文章编号：2096-1472(2016)-12-20-04

基于支持向量机(SVM)的音乐情感分类

陈维华
（河北软件职业技术学院计算机应用工程系，河北 保定 071002）

摘 要：如今数字化信息对人们的生活产生了巨大的影响，对于爱好音乐的人们来说大量的歌曲信息如果进行有效的分类，如何对这些音乐进行分类，从而进行有效的信息检索是信息处理领域的研究热点。通过对音乐的乐理知识的深入分析，能够有效的提取出反应音乐情感的特征向量，运用支持向量机(SVM)实现了对音乐的情感分类。实验结果表明，所运用的分类方法准确有效。

关键词：支持向量机；MIDI文件；音乐情感分类
中图分类号：TP391.4 **文献标识码**：A

Emotion Classification of Music Based on Support Vector Machine

CHEN Weihua
(*Department of Computer Application Engineering*,*Hebei Software Institute*,*Baoding* 071002,*China*)

Abstract: Now the digital information on the people's life has had a huge impact,for people interested in music for many songs if information for effective classification,how to classify these music,thus effectively information retrieval is a hot research in the field of information processing.Through in-depth analysis of the music knowledge,can extract features vector reaction music emotion effectively,we use the Support Vector Machine (SVM) to realize the emotion classification of music.The experimental results show that the classification method is accurate and effective.

Keywords: Support Vector Machine(SVM);MIDI files;music emotion classification

1 引言(Introduction)

音乐作为大众娱乐主要方式之一，不仅为大众提供美妙的听觉享受，同时其本身也具有丰富的情感内涵，具有很强的情感调节功能，满足听众的情感需求。目前很多大型音乐网站都需要高效的检索和推荐功能，而音乐情感则是音乐检索与推荐工作的重要依据之一。目前音乐的感情多数是根据听众或者专家对于音乐的评价作为感情标签，其工作量相当大，而且不同认知水平的人对于同一段音乐有不同的评价。随着计算机与人工智能技术的发展，音乐感情的自动识别成为国内外学者研究的热点问题之一。

现有的音乐情感研究模型，可分为基于文本关键词模型和基于音频参数模型两种。David Torres等人[1]根据乐曲的歌词信息，识别乐曲的情感类型。Yuan-Yuan shi等人[2]通过构建一个调制频谱分析系统，从压缩的音乐文件中提取节奏信息，通过分析音乐节奏的变化实现对音乐感情的分类。M.wang等人[3]通过提取音乐的频谱信息，利用支持向量机实现对于音乐情感的分类。Cyril Laurier和Perfeto Herrera等人利用音乐波形文件的频谱特征，利用支持向量机实现音乐的情感分类。

本文从乐谱创作的角度出发，以音乐网站音乐推荐系统对于音乐情感需求为背景，提出通过对音乐的MIDI文件解析得到音乐的感情特征组成六维感情特征向量，利用支持向量机(SVM)实现对音乐表达的情感进行具体分类。

2 Thayer情感模型(Thayer emotion model)

目前，对于音乐情感分类所采用的模型主要包括Thayer情感模型、威廉.冯特情感模型、Hevner情感模型、Russell情感模型和Plutchik情感模型等，其中Thayer情感模型是二维情感模型[4]纵坐标表示的是能量维度，从“平静的”到“活力的”变化，反映的是主体的情感活跃程度，横坐标表示的是压力维度，从“消极”到“积极”的变化，反映了主体的

基金项目：2016年河北省高等学校科学技术研究项目研究成果之一。项目名称：基于微信企业号的移动办公系统设计与开发；项目编号：QN2016232；项目主持人：陈维华.

www.rjgcz.com

主观感受，从而将音乐情感划分为具有代表性的四类：兴奋的、愤怒的、悲伤的和放松的。如图1所示为Thayer情感模型采用的能量维度和压力维度与乐理理论所对应的情感表达方式具有较好的对应关系，因此本文将采用Thayer情感模型，利用支持向量机将乐曲的表达情感分为四类。

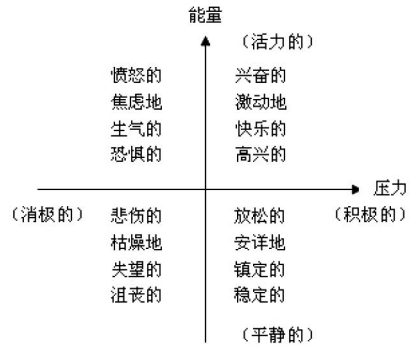

图1 Thayer情感模型

Fig.1 Emotion model

3 MIDI文件的音乐特征的提取(The MIDI music feature extracting of files)

MIDI文件是音乐文件的一种存储形式，它以字节的形式对乐曲进行描述，使用数字编码描述乐谱的基本特征[4]。多数MIDI文件都是多音轨文件，其中的主旋律和伴奏旋律都分布在不同的音轨上，通常将主旋律所在的音轨称之为主音轨，其包含了音乐的主要情感信息。因此通过分析主音轨的情感就可以识别出该段音乐所表达的感情。目前比较成熟的主旋律提取方法是根据信息熵理论定义音轨特征的熵值，然后由MIDI文件的音轨信息熵和其他重要特征组成特征向量构建随机森林分类器，进而提取MIDI文件的主旋律。实验证明，该方法的正确率达到93%以上。

音乐是音符的有机组合，其不同的组合形式不同的旋律，表达不同的情感。每个音符都有自身的音高、音长和音强等信息，但每首音乐并不是由单个的音符所表达，而是通过不同的音符之间的组合表达旋律，抒发感情。因此音乐的旋律、节奏也是感情表达的重要方面。因此，本文通过对MIDI文件的解析，可以得到的基本数据信息包括音高、音长、音强和节拍等信息，本文分别定义为P、Du、In和J。

音乐是由不同的音符组合在一起，其组合规律形成自身的特点，表达着各种感情。包括音程、旋律、节奏等高层次的音乐特征。因此本文定义如下表达音乐情感的向量空间。

音高：音高是音乐的基本要素之一，其高低及其变化是音乐内容和感情的表达的重要方式。本文分别定义音高均值和音高方差表示乐曲的音高水平和音高变化：

音高均值定义为：

$$\overline{P}=\frac{\sum_{i=1}^{n}P_i}{n} \tag{1}$$

其中，n表示乐段音符的个数，P_i表示音符的音高。

音高变化定义为：

$$Var_P=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i-\overline{P})^2} \tag{2}$$

音强：音乐的强度是音乐表达感情的重要方式之一。通过音强的变化可以进一步加强和丰富音乐的情感。本文用音强均值和音强方差来衡量音乐的音强及音强变化：

音强均值定义为：

$$\overline{In}=\frac{\sum_{i=1}^{n}In_i}{n} \tag{3}$$

音强变化定义为：

$$Var_In=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(In_i-\overline{In})^2} \tag{4}$$

旋律：旋律是由连续音符以特定的顺序组织起来形成的，它是音乐情感重要的表现手段之一，反映了音乐的走向信息。通常包括时间信息和走向信息。因此本文用根音的时值作为时间信息，前后的音高的差值作为走向信息，定义如下：

$$Mel=\frac{P_{i+1}-P_i}{Du_i} \tag{5}$$

节奏：节奏是音乐中规律性出现的强弱和长短的现象，是音乐情感的重要表现手段。本文用紧张性和稳定性两个特征来表征音乐的节奏，具体定义如下：

$$Density=\frac{n}{D} \tag{6}$$

$$Rhy=\sum_{i=1}^{n-1}\left|\frac{In_{i+1}}{D_i}\right| \tag{7}$$

因此，对于任何一段乐曲，都可以找到一个7维的向量表示其感情特征。

4 支持向量机理论(Support vector machine theory)

支持向量机是统计学习理论发展的产物，针对有限样本情况，建立了一套完整、规范的基于统计的机器学习理论和方法，大大减少了算法设计的随意性，被广泛的应用于模式识别、函数逼近、概率密度估计、降维等领域。

4.1 线性支持向量机

4.1.1 两类线性可分的情况

设训练样本输入为x_i，$i=1,n,...$；$x\in R^d$，对应的输出为

www.rjgcz.com

$y\in\{+1,-1\}$，其中+1和−1则代表两分类的类别标识，则存在分类的超平面，其方程可表达为$w\cdot x+b=0$。为保证所有样本能够分类正确且存在分类间隔，须满足如下约束条件：

$$y_i[x_i\cdot w+b]-1\geq 0,\quad i=1,2,...,n \tag{8}$$

在约束条件(1)下，最小化函数为：

$$\min\quad \varphi(w)=\frac{1}{2}\|w\|^2=\frac{1}{2}(w\cdot w^T)$$
$$s.t.\quad y_i[x_i\cdot w+b]-1\geq 0,\quad i=1,2,...,n \tag{9}$$

采用拉格朗日(Lagrange)乘数法求解该具有线性约束的二次规划问题。引入(Lagrange)乘子$a_i>0$,得到对偶的最优化问题

$$\max\quad Q(x,a_i)=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i-1}^{n}\sum_{i-1}^{n}a_ia_jy_iy_j(x_i\cdot x_j)$$
$$s.t.\quad a_i\geq 0\quad 及\quad \sum_{i=1}^{n}a_iy_i=0 \tag{10}$$

如果为a_i^*最优解，那么

$$w^*=\sum_{i=1}^{n}a_i^*y_ix_i \tag{11}$$

根据库恩−塔克(Karush−Kuhn−Tucher，KKT)条件可知，当Lagrange乘子与约束条件的乘积等于0时，在最优点，即

$$a_i[y_i(w\cdot x_i+b)-1]=0,\quad i=1,2,...,n \tag{12}$$

对于多数样本a_i的值为0，不为0的a_i^*则对应于支持向量，相应的学习方法称为支持向量机(SVM)。

最优化分类函数为

$$f(x)=sign\{(w^*\cdot x)+b^*\}=sign\left\{\sum_{i=1}^{n}a_i^*y_i(x_i\cdot x)+b^*\right\} \tag{13}$$

其泛化能力最强，$sign(\cdot)$为符号函数，n则为支持向量个数。

4.1.2 两类线性不可分的情况

引入一个松弛项$\xi\geq 0$，表示样本距支持平面的偏差。则广义最优化分类面可表示成如下优化问题：

$$\min\quad \varphi(w)=\frac{1}{2}\|w\|^2+c\left[\sum_{i=1}^{n}\xi_i\right]$$
$$s.t.\quad y_i[(w\cdot x_i)+b]-1+\xi_i\geq 0\qquad i=1,2,...,n \tag{14}$$

c为惩罚函数，表示对错误分类样本的惩罚程度，与线性可分情况相似，可求得a的最优化问题为

$$\max\quad Q(x,a_i)=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i-1}^{n}\sum_{i-1}^{n}a_ia_jy_iy_j(x_i\cdot x_j)$$
$$s.t.\quad 0\leq a_i\leq c\quad 及\quad \sum_{i=1}^{n}a_iy_i=0 \tag{15}$$

最优分类函数为

$$f(x)=sign\{(w^*\cdot x)+b^*\}=sign\left\{\sum_{i=1}^{n}a_i^*y_i(x_i\cdot x)+b^*\right\} \tag{16}$$

4.2 非线性支持向量机

对于非线性分类超平面，支持向量机(SVM)的核心思想是：通过非线性变化Φ将输入空间中的样本X变换到一个高维特征空间中，进而在新的空间中求得最佳线性分类面进行分类。而这种非线性变换是通过定义适当的内积函数——核函数(Kernel Function)实现的。

假设有非线性映射Φ：$R^n\rightarrow H$将输入空间的样本映射到高维特征空间H中，当在特征空间中构造最优超平面时，训练算法仅使用特征空间中的点积，即$\Phi(x_i)\bullet\Phi(y_i)$。所以若能找到一个函数$K(\bullet)$，使$K(x_i,x_j)=\Phi(x_i)\bullet\Phi(y_i)$，则在高维空间中不需要知道变换$\Phi$的形式，只需进行相应的内积运算即可。

由Mercer理论可知，令$K(x_i,x_j)=\Phi(x_i)\bullet\Phi(y_i)$为内积函数的核函数，则相应的关于$a$二次规划问题变为

$$\max\quad Q(x,a_i)=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i-1}^{n}\sum_{i-1}^{n}a_ia_jy_iy_j[\Phi(x_i)\cdot\Phi(y_i)]$$
$$=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i-1}^{n}\sum_{i-1}^{n}a_ia_jy_iy_jK(x_i,x_j) \tag{17}$$
$$s.t.\quad 0\leq a_i\leq c\quad 及\quad \sum_{i=1}^{n}a_iy_i=0$$

最优分类函数为

$$f(x)=sign\{(w^*\cdot x)+b^*\}=sign\left\{\sum_{i=1}^{n}a_i^*y_i(x_i\cdot x)+b^*\right\} \tag{18}$$

式中，n表示支持向量数，该式则被称为支持向量机(SVM)。

由Mercer条件得到的内积函数$K(x_i,x_j)$称为核函数。

5 基于SVM的音乐情感分类(Music emotion classification based on SVM)

5.1 M-ary分类方法

支持向量机方法的提出是针对二分类问题，将二分类拓展为多分类是一段时间内支持向量机算法研究的重要方向之一，假设多分类问题有k个类别$S=\{1,2,...,k\}$，训练样本为$\{(x_i,y_i),i=1,2,...,n\}$，其中$y_i\in S$，实现SVM的多分类问题主要方法有四种方法，其中M−ary多分类方法具有结构简单，使用的分类器最少等优点，因此本文采用M−ary算法实现对音乐情感的分类。本文支持向量机(SVM)采用径向基函数作为核函数，核函数$\Psi_k(x_i,x_j)=\exp\{-\|x_i-y_j\|^2/2\sigma_k^2\}$，$k=1,2,...,M$，取$\gamma=10$，$\sigma^2=0.2$。

将兴奋和激动的情感分为第一类，将愤怒和焦虑的情感分为第二类，将悲伤枯燥的情感分为第三类，放松安详的情感分为第四类。因为$\log_2^k=4$,取$k=2$，所以实现本文的分类由两个SVM分类器即可。第一个分类器，类别一、三相对应

www.rjgcz.com

的样本记为+1，类别二、四相对应的样本记为−1；对于第二个分类器，类别一、四记为+1，类别二、三记为−1。通过标记结果的组合(+1，+1)、(+1，−1)、(−1，+1)、(−1，−1)四类，实现对样本的分类。

5.2 数据预处理

实验选取200首多音轨文件，包括80首港台歌曲、120首大陆歌曲，然后由两名专业教师和三名音乐专业的学生对每首音乐每段表达的感情进行标注。共制作表达不同感情的音乐段240段。然后利用方法六对200首音乐提取主音轨，生成如图2所示的文本文件。将240段音乐文件制作成对应的MIDI文本文件作为实验素材，如图2所示。

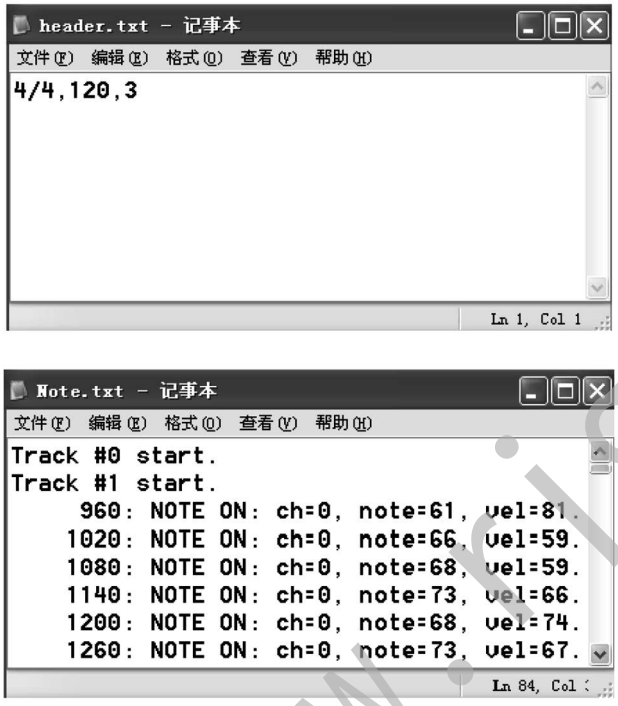

图2 MIDI文件的节拍和音符信息

Fig.2 MIDI file beat and note information

5.3 实验结果与分析

实验选取240段音乐素材中分别表达四类情感的乐段各20段共80段作为训练样本，其余的160段作为测试样本。首先用80段训练样本对支持向量机进行训练，然后对支持向量机进行学习和训练。实验采用matlab7.0进行仿真，分类结果如表1所示。

表1 乐段情感分类结果

Tab.1 Music emotion classification results

样本类型	实验样本数量	识别正确数量	识别正确率
训练样本	80	80	100%
测试样本	160	126	78.8%

通过仿真实验结果表明，该方法取得了较为理想的识别结果，出现误判的原因主要是本身乐段所表达的情感不是很明确，专家给的情感类型也并不同意，因此产生了误判。同时又部分实验样本乐段情感明确，出现误判。

6 结论(Conclusion)

乐曲的情感识别对于音乐的自动化制作、音乐工业化应用、音乐的检索和推荐都具有重要意义。本文通过对于对MIDI文件的解析，提取出能够表现音乐音高及变化、音强及变化、旋律、节奏等能够表达音乐情感的7个向量，通过对支持向量机的训练，进而实现了对于MIDI音乐文件的情感分类，实验证明了该方法的有效性，误差可以控制在合理的范围之内。下步工作将主要针对复杂的音乐，尤其是具有复合性音乐情感的音乐作品，如何进行音乐情感的分类与识别。

参考文献(References)

[1] David Torres,et al.Identifying Words that are Musically Meaningful[J].University of California,San Diego.Austrian Computer Society,2009:143−152.

[2] Yuan−Yuan Shi,et al.A Temp of Feature Via Modulation Spectrum Analysis and its Application to Music Emotion Classification[C].Presented at IEEE International Conference on Multimedia and Expo(ICME 2006),Toronto,CANADA,2006:247−253.

[3] M.Wang,N.Zhang,H.Zhu.User−adaptive Music Emotion Recognition[J].New York,NY,10016−5997,United States,2004:201−209.

[4] R.Thayer.The Biopsychology of Mood and Arousal[M].Oxford University Press,1989:62−69.

[5] 赵健谕.音乐情感识别方法的研究[M].辽宁:辽宁大学,2011:22.

[6] 葛广英.基于SVM的车型检测和识别算法[J].计算机工程,2007(06):11−14.

[7] 肖建华.智能模式识别方法[M].广州:华南理工大学出版社,1998:18−22.

作者简介：

陈维华(1978−)，女，硕士，副教授.研究领域：计算机应用，信息系统.

www.ijccz.com