

# 基于 EMD 和相空间重构的钢琴和弦识别<sup>\*</sup>

刘 婷

(首都医科大学附属北京妇产医院 北京 100026)

**摘 要** 论文建立了特征参量结合分类器的识别流程。在特征参量方面,论文采用相空间重构的方法获取和弦信号的非线性特征用于分类器的构造和相关的识别实验。由于不同和弦信号对应不同的最优相空间重构参数,而在识别未知和弦以及分类器构造过程中,需要固定延迟时间和嵌入维的数值,因此,会造成特征信息丢失的问题,影响识别准确率。为了弥补上述不足,论文在信号预处理环节中引入了 EMD 自适应信号分解的方法,相关实验表明,该方法可以较为准确地识别出具体的和弦种类。

**关键词** 和弦识别; 相空间重构; EMD

**中图分类号** TN912 **DOI**:10.3969/j.issn.1672-9722.2016.03.025

## Piano Chord Recognition Based on EMD and Phase Space Reconstruction

LIU Ting

(Beijing Obstetrics and Gynecology Hospital, Capital Medical University, Beijing 100026)

**Abstract** The process on the basis of characteristic parameters and classifier is built in this thesis. Nonlinear characteristics extracted from the chord are used in constructing classifier and experiments. Although different chords have different phase space reconstruction parameters, the value of delay-time and embedding dimension has to be fixed. And part of the nonlinear characteristics will be lost which lower the recognition accuracy. For making up the disadvantage mentioned above, EMD adaptive signal decomposition methods are used in preprocessing. The result of experiments show that different type of chords can be recognized by this method.

**Key Words** chord recognition, phase space reconstruction, EMD

**Class Number** TN912

### 1 引言

随着计算机网络和多媒体技术的发展,音乐的种类和数量也在快速地增长,随之而来的是对音乐进行有效储存、检索和管理的实际需求,而和弦作为音乐的重要内容之一,也越来越多地应用到这些领域当中<sup>[1]</sup>。对于专业的音乐人士,能够通过听力来进行音乐和弦种类的识别,而在信号分析识别领域,人们希望让计算机也具备这样的能力。已有的音乐识别方法中大多基于时频分析,该方法容易受噪声影响,使识别准确率降低,因此,如何提高音乐信号特征信息的抗噪性和识别准确率一直以来都

是研究人员关注的问题<sup>[2~4]</sup>。

随着混沌理论不断发展,音乐信号的非线性特征从一个新的角度反映了信号的本质特征,因此,可以将利用相空间重构方法获得的特征矩阵应用到和弦分类器的构造和相关的识别实验中<sup>[5~6]</sup>。在提取音乐信号的非线性特征时,直接采用相空间重构的方法往往会造成部分和弦特征信息的丢失,因此,在提取和弦信号非线性特征之前,利用 EMD 自适应信号分解的方法将原始信号分解成若干子信号,然后再对子信号进行相空间重构,该方法在很大程度上能够减少固定相空间重构参数所造成的信息丢失的问题<sup>[7~9]</sup>。

\* 收稿日期:2015 年 9 月 11 日,修回日期:2015 年 10 月 22 日

作者简介:刘婷,女,硕士,工程师,研究方向:网络管理和系统维护。

## 2 和弦信号的 EMD 自适应信号分解

一般情况下,对和弦信号进行 EMD 自适应分解后,原始信号将被分解成十个以上的子信号,其中,绝大部分子信号中所含的信息量都是非常小的,并不能表征原始信号的本质特征,另外,如果对每个子信号都提取非线性特征,就会加大程序的计算量,影响 SVM 分类器构造速度和未知和弦信号的识别速度,因此,需要从分解出的所有子信号中选取包含信息量较大的子信号作为特征信号,用于分类器构造以及和弦识别实验。图 1 显示的是对大三和弦 CMajor 的 EMD 分解后前三个信号分量。

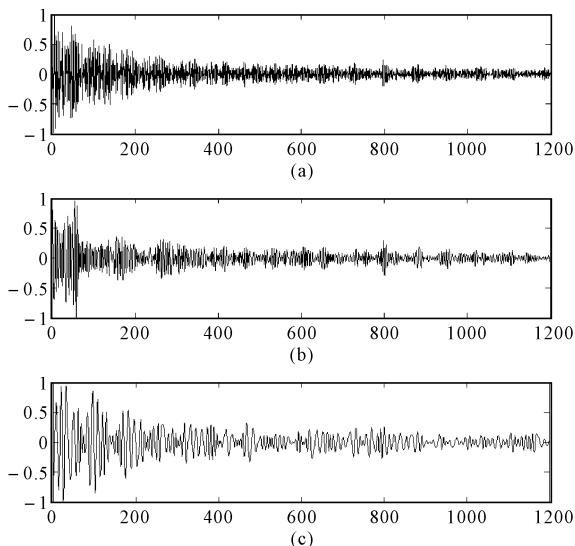


图 1 对 CMajor 进行 EMD 分解的前三个分量

其中,图 1(a)~图 1(c)为对 CMajor 进行 EMD 分解后的前三个子信号。

图 2 显示的是上述三个子信号对应的频谱图像。

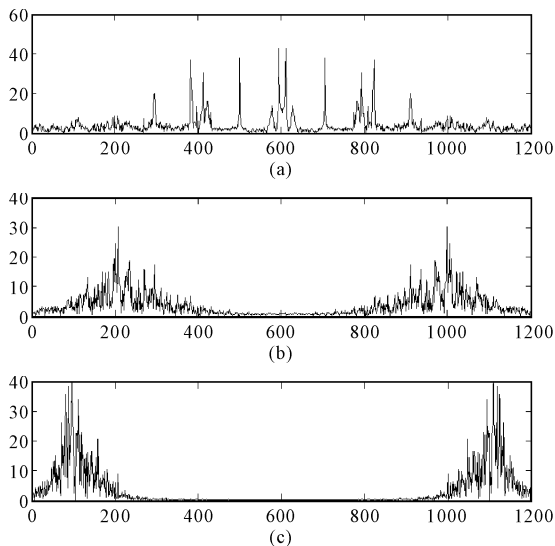


图 2 对 CMajor 的 EMD 分解后前三个分量的频谱

其中,图 2(a)~图 2(c)为前三个子信号对应的频谱。

通过对和弦信号进行 EMD 分解后的各分量图像及对应的频谱图可以看出,前三个子信号包含较多的信息量,无论在信号能量还是频谱上都能够反映出整个和弦信号的主要特征,因此,构造和弦识别分类器以及提取未知和弦特征的过程中,对分解后的前三个子信号进行相空间重构,得到相应的三个特征矩阵,然后整合这三个特征矩阵后即可提取出该和弦信号的主要特征。

## 3 基于 EMD 和相空间重构的和弦识别

利用 EMD 和相空间重构的方法进行和弦识别时,除了要决定特征子信号数量外,还要确定子信号所对应的相空间重构参数,即延迟时间  $\tau$  和嵌入维  $m$ 。

### 3.1 相空间重构参数的确定

在提取这三个子信号的非线性特征时,需要确定前三个分量的相空间重构参数,即延迟时间  $\tau$  和嵌入维  $m$  的值,下列各图显示的是对大三和弦 CMajor 进行 EMD 分解并利用互信息法和伪最近邻点的改进方法对前三个子信号进行相空间重构参数确定的情况。图 3~图 5 是对大三和弦 CMajor 进行 EMD 分解后第一分量延迟时间  $\tau$ 、嵌入维  $m$  的确定情况以及相空间重构的轨迹。

图 3 为利用互信息法求得延迟时间  $\tau=3$ 。

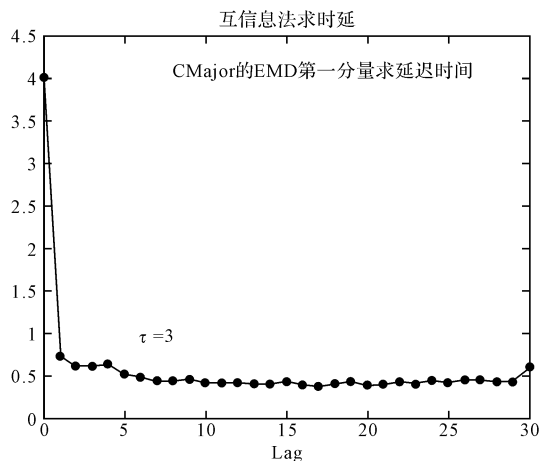


图 3 确定延迟时间

图 4 为利用 Cao 方法求得嵌入维  $m=13$ 。

图 5 为 CMajor 第一分量对应的相空间重构轨迹。

利用同样的方法可以对 CMajor 进行 EMD 分解后第二分量和第三分量确定延迟时间  $\tau$  和嵌入

维  $m$  的值。通过对若干和弦信号进行 EMD 分解, 分析每个和弦前三个 EMD 分量的延迟时间  $\tau$  和嵌入维  $m$  的情况, 确定第一分量的延迟时间  $\tau=3$ , 嵌入维  $m=10$ , 第二分量的延迟时间  $\tau=5$ , 嵌入维  $m=10$ , 第三分量的延迟时间  $\tau=6$ , 嵌入维  $m=10$ , 将这些确定后的相空间重构参数用于特征提取和分类器构造中。

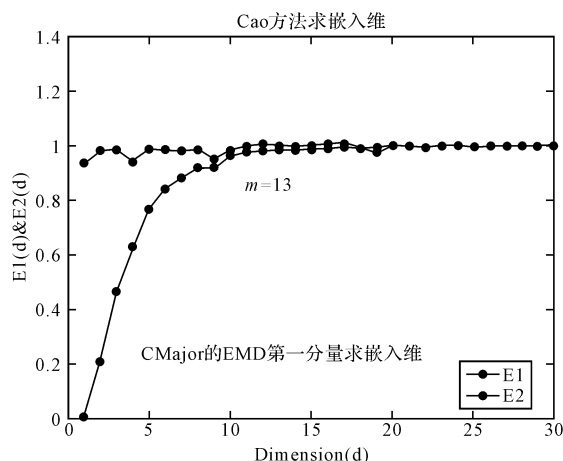


图4 确定嵌入维

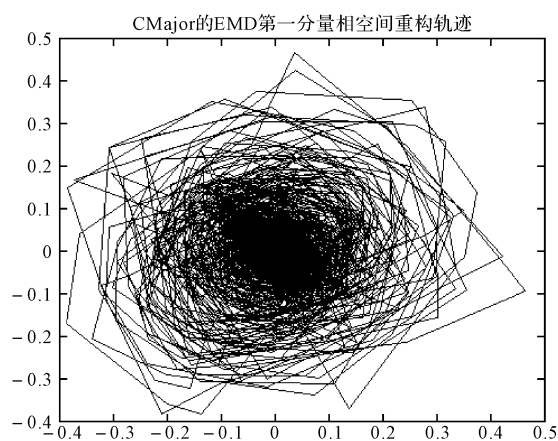


图5 相空间重构轨迹

### 3.2 识别特征量和分类器的构造

在进行 SVM 分类器构造时, 首先对获取的每个样本信号进行 EMD 自适应信号分解, 并且利用前三个子信号分量来提取出能够表征和弦本质的特征参量。按照前面确定的每个子信号参数, 对每个子信号进行相空间重构。最后, 利用重构后的相空间矩阵进行训练信号特征参量和识别特征参量的构造, 其结构如图 6 所示。



图6 基于 EMD 的特征参量矩阵

在进行 SVM 分类器的构造过程中, 除了进行

特征参量矩阵的构造之外, 还需要进行训练参量标签的标识, 根据矩阵行的数目  $N$ , 进行  $N \times 1$  阶标签矩阵的构造, 矩阵中元素设为 1、2、3……, 分别对应不同种类的和弦以及音符。

在进行 SVM 分类器的构造过程中, 除了训练信号特征参量和特征参量对应的标签矩阵以外, 还要选择核函数种类以及惩罚因子  $C$ , 在本次实验中, 核函数选取径向基核函数  $K(x, y) = \exp(-\|x - y\|^2 / 2\delta^2)$ , 其中,  $\delta=0.02$ , 惩罚因子  $C=500$ 。

### 3.3 识别实验及分析

在确定了相应参数以及特征参量的构造方法后, 对三和弦 CMajor、DMajor、AMajor、BMajor、CMinor、DMinor、AMinor 和 BMinor, 七和弦 Major 7th、Minor 7th、Dominant 7th、Diminished 7th 以及 C、E、G 和 A 四个单音音符进行种类识别, 其识别流程如图 7 所示。

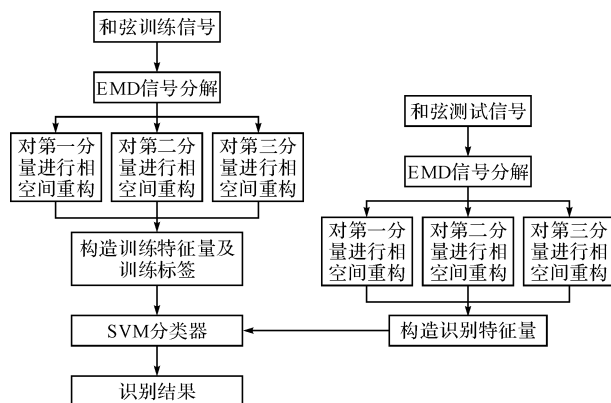


图7 基于 EMD 和相空间重构的和弦识别流程图

按照图 7 所示的流程, 对上述 16 种音乐信号进行识别实验, 表 1 显示的是对十二种和弦及四个单音进行识别实验的结果, 其中

$$\text{识别准确率} = \frac{\text{识别标签中正确标签数}}{\text{预测标签总数}}$$

通过识别结果可以看出, 从整体上, 利用 EMD 结合相空间重构的识别方法可以提取出和弦信号的非线性特征, 准确地识别出具体的和弦种类。在这个过程中, 引入自适应信号分解环节, 主要是为了解决固定相空间重构参数所引起的非线性特征信息丢失的问题。利用 EMD 自适应信号分解的方法将原始信号分解成包含若干特征参量的子信号, 然后再对每个子信号进行相空间重构, 这样, 就在很大程度上弥补了固定参数造成的信息量丢失。然而, 通过识别结果发现, 不同和弦或单音音符的识别准确率也是不相同的, 有些准确率较高, 例如 DMajor, 而有些识别准确率却相对较低, 例如单音 E。产生这样的结果主要是因为自适应 EMD 方法

在进行自适应信号分解时固有的不足造成的,分解后的各个子信号,特别是前几个子信号中包含的特征信息较多,而且发生了各种特征信息混叠的现象,这在某种程度上也影响了最终的识别准确率。

表 1 基于 EMD 和相空间重构的和弦识别准确率

和弦(音符)种类	识别准确率
C	84.43%
E	62.12%
G	90.76%
A	87.35%
CMajor	86.32%
DMajor	97.67%
AMajor	89.31%
BMajor	81.92%
CMinor	81.41%
DMinor	92.63%
AMinor	80.76%
BMinor	83.82%
Major 7th	90.13%
Minor 7th	79.51%
Dominant 7th	81.09%
Diminished 7th	85.63%
平均准确率	84.62%

## 4 结语

为了弥补直接提取和弦非线性特征的不足,本文利用自适应信号分解的方法,将和弦分解成包含若干特征子信号,再对每个子信号进行相空间重构,将得到的若干相空间重构矩阵整合在一起形成表征和弦本质的非线性特征参量。然而,根据对 EMD 分解特性的描述,当数据由白噪声组成,且白噪声的尺度均匀的分布在时间或时间-频率尺度上,EMD 可以作为一个二进滤波组。当数据不是纯的白噪声,一些尺度会丢失,这样就会出现模式混叠现象<sup>[10]</sup>。而实验中获得的音频数据中都不可能地融入了噪声,因此,会出现 EMD 模式混叠现象,造成某些和弦或单音识别准确率的下降,在今后的工作中,需要采用改进后的自适应信号分解方法,使分解后生成的子信号中包含较为单一的特征信息,从而能够更好地反映信号的本质特征。

### 参考文献

- [1] 袁莉萍. 论音乐信息网络资源的设计与开发[J]. 现代图书情报技术, 2002, 3: 61-62, 73.  
YUAN Liping. Exploitation and Construction of Musi-

- cal Information in Internet[J]. New Technology of Library and Information Service, 2002, 3: 61-62, 73.
- [2] Humphrey, E. J., Taemin Cho, Bello, J. P. Learning a robust Tonnetz-space transform for automatic chord recognition[C]//Proceedings of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, 25-30 March 2012, 2012: 453-456.
- [3] Oudre, L., Grenier, Y., Févotte, C. Chord Recognition by Fitting Rescaled Chroma Vectors to Chord Templates[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2011, 19(7): 2222-2233.
- [4] Oudre, L., Févotte, C., Grenier, Y. Probabilistic Template-Based Chord Recognition[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2011, 19(8): 2249-2259.
- [5] Yang, Hui. Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vectorcardiogram Signals[J]. IEEE Transactions on Biomedical Engineering, 2011, 58(2): 339-347.
- [6] 刘树勇, 朱石坚, 俞翔. 确定相空间重构嵌入维数的研究[J]. 哈尔滨工程大学学报, 2008, 29(4): 374-381.  
LIU Shuyong, ZHU Shijian, YU Xiang. Determinating the embedding dimension in phase space reconstruction[J]. Journal of Harbin Engineering University, 2008, 29(4): 374-381.
- [7] Rainer Typke, Agatha Walczak-Typke. Indexing Techniques for Non-metric Music Dissimilarity Measures[J]. Computational Intelligence, 2010, 274: 3-17.
- [8] 陈仁祥, 汤宝平, 马婧华. 基于 EEMD 的振动信号自适应降噪方法[J]. 振动与冲击, 2012, 31(15): 82-86.  
CHEN Renxiang, TANG Baoping, MA Jinghua. Adaptive de-noising method based on ensemble empirical mode decomposition for vibration signal[J]. Journal of Vibration and Shock, 2012, 31(15): 82-86.
- [9] 王彪. 基于 EMD 法的语音信号特征提取方法研究[J]. 科学技术与工程, 2012, 20(10): 2462-2464.  
WANG Biao. Speech Signal Feature Extraction Method Research Based on EMD[J]. Science Technology and Engineering, 2012, 20(10): 2462-2464.
- [10] 胡劲松, 杨世锡, 任达千. 一种基于 EMD 的振动信号时频分析新方法研究[J]. 振动与冲击, 2008, 27(8): 71-73, 154.  
HU Jinsong, YANG Shixi, REN Daqian. Study on the Method of EMD-Based Vibration Signal Time-Frequency Analysis[J]. Journal of Vibration and Shock, 2008, 27(8): 71-73, 154.