

基于 Goertzel 算法的音阶频率分析

丁志中^{1,2}, 夏 杨¹, 杨 萍¹

(1. 合肥工业大学, 计算机与信息学院, 合肥 230009; 2. 中国科技大学, 电子工程与信息科学系, 合肥 230027)

摘 要: 音乐信号音阶频率分析的基本方法是快速傅里叶变换, 但其谱线频率与定义的音阶频率在位置上有误差。对此, 常数 Q 变换可以提供更好的性能, 但其效率并不令人十分满意。本文基于 Goertzel 算法, 给出一种音阶频率分析的方法。它具有简单有效、且容易编程的特点。

关键词: 信息处理技术; 音乐信号分析; Goertzel 算法; 离散傅里叶变换; 快速傅里叶变换
中图分类号: TN912.3 **文献标识码:** A **文章编号:** 1672-2892 (2006)01-0026-04

Analysis of Music Pitch Based on Goertzel Algorithm

DING Zhi-zhong^{1,2}, XIA Yang¹, YANG Ping¹

(1. School of Computer and Information, Hefei University of Technology, Hefei 230009, China

2. Dept. of EELS, University of Science and Technology of China, Hefei 230027, China)

Abstract: Analysis of music pitch can be performed by Fast Fourier Transform. However, some of its spectral line frequencies are different from the defined frequencies of music notes. Constant Q transform can present better performance for music signals. Unfortunately, the efficiency of the algorithm is not very satisfying. A method based on Goertzel algorithm for analysis of music note frequencies is presented in this paper. It is relatively simple and easy to be programmed, and it is efficient as well.

Key words: information processing technology; music signal processing; Goertzel algorithm; DFT; FFT

1 引言

如果采用快速傅里叶变换 (Fast Fourier Transform, FFT) 分析音乐信号的音阶频率, 由于其谱线所表示的频率点是等间隔分布的, 与音乐信号音阶频率的分布规律不一致, 因此 FFT 在某些音阶频率上会产生较大的估计误差^[1]。对此问题, J C Brown 给出了谱线频率分布与音阶频率分布相一致的常数 Q 变换 (Constant Q Transform, CQT)^[2]。然而 CQT 没有非常有效的快速算法。文献[3]给出了一种利用 FFT 的 CQT 快速算法, 但是在此算法中必须固定窗函数的长度。事实上, 这样的处理部分地损失了 CQT 的特点。

在很多应用中, 用于分析的数据长度通常较大 (如几千点数据), 而感兴趣的频率点数只有很少几个。在这种情形下, Goertzel 算法^[4-5]比 FFT 更为有效。如同 FFT 一样, 在 Matlab 中已有 Goertzel 算法实现的函数供使用。因此, 利用 Goertzel 算法分析音乐信号的音阶频率, 在频率估计的准确性、算法速度、以及通用性和易用性等方面, 可以获得很好的综合性能。

2 Goertzel 算法原理

设序列 $x(n)$, $n \in [0, N-1]$, 其离散傅里叶变换 (Discrete Fourier Test, DFT) 为:

$$X(k) = \sum_{l=0}^{N-1} x(l) e^{-j\frac{2\pi}{N}kl} = \sum_{l=0}^{N-1} x(l) W_N^{kl}$$

式中: $W_N = e^{-j\frac{2\pi}{N}}$ 。考虑到 $W_N^{-kN} = e^{j\frac{2\pi}{N}kN} = e^{j2\pi k} = 1$, 得:

$$X(k) = W_N^{-kN} \sum_{l=0}^{N-1} x(l) W_N^{kl} = \sum_{l=0}^{N-1} x(l) W_N^{-k(N-l)} \quad (1)$$

由于 $x(n)|_{n=N} = x(N) = 0$, 所以上式可写为:

$$X(k) = \sum_{l=0}^N x(l) W_N^{-k(N-l)} \quad (2)$$

式(2)的求和具有卷积的形式, 因此可将 $X(k)$ 视为序列 $x(n)$ 激励某一系统后的输出。如定义 $h_k(n)$ 如下:

$$h_k(n) = \begin{cases} e^{j\frac{2\pi}{N}kn} & n \geq 0 \\ 0 & n < 0 \end{cases} = \begin{cases} W_N^{-kn} & n \geq 0 \\ 0 & n < 0 \end{cases} \quad (3)$$

则 $h_k(n)$ 和 $x(n)$ 的卷积和 $y_k(n)$ 为:

$$y_k(n) = x(n) * h_k(n) = \sum_{l=-\infty}^{\infty} x(l)h_k(n-l) = \sum_{l=0}^n x(l)W_N^{-k(n-l)} \quad (4)$$

z 变换表示为:

$$Y_k(z) = X(z)H_k(z) \quad (5)$$

$$\text{式中: } H_k(z) = \sum_{n=0}^{\infty} h_k(n)z^{-n} = \sum_{n=0}^{\infty} W_N^{-kn}z^{-n} = \sum_{n=0}^{\infty} (W_N^{-k}z^{-1})^n, \text{ 即: } H_k(z) = \frac{1}{1-W_N^{-k}z^{-1}} \quad (6)$$

$H_k(z)$ 是一个一阶递归系统。 $y_k(n)$ 是当 $x(n)$ 激励该系统时的输出。另一方面, 比较式(2)和式(4)可以看出, 对于给定的 k 值有:

$$X(k) = y_k(n)|_{n=N} \quad (7)$$

也就是说在给定 k 值下, 式(6)系统在 $n=N$ 时刻的输出即为序列 $x(n)$ 的第 k 个 DFT 样值 $X(k)$ 。该系统可用差分方程表示如下:

$$y_k(n) = x(n) + W_N^{-k}y_k(n-1) \quad N \geq n \geq 0 \quad (8)$$

式中: $y_k(-1)=0$, $x(N)=0$ 。上式经过 N 步迭代可以算得 $y_k(N)=X(k)$ 。若考虑到信号可能为复数的一般情形, 不难推知计算每点 $X(k)$ 共需 $4N$ 次实数乘和 $4N$ 次实数加。但只要将 $H_k(z)$ 稍加变化, 就可得到改进的算法。式(6)两边同乘 $(1-W_N^{-k}z^{-1})$ 得:

$$H_k(z) = \frac{1-W_N^{-k}z^{-1}}{(1-W_N^{-k}z^{-1})(1-W_N^{-k}z^{-1})} = \frac{1-W_N^{-k}z^{-1}}{1-2z^{-1}\cos\frac{2\pi k}{N}+z^{-2}} \quad (9)$$

上式是一个二阶系统, 可用下列差分方程表示:

$$v_k(n) = x(n) + 2\cos\frac{2\pi k}{N}v_k(n-1) - v_k(n-2) \quad (10)$$

$$y_k(n) = v_k(n) - W_N^{-k}v_k(n-1) \quad (11)$$

由于 $X(k) = y_k(N) = v_k(N) - W_N^{-k}v_k(N-1)$, 计算 $X(k)$ 时, 并不需要每步都计算式(11), 只需计算最后一步, 即迭代计算 $X(k)$ 的方程为:

$$v_k(n) = x(n) + 2\cos\frac{2\pi k}{N}v_k(n-1) - v_k(n-2) \quad N \geq n \geq 0 \quad (12)$$

$$X(k) = y_k(N) = v_k(N) - W_N^{-k}v_k(N-1) \quad (13)$$

按上两式计算一点 $X(k)$ 需要 $2(N+2)$ 次实数乘, $4(N+1)$ 次实数加, 即乘法计算量降低了近一半。

3 音阶频率分析

3.1 数据窗长度的选取

为了使谱线频率与音阶频率准确对应, 数据窗的长度应随着音阶频率而变化。窗长度选取主要考虑两个因素: 分辨率和虚假频率的抑制。设第 i 个音符的音阶频率为 f_i , 采样频率为 f_s , 则相应的最小数据窗长度为:

$$N_i = \lceil f_s / (f_i - f_{i-1}) \rceil = Q \cdot f_s / f_i \quad (14)$$

式中: $\lceil x \rceil$ 为不小于 x 的最小整数, $Q = f_i / (f_i - f_{i-1}) = f_j / (f_j - f_{j-1})$ ($i \neq j$), 即对于每个音阶频率, Q 为一常数, 这是因为音阶频率的分布具有常 Q 特性。

如果要检测信号中是否含有某个特定的音阶频率, 对于不同的音阶频率, 可根据式(14)选取不同的窗长度, 这样 $X(k)$ 的计算量达到最小。如果要分析音乐信号的旋律, 则需对一帧数据分析若干个八度音程(如简谱表示的 2 和 $\dot{2}$ 之间是一个八度音程)内的音阶频谱。此时若对于每一个音阶频率, 按式(14)调整数据窗长度, 调用一次 Goertzel 算法, 则算法的效率会降低。考虑到具有倍频关系的频率点可以使用相同的数据窗长度, 因此可以只对低音区第一个八度内的音阶调整窗长度, 记为 N_1-N_{12} , 而对其他音阶频率, 均在 N_1-N_{12} 中取一个对应的窗长度。这样, 不仅提高了算法的效率, 也有利于改善高音区音阶频率估计的鲁棒性。

3.2 谱线位置 k 值

利用 Goertzel 算法分析音阶频谱的另一个关键问题是正确地确定音阶频率所对应的谱线位置, 即 $X(k)$ 中 k 的取值。Goertzel 算法的本质还是计算 DFT, 其谱线位置对应的频率是线性分布的, 而音阶频率是指数分布的。另一方面, 算法中将有倍频关系的音阶 (例如简谱表示的 $\dot{2} \quad \dot{2} \quad \dot{2} \quad \dot{2} \dots$) 作为一组, 调用一次 Goertzel 算法, 各次调用的结果还需组合, 以得到整个的音阶频谱。因此音阶频率与各次调用中谱线位置的对应关系就相对复杂一些。

如果数据窗长度为 N_i , 采样频率为 f_s , 则谱线的频率间隔为 f_s / N_i [6], 则频率 f_i 位于 $f_i / (f_s / N_i)$ 处, 考虑到直流分量为第 1 条谱线, 因此 f_i 对应的谱线位置 k_i 为:

$$k_i = \left[\frac{f_i \cdot N_i}{f_s} \right] + 1 \quad (15)$$

其中 $[\cdot]$ 为四舍五入的取整运算。假设音阶 $\dot{2}$ 的频率为 f_2 , 对应的数据窗长度为 N_2 , 则分析音符 $\dot{2} \quad \dot{2} \quad \dot{2} \quad \dot{2} \dots$ 时, 对应的谱线位置分别为 $[f_2 N_2 / f_s] + 1, [2f_2 \cdot N_2 / f_s] + 1, [4f_2 \cdot N_2 / f_s] + 1 \dots$ 。由式 (14) 知, $f_i N_i$ 为常数, 当分析下一个音符组 $\dot{3} \quad \dot{3} \quad \dot{3} \quad \dot{3} \dots$ 时, 谱线位置 k 的取值基本不变。

3.3 算法步骤

如果是在 Matlab 下分析音阶频率, 算法的基本步骤如下:

- 设定待分析音乐信号的音阶频率范围;
- 按照式 (14) 确定第一个八度内各音阶频率对应的窗长度;
- 按照式 (15) 确定音阶频率位置 k ;
- 调用 Goertzel 函数计算各组音阶频率在指定谱线位置处的 $X(k)$;
- 将各次 Goertzel 函数的调用结果重新组合, 形成被分析数据段的音阶频谱。

如果用其他语言实现 Goertzel 算法, 根据式 (12)、(13) 编制相应的代码即可。事实上, 由于 Goertzel 算法的基本特点及其简单性, 在需用其他编程语言进行音乐信号分析时, 本文的算法才更显示出其易于实现的优点。

4 实验

用 Matlab 同时产生 12 个等幅正弦信号, 仿真同时演奏的一个八度内 12 个音符的音乐信号。图 1 和图 2 是分别直接用 FFT 和本文算法的分析结果, 由图可以清楚看到本文基于 Goertzel 算法的优越性。该算法对虚假频率有很强的抑制, 且幅度均匀, 而用 FFT 进行分析时, 在音阶频率的附近有很密集的虚假频率。

图 3 和图 4 是对铜管作音阶演奏的实际音乐信号进行分析的结果。虽然从视觉上看, 两图中均能反映出“旋律”的变化, 但是在 DFT 谱中, 每条“黑影线”均是由很多条谱线构成的, 这给音阶频率分析带来了较大的困难和计算量, 而在图 4 中的“黑影线”仅由 1-2 条谱线构成, 音阶频率分析会方便得多。另外, 需要解释的是, 图 3 中的纵坐标频率是线性分割的, 而音阶频率是指数分布的, 因此图 3 中的“旋律”轨迹近似指数曲线。如果数据窗的长度是 5351 点, 其纵轴方向将有 5351 条谱线。图 4 纵轴标注的是频率, 但事实上纵轴对应的是音符标号, 因此图中的“旋律”轨迹是线性的。如果频率分析范围为 6 个八度音程, 则共有 72 个音符, 即图 4 的纵轴方向只有 72 条谱线。

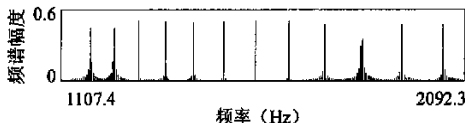


图 1 一个 8 度内 12 个音符仿真信号同时演奏时的直接 FFT 分析

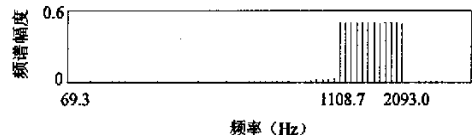


图 2 一个 8 度内 12 个音符仿真信号同时演奏时本文基于 Goertzel 算法的分析

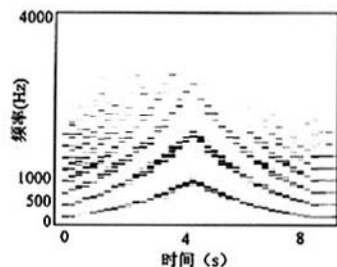


图 3 铜管音阶演奏的实际音乐信号的直接 DFT 时-频分析

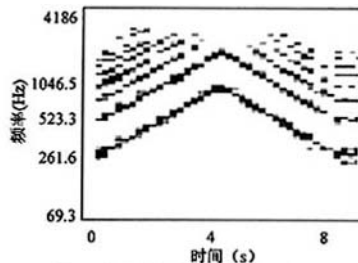


图 4 铜管音阶演奏的实际音乐信号的本文基于 Goertzel 算法的时-频分析

5 结论

如前所述, 由于 Goertzel 算法的简单性, 所以当期望用单片机实现音阶频率检测时, 本文算法更具吸引力, 因为它既有较好的性能, 编程也相对简单 (根据第 2 节介绍的算法原理和公式(12)、(13)编程即可)。此外, 对本文算法作以下几点说明:

a) 由于窗长度 N_l 计算的取整过程, 谱线位置频率和音阶频率之间仍有误差。如果要进一步减小误差, 可对计算 $X(k)$ 的定义式 (2) 稍加修正, 即 $w_N^{-k(N-l)}$ 中的 N 不进行取整, 其基本思想和文献[1]中讨论的常数 Q 变换定义类似。

b) 音阶频率之间的关系为 $f = f_0 \cdot 2^{1/b}$ 。如 $b=12$, 则每个倍频程内有 12 个音阶。但在分析音乐信号时, 也可取 $b=24, 36$ 等。

c) 实际演奏音乐的音高未必准确地等于定义的音高频率, 为使谱分析得到的频率更接近实际信号的音高频率, 可取较大 b 值, 以获得更多的频谱细节。

参考文献:

- [1] 丁志中, 戴礼荣. 音乐信号分析中常数 Q 变换的性能研究[J]. 声学技术, 2005, 24(4): 259-263.
- [2] Judith C Brown. Calculation of a constant Q spectral transform[J]. J. Acoust. Soc. Am, 1991, 89(1): 425-434.
- [3] B Blankertz. The Constant Q Transform[EB/OL]. <http://www.math.uni-muenster.de/math/inst/logic/org/staff/blankertz/constQ/constQ.pdf>, 2004-05-11.
- [4] G Goertzel. An Algorithm for the Evaluation of Finite Trigonometric Series[J]. American Math. Monthly, 1958, 65: 34-35.
- [5] A V 奥本海姆, R W 谢弗, J R 巴克. 离散时间信号处理[M]. 刘树堂, 黄建国, 译. 西安: 西安交通大学出版社, 2001.
- [6] 芮坤生, 潘孟贤, 丁志中. 信号分析与处理[M]. 北京: 高等教育出版社, 2003.

作者简介:

丁志中 (1961-), 男, 安徽省芜湖市人, 在读博士研究生, 副教授, 主要从事信号处理、通信等方面的研究, 以第一作者身份已发表论文 10 余篇. Email: zzding@mail.hf.ah.cn.

夏 杨 (1962-), 女, 江苏省高邮县人, 硕士, 讲师, 从事通信方面的教学与研究。

杨 萍 (1958-), 女, 安徽省寿县人, 学士, 副教授, 从事电子信息技术的教学与研究工作, 已发表专业论文 10 篇。

基于Goertzel算法的音阶频率分析

作者: [丁志中](#), [夏杨](#), [杨萍](#), [DING Zhi-zhong](#), [XIA Yang](#), [YANG Ping](#)

作者单位: [丁志中, DING Zhi-zhong \(合肥工业大学, 计算机与信息学院, 合肥, 230009; 中国科技大学, 电子工程与信息科学系, 合肥, 230027\)](#), [夏杨, 杨萍, XIA Yang, YANG Ping \(合肥工业大学, 计算机与信息学院, 合肥, 230009\)](#)

刊名: [信息与电子工程](#) 

英文刊名: [INFORMATION AND ELECTRONIC ENGINEERING](#)

年, 卷(期): 2006, 4 (1)

参考文献(6条)

1. [丁志中;戴礼荣](#) 音乐信号分析中常数Q变换的性能研究[期刊论文]-[声学技术](#) 2005 (04)
2. [芮坤生;潘孟贤;丁志中](#) [信号分析与处理](#) 2003
3. [A V奥本海姆;R W谢弗;J R巴克;刘树堂, 黄建国](#) [离散时间信号处理](#) 2001
4. [G Goertzel](#) [An Algorithm for the Evaluation of Finite Trigonometric Series](#)[外文期刊] 1958
5. [B Blankertz](#) [The Constant Q Transform](#) 2004
6. [Judith C Brown](#) [Calculation of a constant Q spectral transform](#)[外文期刊] 1991 (01)

本文读者也读过(10条)

1. [田伟](#), [周新力](#), [吴海荣](#) [基于Goertzel算法的DTMF信号解码研究](#)[会议论文]-2005
2. [王卫兵](#), [朱秋萍](#), [徐心毅](#), [戴加良](#), [刘克刚](#), [WANG Weibing](#), [ZHU Qiuping](#), [XU Xinyi](#), [DAI Jialiang](#), [LIU Kegang](#) [Goertzel算法的一种改进计算结构](#)[期刊论文]-[武汉大学学报 \(理学版\)](#) 2007, 53 (3)
3. [刘磊](#), [刘海东](#), [江桦](#), [LIU Lei](#), [LIU Hai-dong](#), [JIANG Hua](#) [戈泽尔算法在信号功率谱计算中的应用与改进](#)[期刊论文]-[信息工程大学学报](#) 2006, 7 (2)
4. [邵明东](#), [SHAO Ming-dong](#) [改进Goertzel算法的DTMF信号检测的仿真与应用](#)[期刊论文]-[电声技术](#) 2009, 33 (12)
5. [曾擎](#), [李玉柏](#), [李恒](#), [ZENG Zheng](#), [LI Yu-bai](#), [LI Huan](#) [基于Goertzel的载波机二次AGC控制方法研究](#)[期刊论文]-[电力系统通信](#) 2005, 26 (12)
6. [赵新](#) [儿童多项睡眠图监测](#)[期刊论文]-[临床神经电生理学杂志](#) 2007, 16 (3)
7. [王乙斐](#), [游舟浩](#), [王颖](#), [青海银](#), [WANG Yi-fei](#), [YOU Zhou-hao](#), [WANG Ying](#), [QING Hai-yin](#) [DTMF信号的合成与识别](#)[期刊论文]-[电子设计工程](#) 2011, 19 (7)
8. [弋今朝](#), [张禄林](#), [钱玉美](#) [一种新的基于 Goertzel算法的DTMF信号检测方法](#)[期刊论文]-[通信技术](#) 2002 (9)
9. [田伟](#), [周新力](#), [吴海荣](#), [李辉](#), [TIAN Wei](#), [ZHOU Xinli](#), [WU Hairong](#), [LI Hui](#) [短波地空通信多普勒频移对DTMF信号解码影响](#)[期刊论文]-[海军航空工程学院学报](#) 2006, 21 (2)
10. [Wumei Wang](#), [Zhibin Gao](#), [Lianfen Huang](#), [Yan Yao](#) [Spectrum Sensing Based on Goertzel Algorithm](#)[会议论文]-2008

本文链接: http://d.g.wanfangdata.com.cn/Periodical_xxydzgc200601007.aspx