

基于小波变换和时域波形的基音检测算法

陈小利<sup>1</sup>, 徐金甫<sup>2</sup>

(1. 中国人民解放军 96550 部队, 河南 洛阳 471031; 2. 信息工程大学 电子技术学院, 河南 郑州 450004)

**摘要:**为了准确地检测语音信号的基音周期,采用小波变换和时域波形相结合的方法,分别用传统的自相关法、平均幅度差法及文中算法对纯净语音和不同信噪比时的含噪语音进行基音检测。实验表明,自相关法易出现半频错误,平均幅度差法易出现倍频错误,且两者随着信噪比的降低,错误帧数呈增加趋势。然而文中算法倍频、半频错误相对较少,基音轮廓清晰、平滑,无大的跳变,符合语音信号慢时变性的一般规律,从而提高了基音检测的精度。

**关键词:**基音检测; 小波变换; 自相关法; 平均幅度差法

中图分类号:TN912-34      文献标识码:A      文章编号:1004-373X(2011)01-0077-03

Pitch Detection Algorithm Based on Wavelet Transform and Waveform in Time Domain

CHEN Xiao-li<sup>1</sup>, XU Jin-fu<sup>2</sup>

(1. Unit 96550 of PLA, Luoyang 471031, China; 2. Institute of Electronic Technology, Information Engineering University, Zhengzhou 450004, China)

**Abstract:** An algorithm is proposed on the basis of wavelet transform and waveform in time domain to improve the accuracy of pitch detection. The pitch period of clear speech and noisy speech in different signal-to-noise ratio is detected by using the proposed method and the traditional methods of autocorrelation function and average magnitude difference function. The experiments show that the half frequency errors may occur in the method of autocorrelation function and double frequency errors may occur in the method of average magnitude difference function. Furthermore, the error frames of the two methods increase with the decrease of signal-to-noise ratio. In the proposed algorithm, there are relatively few double and half frequency errors, and the pitch contour is clear and smooth without big jump. The results are coincident with the common rule as the slow variation with time of speech signal. Therefore, the accuracy of pitch detection can be improved effectively by the algorithm.

**Keywords:** pitch detection; wavelet transform; autocorrelation method; average magnitude difference method

0 引言

发浊音时,声带振动引起的周期性称为音调,声带振动的周期称为音调周期或基音周期。基音检测即基音周期的提取,是语音处理中一个十分重要的问题,是各种语音系统的必要组成部分<sup>[1]</sup>。基音检测的方法大体上可以分为时域法(如自相关法<sup>[2]</sup>、平均幅度差法<sup>[3]</sup>等)和频域法(如倒谱法<sup>[4]</sup>等)。近年来,随着现代信息处理方法的发展和时频分析的进一步发展,又提出了时频联合分析来提取基音周期的方法(如小波变换法<sup>[5]</sup>),它们在某一方面都具有不少的优点。但由于语音信号本身结构比较复杂,以及声道共振峰和噪声的影响等,致使适合不同讲话者、不同要求和环境的基音检测方法还未出现。

小波变换具有多分辨率分析特性<sup>[6]</sup>,可将信号分解

为低频部分的近似分量和高频部分的细节分量。因此,可对语音信号进行小波分解,根据分解得到的各级系数,利用低频系数对语音信号进行重构,得到包含基音信息的低频子带信号,并将其作为基音检测的基础。

传统的自相关法和平均幅度差法只利用了语音信号的幅值信息,对信号的形状重视不够,所以基音提取中不免产生一些误判。本文则综合利用信号波形的多种特征,如波峰的宽度、面积、幅值大小等信息,并根据这些信息的匹配程度检测基音周期<sup>[7]</sup>。

1 基音检测算法原理

基音检测流程如图 1 所示。首先,将语音信号通过一个截止频率为 60~4 000 Hz 的带通滤波器进行滤波,将频率限制在 4 000 Hz 以内。其次,对滤波后的信号进行二级小波分解,提取 1 000 Hz 以下的低频子带信号,并将该子带信号通过数值滤波器加以平滑。最后,对判决为浊音的帧,利用其时域波形信息检测基音周期。

收稿日期:2010-08-19  
基金项目:河南省自然科学基金资助项目(0411010100)

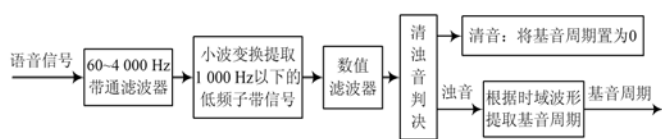


图1 基音检测流程图

### 1.1 预处理

基音检测前,需对语音信号进行预处理,本文采用了带通滤波、小波变换和数值滤波作为语音信号的预处理。

语音信号通过截止频率为 60~4 000 Hz 的带通滤波器,可将其频率限制在 4 000 Hz 以内,这就保证了后续的小波分解能精确提取出 1 000 Hz 以下的低频子带信号。这里将低频端截止频率设为 60 Hz 是为了去除电源信号的干扰。

小波变换具有多分辨率分析特性,利用小波变换不仅可以精确地分离出后续分析所需的子带信号,而且可以提高算法的抗噪性能<sup>[8]</sup>。

从滤波器的角度来看,小波分解相当于将输入序列进行双通道滤波的过程,即将输入信号  $a_0$  分别经过低通滤波器  $\bar{h}[n] = h[-n]$ 、高通滤波器  $\bar{g}[n] = g[-n]$  进行滤波处理。图 2(a) 所示为小波分解图,即每变换一次就把信号分成低频分量(平滑分量)和低频分量(细节分量);且每次输出采样率都减半,以保证总的输出序列长度不变,这样就将原始信号进行了多分辨率分解,从而得到多级包含低频分量的序列  $a_1[n], a_2[n], \dots$  和包含高频分量的序列  $d_1[n], d_2[n], \dots$ 。对应的小波重构如图 2(b) 所示,用对偶低通滤波器  $h$  和对偶高通滤波器  $g$  对插值后的序列滤波,可得到重构信号  $a_0$ 。

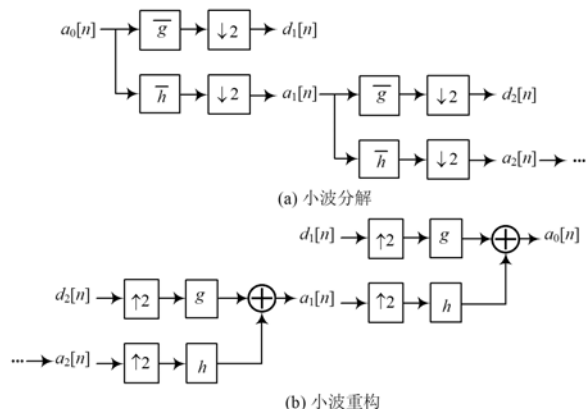


图2 小波分解和重构

Daubechies 小波在给定阶数时,具有最大消失矩和良好平滑度,且分析级数为 6 时,有利于基频的提取<sup>[9]</sup>。因此,本文采用 db6 小波对带通滤波后的信号进行 2 级小波分解,得到低频分量序列  $a_2[n]$  和低频分量序列  $d_1[n], d_2[n]$ 。将  $d_1[n], d_2[n]$  置为零,进行小波重

构,就可以得到 1 000 Hz 以下的低频子带信号。这样就去除了高频噪声和第三、四共振峰的影响,但第一、二共振峰有时依然存在,使得浊音段语音信号的周期性模糊,产生了错误的基音估计。大量实验发现,宽度  $N=9$  的数值滤波器可有效去除这一影响,突出了浊音信号的周期性,使基音估计可靠<sup>[10]</sup>。因此,在小波变换后级联一数值滤波器,其传递函数为:

$$H(z) = \frac{1}{N} \sum_{n=0}^{N-1} z^{-n} \quad (1)$$

### 1.2 利用时域波形信息检测基音周期

从语音信号的时域波形判断基音周期时,并不只是通过其幅值来判断,而是综合利用了波峰的宽度、面积、幅值大小等信息,并根据这些信息的匹配程度检测基音周期。

语音信号经预处理后得到了周期性较好的信号,极大值较清晰。为了减少计算量,设置一阈值对浊音帧进行中心削波处理,以去除幅度较小的峰值来减少要比较波峰的数目,阈值可取为该浊音帧全局最大值的 55%~65%。这里对每个波峰采用四维特征向量  $\mathbf{B}(y_1, y_2, y_3, y_4)$  来描述<sup>[7]</sup>。其中,  $y_1$  表示波峰的正负,  $y_2$  表示波峰的最大幅度,  $y_3$  表示波峰的面积,  $y_4$  表示波峰的底部宽度,利用该特征向量来判断波峰间的相似程度。

对两个不同的特征向量  $\mathbf{B}_i(y_1, y_2, y_3, y_4)$  和  $\mathbf{B}_j(y'_1, y'_2, y'_3, y'_4)$  (其中下标  $i, j$  分别代表第  $i, j$  个波峰),如果下式成立:

$$y_1 = y'_1 \text{ 并且 } |y_i - y'_i| / (y_i + y'_i) < \alpha, i = 2, 3, 4 \quad (2)$$

则这两个波峰是相似的。 $\alpha$  是给定的门限值,由实验确定。

如果两个波峰  $\mathbf{B}_i$  和  $\mathbf{B}_j$  相似,且  $\mathbf{B}_i$  和  $\mathbf{B}_j$  的时间间隔在所有与  $\mathbf{B}_i$  相似的  $\mathbf{B}_j$  中是最小的,则可以说  $\mathbf{B}_i$  与  $\mathbf{B}_j$  是匹配的。匹配的两个波峰之间的距离即为基音周期。

## 2 实验验证

实验所用原始语音波形如图 3 所示。对原始语音分帧(一帧 30 ms, 帧移 10 ms),分别用自相关(ACF)法、平均幅度差(AMDF)法及本文算法进行基音检测,结果如图 4 所示。语音信号以 16 kHz 采样,16 位量化,数据长度为 48 452 点,幅度绝对值小于等于 1。

由图 4 知,对纯净语音进行基音检测时,ACF 法在第 145 帧,146 帧,149 帧,151 帧出现半频;AMDF 法在第 125 帧,126 帧,127 帧,128 帧,129 帧出现倍频;本文算法则无倍频、半频错误,基音轨迹较平滑,效果最好。

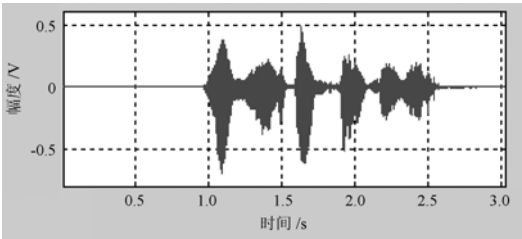


图 3 原始语音信号

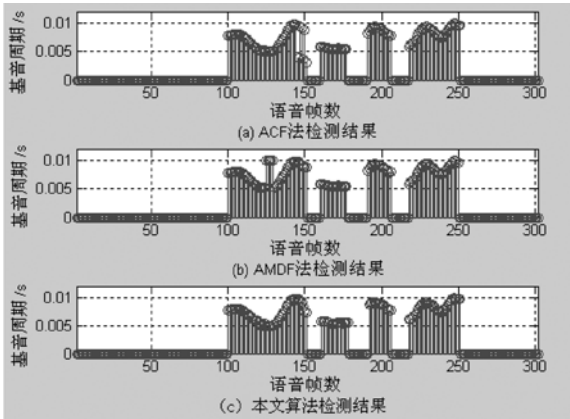


图 4 ACF 法、AMDF 法及本文算法基音检测结果

为了检测本文算法的鲁棒性,在原始语音中加入高斯白噪声,分别用 ACF 法、AMDF 法及本文算法在不同信噪比下对语音进行基音检测,结果如表 1 所示。实验采用上述语音信号(共 302 帧),表中给出了不同信噪比下倍频帧数、半频帧数及无法检测的帧数。

表 1 不同信噪比下 ACF 法、AMDF 法及本文算法的检测结果

信噪比 /dB	ACF 法			AMDF 法			本文算法		
	倍频 /帧	半频 /帧	无法 检测 /帧	倍频 /帧	半频 /帧	无法 检测 /帧	倍频 /帧	半频 /帧	无法 检测 /帧
∞	0	4	0	5	0	0	0	0	0
10	0	6	0	6	0	0	0	0	0
5	0	6	0	7	0	0	0	0	0
0	0	8	0	9	0	0	1	1	0

由表 1 可知,ACF 法易出现半频错误,AMDF 法易出现倍频错误,且两者随着信噪比的降低,错误帧数呈增加趋势。这是因为 ACF 法易受声道共振峰的影响;AMDF 法对语音信号的幅度变化比较敏感,且短时 AMDF 随着滞后时间的增加,求和项减少,峰值幅度逐渐下降,使得谷值点检测比较困难。本文算法中采用了

小波变换和数值滤波处理,有效地去除了共振峰及高频噪声的影响,所以检测效果最好,在信噪比为 0 dB 时仅有 1 帧出现半频,1 帧出现倍频。因此,本文算法具有更好的检测效果和噪声鲁棒性。

3 结 论

本文提出了一种基于小波变换和时域波形信息的基音检测算法。该算法利用小波变换和数值滤波对语音信号进行预处理,有效去除了共振峰和高频噪声的影响;再利用信号波形的多种特征,如波峰的宽度、面积、幅值大小等信息,根据这些信息的匹配程度检测基音周期。实验表明,该算法有效减少了倍频、半频错误,基音轮廓清晰、平滑,无大的跳变,符合语音信号慢时变性的一般规律,其检测性能明显优于自相关法和平均幅度差法。

参 考 文 献

[1] 赵力. 语音信号处理[M]. 北京:机械工业出版社,2003.

[2] RABINER L R. On the use of autocorrelation analysis for pitch detection [J]. IEEE Trans. on Acoustics, Speech, Signal Processing, 1977, 25(1): 24-33.

[3] ROSS M J, SHAFFER H L, COHEN A, et al. Average magnitude difference function pitch extractor [J]. IEEE Trans. on Acoustics, Speech, Signal Processing, 1974, 22 (5): 353-362.

[4] 王晓亚. 倒谱在语音的基音和共振峰提取中的应用[J]. 无线电工程,2004,34(1):57-61.

[5] KADAMBE S, BOUDREAUX-BARTELS G F. Application of the wavelet transform for pitch detection of speech signals [J]. IEEE Trans. on IT, 1992, 38(2): 917-924.

[6] MALLAT S. A theory for multiresolution signal decomposition; the wavelet representation [J]. IEEE Trans. on pattern analysis and machine intelligence, 1989, 11 (7): 674-693.

[7] 李炳照,蔡涛,许天周. 一种小波变换与时域波形相结合的基音周期检测方法[J]. 电声技术,2001(9):3-6.


[8] 许刚,黄冰. 基于小波变换和归一化自相关的基音检测算法 [J]. 桂林电子工业学院学报,2003,23(6):1-5.

[9] 岳东剑,柴佩琪,宣国荣. 应用小波分析技术进行汉语语音基音频率分析的研究[J]. 模式识别与人工智能,2000,13 (2):235-238.

[10] 鲍长春,樊昌信. 基于归一化互相关函数的基音检测算法 [J]. 通信学报,1998,19(10):27-31.

作者简介: 陈小利 女,1979 年出生,河南洛阳人,硕士研究生,工程师。主要研究方向为语音信号处理。  
徐金甫 男,1965 年出生,河南新密人,博士,教授。主要研究方向为语音信号处理、EDA 技术。

# 基于小波变换和时域波形的基音检测算法

作者: 陈小利, 徐金甫, CHEN Xiao-li, XU Jin-fu  
作者单位: 陈小利, CHEN Xiao-li (中国人民解放军96550部队, 河南洛阳, 471031), 徐金甫, XU Jin-fu (信息工程大学电子技术学院, 河南郑州, 450004)  
刊名: 现代电子技术   
英文刊名: MODERN ELECTRONICS TECHNIQUE  
年, 卷(期): 2011, 34(1)

## 参考文献(10条)

1. 赵力 语音信号处理 2003
2. RABINER L R On the use of autocorrelation analysis for pitch detection[外文期刊] 1977(01)
3. ROSS M J;SHAFFER H L;COHEN A Average magnitude difference function pitch extractor 1974(05)
4. 王晓亚 倒谱在语音的基音和共振峰提取中的应用[期刊论文]-无线电工程 2004(01)
5. KADAMBE S;BOUDREAUX-BARTELS G F Application of the wavelet transform for pitch detection of speech signals[外文期刊] 1992(02)
6. MALLAT S A theory for multiresolution signal decomposition:the wavelet representation[外文期刊] 1989(07)
7. 李炳照;蔡涛;许天周 一种小波变换与时域波形相结合的基音周期检测方法[期刊论文]-电声技术 2001(09)
8. 许刚;黄冰 基于小波变换和归一化自相关的基音检测算法[期刊论文]-桂林电子工业学院学报 2003(06)
9. 岳东剑;柴佩琪;宣国荣 应用小波分析技术进行汉语语音基音频率分析的研究[期刊论文]-模式识别与人工智能 2000(02)
10. 鲍长春;樊昌信 基于归一化互相关函数的基音检测算法[期刊论文]-通信学报 1998(10)

## 本文读者也读过(10条)

1. 张起贵, 沈江峰. ZHANG Qi-gui, SHEN Jiang-feng 一种实时基音检测算法[期刊论文]-中北大学学报2008, 29(1)
2. 楼红伟, 赵建伟, 胡光锐 一种小波加权的基音检测方法[期刊论文]-上海交通大学学报2003, 37(3)
3. 蔡润身, 朱耀庭, 陈希, 史绍强, 杨美艳. CAI Run-shen, ZHU Yao-ting, CHEN Xi, SHI Shao-qiang, YANG Mei-yan 基于小波变换和动态验证的基音检测算法[期刊论文]-计算机工程与设计2010, 31(11)
4. 刘永高 语音合成中基音检测算法研究与实现[学位论文]1998
5. 郑继明, 王劲松. ZHENG Ji-ming, WANG Jin-song 语音基音周期检测方法[期刊论文]-计算机工程2010, 36(10)
6. 徐国庆, 杨丹. Xu Guo-qing, YANG Dan 基于Daubechies小波的基音检测[期刊论文]-武汉化工学院学报2005, 27(4)
7. 宋兵, 顾传青, 张建军, Song Bing, GU Chuan-qing, ZHANG Jian-jun 一种新的基于小波变换的基音检测算法[期刊论文]-上海大学学报(英文版) 2005, 9(4)
8. 徐金甫, 陈小利. XU Jin-fu, CHEN Xiao-li 基于线性预测的综合基音检测法[期刊论文]-计算机工程与设计2008, 29(4)
9. 樊桂花, 唐斌 基于小波变换的语音基音周期鲁棒性检测[期刊论文]-信号处理2004, 20(4)
10. 岳东剑, 柴佩琪, 宣国荣. Yue Dongjian, Chai Peiqi, Xuan Guorong 应用小波分析技术进行汉语语音基音频率分析的研究[期刊论文]-模式识别与人工智能2000, 13(2)

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_xddzjs201101024.aspx](http://d.g.wanfangdata.com.cn/Periodical_xddzjs201101024.aspx)