

Talk-off 鲁棒的双音多频信号解码器*

吴泽伟,袁红星,吴少群,余辉晴

(宁波工程学院 电子与信息工程学院,浙江 宁波 315016)

摘要: 双音多频信号被广泛用于音频电话及交互式控制系统中,正确解码是这些应用的关键。然而双音多频信号解码因噪声干扰面临着 Talk-off 的难题。针对该问题,在 Goertzel 算法基础上提出利用频率分辨率及其倍频检测干扰噪声的解决方案。测试结果表明该方法优于二次谐波检测的解决方案。

关键词: 双音多频;Talk-off;Goertzel 算法;二次谐波

中图分类号: TP301

文献标识码: A

文章编号: 1674-7720(2013)06-0061-04

Talk-off resistant dual tone multiple frequency decoder

Wu Zewei, Yuan Hongxing, Wu Shaoqun, Yu Huiqing

(School of Electron and Information Engineering, Ningbo University of Technology, Ningbo 315016, China)

Abstract: Dual tone multiple frequency has been widely used in touch tone phone and interactive control system. It is critical to decode it correctly in such applications. However, it has talk-off issues due to background noise interference. A Goertzel algorithm based approach which considers frequency resolution and its multiple frequency detection to deal with the issues is presented. Many experimental results with test data from real environments show it is superior to harmonic frequency detection methods.

Key words: DTMF; Talk-off; Goertzel algorithm; harmonic frequency

双音多频 DTMF(Dual Tone Multi-Frequency)信号由两个单频正弦信号叠加生成,是音频电话的拨号信号标准^[1]。由于具有良好的抗噪特性,已被广泛用于诸如主叫识别信号传送及显示、电话语音服务、远程控制和 ATM 终端等通信系统拨号传输中^[2]。根据 ITU 的建议标准 Q23^[3],构成 DTMF 信号的两个频率分别来自行频组(697 Hz, 770 Hz, 852 Hz, 941 Hz)和列频组(1 290 Hz, 1 336 Hz, 1 477 Hz, 1 633 Hz),共 16 种组合,分别表示 16 个按键值,即 10 个数字键 0~9 和 6 个功能键 *、#、A、B、C、D^[4]。

DTMF 解码是指从受干扰的语音信号中检测出 DTMF 信号并将其还原成按键值。正确的解码对于基于 DTMF 的应用至关重要。DTMF 解码的实质是从有噪声信号中检测出两个正弦波频率。近年来,已有多个解码算法被提出。最直接的方法是 DFT,通过 DFT 得到输入信号的频域信息,进而分析频谱确定是否存在 DTMF 频率。DFT 通常以 FFT 的运算来实现,由于 FFT 生成从 DC 到半采样率频带内所有频率信息,因此该方法计算量大,且需要较多的存储空间用于暂存计算结果^[5]。针对

该问题,参考文献[6]提出基于 Goertzel 的 DFT 快速运算。与 FFT 不同,Goertzel 算法只对 DTMF 的 8 个频点计算 DFT,因此计算量大大降低。此外,Deosthali 等人提出了非均匀 DFT 算法^[7],Popovic 在牺牲相位信息的情况下提出了改进的 Goertzel 算法^[8],金鑫春等人研究了 Goertzel 算法的参数选择问题,但这些算法都没有考虑 DTMF 的 Talk-off 错误问题。解决该问题的常用方法是检测二次谐波。由于 DTMF 信号只包含基音成分,因此只要检测到较强的二次谐波,就认为该信号不是 DTMF 信号。但在强噪声环境下,该方法仍会出现误判,即发生 Talk-off 错误。

本文针对 Talk-off 错误问题,在 Goertzel 算法的基础上,提出基于频率分辨率及其倍频检测的方案,并利用后续的判断逻辑进一步排除虚假 DTMF 信号。大量真实应用环境下的语音测试结果表明,本文方法在强噪声干扰情况下仍能避免 Talk-off 错误问题。

1 DTMF 解码器

1.1 解码器框架

本文提出的 DTMF 解码器如图 1 所示,主要包括信号预判、自动增益控制 AGC(Auto Gain Control)、基于 Goertzel

* 基金项目:浙江省大学生科技创新活动计划(新苗人才计划)(2012R422008);宁波市自然科学基金(2012A610043)

网络与通信 Network and Communication



图1 DTMF 解码器框架

的频率检测和判断逻辑几个模块。信号预判模块根据输入信号的能量决定其是否包含DTMF信号；AGC将输入信号的动态范围变换到Goertzel算法可以处理的范围；基于Goertzel的频率检测模块计算输入信号在8个DTMF频点的能量、频率分辨率及其倍频的能量，并以此作为下一个模块的输入；判断逻辑模块根据前一个模块的输入，决定某对DTMF频率是否存在，并将其翻译成对应的按键值。

1.2 信号预判

根据ITU的建议标准Q23对DTMF信号的规定，如果输入信号在8个DTMF频点的能量均 ≥ -25 dBm且 ≤ 9 dBm，则认为可能包括DTMF信号。如果输入信号在8个DTMF频点处的能量均 ≤ 55 dBm，则认为该信号肯定不存在DTMF音调。信号预判模块能够有效地排除一些虚假DTMF信号，减小后续模块误判的可能性。

1.3 AGC

AGC从一帧输入数据的前若干个样点中确定幅度模的峰值，将Goertzel算法所要处理的样点都乘以一个系数，使得该峰值等于Goertzel算法所能处理的最大值。

1.4 基于Goertzel算法的频率检测

由于Goertzel算法简单、易于实现，本文采用Goertzel算法进行频率检测。Goertzel算法的处理对象是数据块，一个数据块由 N 个样本构成，每个数据块输出一个检测结果。数据块的大小，即 N 决定了频率分辨率 f_k ：

$$f_k = f_s / N \quad (1)$$

其中 f_s 表示采样频率，对于语音信号， f_s 一般取8 kHz。

Goertzel算法利用一个双极点的二阶IIR滤波器来估计DFT值，其结构如图2所示。离散域上的Goertzel算法如下：

对于 $n=0 \cdots N$ ，递归计算：

$$v_k(n) = 2\cos(2\pi k/N)v_k(n-1) - v_k(n-2) + x(n)$$

其中 $v_k(-1)=0$ ， $v_k(-2)=0$ ， $x(n)$ 表示 n 时刻输入信号的采样值。

每 N 个样本估计一次DFT模值：

$$|X(k)|^2 = y_k(N)y_k^*(N) \\ = v_k^2(N) + v_k^2(N-1) - 2\cos(2\pi f_k/f_s)v_k^2(N)v_k^2(N-1)$$

k 由下式确定：

$$k = \text{floor}(0.5 + Nf_k/f_s) \quad (2)$$

其中 f_k 和 f_s 分别表示待检测的目标频率和采样频率， $\text{floor}()$ 表示向下取整运算。本文用14个如图2所示的IIR滤波器实现14个频率的检测，即8个DTMF频率和6个与频率分辨率相关的频率。这6个频率检测主要是为了解决Talk-off问题，其计算如下：

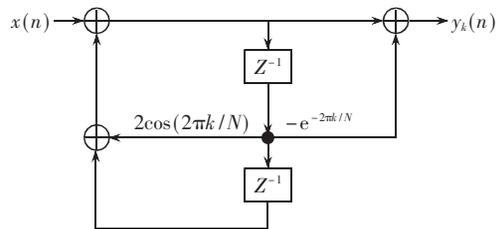


图2 Goertzel算法的IIR结构

$$f_{a1} = \text{floor}(f_s/N), f_{a2} = \text{floor}(3f_s/N), f_{a3} = \text{floor}(4f_s/N)$$

$$f_{a4} = \text{floor}(5f_s/N), f_{a5} = \text{floor}(7f_s/N), f_{a6} = \text{floor}(14f_s/N)$$

当 $N=102$ ，采样率为8 kHz时，这6个频率分别为78 Hz、235 Hz、314 Hz、392 Hz、549 Hz和1 098 Hz。为了与DTMF频率区别开来，本文将这6个频率称为辅频。

由于现有的语音增强模块都是一个语音帧(8 kHz时为20 ms, 160个采样点)处理一次。为了便于将DTMF解码器集成到现有系统中，本文的DTMF解码器每个语音帧调用一次Goertzel算法进行频率检测。根据ITU Q24^[9]，DTMF音调的持续时间介于45 ms~55 ms之间，连续的DTMF信号之间的间隔至少为45 ms。因此一个DTMF音调由多个语音帧组成。这意味着一个DTMF音调会被检测多次。对同一个DTMF音调，理论上每次检测结果都应该是一样的。然而受噪声影响，前后几次检测结果有可能出现不一致的情况。为解决该问题，本文采用表决方法输出最可靠的检测结果，多次检测的表决也增强了本文解码器的抗噪性能。

1.5 判断逻辑

图3给出了判断逻辑模块的工作流程，根据前面14个IIR滤波器的检测结果判断检测到的行频和列频是否有效。对于有效的检测结果，根据图4所示的行频、列频与按键的对应关系，输出按键值。

2 实验结果与分析

为了验证本文方法的有效性，用真实使用环境下的语音数据作为测试数据。测试数据获取方法是，首先在两个手机之间建立通信链接，然后在其中一个手机上随机连续地按下多个按键，另一个手机将接收到的语音信号保存成测试用的wav文件。为了测试Talk-off性能，在发送DTMF信号的手机端分别播放噪音、语音和音乐。在算法实现时，考虑到在定点DSP上的移植，用Q14格式将其定点化(DFT模值用32位表示，其余的均用16位表示)。

程序有3个可调参数，其中Frames between Successive Dials指两个连续按键间隔的时间，该参数用于处理按键防抖动，以帧为单位，对于8 kHz采样率的语音而言，一帧为20 ms。对于真实环境下的测试数据，程序中的3个参数都是固定的，分别设为3、2和5。测试中，将本文方法与参考文献[2]的方法进行了对比。其中，参考文献[2]通过二次谐波检测来排除噪音对检测结果的干扰。参考文献[2]的参数设置如下： $N=201$ ，二次谐波的能量至少比基频能量小20 dB。

《微型机与应用》2013年第32卷第6期

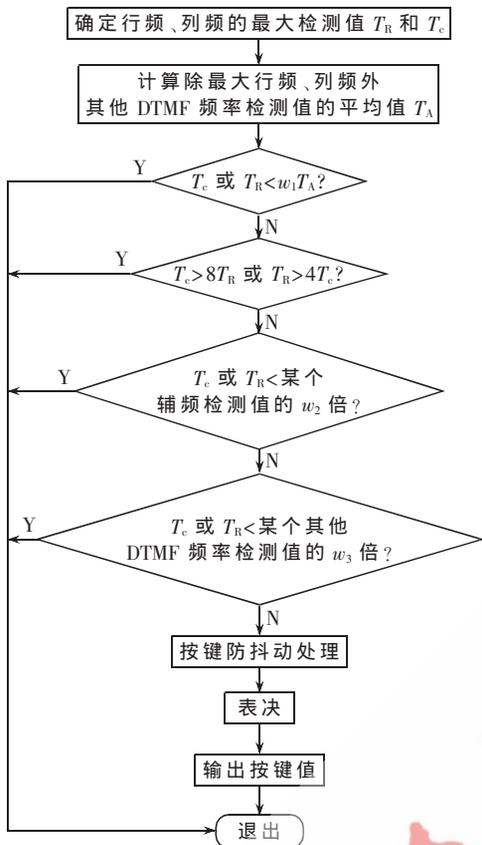


图3 判断逻辑流程图

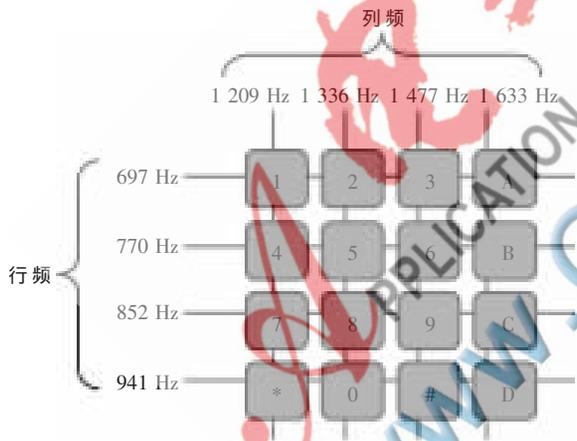


图4 DTMF 频率与按键对应关系

2.1 真实环境下测试 1

在发送 DTMF 信号的手机侧播放如图 5 所示的噪声信号。接收侧手机录制的语音信号如图 6 所示。这个测试信号的 DTMF 按键值序列为: 123456789*0##0*98765423147856280#963547*。参考文献[2]的检测结果为: 1234567899*?##0*98765423147856280#963547* (单下划线表示误检测, 双下划线表示重复检测, ?表示漏检), 重复检测一个 9, 漏检一个 0, 误将 6 检测为 5; 本文方法检测成功率为 100%。

2.2 真实环境下测试 2

在发送 DTMF 信号的手机侧播放如图 7 所示的音乐

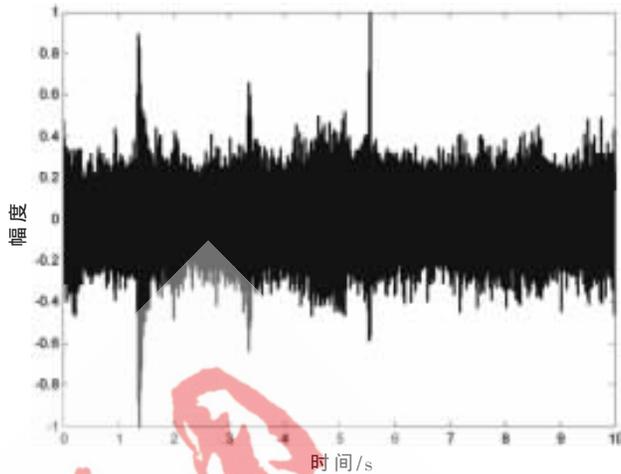


图5 发送端噪声信号

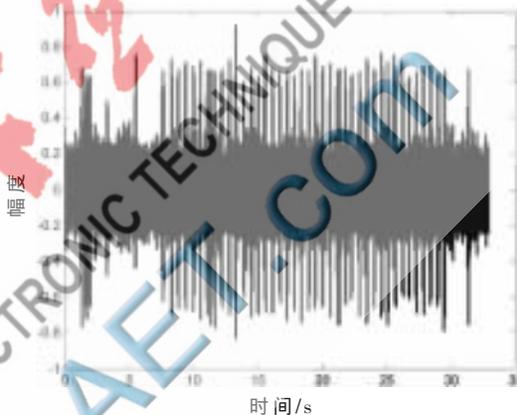


图6 发送端增加噪声时接收到的语音

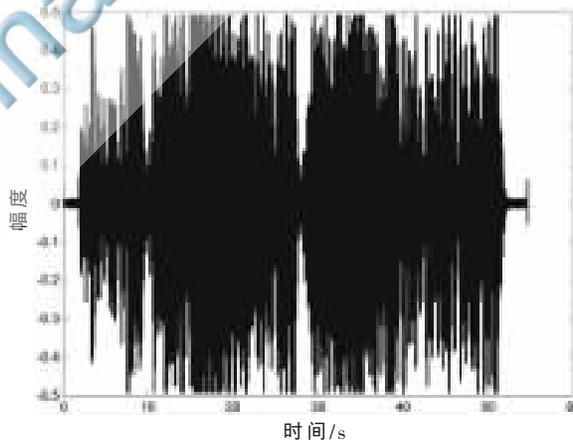


图7 发送端播放的音乐信号

信号。接收侧的语音信号如图 8 所示, 对应的 DTMF 按键值序列为: 123456789*0##0*987654321147*2580369##9630852*741。参考文献[2]的检测结果为: 123456789*0##0*98765?321147*2580369###963085?*741 (单下划线表示误检测, 双下划线表示重复检测, ?表示漏检), 漏检 4 和 2, # 被重复检测一次, 结束时误检测出 7; 本文方法检测成功率为 100%。

2.3 真实环境下测试 3

在发送 DTMF 信号的手机侧, 操作者一边按键一边

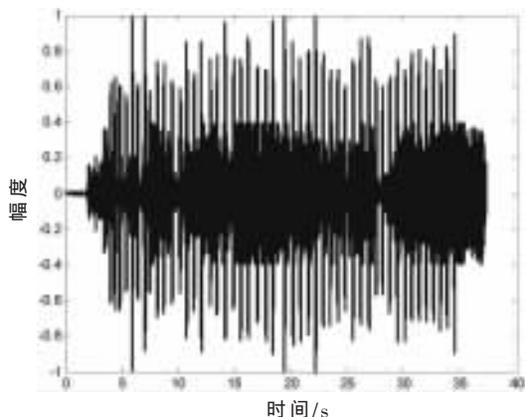


图8 发送端增加音乐信号时接收到的语音

发声。接收侧的语音信号如图9所示,对应的DTMF按键值序列为:5152535455565758595*505#1525354565758595*505#5。参考文献[2]的检测结果为:511?253545555657585595*505#1525354565758595*5005#5(单下划线表示误检测,双下划线表示重复检测,?表示漏检),第2个按键1被重复检测1次,第3个按键5被漏检,第10个按键、第17个按键和第43个按键分别被重复检测1次;本文方法检测成功率为100%。

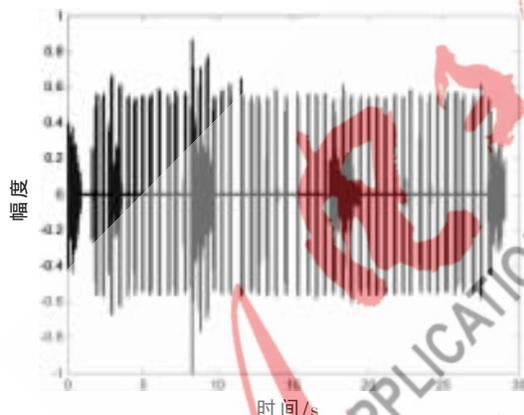


图9 发送端操作者发声时接收到的语音信号

上述真实环境下的测试结果表明,本文方法在处理Talk-off问题上比传统的二次谐波检测更为有效。其原因是,离散域上用Goertzel算法检测二次谐波的准确性依赖于频率分辨率的选择。在强背景噪声环境下,其检测结果更容易受到影响。本文直接利用频率分辨率及其倍频进行干扰语音的检测,可以有效避免该问题。

2.4 ITU标准兼容性测试

ITU对DTMF解码器的性能规定主要体现在4个方面,即Talk-off、扭曲(twist)、频偏和SNR。Talk-off的测试如前文所述。扭曲指的是检测到的行频与列频DFT模值之比,对于有效的DTMF信号,要求在-4 dB~8 dB范围内。扭曲测试是在检测DTMF之后的逻辑判断中进行的,目的是用于排除虚假DTMF信号。图4中判断行频、列频检测值之间关系的步骤即为扭曲测试。ITU规定频偏的容许范围为 $\pm 1.5\%$,当超过 $\pm 3.5\%$ 时则认为是无效

的DTMF信号。在SNR等于或高于15 dB时,ITU要求成功检测率为100%。按照参考文献[10]所述的方法对频偏和SNR进行测试。实验结果表明,在SNR为10 dB的情况下,频偏在 $\pm 1.5\%$ 范围内时,本文解码器的检测成功率为100%;当频偏超过 $\pm 2.0\%$,译码开始出错。当SNR在-3.7 dB以上时,本文检测成功率为100%。其中,测试用的DTMF信号都由Cool Edit Pro软件生成;实验中的3个参数均设为1。

Talk-off是DTMF解码器的重要性能,然而现有算法对此问题的解决方案主要围于二次谐波的检测。本文另辟蹊径,提出频率分辨率及其倍频检测的解决思路。真实使用环境下的测试数据验证了该方法的可行性。ITU标准兼容性测试实验表明本文方法能够满足ITU对DTMF解码器的规定。

参考文献

- [1] 王乙斐,游舟浩,王颖,等.DTMF信号的合成与识别[J].电子设计工程,2011,19(7):71-73.
- [2] 邵明东.改进Goertzel算法的DTMF信号检测的仿真与应用[J].声电技术,2009,33(12):65-69.
- [3] ITU.ITU-T recommendation Q.23 technical features of pushbutton telephone sets[S].1988.
- [4] 金鑫春,汪一鸣.Goertzel算法下DTMF信号检测及参数优化[J].现代电子技术,2010(6):152-155.
- [5] MILOS S T,DUSAN R.Performance analysis of the DTMF detector based on the Goertzel's algorithm[C].Proc.of 14th Telecommunications Forum,2006.
- [6] POPOVIC M.Digital signal processing[M].Belgrade: Academic mind,2003:100-120.
- [7] DEOSTHALI S,MCCASLIN R,EVANS B.A low-complexity ITU-compliant dual tone multiple frequency detector[J].IEEE Trans.on Signal Processing,2000,48(5):1-20.
- [8] POPOVIC M.Efficient decoding of digital DTMF and R2 tone signaling[J].Factual Univ.Ser., Elec.Energ,2003,16(3):389-399.
- [9] ITU.ITU-T recommendation Q.24 multifrequency push button signal reception[S].1989.
- [10] 陈通,曹小强.基于NDFT Goertzel滤波器的DTMF信号检测的改进方法[J].西南大学学报,2008,30(1):152-155.

(收稿日期:2012-11-21)

作者简介:

吴泽伟,男,1991年生,本科生,主要研究方向:电路设计。

袁红星,男,1980年生,博士,高工,主要研究方向:信号与信息处理、3D视频信号获取与处理。