

基于 H-K 算法的 MIDI 音乐主旋律提取

刘 勇, 林景栋, 穆伟力

(重庆大学 自动化学院, 重庆 400044)

摘要:音乐的主旋律音轨包含了很多重要的音乐旋律信息,是音乐特征识别的基础,也是音乐灯光表演方案设计的前提工作。这方面的工作涉及了音乐旋律的表达、旋律特征的抽取以及分类技术等许多内容。针对多音轨 MIDI 文件,介绍一种多音轨 MIDI 音乐主旋律识别方法。通过对表征音乐旋律的特征量的提取,采用 H-K 分类算法,构建音轨分类器模型,对 MIDI 主旋律音轨和伴奏旋律音轨进行分类,从而提取 MIDI 音乐主旋律音轨。实验结果显示取得了较好的效果,为音乐灯光表演方案的自动设计做了必要的前提工作。

关键词:音乐灯光;MIDI;主旋律提取;特征向量

中图分类号:TP273

文献标识码:A

文章编号:1673-629X(2011)06-0154-03

Melody Extraction Method from MIDI Based on H-K Algorithm

LIU Yong, LIN Jing-dong, MU Wei-li

(Automation College of Chongqing University, Chongqing 400044, China)

Abstract: Main melody is the basis of the music feature recognition and the premise of music light show design work by compute, as it contains the most important information about music melody. This work includes many questions, such as melody representation, melody character extraction and classifying technology. This paper represents a model for auto melody extraction method for multi-track MIDI files. Construct the feature vector space by means of extracting the music melody feature. To classify the melody of the main track and the accompaniment tracks with the H-K algorithm in order to extract the main melody track, and achieved good results. It prepares for the program of automatic design the music light show.

Key words: music-light show; MIDI; melody extraction; eigenvector

0 引言

随着城市建设的快速发展,音乐灯光表演是近年来城市灯光景观的一个新兴发展方向,以其艺术性、装饰性以及文化内涵备受瞩目。从目前的研究来看,大多数的音乐灯光表演方案的设计还是借鉴于传统舞台灯光表演的设计方法,即由设计人员对乐曲进行编辑处理,再根据乐曲的特征配以一定的灯光动作,如果有新的曲目,就需要设计人员重新编辑处理,因此系统灵活性差,并且对设计人员的要求较高。鉴于此情况,在重庆“朝天扬帆”大型动感音乐灯光表演工程项目,笔者所在团队开发设计了一个基于音乐特征识别的灯光表演方案的计算机辅助设计系统。该系统通过对音乐特征的提取,灯光动作库的建立,实现音乐与灯光动作

的匹配,最终完成音乐灯光表演方案的自动生成,而音乐特征识别技术是该系统的前提性工作。采用 MIDI 音乐文件作为数据源,但 MIDI 文件几乎都是多音轨的,并通常都有一个表征主旋律的音轨^[1],所以必须先对 MIDI 文件进行音轨分解,得出主旋律音轨才能为其所用。

文献[2]通过提取出音轨的特征向量,如音轨名称、通道号、平均力度、主音量等,将这些特征量加权求和,最高者即为主旋律音轨。文献[3]通过对 MIDI 音乐文件去除每一音轨中的控制音符和不包含旋律信息的音轨,归并到该文件中的具有音高柱状图特征的音符集,从而提取出主旋律音轨。文献[4]通过考察 MIDI 文件的各个音轨,如果该音轨具有旋律特征的片段总长度超过规定阈值,则将其加入到候选音轨集,再通过比较,从中择优选出主旋律音轨。

文中采用一种从 MIDI 文件中提取音轨特征量的方法,通过人为标识一部分数据训练得出主旋律音轨判别函数,训练数据越大,主旋律音轨提取的准确度越高。

收稿日期:2010-11-14;修回日期:2011-02-26

基金项目:重庆市“朝天扬帆”项目支持

作者简介:刘 勇(1985-),四川遂宁人,硕士研究生,研究领域为音乐灯光控制及工业控制;林景栋,副教授,硕士生导师,研究方向为基于数据驱动的复杂工业过程状态识别和优化控制、先进控制理论 research 等。

1 音乐旋律特征量的提取

1.1 音频文件的选择

音长、音高、力度、音色、曲调、节拍是音乐的几个基本特征,音乐特征识别就是通过对几个参数进行分析识别,特别是对音长和音高信息的分析,从而得到音乐的节奏、旋律、曲式特征以及音乐情感特征等信息,为音乐灯光动作匹配作知识准备。

音乐信息的记录主要有三种记录格式:声音文件、MIDI 文件和模块文件。声音文件记录的是对真实声音的波形进行二进制采样的数据,记录了实际的演奏效果。MIDI 文件记录的是音乐演奏指令序列,但并不是真实的声音数据,其中包括音乐的乐谱信息以及音乐演奏的全部过程。模块文件综合了声音文件和 MIDI 格式文件的相关特征,即模块文件当中不仅保存了真实声音的采样数据,同时还保存得有音乐演奏指令序列。

金毅等人通过研究指出^[5]:从声音文件中提取旋律特征处理过程复杂,难度高,精确度不如 MIDI 文件和模块文件;模块文件通用性很差,编码格式又多;对于 MIDI 文件,很多音乐的基本特征信息都可以从中直接提取出来,这就为音乐特征的自动识别提供了极大的方便,所以文中选取 MIDI 乐曲作为提取音乐旋律的数据源。

MIDI 格式的音乐文件记录了音乐的全部乐谱和演奏的全过程,很多音乐的基本特征(音高、音长、音色、速度、力度等)可以直接提取出来,在 MIDI 文件中,每一类特征都是用 0-127 之间的数字描述。图 1 所示是一个 MIDI 音乐文件的格式^[6,7]。

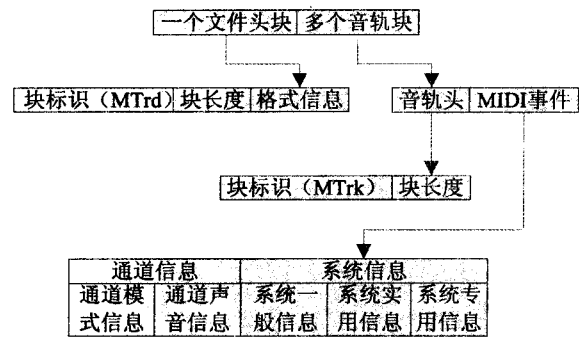


图 1 MIDI 音乐文件的基本结构

1.2 音乐特征量的选取

根据音乐乐理及实例数据分析,主旋律音轨和伴奏音轨主要区别体现在左右声道平衡度、音量、音长、力度、音高和音程等几个方面,因此文中将这几个特征参数作为主辅音轨的判断依据。前面 5 个特征量都能直接从 MIDI 文件中直接提取^[8],经过如表 1 所示的计算公式就可以得到,下面重点介绍音程比的提取方法。

音程(Interval)指相邻两音高差的绝对值。彭琼等对 300 多首乐曲进行音轨音程统计,结果发现旋律音程主要都集中在 0-6 度之间,但八度以上的音程却很少出现^[9]。由此可见音程是可以用来区别伴奏旋律音轨和主旋律音轨的一个重要特征。但是如果仅将音程作为主旋律音轨的判定依据,是不充分的。笔者通过相关实验发现仅仅依靠音程对 MIDI 乐曲的主旋律音轨进行提取,结果显示该方法仅对一部分音乐有效。因此还应该考虑音乐的其它特征对主旋律音轨的影响。参照文献[3,9,10]的方法,首先对音乐的音高信息进行处理,以得到所需的音程统计数据。

设 Interval 是相邻两个音符 Pitch_j 和 Pitch_{j+1} 的音程的绝对值,并根据如下公式计算处理:

Interval_j =
$$\begin{cases} |Pitch_{j+1} - Pitch_j| & |Pitch_{j+1} - Pitch_j| < 25 \\ 25 & |Pitch_{j+1} - Pitch_j| \geq 25 \end{cases}$$

设数组 Count_{*j*} 存放音轨 *i* 的音程统计个数结果,将 Interval_{*j*} ≤ 6 的音程数进行累加并记为 ToneM_{*i*},即 ToneM_{*i*} 为音轨中 0-6 度的音程。将音轨 *i* 的所有音程进行累加统计得到 Tone_{*i*},将音轨 0-6 度的音程所占的比例作为音程特征量,即:

$$x_4 = \frac{ToneM_i}{Tone_i} = \frac{\sum_{j=0}^6 Count_j}{\sum_{j=0}^{25} Count_j}$$

表 1 音乐特征量的选取

音乐特征量	计算公式	简述
左右声道平衡度	$x_1 = 1 - \frac{ Balance_i - 64 }{64}$	表征该音轨的音符在左右声道中音量大小的比例
主音量	$x_2 = \frac{Volume_i}{127}$	声音振幅的大小
平均音长	$x_3 = \frac{1}{n} \sum_{j=1}^n Duration_{ij}$	音符持续时间的平均值,它决定于发音体振动时间的久暂
平均力度	$x_4 = \frac{1}{n} \sum_{j=1}^n \frac{Velocity_{ij}}{127}$	表征音的强弱程度,用它来表征各音轨能量的大小
平均音高	$x_5 = \frac{1}{n} \sum_{j=1}^n \frac{Pitch_{ij}}{127}$	指各种不同高低的音,即音的高度,由发音体的振动频率决定的
音程比	$x_6 = \frac{ToneM_i}{Tone_i} = \frac{\sum_{j=0}^6 Count_j}{\sum_{j=0}^{25} Count_j}$	音程指相邻两音高差的绝对值。定义某一音轨中 0-6 度音程所占的比例为音程比

2 H-K 算法的主旋律提取算法

2.1 识别主音轨的预处理

MIDI 格式的音乐文件可以有 1 至 16 个音轨,一般来说,每个通道都由一种乐器演奏的音符序列指令组成,代表着同一种乐器产生的声音。在文献[9]中提到一些主旋律音轨的判别方法,概括起来如下:

(1)有的 MIDI 音乐文件的音轨头里面会有一些

说明,其中可能会包含乐曲的相关说明,有的可以通过此说明获得音轨通道的相关分配情况;

(2)比较规范的 MIDI 音乐文件通常会将通道 10 分配给打击乐器使用,而打击乐多是用来伴奏的,所以 10 号音轨通道一般不会是主旋律音轨;

(3)主旋律音轨对音乐的表现起着重要作用,故其所用音符数量一般不会太少,所以音符数量小于音符数量平均值的一般不会是主旋律音轨;

(4)音轨总长小于各音轨长度平均值一半一般不会是主音轨。

通过上面这四条规则可以排除一些明显不是主音轨的音轨,同时也提高了定位的准确性。

2.2 基于 H-K 算法的主旋律提取算法

基于对各音轨特征量的统计,设想采用某种算法来构建一种模型,通过此模型可以自动地定位该特征量表征的音轨是主旋律音轨还是伴奏旋律音轨。显然主音轨定位是一个二分类判别问题。解决该问题的主要方法就是构造一个合适的分类器。

关于构建分类模型的问题,有基于规则的分类方法和基于统计的分类方法。基于规则的分类方法理论基础强,合理性高,且分类的规则易于人们理解,但是实用性差,在很多时候往往显得力不从心;基于统计的分类方法往往是依据某种统计后得到的客观规律,并通过训练与类别相关的数据来完成分类器的建立工作的。常见的基于统计的分类器构造方法有经验法、统计法、机器学习和 BP 神经网络^[11]等。H-K 算法针对两类或多类分类问题,在特征空间中寻找一个最优的超平面作为两类的分割^[12]。H-K 算法思想很朴实,就是在最小均方误差准则下求得权矢量。它适用于线性可分和非线性可分得情况,对于线性可分的情况,给出最优权矢量,对于非线性可分得情况,能够判别出来,以退出迭代过程。因此文中将运用 H-K 算法进行主旋律音轨定位。

由左右声道平衡度、音量、音长、力度、音高和音程构成一个 6 维特征空间,音乐特征矢量 $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^T$ 对应于特征空间中的一个点,运用已知的样本进行学习产生一个最优超平面 $d(\mathbf{x}) = w_0\mathbf{x} + w_7 = 0$,其中 $w_0 = (w_1, w_2, w_3, w_4, w_5, w_6)$,将特征空间划分成两个子区域,使不同模式类分布在不同的子区域中。将特征量符号规范化,即将特征矢量加上一维分量 1,并将伴奏音轨的所有特征分量都乘以-1,得特征量增广矩阵 \mathbf{X} ,为了使解可靠,引入 N 维余矢量 $\mathbf{b} > 0$,于是不等式方程组变为:

$$w_0' \mathbf{X} \geq \mathbf{b} > 0$$

式中:

$$w_0' = (w_1, w_2, w_3, w_4, w_5, w_6, w_7)$$

使用平方误差标准函数:

$$J(w, \mathbf{b}) = \|w_0' \mathbf{X} - \mathbf{b}\| = \sum_{i=1}^n (w_i x_i - b_i)^2 \Rightarrow \min$$

H-K 算法将准则函数 $J(\cdot)$ 视作 w 和 \mathbf{b} 的函数,在迭代过程中修正 w 的同时,也对矢量 \mathbf{b} 进行调整,运用最优化技术求得准则函数 J 关于 w 和 \mathbf{b} 的极小值点。在迭代调整过程中,应满足 \mathbf{b} 的各分量均为正值的约束,同时也为 J 使 w 更趋向解区的中心。 $\mathbf{b}(k)$ 的各分量只能向增大的方向调整。

文中以“光辉岁月.mid”为实例运用 H-K 分类算法,所得结果如表 2 所示。由表 2 可知,音轨 3 是音乐“光辉岁月.mid”的主旋律音轨。

表 2 “光辉岁月.mid”的主旋律特征抽取实例

音轨号	音轨 3	音轨 10	音轨 2
发音时间	00:35:300	00:35:300	00:35:400
发音音量	0.7874	1	1
左右声道平衡度	1	0.5	0.5
平均力度	0.9985	0.62992	1
音程比值	0.7	1	0.984848
音长均值	62.45614	65.41936	65.41936
音轨判别值	0.037	-0.022	-0.035

3 实验结果

通过对 MIDI 音乐文件的分析,对表征音轨主旋律的音乐特征量进行提取,采用 H-K 算法来区分主旋律音轨和伴奏旋律音轨。随后选取了 120 首 MIDI 音乐文件,人工标注得到 120 个主旋律音轨和 427 个伴奏旋律音轨,并提取这些音轨的特征量作为分类模型的输入样本向量,对分类模型进行训练,以得到分类模型的准确参数。最后,随机选择 50 首 MIDI 音乐文件对分类器进行测试。研究结果表明,此方法能有效提取 MIDI 音乐文件的主旋律音轨。图 2 给出了随着样本数据的变化,主音轨识别正确率的变化情况,显然训练数据越大,主旋律音轨提取准确度越高。

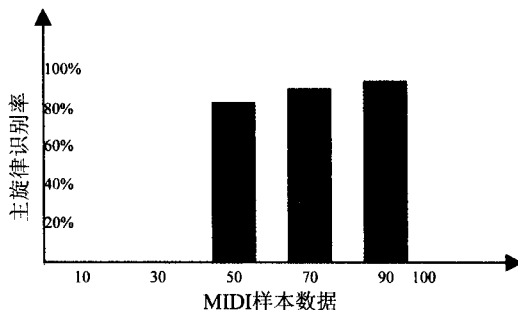


图 2 主旋律正确识别率

4 结束语

基于多音轨的 MIDI 文件,提取了表征旋律的音乐特征量,构建特征向量空间,通过 H-K 分类算法,在此基础上实现了多音轨 MIDI 音乐主旋律的提取。方法

(下转第 160 页)

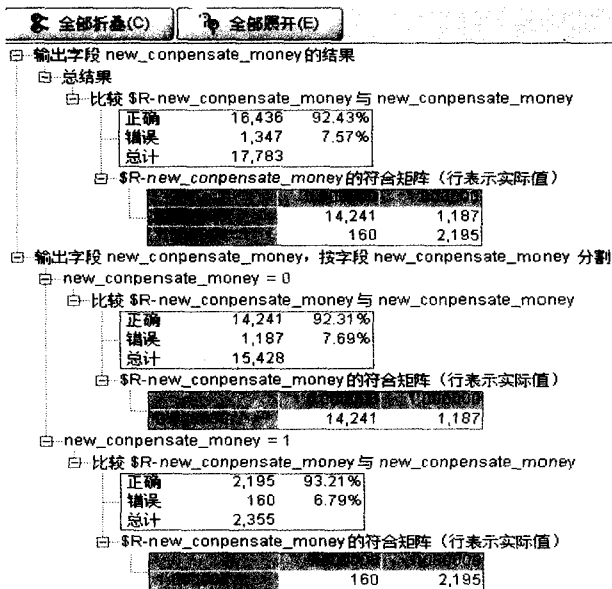


图 8 客户理赔风险挖掘验证集收益表

从上图可以看出,在训练集和验证集上的错误率分别是 7.57% 和 7.65%,即正确率分别是 92.43% 和 92.21%,这说明本研究中客户理赔风险特征模型的正确率是非常令人满意的。

3 结束语

通过客户流失模型和客户理赔风险模型,从中得出了客户流失的特征、客户理赔风险大小的特征等结论。在面向分析的数据仓库的基础上,可以利用数据挖掘技术来设计险种绑定销售,制定挽留客户的措施,并且控制保险公司理赔风险,为保险产品定价提供合

理依据。

参考文献:

- [1] 闫娜娜,刘 锋,李锡娟,等. 支持 CRM 分析的数据仓库多维启动模型[J]. 计算机技术与发展,2008,18(5):21-22.
- [2] 于红蕾,华庆一,刘燕玲,等. 数据仓库在电信统计分析中的应用[J]. 计算机技术与发展,2007,17(8):59-60.
- [3] 张 宁,贾自艳,史忠植. 数据仓库中 ETL 技术的研究[J]. 计算机工程与应用,2002(24):214-215.
- [4] 连立贵,金 凤,蔡家媚. 数据仓库中的数据提取[J]. 计算机工程,2001(9):61-62.
- [5] Lingand R, Yen D C. Customer Relationship Management: An Analysis Framework and Implementation Strategies[J]. Journal of Computer Information System, 2001(3):82-97.
- [6] Ruggieri S. Efficient C4.5[J]. IEEE Transactions on knowledge and Data Engineering, 2002,14(2):438-444.
- [7] 梅 强,张冬荣. 数据挖掘在保险分析中的应用[J]. 计算机工程,2004(12):37-38.
- [8] 桂现才,彭 宏;王小华. 基于决策树的保险客户流失分析[J]. 计算机工程与设计,2005(8):59-60.
- [9] 马建红,王万森. 基于数据仓库的保险管理系统的设计与实现[J]. 微机发展,2004,14(7):64-66.
- [10] 梁 循. 数据挖掘:建模、算法、应用和系统[J]. 计算机技术与发展,2006,16(1):86-87.
- [11] 王爱平,王占凤,陶嗣干,等. 数据挖掘中常用关联规则挖掘算法[J]. 计算机技术与发展,2010,20(4):17-20.
- [12] 姚毓才,王本年. 数据挖掘工具的分类与挖掘[J]. 计算机技术与发展,2006,16(8):25-27.

(上接第 156 页)

计算简单,易于实现,适用于主旋律只布在一个音轨的音乐。通过对 50 多首 MIDI 音乐进行分析统计,92.8% 都能准确地提取出主音轨。结果令人满意,同时也表明了该方法的有效性及可行性,从而为音乐灯光表演方案辅助设计系统的构建进行了很好前提准备。

参考文献:

- [1] Liu Li, Cai Junwei, Wang Lei, et al. Melody Extraction from Polyphonic MIDI Files Based on Melody Similarity [C]//2008 International Symposium on Information Science and Engineering (ISISE '08). [s. l.]:[s. n.], 2008:232-235.
- [2] 赵 芳,吴亚栋,宿继奎. 基于音轨特征量的多音轨 MIDI 主旋律抽取方法[J]. 计算机工程, 2007,33(2):165-167.
- [3] 冯国杰,王吉军. 基于分层次聚类的 MIDI 音乐主旋律提取方法[J]. 计算机工程与应用,2009,45(26):233-235.
- [4] 叶 霖,李雄飞,刘丽娟,等. 一种有效识别 MIDI 文件中主旋律音轨的方法[J]. 计算机应用与软件,2010,27(1):48-50.

- [5] 金 毅,黄 敏. 基于旋律的音乐检索研究——旋律特征的表达和提取[J]. 信息检索技术,2003,4:49-51.
- [6] 杨 军. MIDI 消息和标准 MIDI 文件格式剖析及应用[J]. 中南民族大学学报(自然科学版),2009,22(Sup):62-64.
- [7] 刘嘉欣. 嵌入式 MIDI 文件格式解析设计与实现[J]. 微计算机信息,2006,22(11-2):66-67.
- [8] 秦 丹. 利用 C# 从 MIDI 文件中获取音乐旋律[J]. 电脑知识与技术,2009(7):4281-4284.
- [9] 彭 琼,支 琤. 计算机自动识别音乐情感的关键技术研究[J]. 电声基础,2008,32(4):35-38.
- [10] Zhu Bin. Music Features Recognition and its Application in National Music Protection [C]//7th International Conference on Computer-Aided Industrial Design and Conceptual Design, 2006 (CAID-CD '06). [s. l.]:[s. n.], 2006:1-6.
- [11] Li Jiangtao, Yang Xiaohong, Chen Qingcai. MIDI melody extraction based on improved neural network [C]//2009 International Conference on Machine Learning and Cybernetics. [s. l.]:[s. n.], 2009:1133 - 1138.
- [12] 孙即祥. 现代模式识别[M]. 长沙:国防科技大学出版社, 2001:46-70.