

复杂声音信号基频检测和处理的 一种方法及实例

陈学煌[†]

(青海师范大学 西宁 810008)

摘要 周期性信号基频的检测具有重要的意义, 通常的硬件检测方法是采用过零比较器或施密特触发器等电路对输入信号进行整形, 从而获得基频脉冲。当信号波形复杂时, 这种检测方法就会失准。本文提出了一种新的硬件检测方法有效转变点实时基频检测法, 它用两个适当电平切割被测信号, 从而获得有效的零点, 解决了信号复杂时过零点增多的现象, 文内并给出了一个检测人声基频进而实现人声向乐器声转化的应用实例。

关键词 周期性信号, 基频检测, 人声, 乐器声

Detection and processing of the fundamental frequency of complicated signal and an illustration of its application

CHEN Xue-Huang

(Qinghai Normal University, Xining 810008)

Abstract The fundamental frequency detection of a periodic signal is of great significance. Often, after reshaping the input signal using a zero comparator, the frequency count is accomplished. When the signal waveform is complicated, however, this kind of detection will be inaccurate. The paper proposes a new hardware detection method—real-time detection by an effective change point. The method cuts the signal examined with two appropriate levels and thus gets the effective zero point. It solves the problem of measuring of the zero points when the signal is complicated. An example of application is provided that realizes the transformation from human voice to the sound of musical instrument after detecting the fundamental frequency of the human voice.

Key words Periodic signal, Fundamental frequency detection, Human voice, Sound of musical instrument

2006-12-30 收稿; 2007-08-14 定稿

作者简介: 陈学煌 (1948-), 男, 湖南蓝山人, 青海师范大学物理系教授, 硕士生导师, 研究方向: 检测与自动化。

[†] 通讯联系人 Email: cxhgz@163.com

1 引言

对各类信号进行测量,是人们同物质世界打交道的一种基本方式。人声的元音和浊辅音是一种周期性的振动信号,在语音识别技术中,人声基频作为基本参数之一具有重要的地位。目前主要是采用时域法(如短时平均幅度差函数 AMDF)、频域法(如倒谱法 CEP)或时频域混合法(如小波变换法)^[1,2]等建立语音模型来计算基频,这些方面的研究已取得了很多的进展,但由于数据获取及计算过程中较大的软件开销都需要一定时间,在某些场合的应用受到限制。同时由于人声信号的复杂性(声音随发音人的性别、年龄、个体差异及语音情感因素而有极大变化),有关文献指出:“至今也还没有一种通用的方法能够准确可靠地估计出任何人在任何情况下的基音周期”^[1]。

多年来,作者致力于一种新型的电子音乐装置研究,试图把人的歌声与乐器声联系起来,这二者之间的桥梁就是基频^[3]。这种音乐装置对基频检测及其它环节都有很高的实时性要求,宜尽可能采用没有时间延迟的方案。但是通常用于仪表内的过零比较器或施密特触发器等基频检测硬件电路,只适合在简单信号情况下使用。因此,有必要研究能准确测量复杂周期性信号基频的硬件方法。作者找到了这样一种方法,它完全采用硬件,具有实时性(不计相位变化),并且对基频检测的准确性远较过零比较器为高,根据其特征,称之为“有效转变点”基频检测法。这个方法用于歌声这种相对较为纯净的人声或其它周期性电信号基频之检测,具有实用的价值。

2 “有效转变点”实时基频检测法概述

众所周知,周期性信号无论其复杂程度如何,都可以分解为基频与若干高次谐波的组合,谐波的数量和幅度描述了该周期性信号的特性。当谐波的幅度较大时,信号波形的特点

是:在一个周期内,信号会出现很多过零点和峰点(正弦信号一周期只有两个过零点,一个正峰点,一个负峰点),这就是使得过零比较器、峰值检测器等都不能得出信号基频的原因。因此,要设法剔除那些并不代表基频的过零点或峰点的干扰,找出准确的基频。“有效转变点”基频检测法就是在这样的思路下形成的一种基于波形分析的硬件检测方法,它是一种有效的硬件基频解决方案。

2.1 “有效转变点”基频检测的原理

周期性信号的一例如图 1(a) 所示,这是一种人的声音信号波形示意图,其峰谷较正弦信号为多,正负方向的峰值分别为 u_{\max} 和 u_{\min} 。下面以其为例说明新的基频检测方法的原理。

为说明方便起见先给出如下的定义:用水平直线 $y = \pm U$ 去截人声信号波形图,求出它们之间交点的时间序列,当 $y > 0$ 时得到正交点序列,当 $y < 0$ 时得到负交点序列,当 $y = 0$ (即 $U = 0$) 时为零点序列。如果正交点后面紧跟一个负交点或负交点后面紧跟一个正交点则称为出现一个“有效转变点”。

上式中 U 的值从 0 开始,最大值不超过 $|u_{\min}|$ 和 $|u_{\max}|$ 中较小的一个。

常用的过零检测方法即相当于这里用 $y = 0$ (即时间轴) 的水平线去找零点,可以从图中看出,在一个信号周期内会有 5 个以上过零点(如 t 轴上的小白点所示)。一般来说,对于简单信号(如正弦信号等),在一个周期内只会出现两个零点,而对于复杂的信号过零点总是较多,因此要设法排除掉那些与基频无关的零点。

现在来看不用 $y = 0$ 去求交点,而是改用 $y = \pm U$ 两条线来求交点的情况。随着 $\pm U$ 值的不同,正交点和负交点的数量都会发生变化, $\pm U$ 值向 u_{\max} 和 u_{\min} 靠近时交点数量虽然大多呈现减少的趋势,但并没有确定的规律可循,因此也不能用正交点或负交点数来代表信号的基频,在本例中这两种交点分别如图

1(b), 图 1(c) 所示, 每个周期内交点脉冲数都较多, 并且会随着 U 值的改变而变化。

如果把视线放在观察正负交点序列之间的关系上, 即观察本文所定义的“有效转变点”之运动变化上, 却能够找到这样一个重要的规律, 即: 随着 $\pm U$ 的值向 u_{\max} 和 u_{\min} 靠近, 这种“有效转变点”的数量会逐渐地减少, 直至最终减少到 2, 与同频正弦波的过零点相同了, 这也就意味着找到了信号的基频 (见图 1(d))。事实上也可以这样来理解, 即: 在时间序列中, 当一个可检的正峰信号后面出现一个负峰信号时, 其中间肯定存在着一个真正有用的信号过零点, 也就是这里所说的存在一个“有效转变点”; 而当一个正峰信号后面仍是一个正峰信号 (或一个负峰信号后面仍是一个负峰信号) 时,

其中间则不存在对基频检测有用的过零点。因此, 只要设法实现上述甄别, 就能够得到复杂信号的基频了。

2.2 “有效转变点”基频检测法的技术实现

根据上述甄别原理, 设计了如图 2 所示的电路。下面对这个电路方案加以说明。

在该方案中, 将 $\pm U$ 的值分别取为 u_{\max} 和 u_{\min} , 而不是采用某个固定的切割电平, 这是因为 $|u_{\max}|$ 和 $|u_{\min}|$ 通常并不相等, 并且由于实际的信号常受到幅度调制等因素的影响, 其幅值大小也不断波动, 采用固定的电平去切割会出现漏检现象, 这里采用顶部钳位的方法来检测峰脉冲, 可靠性就较高。图中使用了指数放大器, 这是一种非线性放大器, 使用它的目的是使信号的峰值部分较中间部分得到更多

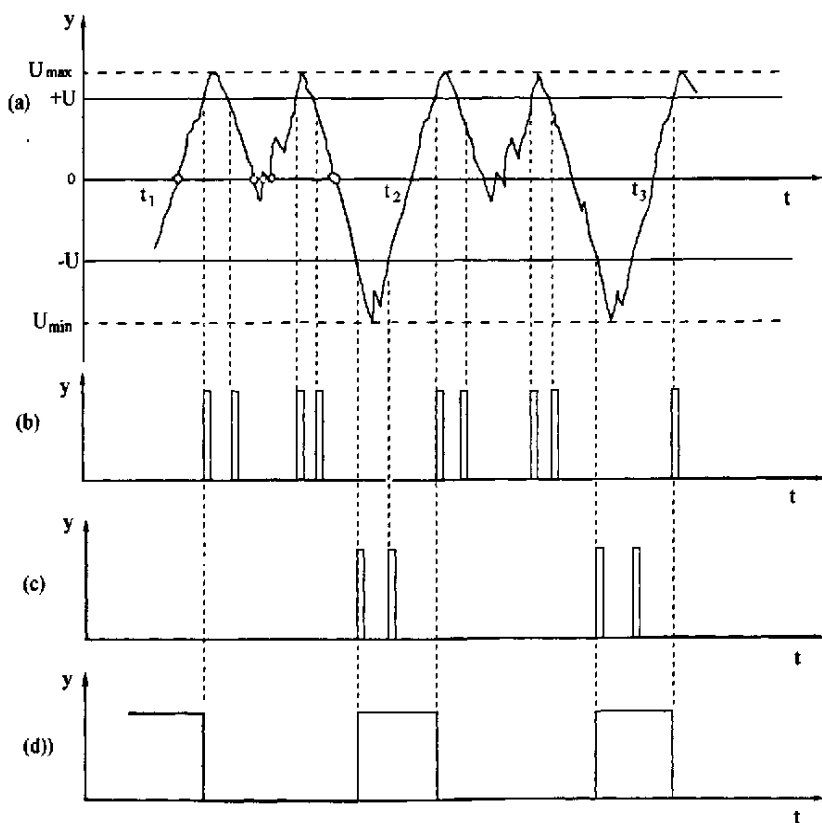


图 1 信号波形图 (a) 原始信号, $t_1 - t_2$, $t_2 - t_3$ 间均为一个周期 (b) 正交点序列脉冲 (c) 负交点序列脉冲 (d) 基频脉冲 (或宽或窄的矩形脉冲波)

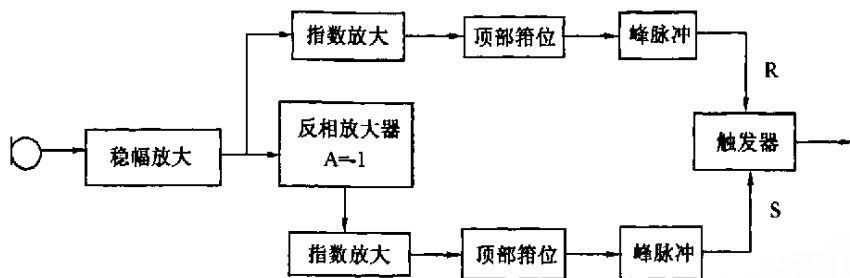


图2 “有效转变点”基频检测器原理框图

的放大,以有利于峰检测,但这种非线性放大作用也不应过强,以免漏掉较弱信号的峰。指数放大电路由运算放大器和非线性反馈元件构成,非线性元件用半导体二极管或三极管的一个PN结,在简单情况下也可用不加直流偏置或低偏置的晶体管放大器构成之。

图中还使用了反相放大器,使得总体的电路得以简化,因为经过反相后,信号的负半周变成了正半周,因而下面的一路指数放大和箝位电路就可以使用与上面一路完全相同的规格,但得到的脉冲序列为负峰脉冲序列。因此两路峰检测电路分别输出了正、负峰脉冲序列。这两列脉冲间产生的有效零点最后由一个触发器鉴别出来,这是一个普通的双稳态触发器,从双稳态触发器真值表可以查出,该触发器只有当R和S端的脉冲出现交替时输出的状态才会改变,从而形成基频脉冲的一个上升(或下降)沿。该检测电路各点的信号波形变化可以参考图1来理解。其中两路峰检测器的输入信号分别相应于图1(a)和图1(a)经倒相后的波形,而输出脉冲则相应于图1(b)和图1(c),最后在触发器的输出端得到信号的基频脉冲,即形如图1(d)所示的矩形脉冲序列。

2.3 与过零比较器检测效果之比

采用上述基频检测器对10名受试者(男女学生各5人分别标记为A_i和V_i)进行不同元音的检测试验,与一个没有回差的过零比较器(图3)的检测结果进行了比较,如表1、2所示。表中○表示检测准确,×表示检测失误。

测试中,除了发个别音时会出现不稳定

外,有效零点基频检测器的总体效果还是比较好的,而过零比较器只在少数元音的检测中出现正确值,在大多数情况下会产生错误的结果。

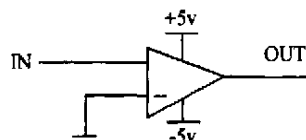


图3 过零比较器

表1 有效零点检测器的结果

	A1	A2	A3	A4	A5	V1	V2	V3	V4	V5
a	○	×	○	○	○	○	○	○	×	○
o	○	○	○	○	○	○	○	○	○	○
e	○	×	○	×	○	○	○	○	○	○
i	○	○	○	○	○	○	○	○	○	○
u	○	○	○	○	○	○	○	○	○	○
v	○	○	○	○	○	○	○	○	○	○

表2 过零比较的结果

	A1	A2	A3	A4	A5	V1	V2	V3	V4	V5
a	×	×	×	×	×	×	×	×	×	×
o	×	×	×	×	×	×	×	×	×	×
e	×	×	×	×	×	×	×	×	×	×
i	×	×	×	×	×	○	×	○	×	○
u	×	×	×	×	×	×	×	○	×	○
v	○	×	○	×	×	○	×	×	○	○

3 基频脉冲频谱的归一化

信号的基频经过上述检测器检测出来后,得到了图1(d)所示的矩形脉冲序列,即基频脉

冲波。可以把这种基频脉冲送往频率计去进行计量，不会影响计数的准确性。但由于这种基频脉冲的波形极不统一，在其它的应用中就不一定合用，例如在将人声基频做为音源信号使用时，会发现它的音色很不稳定，与乐音的音色相去甚远，因此有必要把这样的基频脉冲进行归一化处理。下面以此为例说明处理方法。

这里所说的归一化，是指处理后的脉冲具有统一的波形，例如，使脉冲具有某一个固定的占空比。在谱分析仪上可以看到，频率相同而占空比不同的基频脉冲其高次谐波成分完全不一样，这种现象利用数学中将周期函数展开为富氏级数的方法可以得到理论解释。数学分析指出，周期函数（满足狄义赫利条件）总是可以分解为常数及一系列重复周期为原函数周期整数倍的简谐函数之和。对于幅度为 A 的矩形脉冲波（基频率为 ω ），可利用富氏级数将它分解为直流成份与一系列离散的简谐波（正弦波）之和：

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\omega t)$$
$$= A \frac{tu}{T} + \frac{2A}{\pi} \sin\left(\pi \frac{tu}{T}\right) \cos(\omega \cdot t) + \dots$$
$$+ \frac{2A}{n\pi} \sin\left(n\pi \frac{tu}{T}\right) \cos(n\omega \cdot t) \quad (1)$$

上式中 tu/T 为脉冲宽度与周期之比，可以看出，对于同样的周期 T ，随着脉冲宽度 tu 的不同， $n\pi tu/T$ 的正弦值亦不相同，各偶次项和

奇次项的系数（即相应谐波的幅度）会有很大的变化，因而基频虽相同而音色可以大不相同。

目前所实施的方法，是使得基频脉冲全部变成占空比为 1:1 的矩形波，即方波。

当脉冲的占空比为 1:1 时，上述 (1) 式中的各偶次项系数 $\frac{2A}{n\pi} \sin\left(n\pi \frac{tu}{T}\right) = 0$ ，整个级数展开式也具有最简单的形式：

$$f(t) = \frac{A}{2} + \frac{2A}{\pi} \cos(\omega \cdot t) - \frac{2A}{3\pi} \cos(3\omega \cdot t)$$
$$+ \frac{2A}{5\pi} \cos(5\omega \cdot t) - \dots \quad (2)$$

这时脉冲信号的谐波群中只有一系列奇次谐波，各次谐波幅度随 n 成反比地减小（图 4），具有完全统一的频谱结构。

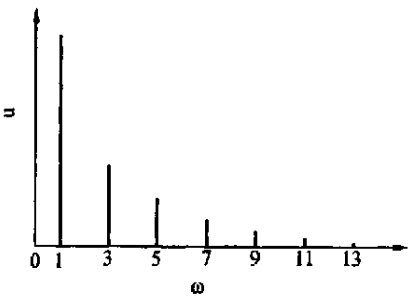


图 4 方波频谱

在技术上为实现此目的，设计了如附图 5 所示的电路。基频脉冲通过鉴相器进入锁相环（采用集成锁相环器件 CD4046，片内包含了虚线框内的部分^[4]），锁相环的输出则经由一组二

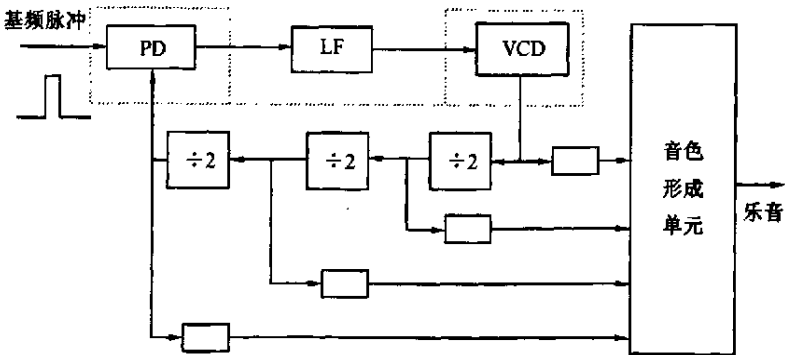


图 5 乐音形成单元

分频器后再送往鉴相器, 相位将被锁定在输入脉冲前沿 (或后沿), 而锁相环的输出脉冲占空比恒为 50%, 这样在各个分频器的输出端得到了 N 倍频的方波信号^[5]。

然而从 (2) 式可以看出, 方波的谐波群中只有奇次谐波, 缺少了听感和谐的偶次谐波成分。这种谐波成分的缺陷使得方波并不能作为最理想的、会产生多种丰富音色的音源信号源。为了弥补这个缺憾, 在图 5 中使多个倍频信号同时输出, 即相当于音乐中的 $2', 4', 8', 16'$ 等多个尺音信号共同作为音源信号, 它们的谐波成分互相交错, 使偶次谐波成分得到了补充。这些信号在经过适当的音色合成电路后, 即可产生出预期的音色 (图中未标注的小方框为缓冲器)。

4 应用效果一例

采用本文所述基频检测方法制成一种将人声转换成乐音的样机, 总体的效果还是较为理想的。曾对近百名受试者进行了声音转化测试, 准确度达到 95% 以上, 只有个别发音者在测试时会出现输出不稳定的情况 (音高的跳动), 观察其声音的波形和频谱, 发现这个别人的声音信号谐波均极为丰富, 基频能量已经很小, 其时域波形有时形似于脉冲串, 这正是本技术中

所采用的有效转变点基频检测方案所忌讳的。好在发出这种声音的人在人群中较少见, 只是在采用美声唱法的专业歌唱家中才有发现, 在所测试的普通人及业余歌手中尚未能见到。因此本技术方案不失其广泛的适应性。图 6 为这种将人声转换成乐音声的伴奏机成品照片。



图 6 样机照片

参 考 文 献

- 1 李晋, 王玲. 基于线性预测和最大似然的基音检测算法. 计算机应用, 2006, 26(5):1232~1236.
- 2 顾良, 刘润生. 高性能汉语语音基音周期估计. 电子学报, 1999, 27(1):8~11.
- 3 陈学煌. 人声、乐器声及其转化. 电声技术, 1990, 02, 5~9.
- 4 中国集成电路大全编写委员会. CMOS 集成电路. 北京: 国防工业出版社, 1985, 566~575.
- 5 郑继禹, 万心平, 张厥盛. 锁相环路原理与应用. 北京: 人民邮电出版社, 1994, 232~248.

复杂声音信号基频检测和处理的一种方法及实例

作者: [陈学煌](#), [CHEN Xue-Huang](#)
作者单位: [青海师范大学, 西宁, 810008](#)
刊名: [应用声学](#) [ISTIC](#) [PKU](#)
英文刊名: [APPLIED ACOUSTICS](#)
年, 卷(期): 2007, 26(6)
被引用次数: 0次

参考文献(5条)

1. [李晋](#), [王玲](#) [基于线性预测和最大似然的基音检测算法](#)[期刊论文]-[计算机应用](#) 2006(05)
2. [顾良](#), [刘润生](#) [高性能汉语语音基音周期估计](#) 1999(01)
3. [陈学煌](#) [人声、乐器声及其转化](#) 1990(02)
4. 《中国集成电路大全》编写委员会 [CMOS集成电路](#) 1985
5. [郑继禹](#), [万心平](#), [张厥盛](#) [锁相环路原理与应用](#) 1994

本文链接: http://d.g.wanfangdata.com.cn/Periodical_yysx200706004.aspx

授权使用: 深圳大学城图书馆(szdxc), 授权号: 1b33d853-9daa-44d9-9c3f-9e23009ac931

下载时间: 2010年11月3日