

峰值频率大相关-小残差的弦乐器校音方法*

唐骏^{1,2}, 刘虹¹

(1. 厦门理工学院 通信工程系, 福建 厦门 361024;

2. 厦门大学 信息科学与技术学院, 福建 厦门 361004)

摘要: 针对弦乐器校音中如何确定音的基频, 提出了一种时频域相结合的峰值频率大相关、小残差(Peak-Frequency Maximum-Correlation Minimum-Residual, PFMCMR)的弦乐器校音新方法。先对音频进行频谱分析, 并利用频谱细化算法计算峰值频率, 时域中相关局部最大与残差局部最小相结合, 确定峰值频率对应的谐波次数, 进而计算出基频。实验结果表明, 该算法精度可达 0.1 Hz 或更高, 其精度较高, 稳定性良好。

关键词: 弦乐器; 校音; 基频; 频谱细化

中图分类号: TN912

文献标识码: A

文章编号: 0258-7998(2014)12-0140-03

Peak-frequency maximum-correlation minimum-residual tuning algorithm

Tang Jun^{1,2}, Liu Hong¹

(1. Department of Communication Engineering, Xiamen University of Technology, Xiamen 361024, China;

2. School of Information Science and Technology, Xiamen University, Xiamen 361004, China)

Abstract: The basic issue of tuning stringed instrument is to determine the fundamental frequency of the sound. A novel method of stringed instruments tuning in time-frequency domain, named PFMCMR (Peak-Frequency Maximum-Correlation Minimum-Residual) is presented. Firstly, spectrum zoom algorithm is used to calculate the peak frequency. Secondly, in time domain, local maximum correlation and local minimum residual are combined to determine the harmonic number of peak frequency. Finally, the fundamental frequency is calculated based on peak frequency and corresponding harmonic number. The algorithm accuracy is up to 0.1 Hz or higher. The experimental result show that PFMCMR method is high accuracy and good stability.

Key words: stringed instrument; tuning; fundamental frequency; spectrum zoom

0 引言

随着多媒体技术的发展及网络的普及, 人们欣赏音乐、获取音乐资源变得十分便捷, 而各种乐器也快速走进音乐爱好者的家庭, 如钢琴、吉它等。

弦乐器因其物理特性, 需要进行定期校正^[1-2], 以往这项工作一直由经过专业训练的调音师来承担。即便是专业的调音师, 也会受生理、心理以及客观环境的影响, 从而出现对音准判断的偏差; 另外, 随着乐器快速普及, 少量的专业调音师难以满足广大的需求, 乐器校音成为难题^[3]。因此, 迫切需要一种仪器, 可以完全排除调音过程中的主观因素, 能够客观准确地校准乐器。

1 基音检测原理与常用算法

校音的根本目的就是精确确定乐音的基频。信号基频的检测根据处理域或方法的不同, 主要可分为时域方法和频域方法。

1.1 时域方法

将信号看作时间的函数, 其波形反映依时间变化的特性, 通过观察信号波形, 确定其基本周期, 从而获得基频^[4]。最常用的有自相关函数法。

自相关函数是信号自身的相关函数, 可以度量信号自身的相似性。对于无限长的离散信号 $x[n]$, 自相关函数的定义为:

$$R_{xx}[m] = \sum_{n=-\infty}^{+\infty} x[n]x[n+m] \quad (1)$$

式中 m 为信号的延迟。

《电子技术应用》2014年 第40卷 第12期

* 基金项目: 福建省教育厅 A 类项目(JA13235); 国家自然科学基金项目(61202013)

对于长度为 N 的离散信号 $x[n]$, 自相关函数的定义为:

$$R_{xx}[m] = \sum_{n=0}^{N-m-1} x[n]x[n+m] \quad (2)$$

如果序列 $x[n]$ 是周期的, 则其自相关函数也是周期的, 且周期相等。

1.2 频域方法

频域有更多的与基频相关的信息。具有基频的信号往往是由频率具有谐波关系的信号组成, 因此有很多利用频域信息提取基频的方法^[5-6]。

谐波峰值法。谐波峰值法是基于离散傅里叶变换(DFT)的分析法, 将信号通过 FFT 变换得到离散的频谱, 确定峰值频率。对于周期信号而言, 峰值频率是基波频率的整数倍, 如果能确定峰值频率对应的谐波次数, 便可以求出基波频率。

离散小波变换法。离散小波变换允许在连续的尺度上将信号分析为高频成分和低频成分, 它是时间和频率的局部变换, 能有效地从信号中提取信息。

2 PFMCMR 算法流程

PFMCMR 算法框图如图 1 所示。输入音频信号先进行端点检测^[7-8], 目的是去除噪声和静音段, 以便仅对有用信号段进行分析, 有用信号段加窗后进行傅里叶分析。设采样频率为 F_s , 窗长为 N , 则 FFT 的谱线间隔为 $\Delta f = F_s/N$, 一般这样的精度离实际要求相差甚远, 因此需要通过频率细化算法, 以求得精度更高的峰值频率。细化的频率间隔可由具体算法的参数控制, 根据实际需要精确到约千分之几赫兹即可。

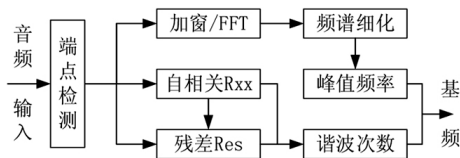


图 1 PFMCMR 算法框图

由前述可知, 利用信号的自相关函数可以确定其基本周期, 但由于实际信号的衰减、起伏及受噪声影响等诸多因素, 相关函数的峰值未必出现在基波周期处, 更多时候出现在基波周期整数倍附近^[9]。因此, 直接将相关函数峰值的位置确定为基本周期既不可靠, 精度也无法满足实际需求。尽管如此, 自相关函数为确定峰值频率对应的谐波次数提供了有用的信息。

自相关反映的是信号结构的某种自相似程度, 弦乐器音色丰富, 音的结构多变, 所以, 仅依据自相关函数来确定基波周期往往会导致倍频或分频。为了更准确地确定峰值频率对应的谐波次数, PFMCMR 算法结合信号移位、相减后的残差幅度值。自相关函数为计算残差时的移位量提供了有用信息。

自相关函数的局部最大值和残差幅度局部最小值相结合, 可以有效地确定峰值频率对应的谐波次数, 从

而最终计算出基频。

3 PFMCMR 算法的具体实现

为了方便定量叙述, 取定一些参数如下: 采样频率 $F_s = 44\ 100$ Hz, 窗长 $N = 4\ 096$ 点。FFT 谱线间隔为 $\Delta f = F_s/N = 44\ 100/4\ 096 \approx 10.77$ Hz。为了得到高精度的峰值频率, 可以采用频率局部细化算法, 如 ZoomFFT 或 CZT^[10], 本文采用后者。PFMCMR 算法的具体实现步骤如下:

(1) 先计算 FFT 谱线幅度的最大值, 其对应的索引记为 idxMaxFFT , 从而可以确定峰值频率一定在 $\text{idxMaxFFT}-1 \sim \text{idxMaxFFT}+1$ 对应的频率范围内;

(2) 取 $\text{idxMaxFFT}-1 \sim \text{idxMaxFFT}+1$ 频率范围进行频率细化, 细化的倍数由 CZT 计算的频率间隔数 M 决定, 一般取 $M=N$, 因此, 细化后的频率间隔为 $\Delta f_M = 2F_s/N^2 \approx 0.005$ Hz;

(3) 计算细化后的峰值频率 F_p , 与之对应的周期记为 T_p ;

(4) 计算一帧信号的自相关函数 $R_{xx}[m]$, 并求其峰值 R_{max} , 对应的索引 m 记为 idxMaxR , 根据实际情况, 在搜索相关函数最大值时, 需要排除 $R_{xx}[0]$ 附近的数点;

(5) 计算 $R_{xx}[kT_p]$, 其中 k 为正整数, 且 $kT_p \leq \text{idxMaxR}$, 如果 $R_{xx}[kT_p]/R_{\text{max}} > \text{THR}$, 则 kT_p 作为基波周期的候选值, 记为 $k_i T_p$, i 为整数, 其最大值为基波周期的候选值的数目, 其中 THR 为一阈值, 本文取 0.85;

(6) 以 $k_i T_p$ 为延迟点数, 计算残差幅度的平均值:

$$\text{Res}[k_i T_p] = \frac{\sum_{n=N_1}^{N_2} |x[n] - x[n - k_i T_p]|}{N_2 - N_1 + 1} \quad (3)$$

其中 N_1, N_2 为非负整数。

因为峰值频率一定是基波的整数倍, 所以基波周期 T_0 一定是 $k_i T_p$ 中的某一个值, $R_{xx}[k_i T_p]$ 越大、 $\text{Res}[k_i T_p]$ 越小, 则 $k_i T_p$ 为 T_0 的可能性越大。所以引入参数:

$$\text{RR}[i] = \frac{R_{xx}[k_i T_p]}{R_{\text{max}}} \times \{1 - \text{Res}[k_i T_p]\} \quad (4)$$

$\text{RR}[i]$ 越大, $k_i T_p$ 为 T_0 的可能性越大, 但由于实际信号的起伏多变, 直接利用 $\text{RR}[i]$ 作为参考值会造成一定概率的误判, 需要作一些置信修正, 即 $\text{RRM}[i] = \text{RR}[i] \cdot T[i]$, 其中 $T[i]$ 为置信因子, 随着 i 的增大而减小。最后求 $\text{RRM}[i]$ 最大值, 其对应的 i 记为 idxMaxRRM 即为峰值频率对应的谐波次数, 所以, 基频 $F_0 = F_p / \text{idxMaxRRM}$ 。

4 测试与性能分析

校音软件两个最重要的指标是精确度和准确度, 为了测试精确度, 采用单音正弦信号和定音器生成的标准音作为测试对象。用吉他和钢琴音进行整体性能测试。正弦单音由软件生成, 定音器的音、吉它音和钢琴音采用现场录制的方式。定音器 A3、吉它 G3、钢琴 B3 音的时域波分别如图 2(a)、2(b)和 2(c)所示。

表 1、表 2 分别为单音正弦和定音器的测试结果。可以看出, 对于单音正弦信号, 实测精度和理论分析相符

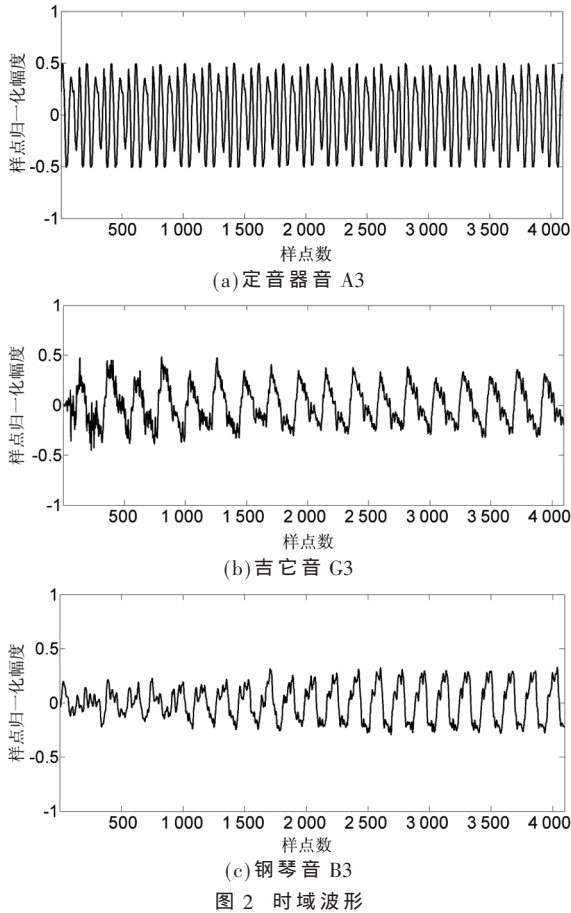


图2 时域波形

表1 单音正弦测试结果

帧号	频率/Hz		
	220	440	880
1	220.021 1	440.006 0	880.001 5
2	220.032 1	440.007 0	880.001 5
3	220.030 2	440.006 0	880.001 5
4	219.902 1	440.006 0	880.001 5
5	220.018 5	439.995 0	880.001 5
平均值	220.000 8	440.004 0	880.001 5
误差	0.000 8	0.004 0	0.001 5

表2 定音器测试结果

帧号	音名/(频率/Hz)		
	A3/220	A4/440	A5/880
1	219.981 1	439.974 4	879.990 9
2	219.981 1	439.963 9	880.012 0
3	219.996 9	439.995 5	879.969 9
4	220.002 1	440.006 0	879.990 9
5	219.988 1	439.963 9	879.954 1
平均值	219.989 9	439.980 7	879.983 6
误差	-0.010 1	-0.019 3	-0.016 4

合,误差均在 0.005 Hz 以内。相比而言,定音器的测试误差要略大,这是因为音频信号在采集过程中会发生失真。即便如此,精度仍在 0.1 Hz 以内。

表3、表4分别为吉它和钢琴3个音的测试结果。可以看出,每个音的误差和标准差均未超过 0.2 Hz。

表3 吉它音测试结果

帧号	音名/(频率/Hz)		
	E2/82.406 9	G3/195.997 7	B3/246.941 7
1	82.512 4	195.995 7	247.029 2
2	82.526 4	195.901 1	247.029 2
3	82.476 0	196.027 3	247.005 6
4	82.266 2	196.079 9	246.616 5
5	82.121 6	195.974 7	246.679 6
平均值	82.380 5	195.995 7	246.872 0
误差	-0.026 4	-0.002 0	-0.069 7
标准差	0.178 8	0.066 1	0.205 9

表4 钢琴音测试结果

帧号	音名/(频率/Hz)		
	E2/82.406 9	G3/195.997 7	B3/246.941 7
1	82.493 9	195.961 2	246.692 4
2	82.509 7	196.070 9	246.834 2
3	82.467 6	195.881 6	246.750 0
4	82.380 0	195.949 5	246.960 4
5	82.285 4	196.028 7	247.028 6
平均值	82.427 3	195.978 4	246.853 1
误差	0.020 4	-0.019 3	-0.088 58
标准差	0.093 8	0.073 5	0.113 1

综合表1~表4的测试结果,表明PFMCMR算法具有高精度和良好的稳定性,完全满足实际应用的要求。

5 结束语

本文针对弦乐器校音的基本问题,介绍了乐音基频检测的常用算法,并分析了这些算法的局限性。为了解决基频检测的精度和稳定性问题,本文提出了一种基于弦乐器音频信号时频特性的校音新算法。利用频谱细化方法得到高精度的谐波峰值频率,时域中结合信号的相关函数和信号残差,基于相关局部最大和残差局部最小原则确定基波周期,从而计算出精确的基波频率。测试结果表明,该方法精度高,稳定性好,满足实际应用。

参考文献

- [1] FLETCHER N H, ROSSING T D. The physics of musical instruments[M]. Springer, 1998.
- [2] SETHARES W A. Tuning, timbre, spectrum, scale[M]. London: Springer, 2005.
- [3] 潘阳. 钢琴校音计算机软件的设计[D]. 合肥: 安徽大学, 2011.
- [4] AMADO R G. Pitch detection algorithms based on zero-cross rate and autocorrelation function for musical notes[C]. Audio, Language and Image Processing, 2008, ICALIP 2008, International Conference on, IEEE, 2008: 449-454.
- [5] BENETOS E, DIXON S. Joint multi-pitch detection using harmonic envelope estimation for polyphonic music transcription[J]. Selected Topics in Signal Processing, IEEE Journal of, 2011, 5(6): 1111-1123.
- [6] RIGAUD F, DAVID B, DAUDET L. A parametric model

(下转第 146 页)

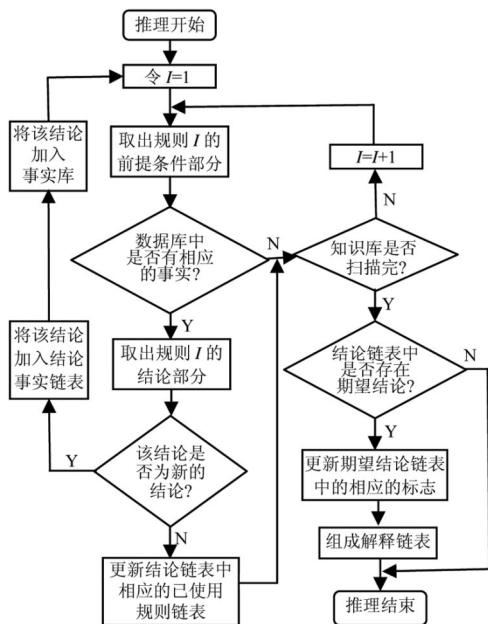


图 5 安控推理流程图

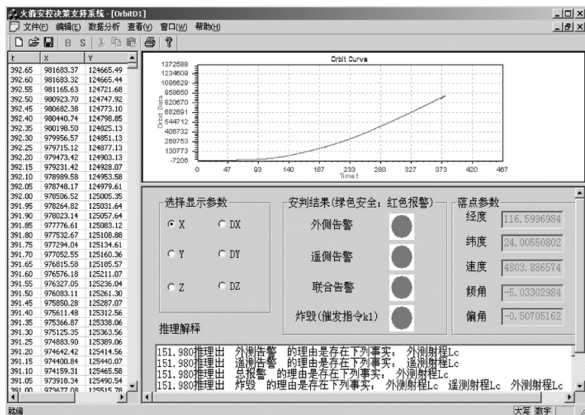


图 6 系统仿真运行界面

4 结论

本文针对现有安控系统存在的缺陷和不足,结合数据库技术和人工智能中的推理技术,建立和完善了火箭飞行安全控制决策系统,它是智能决策支持系统在理论和实践上的一个新突破。实验和仿真表明,该系统很大程度上提高了火箭安控系统的准确性和可靠性,能对其进行有效的管理。

(上接第 142 页)

and estimation techniques for the inharmonicity and tuning of the piano[J].The Journal of the Acoustical Society of America, 2013, 133(5): 3107-3118.

[7] KHOA P C.Noise robust voice activity detection[D].Singapore: Nanyang Technological University, 2012.
 [8] MA Y, NISHIHARA A.Efficient voice activity detection algorithm using long-term spectral flatness measure[J]. EURASIP Journal on Audio, Speech, and Music Processing, 2013, 2013(1): 1-18.
 [9] BROWN J C, VAUGHN K V.Pitch center of stringed

参考文献

[1] 张庆振,李清东,任章.基于故障模式分析的运载火箭发射决策系统推理技术研究[J].航天控制,2006,24(3): 81-83.
 [2] Probabilistic Risk Assessment of the Space Shuttle.A study of the potential of losing the vehicle during nominal operation[R].(NASA-CR-197808).New York: Science Application International Corporation, Advanced Technology Division, 28 February, 1995.
 [3] 李娅,张建勋,杨欣.不确定推理模型的火箭安全决策系统[J].重庆大学学报(自然科学版),2010,33(10): 118-122.
 [4] Zhang Xiaokun, Gong Yingkui, Qu Jianghua.An optimal selection model of the satellite lurk orbit[C].Applied Mechanics and Materials, 2013, 249-250: 270-273.
 [5] 蔡自兴,徐光佑.人工智能及其应用(第二版)[M].北京:清华大学出版社,1996.
 [6] 黄洪,任卫红,余达太,等.基于故障树的等级测评专家系统模型研究[J].计算机应用研究,2010,27(1): 204-208.
 [7] Liao Ruijin, Zheng Hanbo, STANISLAW G.An integrated decision-making model for condition assessment of power transformers using fuzzy approach and evidential reasoning[J].IEEE Transactions on Power Delivery, 2011, 26(2): 1111-1118.
 [8] Wang Jianqiang, Nie Rongrong, Zhang Hongyu.Intuitionistic fuzzy multi-criteria decision-making method based on evidential reasoning[J].Applied Soft Computing Journal, 2013, 13(4): 1823-1831.
 [9] ONG D C C, KHADDAJ S, BASHROUSH R.Logical reasoning and decision making[J].Proceedings of 2011, 10th IEEE International Conference on Cybernetic Intelligent Systems, CIS 2011, 2011: 26-31.

(收稿日期:2014-09-14)

作者简介:

李娅(1979-),通信作者,女,硕士,讲师,主要研究方向:专家系统、算法控制, E-mail: ly_ppx@cqut.edu.cn。

张建勋(1971-),男,博士,教授,主要研究方向:图形图像、算法控制。

肖朝晖(1970-),男,硕士,副教授,主要研究方向:嵌入式系统。

instrument vibrato tones[J].The Journal of the Acoustical Society of America, 1996, 100(3): 1728-1735.

[10] MARTIN G.Chirp Z-transform spectral zoom optimization with MATLAB[Z].Sandia National Laboratories Report SAND2005-7084, 2005.

(收稿日期:2014-09-26)

作者简介:

唐骏(1977-),男,博士研究生,讲师,主要研究方向:多媒体信号处理。

刘虹(1978-),女,硕士,讲师,主要研究方向:嵌入式系统设计。

《电子技术应用》2014年第40卷第12期