及,听觉信息资源(如歌曲、音乐等)被迅速传播,人们可以方便地从互联网中获取丰富的歌曲或音乐,网络 KTV 点歌平台就是一种具体的应用形态。基于视频点歌平台的自动评分系统利用音频处理技术对演唱者的歌唱进行自动评价,它不仅丰富了多媒体内容分析技术,也构建了一种更为和谐的人机交互环境。

针对视频点歌评分系统,目前国内外研究者已经取得了一系列的研究成果。例如, Pardo 等人使用隐马尔科夫模型对输入音频片段进行建模,采用编辑距离衡量输入音频片段和数据库中模板的相似度<sup>[1]</sup>; Nakano 等人使用音色、音程和颤音,设计了一个两层支持向量机分类器来对歌曲样本进行分类<sup>[2]</sup>; Mayor 使用 MIDI 音源制作旋律模板,分析并描绘出 CD 中歌曲样本的音调轮廓,比较演唱者歌声与 CD 原声的相似性<sup>[3]</sup>,等等。现有的系统基本是采用音调、音色等声学特征,实现了基于音频的内容分析,但仍存在不足,具体表现在:①需要准备语料模板并制作每个音频的特征模板,模板制作不仅工作量巨大而且花费成本也较高,而对 Internet 上所有音频提取特征模板也缺乏可行性;②普遍采用单一的音频特征,如音调或声强来描述音频,并未考虑多种特征组合的相似性计算;③缺乏实时性,对输入歌曲进行整体特征提取后,再作分类或检索,未能实现对歌曲的实时自动分析。

本文综合音频处理技术和乐理两方面的知识,提取响度和音调两种声学特征,并采用粒子群算法优化两种特征在系统中的权重,构建并实现了一个视频点歌实时评分系统。实验结果表明,该系统能够精确地描绘两段歌曲的特征,并对两者的相似度作出合理评价。

## 2 音频特征提取方法

人耳对声音的主观评价可以由声学感知特征来表征,听觉感知特征是综合了声学 and 乐理知识的音频特征,它考虑人耳听觉感知特点,结合乐理知识,体现了人对音频特别是音乐的一种主观感受。常用的听觉感知特征有:响度、音调、节奏和音色。这些特征通过时域、频域特征及其组合来计算。而歌唱音准主要的影响因子是音调,其次是响度<sup>[4]</sup>。下面将提取响度和音调两种音频特征,并计算两段输入音频的相似性。

### 2.1 响度特征的提取

响度计算模型最早由 Fletcher 和 Munson 提出<sup>[5]</sup>,主要根据信号频谱计算声音响度,而后 Zwicker 和 Moore 等人在此基础上提出各自的模型,但 Zwicker 模型使用图、表的插值方式,故计算不便<sup>[6]</sup>; Moore 模型在噪声环境下需要进行大量的运算<sup>[7]</sup>。归一化响度计算方法是基于人耳听觉模型提出的,其计算简便,实时性较强<sup>[8]</sup>。本文在归一化响度的基础上,提出频域峰值-归一化响度来描述歌唱音频片段的响度。

到稿日期:2013-07-20 返修日期:2013-09-03 本文受广东省自然科学基金项目(7301329),深圳市科技计划项目(200740)资助。

王志强(1963-),男,硕士,教授,主要研究方向为多媒体信息处理、Petri 网等;郭 宁(1985-),男,硕士,主要研究方向为多媒体信息检索;傅向华(1977-),男,博士,副教授,主要研究方向为互联网信息检索、软件工程等。

一帧音频信号经 FFT 后的频谱  $X(i) (i=1, \dots, N/2)$  在形式上表现为  $i$  次谐波的声波震动幅度与声压的正比关系, 具体公式为:

$$Loudness(i) = 2^{C(i)/10 + 2\log X(i) + 2\log X_{1K}} \times Loudness_{1K} \quad (1)$$

其中,  $L(i)$  表示  $i$  次谐波的响度,  $C(i)$  表示  $i$  次谐波的衰减量, 其值通过二次插值的方法获得。  $Loudness_{1K}$  表示 1kHz 谐波分量的响度。

那么第  $i$  次谐波的归一化响度可描述为:

$$Loudness_{norm}(i) = \frac{2^{C(i)/10 + 2\log X(i)}}{\sum_{i=1}^{N/2} 2^{C(i)/10 + 2\log X(i)}} \quad (2)$$

考虑歌唱时, 人声的频率范围为 50~1000Hz, 将此频段范围划分为若干个子频带, 对应于每个子频带的归一化响度定义为:

$$Loudness_{Band}(i) = \sum_{i \in band_j} Loudness_{norm}(i), j=1, 2, \dots, n_b \quad (3)$$

其中,  $n_b$  是频带数量。

在实际系统中, 主要考虑 1 次谐波(基频一倍)的响度, 频域峰值-归一化响度定义为:

$$X_{Peak-L}(i) = X_{max}(i) + X_{max}(i) \quad (4)$$

其中, 采用 FFT 变换后频谱的最高  $X_{max}(i)$  与次高峰峰值能量  $X_{max}(i)$  的和表示  $X(i)$ 。

图 1(a)(c)(e) 的 3 个波形由上至下分别对应的是原唱、跟唱者、乱唱者的原始波形, 横坐标表示帧数, 纵坐标表示归一化响度值。图 1(b)(d)(e) 的 3 条曲线分别是对应的归一化响度, 横坐标表示帧数, 纵坐标表示归一化响度。图中凹陷部分曲线表示停顿静音、气音或噪声。由图 1 的归一化响度轨迹可以看出, 跟唱的归一化响度轨迹与原唱的归一化响度轨迹相似度非常高, 乱唱的归一化响度轨迹则与原唱和跟唱的偏差很大。

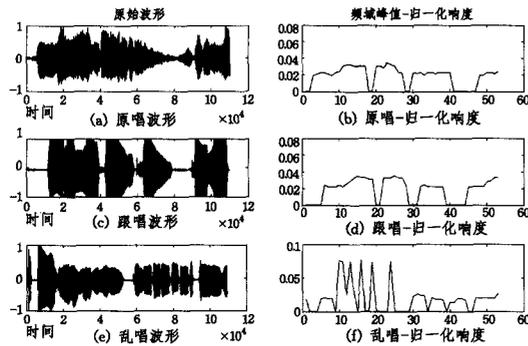


图 1 频域峰值-归一化响度

## 2.2 音调特征的提取

影响音调的主要因素是音高特征(基本频率), 而 Mel 频率是基于人耳听觉模型提出的, Mel 倒谱系数 MFCC 已广泛应用于语音和音频识别领域。本文在基频提取的基础上, 结合音乐乐理知识、人耳听觉模型和语音信号处理技术提出了 Mel 轨迹, 来准确并实时地描绘歌曲片段音调和旋律的相似程度。

在获取歌曲片段 Mel 轨迹之前, 需对语音信号进行预处理, 如分帧、加窗和去噪等, 并对分帧后的语音信号进行 FFT 变换转到频域, 从而提取其基频数据。

低频能量化是指对语音信号 FFT 变换后进行低频能量比处理, 由于人的基频范围为 50~1000Hz, 基频超出此范围

的则为非人声。基频范围为 50~1000Hz 的能量与整帧的能量比值称为低频能量比, 通过求取低频能量比并设定阈值来判断语音含有噪音的大小以及是否静音, 低频能量比高的语音帧为正常人的歌唱声音, 低频能量比低的语音则为噪声或静音。

低频能量化主要是针对歌曲中人声停顿的静音部分, 主要表现在 Mel 轨迹的凹陷部分。

低频能量化的处理过程如图 2 所示。

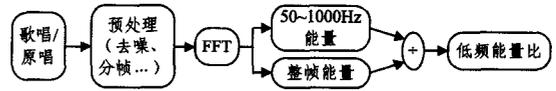


图 2 低频能量化处理过程

Mel 频率与频率的转换关系如式(5)所示:

$$Mel = 2595 * \log_2(1 + frequency/700) \quad (5)$$

其中,  $Mel$  表示 Mel 频率,  $frequency$  表示频率, 2595 和 700 是根据响度曲线和听觉模型确立的转换因子。

人声歌唱的基频范围为 50~1000Hz, 对于一帧语音信号, 进行 FFT 变换和低频能量映射处理后, 选取有效语音在其频率为 50~1000Hz 内按等分间隔分割成 10 段, 并求得每段能量, 选取能量最大者的频率作为基频。

通过式(5)将基频转换成相应的 Mel 频率, Mel 频率随音频帧的推移呈变化的曲线, 称为 Mel 轨迹。Mel 轨迹不仅反映了音频信号基频的变化趋势, 而且反映了歌唱的节奏与时值, 能够实时、准确、有效地描述歌唱的语音特征。Mel 轨迹算法的流程图, 如图 3 所示。

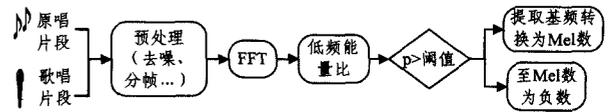


图 3 Mel 轨迹算法流程

## 3 基于 PSO 的音频特征权重计算

音频特征的维度较高, 且特征之间的权重往往通过经验或大量观测实验得出。特征权重的计算方法有机器学习、多层神经网络、智能算法等。针对视频点歌评分系统, 群体智能算法——粒子群优化算法 (Particle Swarm Optimization, PSO) 计算复杂度低, 特别适合于大量人声音频片段的训练, 耗时较短, 收敛迅速。本文通过 PSO 算法计算音频特征权重, 结果表明 PSO 选择的权重避免了人为设定权重的盲目性, 且更为准确。

### 3.1 粒子群优化算法

粒子群优化算法是对鸟群觅食过程中的迁徙和聚集的模拟, 由简单个体组成群落以及个体之间的互动行为模拟搜索全局最优解<sup>[9]</sup>。

具体描述为: 在一个  $M$  维的目标搜索空间中, 由  $n$  个粒子构成一个群体, 其中第  $i$  个粒子的位置可以表示为  $M$  维的位置  $x_i = (x_{i1}, x_{i2}, \dots, x_{im}, \dots, x_{iM})$ 。根据一定的标准(适应值函数)计算  $x_i$  当前的适应值, 以此衡量粒子位置的优劣。每次迭代中粒子的飞行速度或者速度矢量表示为  $v_i = (v_{i1}, v_{i2}, \dots, v_{im}, \dots, v_{iM})$ , 粒子迄今为止搜索到的最优位置为  $p_i = (p_{i1}, p_{i2}, \dots, p_{im}, \dots, p_{iM})$ , 整个粒子群迄今为止搜索到的最优位置为  $p_g = (p_{g1}, p_{g2}, \dots, p_{gm}, \dots, p_{gM})$ 。每次迭代中, 粒子根据以下式子更新速度和位置:

$$v_{im}^{k+1} = \omega v_{im}^k + \lambda_1 r_1 (p_{im} - x_{im}^k) + \lambda_2 r_2 (p_{gm} - x_{im}^k) \quad (6)$$

$$x_{im}^{k+1} = x_{im}^k + v_{im}^{k+1} \quad (7)$$

其中,  $i=1, 2, \dots, n, k$  是迭代次数,  $r_1$  和  $r_2$  为  $[0, 1]$  内的随机数组成的矢量。  $\lambda_1$  和  $\lambda_2$  为学习因子, 也称加速因子, 作用是使粒子具有自我总结和向群体中优秀个体学习的能力, 从而向自己的历史最优点和群体内最优点靠近。

### 3.2 相似性计算

设一个歌曲片段表示为一个由  $M$  维特征描述的向量  $x = (f_{i1}, f_{i2}, \dots, f_{im}, \dots, f_{iM})$ , 待检索片段为  $q = (f_{j1}, f_{j2}, \dots, f_{jm}, \dots, f_{jM})$ , 在第  $i$  个歌曲特征  $f_i$  通过归一化后, 两者相似性可以用欧几里德距离表示:

$$dis(x_i, q_j) = \sqrt{\frac{1}{N_i} \sum_{j=1}^{N_i} (x_{ij} - q_{ij})^2} \quad (8)$$

其中,  $N_i$  是第  $i$  个音频特征的特征值数目, 同时  $dis(x_i, q) \in [0, 1]$ 。

若加入音频特征的权重值, 那么两者相似性用加权的欧几里德距离可以表示为:

$$Sim_w(x_i, q) = \sqrt{\sum_{j=1}^M (x_{ij} - q_{ij})^2 (W_j)^2} \quad (9)$$

其中,  $W_j$  为第  $j$  个特征在音频检索模型中的权重。

### 3.3 音频特征权重计算

根据式(11)的相似性计算方法, 适应度函数设计如下: 对于一个表示为一个由  $M$  维特征描述  $x = (f_{i1}, f_{i2}, \dots, f_{im}, \dots, f_{iM})$  的向量的歌曲片段, 其  $M$  维特征对应的权重向量为  $w = (w_{i1}, w_{i2}, \dots, w_{im}, \dots, w_{iM})$ , 适应度函数为:

$$f(w) = 1 - \frac{n(w)}{E} \quad (10)$$

其中,  $\frac{n(w)}{E}$  表示第  $E$  次检索中查询歌曲匹配的精确度,  $n(w)$  是权重  $w$  向量的函数, 表示特征权重为  $w$  向量时内容相似的歌曲数目。显然, 粒子群的最优值应为使第  $E$  次  $n$  最大, 而使  $f(w)$  最小。

#### 算法 1 基于 PSO 的特征权重算法

- ① 粒子群中每个粒子位置  $x_i$  和速度  $v_i$  初始化为一个随机值, 得到一个解  $p_i$  和  $p_g$ , 此时迭代数为 1。
- ② 根据式(10)计算每个粒子的  $f(x_i)$ 。
- ③ 对于每一个粒子, 若  $f(x_i(w)) < f(p_i)$ , 则  $p_i = x_i$ , 更新粒子位置  $p_i$ 。
- ④ 对于整个粒子群, 若  $f(x_i(w)) < p_g$ , 则  $p < x_i M$ 。
- ⑤ 通过式(6)更新粒子的速度。
- ⑥ 通过式(7)更新粒子的位置。
- ⑦ 如果  $x = T_M$ , 记录  $p_g$  的位置, 迭代结束。否则,  $x = x + 1$ , 执行步骤②。

本文采用 PSO 算法计算响度和音调的特征权重。训练数据集为 2006—2009 年 MIREX 评测中 QBSH (Query By Singing and Humming) 评比的标准数据库<sup>[10]</sup>。数据库中共有 48 首歌曲, 4813 个人声片段。从中选取 5 男 5 女共 247 个人声片段作为训练片段。

PSO 算法的初始值设置为最大迭代次数 300, 粒子总数 70, 随机数距离  $[-1, 1]$ , 学习因子为 2。采用算法 1 进行运算, 选取其中 20 次令式(11)适应值最优的权重值, 得到的权值分布如图 4 所示。

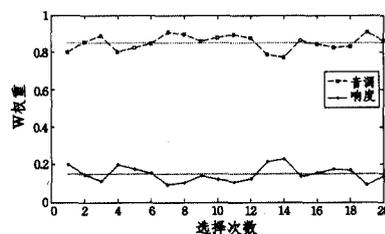


图 4 基于 PSO 的特征权重计算

实验中, 在迭代 300 次内, PSO 算法均收敛, 由图 4 可以看出, 在 20 个最优收敛值的权重中, 音调权重约为 0.85, 响度权重约为 0.15。本文将此权重值定为歌曲评分系统加权相似性评分中音调和响度的权重。

## 4 应用示例

### 4.1 视频点歌评分系统的构建

视频点歌评分系统的关键问题是提供一种有效的方法来衡量歌唱者的音频与原唱在内容上的相似度, 为此, 本文将提出的频率峰值-归一化响度和 Mel 轨迹应用于点歌实时评分系统中, 以响度和音调作为评分依据来评价歌唱者的演唱水平。系统的工作原理如图 5 所示。

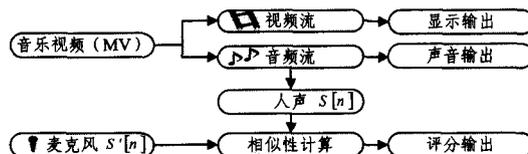


图 5 点歌实时评分系统工作原理

图 5 中播放的歌曲片段是一种能用计算机播放的音乐视频 (Music Video), 以 MPEG2 格式存储。MPEG2 格式的歌曲 MV 包括视频轨和 2 个音轨, 一个音轨包含纯人声原唱信息, 另一个音轨包括伴奏。数据流形式的歌曲片段, 通过两个音轨同步频率相减算法滤除伴奏, 将得到待处理的人声片段  $S[n]$ , 这里使用不包括伴奏的人声  $S[n]$  与麦克风录入的演唱者声音  $S'[n]$  作对比, 将排除伴奏等噪声对响度、音调特征的计算干扰, 使得评分更为准确。在相似性计算模块中, 本文采取基于响度和音调特征的加权相似性评分。相似性计算原理如图 6 所示。

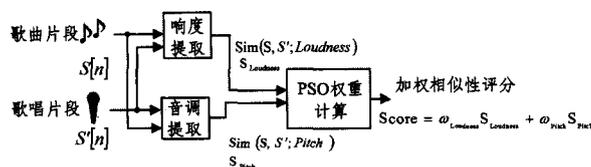


图 6 相似性计算原理

对麦克风录入的演唱者歌曲片段  $S'[n]$  与歌曲片段中的人声原唱信息  $S[n]$  分别作响度特征提取和音调特征提取, 进行响度和音调的相似性计算并得到评分  $S_{Loudness}$  和  $S_{Pitch}$ , 经过 Z-score 标准化后, 使用 PSO 算法计算权重值, 最后得到加权的相似性评分:

$$Score = \omega_{Loudness} S_{Loudness} + \omega_{Pitch} S_{Pitch}$$

### 4.2 系统实验

为了验证归一化响度和 Mel 轨迹计算相似度的有效性, 设计了两个实验。

#### 4.2.1 仿真实验

实验截取流行歌曲《童话》中的一段, 通过 Matlab 编程分

别提取原唱语音信号和歌唱者演唱声音,并模拟系统的伴奏分离过程,得到原唱、跟唱、乱唱的归一化响度、音调(Mel)轨迹,如图7所示。

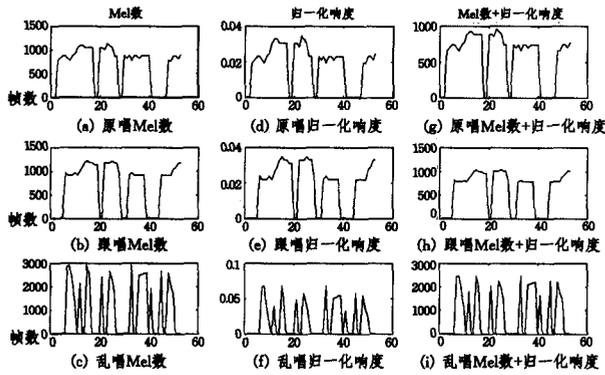


图7 原唱、跟唱、乱唱的3种轨迹

图7(a)(b)(c)的3个波形分别对应的是原唱的Mel轨迹、歌唱者的Mel轨迹、歌唱者的Mel轨迹,横坐标表示帧数,纵坐标表示Mel值。图7(d)(e)(f)的3条曲线分别是对应的归一化响度,横坐标表示帧数,纵坐标表示归一化响度。图7(g)(h)(i)的3条曲线分别是使用图4算出的权值,以及采用Mel轨迹和归一化响度结合的方法算出的波形,横坐标表示帧数,纵坐标表示综合相似度。图中凹陷部分曲线表示停顿静音、气音或噪声。由图7的Mel轨迹可以看出,跟唱的Mel轨迹、归一化响度轨迹与原唱的Mel轨迹、归一化响度轨迹的相似度非常高,乱唱的Mel轨迹、归一化响度轨迹则与原唱和跟唱的偏差很大;并且,结合了PSO算出的权值后,乱唱区别更为显著,同时加权后的归一化响度和音调轨迹更加贴近原唱轨迹。

#### 4.2.2 视频点歌评分系统实验

将基于响度Mel轨迹的相似性计算方法应用于歌曲评分系统,并设计一个实验。实验采集5男5女共10人的演唱音频数据,录音人均没有音乐背景知识,测试者从流行音乐歌曲库(14312首)中自选一首曲目,系统录音采样率为44.1kHz,16bit量化。其中组I的测试者选择了节奏较慢的经典老歌,组II的测试者选择中等节奏的流行歌曲,组III的测试者选择了节奏较快的流行歌曲。按照情感、音准、节奏、旋律等对测试演唱进行主观评分,主观评分为十分制。为了更好地比较主客观评分的结果,将主观评分的结果从十分制映射到百分制,比较结果如表1所列。其中,Singer为10位演唱者(编号A-J),S为主观评分(Subjective),O为客观评分(Objective),F为分数,R为排名。

表1 基于归一化响度+音调的评分实验结果

Singer	A	B	C	D	E	F	G	H	I	J	
Team	I	I	II	II	II	II	II	II	III	III	
S	F	86	92	75	78	84	79	68	82	65	67
	R	2	1	7	6	3	5	8	4	10	9
O	F	83	90	77	75	82	76	66	78	63	65
	R	2	1	5	7	3	6	8	4	10	9

为了进一步验证基于归一化响度和音调的歌曲相似性计算算法的有效性,本文构建了基于单一特征的歌曲相似性计算算法。分别选取响度、音调两个特征,实验数据集不变。得到的比较结果如表2、表3所列。

表2 基于响度的评分实验结果

Singer	A	B	C	D	E	F	G	H	I	J	
Team	I	I	II	II	II	II	II	II	III	III	
S	F	91	89	88	72	83	82	79	77	68	70
	R	1	2	3	8	4	5	6	7	10	9
O	F	84	87	83	76	79	70	72	75	64	69
	R	2	1	3	5	4	8	7	6	10	9

表3 基于音调的评分实验结果

Singer	A	B	C	D	E	F	G	H	I	J	
Team	I	I	II	II	II	II	II	II	III	III	
S	F	80	90	87	83	84	78	68	75	66	65
	R	5	1	2	4	3	6	8	7	9	10
O	F	85	86	83	78	79	74	64	72	63	60
	R	2	1	3	5	4	6	9	7	8	10

图8为基于加权的响度和音调、单一归一化响度、单一音调评分的实验对比图。

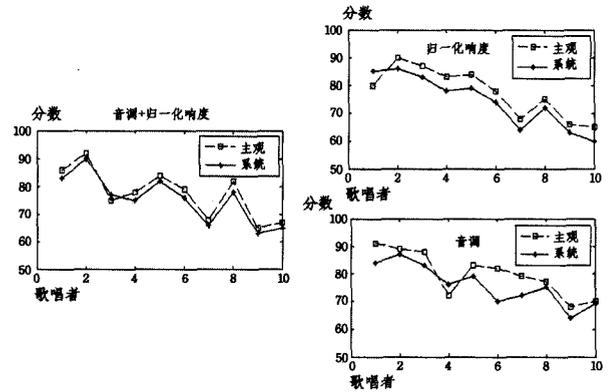


图8 基于归一化响度+音调、单一响度、单一音调的3种算法评分结果

#### 4.3 实验结果分析

①本实验对歌曲片段构建了3种相似性计算算法,由表1、表2、表3可知,在点歌实时评分系统中,基于加权的归一化响度和音调的评分最接近人的主观评分,这与Matlab仿真实验得到的结果吻合。

②基于响度和音调的相似性计算算法在分数和排名走势上,系统与主观的评分均较为接近,但是仍存在一定数量的误判。

③基于加权的归一化响度和音调的评分效果较好,一方面由于响度和音调体现了人耳的听觉感知特性。另一方面利用PSO计算归一化响度和音调特征值的权重,得到了较切合实际系统的有效权重,较好地解决了穷举或人为设定特征权重的问题。

**结束语** 本文在深入研究歌曲声学乐理的基础上,提取改进的频域峰值-归一化响度和音调特征,并使用粒子群算法优化特征权重,用于视频点歌评分系统。实验表明,经过优化权重的频域峰值-归一化响度和音调,能有效地描述歌唱者演唱声音和歌曲原唱的相似程度,给出较切合主观评分的评价分数。下一步研究主要是对歌曲的其他声学特征,如音色、节奏特征进行研究,提取其对应的声学特征并应用到视频点歌系统中。

#### 参考文献

[1] Chechil G, Le E, Rehn M, et al. Large-scale Content-based Audio Retrieval from Text Queries [C]//1st International ACM Con-

- ference on Multimedia Information Retrieval. Canada, August 2008;105-112
- [2] Nakano T, Goto M, Hiraga Y. MiruSinger: A Singing Skill Visualization Interface Using Real-time Feedback and Music CD Recordings as Referential Data [C]//9th IEEE International Symposium on Multimedia-Workshops. Taiwan, December 2007;75-76
- [3] Mayor O, Bonada J, Loscos A. Performance Analysis and Scoring of the Singing Voice [C]// AES 35th International Conference: Audio for Games. London, UK, February 2009;1-7
- [4] Ruiz-Reyes N, Vera-Candeas P, Muñoz J E, et al. New Speech/music Discrimination Approach Based on Fundamental Frequency Estimation [J]. Multimedia Tools and Application, 2009, 41(2): 253-286
- [5] 马元峰, 陈克安, 王娜. Moore 响度模型的数值计算方法[J]. 声学技术, 2008, 27(3): 390-395
- [6] Yuan Ji-xuan, Song Zhu-mei. Sound Loudness Research on Acoustic-structure Coupling Field in an Enclosed Cavity[C]// 2nd Pacific-Asia Conference on Circuits, Communications and System. Beijing, China, August 2010;456-462
- [7] Rao Dan, Xie Bo-sun, Liang Zhi-qiang. Binaural Loudness Spectra Analysis of Individualized Binaural Room Impulse Responses [C]//2nd International Congress on Image and Signal Processing. Tianjin, China, October 2009;1-4
- [8] Zheng Gui-bin, Han Ji-qing. Real-time Audio Retrieval Method and Automatic Commercial Detecting System [J]. Journal of Computer Science, 2006, 2(3): 297-302
- [9] Nozaki T, Kameyama K. Feature Selection for User-adaptive Content-based Music Retrieval using Particle Swarm Optimization [C]//10th International Conference on Intelligent Systems Design and Applications. 2010;941-946
- [10] You J, Park S, Kim I. An Efficient Frequent Melody Indexing Method to Improve the Performance of Query-by-humming Systems [J]. Journal of Information Science, 2008, 34(6): 777-798

(上接第 44 页)

采集, 并使用 GEP 对实验结果进行分析和测试, 得出如下结论: (1) 足底压力和膝关节角度随时间周期性变化; (2) 步行速率对足底垂直压力和膝关节角度变化未见明显影响; (3) 使用 GEP 能有效地对膝关节角度进行识别, 说明足底压力数据可作为外骨骼机器人控制信号; (4) 该方法为提出不相关的自变量提供了一种参考。

本步态数据获取系统采用足底压力信号和膝关节角度同步采集, 进行步态曲线的连续分析, 相比传统通过步态相位划分来分析的方法有所改进。由于采样器材原因, 本文实验只进行了右脚(腿)压力和角度的采集, 导致每个步态周期缺失大约 40% 的压力数据, 这虽然对整体建模影响甚微, 但通过双脚压力值计算会使模型鲁棒性更高。GEP 训练结果提示, 未来可以使用更少的压力点进行识别。另外, 今后的实验还需考虑对样本一段时间无序的自然的步态行为进行采样分析, 以提高模型的普适性和通用性。

## 参 考 文 献

- [1] 杨灿军, 陈鹰, 路甬祥. 人机一体化智能系统理论及应用研究探索[J]. 机械工程学报, 2000, 36(6): 42-47
- [2] Kazerooni H, Racine J L, Huang L, et al. On the control of the berkeley lower extremity exoskeleton (BLEEX) [C]// Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA 2005). IEEE, 2005; 4353-4360
- [3] Strausser K A, Kazerooni H. The development and testing of a human machine interface for a mobile medical exoskeleton [C]// 2011 IEEE/RSJ International Conference on Intelligent Robot and System. San Francisco, CA, USA, Sep. 2011; 4911-4916
- [4] 王楠, 王建华, 周民伟. 人体下肢外骨骼机器人的步态研究现状 [J]. 中国骨科临床与基础研究杂志, 2012, 4(1): 62-67
- [5] 蔡付文, 王人成, 李广庆, 等. 低成本人体步态分析系统的研究 [J]. 中国康复医学杂志, 2008, 23(1): 49-53
- [6] Zhang Xiao-dong, Choi H. Pattern Recognition of Human Grasping Operations Based on EEG [J]. International Journal of Control Automation and Systems, 2006, 4(5): 592-600
- [7] Ferris D P, Czerniecki J M, Hannaford B. An Ankle-Foot Orthosis Powered by Artificial Pneumatic Muscles [J]. Journal of Applied Biomechanics, 2005, 21(2): 189-197
- [8] 蔡春风. 人体表面肌电信号处理及其在人机智能系统中的应用研究 [D]. 杭州: 浙江大学, 2006
- [9] 吴剑锋, 吴群, 孙守迁. 简约支持向量机分类算法在下肢动作识别中的应用研究 [J]. 中国机械工程, 2011, 22(4): 433-438
- [10] Kawamoto H, Lee S, Kanbe S, et al. Power assist method for HAL-3 using EMG-based feedback controller [C]// IEEE International Conference on Systems, Man and Cybernetics, 2003. 2003; 1648-1653
- [11] 孙建, 余永, 葛云建. 基于接触力信息的可穿戴下肢助力机器人传感系统研究 [J]. 中国科学技术大学学报, 2008, 38(12): 1432-1438
- [12] Pratt J, Krupp B, Morse P G, et al. The RoboKnee, an exoskeleton for enhancing strength endurance during walking [C]// Proceedings of the IEEE International Conference, ICRA 2004. Robotics and Automation, 2004; 2430-2435
- [13] Cordero A F B, Koopman H J F M, van der Helm F C T. Use of pressure insoles to calculate the complete ground reaction forces [J]. Journal of Biomechanics, 2004, 37: 1427-1432
- [14] Savelberg H, De Lange A L. Assessment of the horizontal, fore-aft component of the ground reaction force from insole pressure patterns by using artificial neural networks [J]. Clinical Biomechanics, 1999, 14(8): 585-592
- [15] ortolero X, Masani K, Popovic M R. Step Prediction During Perturbed Standing Using Center of Pressure Measurements [J]. Sensors, 2007, 7(4): 459-472
- [16] Zhang Kuan, Sun Ming, Lester D K, et al. Assessment of human locomotion by using an insole measurement system and artificial neural networks [J]. Journal of Biomechanics, 2005, 38(11): 2276-2287
- [17] 郑成闻. 基于柔性双足信息的助力机器人行走控制方法研究 [D]. 合肥: 中国科学技术大学, 2011
- [18] Candidia F. Gene Expression programming; mathematical modeling by an artificial intelligence [M]. Springer-Verlag, 2006
- [19] 张立佑. 成人足底压力中心前进路径模式之建立 [D]. 台湾朝阳科技大学, 2009