

基于改进谱减法的语音增强

李海忠, 杨燕翔, 王娟, 梁俊
(西华大学 电气信息学院, 四川 成都 610039)

摘要:针对传统谱减法具有残余音乐噪声过强,清音部分损失严重的缺点,提出了一种利用语音信号的短时平均幅度差特征(AMDF)并结合短时平均幅度(AM)的语音检测算法,在原始语音估计式中引入了参数 α 和 β ,对传统的谱减法进行改进。根据采集的真实带噪语音数据,将传统谱减法和改进的谱减法结果进行了比较分析。在研究基于改进的谱减法的语音增强算法基础上,构建了以 TMS320VC5509 和 TLV320AIC23 Codec 为核心器件的实时系统,能够有效地提高语音信噪比,信号的信噪比由 12.2 dB 提高到了 4.0 dB,改善了语音质量。

关键词: 语音增强; 改进的谱减法; DSP; TMS320VC5509; TLV320AIC23

中图分类号: TP391

文献标识码: A

文章编号: 1674-6236(2010)04-0139-03

Speech enhancement based on improved spectrum subtraction

LI Hai-zhong, YANG Yan-xiang, WANG Juan, LIANG Jun

(School of Electrical and Information Engineering, Xihua University, Chengdu 610039, China)

Abstract: Aiming at reducing musical noise and the loss of unvoiced speech in the spectral subtraction method, a reliable voice active detection algorithm which utilizes the characteristics of short-time AMDF (Average Magnitude Difference Function) and short-time AM (Average Magnitude) to detect the speech endpoint is introduced. And the spectral subtraction is improved by introducing two parameters α and β . According to collecting real speech data of noise, the outcomes of classical spectral subtraction and modified algorithm were compared. On the research of speech enhancement based on improved spectrum subtraction in the thesis, a real-time system centered on TMS320VC5509 and TLV320AIC23 Codec was built, and it increased SNR which was increased from 12.2dB to 4.0dB effectively and improved the quality of speech.

Key words: speech enhancement; improved spectrum subtraction; DSP; TMS320VC5509; TLV320AIC23

随着语音技术的深入研究和实际应用增多,各种语音处理系统需要进一步提高其性能。语音增强是关键技术之一,采用语音增强算法提高语音处理系统前端预处理的抗噪声能力和输入信号的信噪比,因此,语音增强作为预处理或前端处理模块存在于语音处理系统中,并成为语音处理与识别系统的重要组成部分。语音增强可改善音质,提高清晰度和可懂度,减少听觉疲劳。由于信号的相位对语音感知并不重要,所以基于短时谱幅度(STSA)的增强方法应用最为广泛,谱减法及其改进型算法是其中的典型代表。采用 TMS320VC5509 DSP 和 TLV320AIC23 Codec 组成实时系统,实时实现语音增强。

1 原理介绍

1.1 基本谱减法

基本谱减法^[1]即在频域将带噪语音的功率谱减去噪声的功率谱得到语音功率谱估计,开发后就得到语音幅度谱估计,将其相位恢复后采用逆傅里叶变换恢复时域信号。考虑到人耳的感觉不灵敏,相位恢复时所采用的相位是带噪语音

的相位信息。谱相减法的原理图如图 1 所示。

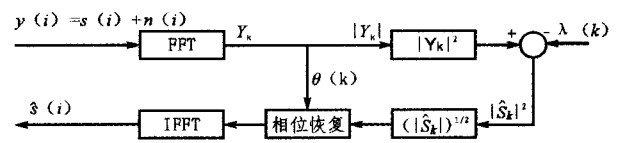


图 1 谱相减法的基本原理

该方法的缺点是增强后的语音中含有明显的“音乐噪声”,这是由频谱相减而产生的一种残留噪声,具有一定的节奏起伏感,故而得名“音乐噪声”。“音乐噪声”产生的原因是因为在谱相减法过程中,是以无声期间统计平均的噪声方差代替当前分析帧的噪声频谱分量。而噪声频谱具有高斯分布,因此相减时,若该帧某频率点噪声分量较大,就会有很大一部分保留,在频谱上呈现随机出现的尖峰,在听觉上形成有节奏性起伏的类似音乐的残留噪声。

1.2 改进谱相减法

在实际的增强过程中,更多使用谱相减法的改进形式^[2]:

$$|\hat{S}_k| = [|Y_k|^a - \beta \lambda_n^a(k)]^{1/a} \quad (1)$$

也就是将增益写成：

$$G_k = (1 - \beta/\gamma_k^\alpha)^{1/\alpha} \quad (2)$$

式中， α 和 β 为调节增益参数。显然，当 $\alpha=2, \beta=1$ 时为普通谱相减法。增大 β ，可增大去噪程度，这样就能减少剩余噪声，从而减弱“音乐噪声”。调节 α 也会达到相似结果。通过适当调节这 2 个参数可使去噪效果达到比较好增强效果。但过多增加去噪程度会使增强后的语音失真增大。

1.3 仿真实验结果分析

实验表明，本系统能够实时、有效提高语音信噪比，增强语音可懂度，改善语音质量达到实时语音增强的目的，用于语音识别、语音编码等数字语音处理系统的前端预处理系统。其去噪结果如图 2 所示。

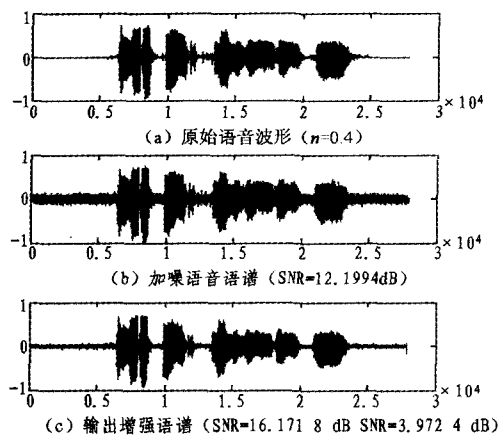


图 2 改进谱减法处理前后信号波形

2 系统硬件设计

该系统硬件设计采用由 TMS320VC5509 型 DSP^[3] 和 TLV320AIC23 型 Codec 组成的实时系统实现语音增强。TMS320VC55x DSP 是德州仪器(TI)公司 C5000 DSP 系列里最新的一代产品。TMS320VC55x DSP 系列主要特点是低功耗，非常适用于音频处理。TLV320AIC23 是一种高性能立体声音频编解码器(Codec)，该器件的数字传输字长可以是 16、20、24、32 bit，支持 8~96 kHz 的采样率。在数模转换器中的二阶多比特结构还可在采样率为 96 kHz 的情况下使信噪比达到 100 dB，从而实现高质量的数字音频回放。该器件在回放中的功耗大于 23 mW，因此可应用于移动的数字音频播放和录音的模拟输入输出等应用系统。图 3 是 TMS320VC5509 DSP 与 TLV320AIC23 Codec 的接口电路连接。

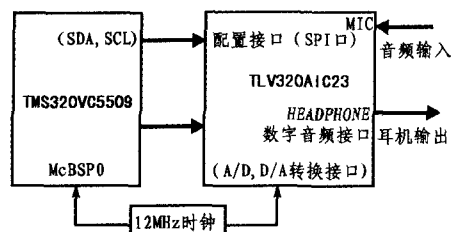


图 3 TMS320VC5509 DSP 与 TLV320AIC23 Codec 的接口电路连接

TLV320AIC23B 的数据接口有 4 种工作方式，分别为：Right justified、Left justified、I²S Mode 和 DSP Mode。其中后 3 种可方便与 DSP 的 McBSP 串口连接，这里是以 DSP Mode 模式进行数据口连接。与 DSP 的 McBSP 无缝连接，唯一需要注意的是：McBSP 的接收时钟和 TLV320AIC23 的 BCLK 都由 McBSP 的发送时钟提供；当 TLV320AIC23 为主设备时，McBSP 的发送与接收时钟均由 TLV320AIC23 提供。其相关电路连接图如图 4 所示。

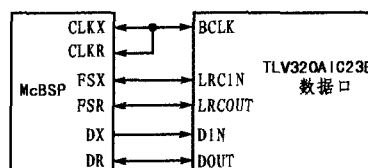


图 4 TMS320VC5509 DSP 和 TLV320AIC23 电路连接

系统中 TLV320AIC23 的主时钟 12 MHz 由外部晶振直接提供。MODE 接数字地，表示利用 I²C 控制接口对 TLV320AIC23 传输控制数据。CS 接数字地，定义 I²C 总线^[4]上 TLV320AIC23 的外设地址，通过将 CS 接到高或低电平，选择 TLV320AIC23 作为从设备在 I²C 总线上的地址。SCLK 和 SDIN 是 TLV320AIC23 控制端口的移位时钟和数据输入端，分别与 VC5509 的 IC 模块端口 SCL 和 SDA 相连。收发时钟信 CLKX1 和 CLKR1 由 TLV320AIC23 的串行数据输入时钟 BCLK 提供，并由 TLV320AIC23 的帧同步信号 LRCIN、LRCOUT 启动串口数据传输。DX1 和 DR1 分别与 TLV320AIC23 的 DIN 和 DOUT 相连，从而完成 TMS320VC5509 与 TLV320AIC23 间的数字信号通信。

3 系统软件设计

图 5 是语音信号在基于 DSP 的语音处理系统处理的软件流程。

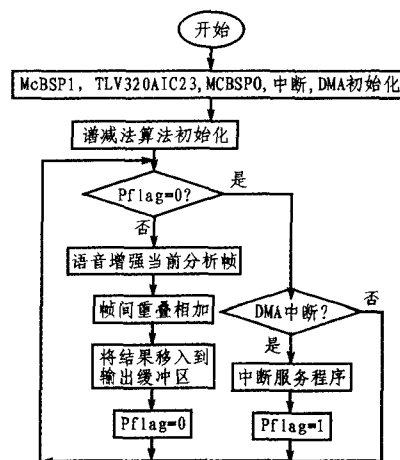


图 5 软件程序流程

3.1 初始化 CPU 频率

首先正确配置 DSP 的时钟发生器，产生 DSP 工作的正常时钟频率。该步骤的关键是数字锁相环 PLL 的应用。为了

在程序中改变 DSP 的工作频率,需设置时钟发生器的设置模式寄存器 CLKMD 中的倍频值 PLLMULT 位和分频值 PLLDIV 位,使 PLL 工作在锁定模式下。而在进入锁定模式之前,必须屏蔽 PLL,使时钟发生器工作在旁路模式,这样才能改变 PLLMUTL 位和 PLLDIV 位的值,从而实现分频和倍频。因此 PLL 初始化应当遵循以下步骤:1)关闭 PLL,时钟发生器进入旁路模式,PLL 开始锁相;2)置 CLKMD 中的相应标志位;3)设置倍频值和分频值;4)使能 PLL,PLL 工作在锁定模式。 $f_{OUT}=[PLLMULT/(PLLDIV+1)]f_{IN}^{[5]}$ 。

3.2 通过 PC 对 TLV320AIC23 编程控制

I²C 模块是 TMS320VC55xx 新增的片内集成外设,可以使 DSP 与 PC 兼容设备通过该接口进行数据通信。I²C 模块接口由串行数据信号 SDA 和串行时钟信号 SCL 组成,SDA 和 SCL 均为双向接口。连接在同一条总线上的 PC 设备工作在多主线性工作模式下。包括 TMSVC55xx DSP 在内的每个 PC 设备都有唯一的设备地址,可供软件寻址。其中,主设备发送时钟并启动数据传输,被主设备寻址的则为从设备。这些设备根据各自的功能,既可以作为发送器,也可作为接收器。这里应注意:

1)在写入 I²C 地址后,只能对一个寄存器进行写操作,而不能对所有寄存器进行写操作。也就是说,对每个寄存器写之前都要写入 I²C 地址。通过在 TLV320AIC23 的写入函数中调用函数 I2C_Write()来实现;

2)I²C 模式下,数据分 3 个 8 位写入。而 TLV320AIC23 有 7 位地址和 9 位数据,也就是说,需要把数据项上面的最高位补充到第 2 个 8 位中的最后 1 位;

3)在对控制寄存器的编程过程中,对应于每一次工作状态的变化,不能仅仅修改某个寄存器的值,而是要对这 10 个寄存器都重新写入,否则系统将无法正常工作,而且应首先写 Rgis2ter10,同时还应对所有寄存器进行复位处理。

3.3 设置 McBSP

一般 McBSP 通过 6 个引脚使数据通路和控制通路与外部设备相连。数据经 McBSP 串口与外设的通信是通过 DR 和 DX 引脚传输的,控制同步信号则由 CLKX、CLKR、FSX、FSR 等引脚实现。McBSP 通道一次可移入或移出一串字。每个 McBSP 最多支持 128 个发送数据通道和 128 个接收数据通道。在 DSP 复位后,RMCM=0,XMCM=0,表示不使用多通道选择模式,使能所有的 128 通道。TMS320VC5509 的

McBSP0 应工作在 SPI 模式下,以便使 McBSP0^[6]的接收器和发送器同步。这时,McBSP0 作为从设备,主时钟和从设备使能信号,由外部的设备产生(TLV320AIC23)。相应的,CLKX 和 FSX 应设置为输入引脚。这样,发送和接收电路都由外部主时钟定时。FSX 引脚内部与 FSR 信号连接。尽管 CLKX 信号由外部主设备产生,且与 McBSP0 异步,仍应使能 McBSP0 的采样率发生器,才能使 SPI 从设备正常运行。采样率发生器的最大频率应设为 CPU 时钟频率的 1/2。内部采样频率时钟用来将 McBSP0 逻辑与外部主时钟和从设备使能信号同步。由于 McBSP 串口的数据线 DR 和 DX 带有缓存寄存器,而帧同步信号 FSX、FSR 以及时钟信号 CLKX、CLKR 具有可编程性,因此它与 TLV320AIC23 之间的接口设计非常灵活。由 TLV320AIC23 的帧同步信号 LRCIN、LRCOUT 启动串口传输,同时将发送接收的数据字长设定为 32 位(左声道 16 位,右声道 16 位)单帧模式,就可方便地实现与 TLV320AIC23 之间的无缝连接。

4 结论

通过实验验证,本文所提出的语音增强算法对噪声背景下的语音增强具有一定的效果,尤其对非平稳噪声污染的语音,其效果更明显,该算法实时性强,计算复杂性低,从而使该系统能够实时、有效提高语音信噪比,增强语音可懂度,改善语音质量,主观测听 MOS 值也有明显的提高,达到实时语音增强的目的。

参考文献:

- [1] 赵胜辉,刘家康,谢湘,等.离散时间语音信号处理[M].北京:电子工业出版社,2004.
- [2] 邓克岩.基于谱减法的语音增强算法在 DSP 环境下的实时实现[J].西北民族大学学报:自然科学版,2007,28(4):150-153.
- [3] 彭启琮,武乐琴,张舰,等.TMS320VC55x 系列 DSP 的 CPU 与外设[M].北京:清华大学出版社,2005.
- [4] 汪安民,徐保根,程昱,等.DSP 嵌入式系统典型开发案例[M].北京:人民邮电出版社,2007.
- [5] 彭启琮.DSP 技术的发展与应用[M].北京:高等教育出版社,2002.
- [6] 胡璞.一种基于 TMS320VC5509 的语音处理系统的设计[J].电子科技,2009,22(5):118-121.

(上接第 138 页)

参考文献:

- [1] OKI Inc. ML7204 data sheet [DB/OL]. (2004)[2009-08-02]. http://www2.okisemi.com/site/products/catalog/telecomics/VoIPCODECs/Intro-3266.html/ML7204_DS.pdf.
- [2] OKI Inc. ML7204_EvaBrd_UM[DB/OL]. (2004)[2009-08-02]. http://www2.okisemi.com/site/products/catalog/telecomics/VoIPCODECs/Intro-3266.html/ML7204_EvaBrd_UM.pdf.
- [3] Xilinx Inc. PicoBlaze 8-bit embedded microcontroller [DB/

OL]. (2008)[2009-08-02]. http://www.xilinx.com/support/documentation/ip_documentation/ug129.pdf.

- [4] 温广翔,徐欣.PicoBlaze 处理器 IP Core 的原理与应用[J].单片机与嵌入式系统应用,2003(12):37-39.
- [5] 徐运涛.4kbit/s 语音编码的标准化的最新进展[J].解放军理工大学学报:自然科学版,2000,1(3):10-16.
- [6] 姚彦,梅顺良,高葆新,等.数字微波中继通信工程[M].北京:人民邮电出版社,1990.