

中国民歌地域风格分类中的特征选择

刘 怡, 蔚 磊, 刘子利, 王 鹏

(中国人民大学信息学院, 北京 100872)

摘 要: 歌曲风格的自动分类技术研究, 是音乐信息检索领域中一个重要课题. 本文主要讨论了对中国民歌地域风格自动分类中不同特征选择方法对于分类性能的影响. 论文选用 10 个不同地域的 1392 首原生态中国民歌, 进行了地域风格的分类实验. 实验结果表明: 在多种分类器的试验中, SVM 分类器的分类准确率最高; 在多种特征选择实验中使用 SVM 与 Active Feature Selection 的特征选择方法的分类准确率最高, 为 83%, 且选择出的有效特征参数从 74 维降为 35 维, 更便于进行参数分析.

关键词: 音乐信息检索; 音乐数据挖掘; 音乐风格分类; 特征选择

中图分类号: TP37, TP181 **文献标识码:** A **文章编号:** 0372-2112 (2008) 12A-152-05

The Feature Selection of Regional Style Classification of Chinese Folk Songs

LIU Yi, WEI Lei, LIU Zi-li, WANG Peng

(Information School, Renmin University of China, Beijing 100872, China)

Abstract: Automatic music style classification is an important topic in the area of music information retrieval. In this paper, we present a study on automatic classification of Chinese folk songs by regional style which mainly focuses on performance of different feature selection method. We did music style classification experiments on 1392 original Chinese folk songs which are collected from 10 different regions. The experiment results show that support vector machine classification performance has a certain advantage among different classifiers without feature selection. Simultaneously, support vector machine get the highest classification accuracy of 83% with active feature selection method, the feature vector dimensions are reduced from 74 to 35 using active feature selection. Therefore, the selected feature set is more convenient for music analysis.

Key words: music information retrieval; music data mining; music style feature selection

1 引言

1.1 背景与动机

随着计算机软硬件技术的进步,越来越多的多媒体数据应用到社会生活的各个领域. 音乐数据是多媒体数据的重要组成部分,在因特网上每天都在产生大量的数字音乐数据,对这些数据的使用包括数字图书馆、网上音乐销售、歌厅点歌服务、个人音乐欣赏、音乐研究、乐曲版权裁决等. 所有这些,都要求相应的音乐信息管理系统能够支持用户对于非格式化音乐数据的各种处理要求. 这一新的应用需求,引起了学术界和工业界的极大兴趣,吸引了包括数据库技术、数字信号处理、模式识别、机器学习等众多领域的科学家共同探讨这一新的挑战中提出的新的理论与技术问题.

音乐的自动分类与检索技术是支持用户对音乐数据处理要求的基础技术,可以直接用于基于高级语义的检索和推荐,与之相关的研究课题受到学术界的广泛重视. 由于西方歌曲中,最重要的特点是流派,因此,对音乐自动分类的研究中,大多数是对于流派分类的研究. 而中国民歌地域风格特点更加鲜明,有不少的音乐爱好者喜欢欣赏民族音乐. 在对中国民族歌曲查询要求中,通过地域风格检索也是重要检索方式. 如:“查找两首新疆民歌”,“查找三首江南小调”,等等. 研究中国民歌的地域风格分类,可以使音乐网站对中国民歌自动标识,以支持用户基于地域风格的查询要求和网站管理中的歌曲推荐要求.

另外,虽然中国民歌地方色彩之丰美,世界各国难与相比. 但至今为止,对于中国民歌的地域特点,音乐界

从理论上的研究还不够充分, 在中国的音乐学领域还是一个研究的课题. 因此, 对中国民歌地域风格自动分类方法的研究, 还可以给音乐界对民歌地域色彩分类课题的研究提供计算机化的分析手段和方法, 促进音乐界对此课题的深入研究.

本文着重讨论在中国民歌地域风格自动分类方法中不同的特征选择方法对于分类性能的影响. 文中首先讨论了中国民歌的地域风格成因和分类, 并通过实验研究了特征选择方法对于分类结果的影响. 最后给出了进一步工作的考虑.

1.2 当前研究现状

民歌的地域风格是中国民歌的独有特点, 而西方民歌中的流派特点更为突出, 因此对于音乐自动分类的研究中, 讨论比较多的是流派的分类方法^[1-3], 很少看到对于基于地域风格的分类论文.

音乐界对于不同的流派的歌曲, 在流派形成的初期, 就有人对其特点作了充分的研究, 后来创作的乐曲一般是遵循这样的特点进行的. 这种创作方式, 使得同一流派的乐曲特点相对规范、明确, 在使用计算机对其分类时, 一般只根据音乐学的研究成果, 人工选取相应的声学特征参数即可, 而无需计算机进行特征选择. 而民歌的地域风格分类就不同了, 由于音乐界对于地域风格的研究还不充分, 分类时不能在音乐理论指导下选择分类特征, 而歌曲的声学特征有几十维至上百维, 难免有无关和冗余的情况, 不仅影响分类的准确率, 也难于进行分析. 因此, 特征选择在基于地域风格的分类中具有重要的意义. 在文献[4, 5]中我们只讨论了中国民歌基于地域风格的分类中, 采样率、分片长度、特征抽取、分类器、以及分类结果后处理对于分类效果的影响, 至今为止还没专门涉及地域风格分类中特征选择的论文.

2 中国民歌的地域风格形成与类别

中国是一个地域辽阔的国家多民族的国家. 在中国, 历史上, 不同的民族是按照地域聚集居住的. 每个民族都有自己与众不同的语言和特殊风格的音乐. 汉族是中国人口最多的民族, 居住的地域也很广, 但也依居住地域, 有不同风格的民间音乐. 对于中国民歌, 按照地域模式, 可以分为 10 大类别: 新疆、内蒙古、西藏、西南、华中、东南、华南、西北、东北、华北^[6], 也可以按照各省地理位置做更细致的划分, 如: 陕北、陕南、山西、四川等.

在中国, 原生态民歌的创作者和演唱者通常都是生活在当地的农民. 原生态民歌与生活紧密相连, 散发着田野的气息和草根的馨香, 是民族、地域文化的一个组成部分, 是中国民间音乐的创作源泉, 是中华民族

“口头非物质文化遗产”的重要部分. 随着经济的发展, 交通的发达, 生活节奏的加快, 原生态民歌赖以生存和发展的外部环境正在消亡, 挽救中华民族优秀的文化遗产—原生态民歌成为文化领域一个重要任务. 为此, 科技部设立了资金达千万元的重点项目—“中国民族民间文艺基础资源拯救”. (项目编号: 2002DEB10048), 从全国录制了不同地域的原生态民歌. 对于这一类民歌, 能提供进一步声学分析是非常重要的应用要求. 为满足这一要求, 特征选择是必不可少的.

本文所采用的数据集, 就是从此项目所采集的原生态民歌中, 选取的录制质量较好的 10 个地域的 1392 首歌, 进行地域风格分类的特征选择研究.

3 分类实验的设计

为了研究特征选择对中国民歌地域风格分类的影响, 我们选取了 10 个地域的 1392 首原生态民歌, 首先做了分类器的选择实验, 选择分类准确率最高的 SVM 分类器, 分别采用 ReliefF, Fisher rule, Active Feature Selection 等特征选择方法以及不进行特征选择的方法, 进行了多组分类实验, 我们希望通过对各种特征选择方法的对比实验, 发现对民歌数据集相对有效的特征选择方法.

3.1 数据集

我们的原生态民歌数据库中共有 10 个地域的民歌 1392 首, 通过文献[4]中对采样率和采样精度和切分长度的讨论, 我们将所有曲目转换成采样率为 16kHz, 采样精度为 8 位, 单声道的 wave 格式, 然后按 10s 长度进行切分, 共得到 10226 个长度为 10s 的原生态民歌片段, 最后从每个民歌片段中提取 74 维特征参数, 得到 10226 个样本数据. 取每个地域中样本数据的 80% 组成分类器训练集数据 (8180 个样本), 取每个地域剩余的 20% 样本组成分类器测试集数据 (2046 个样本).

表 1 原生态民歌数据库		
地域类别	歌曲数目	10s 片断数
内蒙古	149	922
吉林	1291	092
陕北	116	454
青海	92	638
山西	159	1455
湖北	119	936
湖南	68	313
广西	73	336
四川	167	1953
贵州	320	2127
总计	1392	10226

原生态民歌数据库曲目及 10s 片段数目如表 1 所示.

3.2 分类器的选择

不同的分类器对于不同的数据集分类效果是不同的, 我们首先希望通过试验, 选择最适用于民歌地域模式分类的分类方法. 表 2 是使用 5 种不同的分类算法对于原生态民歌数据集进行 10 类地域风格分类的结果.

在多种分类方法中, SVM 方法可以用少量训练集取得较好的分类效果, 并且得到的解与样本的具体分

布无关,具有很好的泛化能力,在音乐地域风格分类中取得了较好的分类效果.从表2的结果中可以看到,在5种分类方法中,SVM方法获得最高的分类准确率82.4%.因此,我们选择了SVM方法进行地域风格分类的特征选择方法研究.本文中以后的实验结果均是采用SVM方法得到的结果.SVM方法最初是为二类分类设计的,以后又发展了多种SVM的多类分类算法.文献[7]中Libsvm中有多种多类分类选项,由于一对一的方法训练精度和测试精度高,训练时间也更短,我们选用Libsvm的一对一方法实现10类分类.

表2 不同的分类器的对原生态民歌的分类准确率

分类器	KNN(k= 1)	BP NN	Nä ve Bayes	Fisher	SVM
分类准确率(%)	71.90	76.00	59.80	70.10	82.40

3.3 特征参数选择算法

特征选择是模式识别中的关键技术,一般情况下,只有在特征向量中包含了足够的类别信息,才能通过分类器实现正确分类,而特征中是否包含足够的类别信息却很难确定.为了提高识别率,总是最大限度的提取特征信息,结果不仅使特征维数增大,而且其中可能存在较多的不相关和冗余特征,因而选择合适的特征来描述模式对模式分类的精度、需要训练时间和需要的实例等许多方面影响很大,并且对分类器的构造也起着非常重要的作用.

我们从每一个音乐片段中提取出74维特征^[4]来表征该片段的地域风格,为了去除74个特征中冗余以及不相关的特征,减少支持向量机分类器的训练时间,提高分类器的分类精度,发现最能反映地域风格的声学特征,我们设计使用下面三种方法进行特征选择.

(1) FishRule 方法

Fisher 准则是特征选择的有效方法之一,其主要思想是鉴别性能较强的特征表现为类内距离尽可能小,类间距离尽可能大,采用单个特征的 Fisher 比作为准则,对特征进行排序,选出哪些鉴别性能较强的特征,可以达到降维的目的并得到较优的识别性能.

(2) ReliefF 方法

ReliefF 是公认的效果较好的 Filter 式特征评估方法,其要点是根据对近距离样本的区分能力来评估特征,ReliefF 评估效率高,对数据类型没有限制,可以较好的去除无关特征.Fisher 和 ReliefF 都可以较好的去除无关特征,但不能去除冗余特征.

(3) ActiveSelection 方法

ActiveSelection 是一种 Filter 和 Wrapper 相结合的主动特征选择的方法.算法思想可简述如下:

(a) 利用 ReliefF 算法对所有 74 维特征计算权重,按照权值由大到小进行排序,选取该数列中权值最高

前 N 维特征作为初始特征子集 A (在本文的实验中我们选取了最高的前 8 维特征作为初始特征子集);

(b) 使用 SVM 分类器在初始特征训练数据集上进行训练,模型训练完成后在测试集上进行分类测试,将测试集中分类错误的样本分离出来,

(c) 对包含 74 维特征的错误样本再次使用 ReliefF 算法进行特征权值排序,选取不包含在初始特征子集 A 中,且权值最大的特征 f 加入特征子集 A 形成新的特征集合,重新使用 SVM 分类器进行训练,并在测试集上进行测试,如果分类器性能提高,则重复上述过程,直至分类器性能不再提高为止.

算法最突出的特点是,每次从分错的样本中提取最能区分地域类别的特征加入特征子集,使其比前向后向的特征选择方法能更准确、迅速地选出区分能力最好的特征集.

我们采用的 ActiveSelection 算法步骤如下:

输入:

训练集 $DS_{tr} = [ds_{tr1}, ds_{tr2}, \dots, ds_{trn}]^T$,

测试集 DS_{test}

训练集的类别标号 $L_{tr} = [l_{tr1}, l_{tr2}, \dots, l_{trn}]$,

测试集的类别标号 L_{test}

向特征子集中加入新特征时最多搜索的特征个数: M

初始化特征子集:

对训练集 DS_{tr} 中的每个特征属性使用 ReliefF 算法计算权重 W , 对每个特征按权重由大到小排序后选取权重较大的前 n 个特征作为初始特征子集 FS , 一般取第 n 个特征与第 $n+1$ 个特征的权重相差较大.

初始化分类准确率: $ACU = 0$.

初始化准确率提升大小: $ACU_{improve} = 0$.

在初始特征子集 FS 上对训练集样本 DS_{tr} 使用支持向量机分类器训练模型,使用训练所得的模型对测试集样本 DS_{test} 进行测试获得分类准确率 ACU_{test} .

$ACU_{improve} = ACU_{test}$.

while ($ACU_{improve} > 0.0$) {

$ACU = ACU_{test}$;

$ACU_{improve} = 0.0$;

取出测试集中分类错误的样本形成错误样本集 $DS_{error} = [ds_{error1}, ds_{error2}, \dots, ds_{errorn}]^T$; 对错误样本集 DS_{error} 中的所有特征属性使用 ReliefF 算法计算权重 W_{error} .

对每个特征按权重由大到小排序,

将不属于 FS 集合且权重最高的特征 f_{new} 加入特征子集 FS 形成新的特征集合.

$m = 1$;

while ($m < M$) { //

在新的特征子集 FS 上对训练集样本 DS_{tr} 使用支持向量机分类器训练模型,使用训练所得的模型对测试集样本 DS_{test} 进行测试获得分类准确率 ACU_{test} .

if ($ACU_{test} \leq ACU$) {

除去特征子集 FS 中新加入的特征 f_{new} ,

将按 W_{error} 排序后不属于 FS 集合且权重仅次于 f_{new} 的特征加

```
    入 FS 形成新的特征集合;  
    m = m + 1;  
}  
else {  
    ACUimprove = ACUbest * ACU;  
    // 准确率提高, 进入下一轮循环  
    Break;  
}  
/  
/  
/  
输出:  
    结合分类错误的样本进行主动选择后的特征子集 FS
```

由算法可以看出, 在最好情况下其时间复杂度能够达到 $O(n)$, 最坏情况下的时间复杂度和顺序搜索算法相同, 为 $O(n^2)$.

4 实验结果与分析

使用 SVM 分类器分别与各种特征选择方法进行的基于地域风格的分类实验结果如表 3 所示.

表 3 采用不同特征选择方法的分类准确率对比

特征选择方法	维 数	准 确 率
FishRule	34	74.9%
ActiveSelection	35	83.0%
不进行特征选择	74	82.4%
Relief	72	82.7%

从表中可以看到, 使用 SVM 方法 FishRule 的特征选择方法, 虽然特征的维数降低了近二分之一, 但分类的准确率很不理想, 分别只是 74.9%.

使用 SVM 方法与 ReliefF 的特征选择方法, 对民歌的地域风格分类准确率达到 82.7%. 但是 ReliefF 所选择特征参数为 72 维, 仅比全部特征降低 2 维. 不进行特征选择, 直接用 SVM 方法分类的分类准确率为 82.4%, 比先使用 ReliefF 进行特征选择再分类的方法, 只低 0.3 个百分点, 但省去了特征选择的工作. 如果只考虑分类的准确率, 例如: 在音乐检索系统中, 对乐曲进行自动分类, 这两种方法都是比较好的方法. 但是, 这两种方法所使用的特征参数达到 70 多维, 难于进行参数语义分析.

使用 SVM 与 Active Feature Selection 的特征选择方法的分类准确率最高, 为 83.0%, 而且分类所用特征参

民歌10类数据主动挖掘特征选择准确率折线图

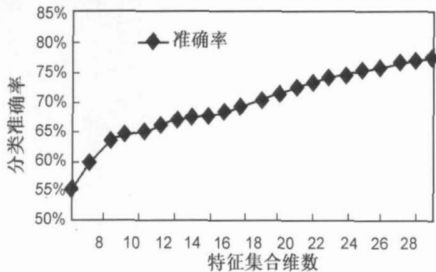


图1 采用主动特征选择方法的分类准确率

数却只有 35 维, 比前两种方法降低了 50% 多, 更易于分析参数所能表示的语义成分, 不仅适合作音乐检索也适合作音乐分析的工具.

从图 1 可以看到, 由于算法设计, 采用主动特征选择算法, 使每次新加入所选特征集的特征, 都能提高分类的准确率. 而且算法在每次从分错的样本中选取新的最具区分能力的特征时, 由于是在 ReliefF 算法进行特征权值排序后得到的序列中, 选择不包含在初始特征子集中, 且权值最大的特征, 选择时, 还可以在权值最大且相近若干相邻特征值中选取更具语义表现力的特征, 以使所选特征更便于进行语义分析.

5 结论与进一步的工作

我们在自己的音乐数据库中, 对音乐的地域风格挖掘做了初步的探讨, 并根据一系列实验的结果, 讨论了各类特征选择方法对地域风格分类的影响. 实验证明, 恰当的特征选择方法不仅可以提高民歌地域风格分类的准确率, 还可以有效特征向量的维数. 使用 Active Feature Selection 的特征选择方法不仅提高了分类准确率, 而且将分类所用特征维数降低 50% 多, 使分类结果更好的用于音乐检索与音乐分析.

今后, 我们还将进一步研究对于民歌数据集的有效冗余特征替换方法, 考虑揭示原生态民歌与创作民歌声学特征之间的联系, 为音乐分析提供更有用的工具.

参考文献:

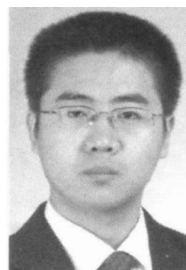
[1] Anders Meng, John Shawe Taylor. An investigation of feature model for music genre classification using the support vector classifier[A]. The 6th International Conference on Music Information Retrieval[C]. London, UK, 2005. 604– 609.
[2] Noris Mohd Norowi, Shyamala C. Doraisamy, Rahmita Wirza, O K Rahmat. Factors affecting automatic genre classification: an investigation incorporating non western musical forms[A]. The 6th International Conference on Music Information Retrieval [C]. London, UK, 2005. 13– 20.
[3] Changsheng Xu, Namunu C Maddage, Xi Shao, Fang Cao, and Qi Tian. Musical genre classification using support vector machines[A]. International Conference on Acoustics, Speech, and Signal Processing[C]. IEEE Xplore, 2003. 429– 432.
[4] 刘怡, 蔚磊, 许洁萍. 中国民歌的地域风格分类[A]. 第 15 届中国多媒体学术会议(N 论文集[C]. 杭州, 清华大学出版社. 2006. 80– 87.
[5] Yi Liu, Jie P Xu, Lei Wei, Yun Tian. The study of the classification of chinese folk songs by regional style[A]. International Conference on Semantic Computing [C]. Irvine, CA, USA. IEEE Computer Society. 2007. 657– 662.

- [6] Chinese folk song[Z] . <http://www.china.org.cn/chinese/minge/435499.htm>
- [7] Chih chung Chang, Chih jen Lin. LIBSVM: a library for support vector machines[R] . 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [8] Feng Chu, Lipo Wang. Active mining discriminative gene sets [A] . The 8th International Conference[C] . Zakopane, Poland. Springer Berlin/Heidelberg. 2006. 880– 889.
- [9] Changsheng Xu, Namunu C. Maddage, Xi Shao. Automatic music classification and summarization[J] . IEEE Transactions on Speech and Audio Processing, 2005, 13(3) : 441– 450.
- [10] Tao Li, George Tzanetakis. Factors in automatic musical genre classification of audio signals[A] . IEEE Workshop on Applications of Signal Processing to Audio and Acoustics[C] . New Paltz, NY, 2003. 143– 146.

作者简介:



刘 怡 女, 1949 年出生于北京, 中国人民大学信息学院副教授. 研究方向: 数据库技术、多媒体数据挖掘、多媒体信息智能检索.
Email: liuyilee@ruc.edu.cn



蔚 磊 男, 1983 年出生于山东东平, 中国人民大学信息学院硕士研究生, 研究方向: 多媒体数据挖掘、音乐信息智能检索.
Email: sunraising@163.com

(上接第 170 页)

- [10] Franzen K, Karlgren J. Verbosity and interface design[A] . Technical Report T2000: 04[C] . Swedish Institute of Computer Science, 2000. 61– 69.
- [11] Kosala R, Blockeel H. Web mining research: A survey[J] . ACM SIGKDD Exploration, 2000, 2(1) : 1– 15.
- [12] Ting Liu, Jinshan Ma, Huijia Zhu, Sheng Li. Dependency parsing based on dynamic local optimization[A] . Proceedings of Tenth Conference on Computational Natural Language

Learning[C] . CoNLL shared task, New York, 2006. 111– 115.

- [13] 梅家驹, 竺一鸣, 高蕴琦, 殷鸿翔. 同义词词林[M] . 上海: 上海辞书出版社, 1996.
Mei J, Zhu Y, Gao Y, Yin H. Tong Yi Ci Ci Lin[M] . Shanghai: Shanghai Lexicographical Publishing House, 1996. (in Chinese)