

基于小波变换和空域相关函数的基音周期检测及其在 DSP 上的实现

陈 宁, 胡安峰, 赖旭芝

(中南大学 信息科学与工程学院, 湖南 长沙, 410083)

摘 要: 利用 TI 公司制造的 DSP 芯片 TMS320VC5509 和 CODEC 芯片 TLV320AIC23 构建一个语音采集系统, 为语音信号处理搭建一个有效的平台。具体方法是: 采用 TMS320VC5509 的 I²C 模块对 TLV320AIC23 进行配置, 并通过 TMSVC5509 的 McBSP 实现两者之间的数据通信; 利用小波变换在信号突变检测方面的原理在不同尺度下对采集到的语音信号进行小波变换, 并引入空域相关函数, 利用相邻尺度的相关系数代替传统小波变换方法中的小波系数寻找局部极大值之间的距离。实验结果表明, 系统以 22 kHz 的采样频率采集到语音信号; 通过计算不同尺度下的相关系数局部极大值之间的距离, 能估计出基音周期; 与传统的小波变换方法相比, 基于空域相关函数的小波变换方法在锐化信号边缘的同时抑制了噪声, 对噪声的鲁棒性更强。

关键词: TMS320VC5509 芯片; TLV320AIC23 芯片; 小波变换; 基音周期; 空域相关函数

中图分类号: TP 311

文献标识码: A

文章编号: 1672-7207(2008)04-0805-05

Speech pitch period detection and it's implementation on DSP based on wavelet transform and spatial correlation function

CHEN Ning, HU An-feng, LAI Xu-zhi

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: A speech acquisition system based on DSP chip TLV320AIC23 and CODEC chip TMS320VC5509 used as an efficient platform of speech signal processing was constructed. The detail procedures were as follows: I²C module of TMS320VC5509 was used to configure the TLV320AIC23 and data transfer between the two chips was implemented through McBSP on TMS320VC5509. Pitch period of the speech acquired was detected based on the principle of singularity detection of signals with wavelet transform at different scales. An arithmetic based on spatial correlation function which used the correlation coefficients to replace the wavelet transform coefficients to look for the distance between the local maximums was proposed. The experimental results show that speech signal is acquired at sample frequency of 22 kHz and the pitch period is estimated by finding the distance between the local maximums. Compared with the classic arithmetic using wavelet transform, the arithmetic based on spatial correlation function can enhance sharp edges and suppress noise, and so it has stronger robust to noise.

Key words: TMS320VC5509 chip; TLV320AIC23 chip; wavelet transform; pitch period; spatial correlation function

基音周期是语音信号的重要参数之一, 准确而快速的基音周期检测对语音信号的合成、编码、识别具有重要意义^[1]。由于语音信号变化十分复杂, 基音周期的变换范围又非常大(2~5 ms), 传统的检测方法(如

平均幅度差函数法、倒谱法等)都存在比较明显的缺陷^[2-3]。近年来发展迅速的小波变换, 由于具有良好的空间域和频率域局部化特性, 特别适合于处理非平稳信号^[4]。特别是小波变换的快速算法 Mallat 算法被提

收稿日期: 2007-11-15; 修回日期: 2008-03-02

基金项目: 中国博士后科研基金资助项目(20060390883); 国家杰出青年科研基金资助项目(60425310)

通信作者: 陈 宁(1970-), 女, 湖南长沙人, 博士, 副教授, 硕士生导师, 从事数字信号处理、非线性电路和分散控制研究; 电话: 13875915950;

E-mail: ningchen@mail.csu.edu.cn

出以后,小波变换在工程领域的应用日益广泛。从语音的产生过程可知,声门的开、闭,使得语音信号的浊音段呈现准周期变化。在声门闭合的瞬间,声道受到较强烈的刺激,在语音信号中引起一个锐变。如果能检测到这一锐变,等同于检测出声门的闭合时刻,那么,相邻 2 次声门闭合的时间间隔就是基音周期^[5]。利用小波变换在信号突变检测方面的原理,在不同的扩展因子下对信号进行小波变换,就可以检测出声门闭合时的锐变信息^[6]。对相邻尺度小波系数计算空域相关系数,可以在锐化信号边缘的同时抑制噪声,与传统的小波变换检测方法相比,对噪声有更好的鲁棒性。TMS320VC55X 是德州仪器(TI)C5000 DSP 系列中最新的一代产品。C5000 系列的主要特点是低功耗,非常适合在音频处理方面的应用^[7]。TLV320AIC23(以下简称 AIC23)是 TI 公司生产的一款高性能立体声音频编解码器,支持频率为 8~96 kHz 的采样率。该芯片在回放中的功率消耗小于 23 mW,因此,对于可移动的数字音频播放和录音使用中的模拟输入输出等应用系统, AIC23 无疑是十分理想的选择^[8]。在此,本文作者利用 AIC23 采集语音信号,传送给 TMS320VC5509(以下简称 VC5509)进行小波变换,之后计算相邻尺度的相关系数,通过相关系数局部极大值之间的距离来判断基音周期。

1 小波变换用于信号突变检测的原理

函数 $f(t)$ 的连续小波变换定义为:

$$WTf(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \varphi\left(\frac{b-t}{a}\right) dt = f * \varphi_a(b) \quad (1)$$

其中: $WTf(b, a)$ 表示 $f(t)$ 在尺度为 a , 平移参数为 b 时的小波变换结果。 $\varphi(t)$ 满足

$$\int_{\mathbb{R}} \varphi(t) dt = 0, \quad (2)$$

称 $\varphi(t)$ 为母小波函数。

若实函数 $\theta(t)$ 满足

$$\int_{-\infty}^{\infty} \theta(t) dt = 0, \quad \lim_{x \rightarrow \pm\infty} \theta(t) = 0, \quad (3)$$

则称 $\theta(t)$ 为光滑函数。若令

$$\varphi(t) = \frac{d\theta(t)}{dt}, \quad (4)$$

则 $\varphi(t)$ 就可以作为母小波函数。此时,式(1)可以变换为:

$$WTf(b, a) = f * \varphi_a(b) = f * \left(a \frac{d\theta_a}{dt} \right) (t) = a \frac{d}{dt} (f * \theta_a)(t) \quad (5)$$

从式(5)可以看出,小波变换 $WTf(b, a)$ 是信号 $f(t)$ 被带有压扩因子 a 的 $\theta(t)$ 平滑后的 1 阶倒数。所以, $WTf(a, b)$ 的极值点对应于信号 $f * \theta_a(t)$ 的锐变点^[9]。这就是小波变换用于信号突变检测的原理,也是本文用小波变换进行基音周期检测的依据。

2 用于基音检测的小波变换

通过 AIC23 采集来的语音信号是以离散序列的形式存储的,要用 VC5509 对其进行分析、处理,采用连续小波变换显然不能满足要求。若令尺度参数 a 取 2 的整数次幂,平移参数 b 均匀离散取值为 $n(n \in \mathbb{Z})$, 则式(1)可化为:

$$WTf(n, 2^j) = 2^{-j/2} \int_{-\infty}^{\infty} f(t) \varphi\left(\frac{n-t}{2^j}\right) dt = f * \varphi_{2^j}(n) \quad (6)$$

这时得到的小波变换称为二进小波变换。二进小波变换正是本文用于基音周期检测的小波变换。

理论上说,必须计算所有整数 j 下压扩因子为 $a=2^j$ 的小波变换。但是,语音信号有各自的特性。下面通过经验知识大体确定分解尺度的范围。

Mallat 算法的实质就是用级联滤波器组实现小波变换而不是用传统的卷积方法计算。级联滤波器组中用到的为理想半带低通滤波器和理想半带高通滤波器^[10]。每经过 1 次变换,送入下一级分解的原始信号的频率范围就变为原来的一半。事实上,只要每次进行小波变换的信号包含基音的信息,就可以通过这段信号的小波变换检测出信号的突变点。

对于语音信号,清音主要占据高频段,无基音周期,包含基音周期的只有浊音。浊音的频段为 30~125 Hz。采集的原始语音信号频段为 0~4 kHz。因此,若 $j=5$, $WTf(n, 2^5)$ 仍可以用于检测基音周期。本文采用 $j=3, 4, 5$ 共 3 个尺度进行小波变换。

3 在基音检测中引入空域相关函数

采用经典的小波变换算法提取基音周期仍存在不足:一是在低尺度的小波系数中会存在许多由突发噪声以及高次谐波引起的伪极大值点;二是由同一个声门关闭时刻引起的局部极值点在不同的尺度下位置不尽相同。

Witkin 提出了利用尺度空间相关性来对信号滤波的思想^[11]:对含噪信号经过子带分解后,从粗尺度到

细尺度逐步搜索信号的主要边缘, 最终从噪声背景中得到真实信号。信号的突变点在不同尺度的同一位置都有较大的峰值出现, 噪声能量却随着尺度的增大而减少, XU 等^[12]基于上述原理提出了利用空域相关函数对信号去噪。对相邻尺度的小波系数直接相乘进行相关计算, 这将在锐化信号边缘的同时抑制噪声。声门闭合时刻表现在语音波形上就是瞬间幅度剧增, 产生突变。因此, 空域相关函数非常适合在强噪声下对信号提取基音周期。

设分解的最大尺度为 2^j , $WTf(n, 2^j)$ 表示尺度 2^j 上位置 n 处含噪信号 f 的离散小波变换, 取相邻尺度的变换值进行相关计算, 定义

$$\text{corr}(n, 2^j) = \prod_{i=0}^{j-1} WTf(n, 2^{(j+i)}) \quad (7)$$

其中: $j < J-1$ 。为了使相关系数与小波系数具有可比性, 将 $\text{corr}(n, 2^j)$ 的能量归一化到 $WTf(n, 2^j)$ 上。定义归一化相关系数为:

$$\text{res corr}(n, 2^j) = \text{corr}(n, 2^j) \sqrt{P_w(2^j) / P_{\text{corr}}(2^j)} \quad (8)$$

其中: $n=1, 2, \dots, N$; $P_{\text{corr}}(2^j) = \sum_{n=1}^N \text{corr}(n, 2^j)^2$;

$$P_w(2^j) = \sum_{n=1}^N WTf(n, 2^j)^2$$

通过比较 $\text{res corr}(n, 2^j)$ 与 $WTf(n, 2^j)$ 便可以鉴别信号的突变点。相关量 $\text{corr}(n, 2^j)$ 中对应信号边缘点的数据较对应噪声的数据大很多, 且经归一化后也应大于 $WTf(n, 2^j)$ 中由信号边缘引起的系数; 然而, 在信息检测过程中, $\text{corr}(n, 2^j)$ 中某些对应噪声的数据归一化后也可能会大于相应的 $WTf(n, 2^j)$ 。因此, 本文中,

$$\text{new corr}(n, 2^j) = c(2^j) \cdot \text{corr}(n, 2^j) \sqrt{P_w(2^j) / P_{\text{corr}}(2^j)} \quad (9)$$

通过大量试验, 发现当 $c(2^j)=0.85$ 时, 基音检测效果比较理想。

4 语音信号的获取

为了使 AIC23 在预期的工作模式下工作, 必须对其内部相应的寄存器进行配置。AIC23 内部提供了 11 个配置寄存器。利用 VC5509 的 I²C(Inter-Integrated Circuit, 内部集成电路)模块可以对这些寄存器进行编程配置, 使 AIC23 工作在要求状态下。在本文中配置

AIC23 采样频率为 22 kHz, 采样位为 16 位, 单声道采样。具体的配置方法见文献[8]。

通过 VC5509 的 McBSP(Multichannel Buffered Serial Ports, 多通道缓存串口)接收和发送采样的音频数据。控制通道只在配置 AIC23 的内部寄存器时工作, 而当传输音频数据时则处于闲置状态。AIC23 通过麦克风输入或者立体声音频输入采集模拟信号, 并把模拟信号转化为数字信号, 存储到 DSP 的内部 RAM 中, 以便 DSP 的处理。

AIC23 与 VC5509 接口如图 1 所示。系统中 AIC23 的主时钟频率 12 MHz 直接由外部的晶振提供。MODE 接数字地, 表示利用 I²C 控制接口对 AIC23 传输控制数据。CS 接数字地, 定义了 I²C 总线上 AIC23 的外设地址(通过将 CS 接到高电平或低电平, 可以选择 AIC23 作为从设备在 I²C 总线上的地址。具体情况见表 1)。SCLK 和 SDIN 是 AIC23 控制端口的移位时钟和数据输入端, 分别与 VC5509 的 I²C 模块端口 SCL 与 SDA 相连。收发时钟信号 CLKX1 和 CLKR1 由 AIC23 的串行数据输入时钟 BCLK 提供, 并由 AIC23 的帧同步信号 LRCIN 和 LRCOUT 启动串口数据传输。DX1 和 DR1 分别与 AIC23 的 DIN 和 DOUT 相连, 从而完成 VC5509 与 AIC23 间的数字信号通信^[13]。

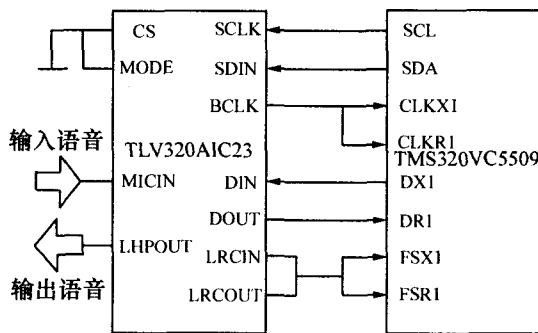


图 1 AIC23 与 VC5509 接口原理图

Fig.1 Interface principle between AIC23 and VC5509

表 1 从设备地址和 CS 端状态的对应关系

Table 1 Address of slave device corresponded to state of CS pin

CS 状态(default=0)	地址
0	0011010
1	0011011

通过以上配置, 从麦克风输入自然语音, AIC23 就按照指定的格式将模拟语音信号转换为数字信号, 并存储在 DSP 的内部 RAM 中, 为后面的小波变换提供了样本。

5 小波变换检测基音周期的步骤

小波变换检测基音周期的步骤如下。

a. 选取小波函数。本文选择了 DB4 小波作为小波函数。需要说明的是, 选择不同的小波函数, 得到的结果不尽相同, 甚至差距比较大。至于选取什么样的小波函数, 目前尚无普遍适用的准则, 要靠不断实验予以验证。经过验证, 本实验中用 DB4 小波函数进行小波变换得到的结果比较理想。对应于 Mallat 算法, DB4 小波函数对应的高通滤波器系数为: $h_0=-0.23038$, $h_1=0.71485$, $h_2=-0.63088$, $h_3=-0.27984$, $h_4=0.18703$, $h_5=0.030841$, $h_6=-0.32883$, $h_7=-0.010597$ 。低通滤波器系数为: $I_0=-0.01597$, $I_1=0.032883$, $I_2=-0.030841$, $I_3=-0.18703$, $I_4=0.027984$, $I_5=0.63088$, $I_6=0.71485$, $I_7=0.23038$ 。

b. 取出定长的语音信号片段, 分别以压扩因子 $2^3, 2^4, 2^5$ 进行二进小波变换^[14]。

c. 在每个压扩因子下, 将系数全局极大值的 0.8 倍作为门限^[12], 小于门限的为清音, 基音周期为 0。

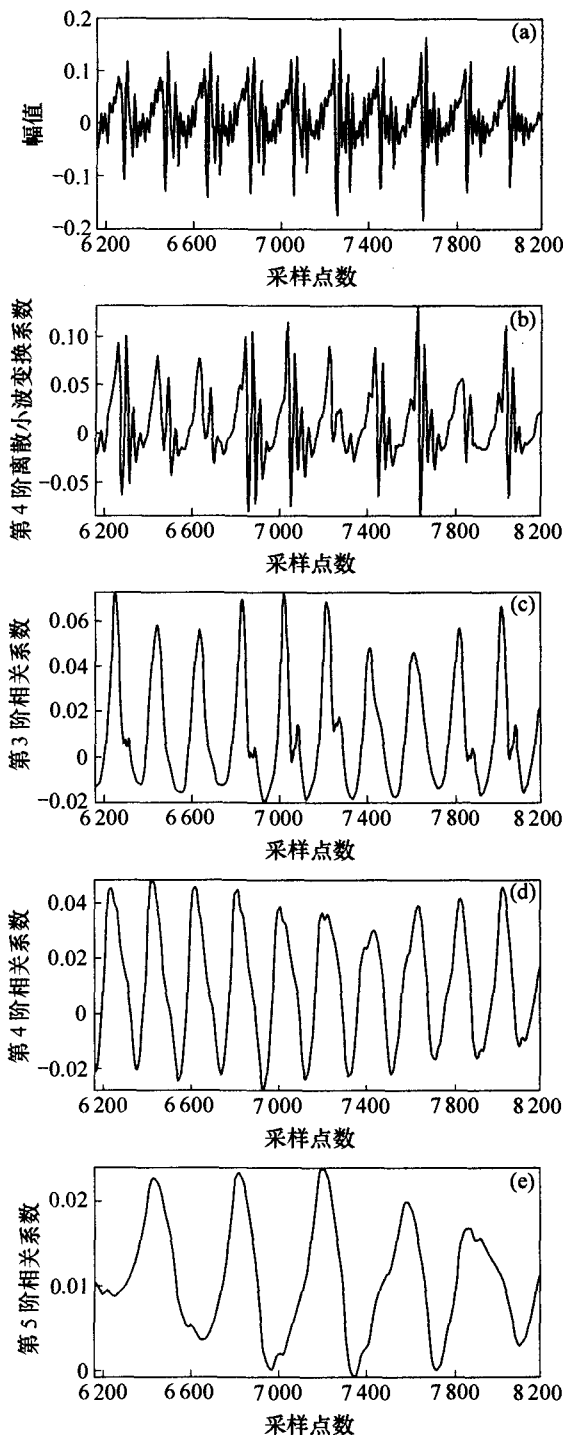
d. 求 $\text{newcorr}(n, 2^j)$ 。比较不同尺度下 $\text{newcorr}(n, 2^j)$ 局部极大值的位置。若在相邻 2 个压扩因子下局部极大值的位置接近或重合, 则较小尺度因子下局部极大值的位置对应于声门的闭合时刻。若相邻压扩因子下局部极大值位置相差很大, 则说明较大压扩因子下的分解已经不能完全包含基音的信息, 不能作为检测结果。

6 实验结果与分析

图 2 所示为一男音“搭”的基音周期检测结果。

图 2(a)所示为在 CCS(Code composer studio, 代码集成开发环境)界面中观察到的原始语音信号。可以看出, 语音采集平台确实采集到了从麦克风输入的自然语音。在这里截取样本的 2 000~6 200 点。本文中以 22 kHz 的频率采样, 故这段语音的长度大约为 91 ms, 而基音周期的范围为 3~25 ms。因此, 这段语音信号包含足够的基音周期, 可以用于进行基音分析。这段语音信号呈现准周期性, 但由于受噪声等因素的影响, 周期性并不明显。

图 2(b)所示为尺度 2^4 下对信号进行小波变换的结果。可以看出, 经过小波变换, 信号的锐变点比较明显, 具有比图 2(a)具有更明显的周期性。但是, 由于受噪声的影响, 变换后的波形中含有许多伪极大值点,



(a) 原始语音信号; (b) 尺度 2^4 时的小波变换; (c) 尺度 2^3 时的相关系数; (d) 尺度 2^4 时的相关系数; (e) 尺度 2^5 时的相关系数

图 2 基音周期检测结果

Fig.2 Results of pitch period estimation

为基音周期的识别带来了困难, 检测错误的概率较大。

图 2(c)所示为尺度 2^3 时相关系数的计算结果。从图 2(c)可以看出, 经过相关系数的计算, 信号的锐变

点得到加强,而同时由噪声所引起的伪极大值点得到较大抑制。在图2(c)中观察的基音周期比图2(b)中的基音周期更加明显。图2(d)所示为尺度 2^4 下相关系数的计算结果。与图2(c)相比,噪声的影响基本消除,可以很明显地看出相邻2个局部极大值之间的距离约为200个采样间隔,即大约9ms。从图2(b),2(c)和2(d)可以看出,通过小波变换后,在相邻尺度上计算相关系数来寻找局部极大值,比采用传统的小波变换方法对噪声的鲁棒性更强。

图2(e)所示为尺度 2^5 下的相关系数计算结果。显然,与图2(c)和图2(d)相比,局部极大值的位置已经出现严重的偏差。限于语音信号本身的特性,分解尺度的上界为 2^5 ,尺度 2^5 下相关系数的计算用到了尺度 2^6 下小波变换的结果,而尺度 2^6 下小波变换的结果不能用于判断声门的闭合时刻。

7 结论

a. 基于TI公司制造的DSP芯片TMS320VC5509和CODEC芯片TLV320AIC23构建了一个语音采集系统。实验结果表明,基于AIC23和VC5509的语音信号采集系统工作正常,采集到了从麦克风输入的语音信号。

b. 利用从芯片TLV320AIC23采集到的一段语音信号,在DSP芯片TMS320VC5509中对其运用Mallat算法进行小波变换,并引入空域相关函数,利用计算不同尺度下的相关系数来代替传统的小波变换系数,通过判断不同尺度下相关系数局部极大值的位置来检测基音周期。

c. 引入空域相关函数的小波变换检测方法比传统的小波变换方法具有更好的抗噪声性能,基音周期检测的正确性也得到很大提高。

d. 在基音周期检测过程中,小波函数的选择非常重要。如何选择适合于基音周期检测的小波函数,有待进一步研究。

参考文献:

[1] Rabiner L R, Cheng M J, Rosenberg A E, et al. A comparative performance study of several pitch detection algorithms[J]. *IEEE Trans on Acoustics, Speech, and Signal Processing*, 1976, 24(5): 399-418.

[2] 刘建,郑方,吴文虎. 基于幅度差平方和函数的基音周期提取算法[J]. *清华大学学报:自然科学版*, 2006, 46(1): 74-77.

LIU Jian, ZHENG Fang, WU Wen-hu. Real-time pitch tracking

based on sum of magnitude difference square function[J]. *Journal of Tsinghua University: Science and Technology*, 2006, 46(1): 74-77.

- [3] 柏静,韦岗. 一种基于线性预测与自相关函数法的语音基音周期检测新算法[J]. *电声技术*, 2005, 29(8): 43-46.
- BAI Jing, WEI Gang. A new method for speech signals pitch detection based on LPC and autocorrelation[J]. *Voice Technology*, 2005, 29(8): 43-46.
- [4] 李硕,李冰洋,王蜜. 小波变换及其在语音信号处理中的应用[J]. *哈尔滨师范大学自然科学学报*, 2006, 22(4): 20-24.
- LI Shuo, LI Bing-yang, WANG Mi. Wavelet transform and its application in analyzing speech signals[J]. *Natural Science Journal of Harbin Normal University*, 2006, 22(4): 20-24.
- [5] 王长富,戴蓓倩,李辉. 基于声门闭合时刻的语音基音周期的提取[J]. *中国科学技术大学学报*, 1998, 28(3): 362-367.
- WANG Chang-fu, DAI Bei-qian, LI Hui. Speech pitch period detection based on GCI[J]. *Journal of China University of Science and Technology*, 1998, 28(3): 362-367.
- [6] 刘娟花,李福德. 基于小波变换的信号去噪研究[J]. *西安理工大学学报*, 2004, 20(3): 289-292.
- LIU Juan-hua, LI Fu-de. Research on signal denoising based on wavelet transformation[J]. *Journal of Xi'an University of Technology*, 2004, 20(3): 289-292.
- [7] Texas Instruments Incorporated. TMS320VC55x 系列 DSP 的 CPU 与外设[M]. 彭启淙,武乐琴,张舰,等编译. 北京:清华大学出版社,2005: 151-388.
- Texas Instruments Incorporated. The CPU and peripheral of TMS320VC55x[M]. PENG Qi-cong, WU Le-qin, ZHANG Jian, translate and edit. Beijing: Tsinghua University Press, 2005: 151-388.
- [8] Texas Instruments. Stereo audio CODEC, 8 to 96 kHz, with integrated headphone amplifier T LV320AIC23 data manual [EB/OL]. [2001-06-01]. <http://www.ti.com>.
- [9] 胡瑛,陈宁. 一种基于小波变换的清浊音分类基音检测方法[J]. *电声技术*, 2006, 30(11): 63-66.
- HU Ying, CHEN Ning. A method of unvoiced/voiced classification and pitch detection based on wavelet transform[J]. *Voice Technology*, 2006, 30(11): 63-66.
- [10] Mallat S, Zhong S. Characterization of signals from multiscale edges[J]. *IEEE Trans on PAMI*, 1992, 14(7): 710-732.
- [11] WANG Yi-yuan, ZHAO Li-ming. Pitch frequency estimation of chinese multi-syllable words based on wavelet transform[C]// *Proceedings of the Third International Conference on Machine Learning and Cybernetics*. Shanghai, 2004: 26-29.
- [12] XU Yan-sun, Weaver J B, Healy D M, et al. Wavelet transform domain filters: a spatially selective noise filtration technique[J]. *IEEE Trans Image Processing*, 1994, 3 (6): 747-758.
- [13] 王嘉,欧建平. 利用 I2C 总线实现 DSP 与音频采样芯片 TLV320AIC23 的接口控制[J]. *微处理机*, 2003, 4(8): 10-12.
- WANG Jia, OU Jian-ping. The implementation of interface control between DSP and audio codec TLV320AIC23 using I²C bus[J]. *Microprocessors*, 2003, 4(8): 10-12.
- [14] Kadambe S, Boudreaux-Bartels G F. Application of the wavelet transform for pitch detection of speech signal[J]. *IEEE Trans Information Theory*, 1992, 38(2): 917-924.