

基于 STM32 的嵌入式语音增强系统*

江伟华¹, 童峰¹, 洪青阳², 李琳²

(1. 厦门大学水声通信与海洋信息技术教育部重点实验室, 厦门 361005;

2. 厦门大学信息科学与技术学院, 厦门 361005)

文 摘: 噪声背景下各类语音信号处理方法的性能都明显下降。二步噪声消除算法 (TSNR, TWO-STEP NOISE REDUCTION) 是一种由传统维纳滤波语音增强算法改进的语音增强算法, 在语音增强研究中得到广泛应用。本文基于 STM32F407 嵌入式处理器进行 TSNR 语音增强算法的实时实现, 通过对嵌入式处理器资源进行调配和复用实现嵌入式实时语音增强。噪声环境下的实测实验表明, 利用本文设计方案的语音增强系统具有有效的降噪性能。

关键词: HSNR; 语音增强; 嵌入式;

中图分类号: TN912.3

环境噪声所引起的畸变严重影响着声纹识别、语音识别等各类语音信号处理系统的性能。因此有必要对语音进行消噪来提高信噪比, 提高声纹识别的性能。谱减方法由于实现简单方便、约束条件少、物理意义直接, 在语音信号的抗噪声处理中得到了广泛的研究^[1-2]。

维纳滤波是在最小均方误差意义下的最优估计, 可通过递推生成一个基于最小均方误差准则的线性滤波器, 使带噪语音经过该滤波器后达到从频谱上减去噪声频谱成分, 恢复原始纯净语音信号频谱的效果^[3]。但在实际应用中, 背景噪声并不是一直平稳的, 传统维纳滤波算法中的噪声估计并不能实时更新, 具有一定延迟, 对算法降噪性能造成影响。

二步噪声消除算法 (TSNR) 利用前一帧语音的后验信噪比来估计当前语音帧的先验信噪比, 从而改善对时变噪声的适应性能^[4]。目前对 TSNR 算法进行硬件实现的工作尚不多见。本文针对实际语音处理系统的需要, 基于 STM32 嵌入式处理器进行 TSNR 语音增强算法的实时实现, 通过对嵌入式处理器资源进行调配和复用实现解决了算法移植中硬件资源不足的问题。实测降噪试验表明了本文嵌入式语音增强算法的有效性。

1 TSNR 算法

维纳滤波是在最小均方误差意义下的最优估

计, 其基本原理如下:

假设 $y(n)$ 为带噪语音信号, $s(n)$ 为纯净语音信号, $b(n)$ 为加性噪声信号, 则^[1]

$$y(n) = s(n) + b(n) \quad (1)$$

分别用 Y_k 、 S_k 、 B_k 表示 $y(n)$ 、 $s(n)$ 和 $b(n)$ 的 FFT 变换, 则可得:

$$Y_k = S_k + B_k \quad (2)$$

又因为语音信号与噪声信号互不相关, 由此可得:

$$P_Y(k) = P_S(k) + P_B(k) \quad (3)$$

其中 $P_Y(k)$ 、 $P_S(k)$ 、 $P_B(k)$ 分别表示 $y(n)$ 、 $s(n)$ 、 $b(n)$ 的功率谱。

设维纳滤波增益为 G_k , \hat{S}_k 为最优估计的 S_k 的 FFT 变换, 则两者关系为:

$$\hat{S}_k = G_k \cdot Y_k \quad (4)$$

传统的维纳滤波算法的性能取决于两点, 第一点是噪声估计的准确性, 即 $\hat{P}_B(k)$ 与实际噪声功率谱的匹配程度; 第二点是增益函数 G_k 的设定, 根据增益函数 G_k 的不同设置衍生出了诸如谱减 (PE)、最大似然估计 (ML) 等频域语音增益函数, 增益函数 G_k 的选取决定了所用的频域增强方法的特性, 也决定了最终的增强效果。

传统维纳滤波算法中的噪声估计并不能实时更新, 具有一定延迟, 对算法性能造成影响, 为了改善语音增强算法的性能, 在传统维纳滤波算法的

*基金项目: 福建省高校产学研合作重大项目(No2015H6019), 厦门市科技计划 (3502Z20153002)

作者简介: 江伟华 (1991-), 男 (汉), 福建龙海, 硕士生。

通讯联系人: 童峰, 教授, E-mail: ftong@xmu.edu.cn

基础上提出二步噪声消除语音增强算法 (TSNR) [4]。

二步噪声消除算法 (TSNR) 实际上是利用第

(m+1) 帧的后验信噪比 $SNR_{post}(m+1, k)$ 来估计

当前帧的先验信噪比 $SNR_{prio}(m, k)$ 。

即 TSNR 算法的过程分为两步:

第一步, 利用 DD 法估计先验信噪比

$SNR_{prio}^{DD}(m, k)$;

第二步, 基于 $SNR_{prio}^{DD}(m, k)$ 与 G_k^{DD} 联合估计

当前帧语音的先验信噪比 $SNR_{prio}^{TSNR}(m, k)$ 。

2 STM32F407 处理器

2.1 芯片介绍

STM32F407 系列芯片是意法半导体集团公司在 2013 年推出的新一代低功耗高性能的集成系统, 其搭载内核为具有 FPU 的 Cortex-M4 内核。该系列芯片工作主频最高可达 180MHz, 内核指令集包含 ARM 精简指令集和浮点 FPS 指令集, 可以对数据进行 FFT 运算以及 FIR 滤波等浮点型数据运算, 经过大量的对比数据可以发现 Cortex-M4 内核的计算性能超越了 TI 同级别的 DSP。下面介绍硬件设计中 STM32F4 的部分功能[ii]。

其系统架构由 32 位多层 AHB 总线矩阵构成, 可实现以下部分的互连:

- 八条主控总线:

- Cortex™-M4F 内核 I 总线、D 总线和 S 总线

- DMA1 存储器总线

- DMA2 存储器总线

- DMA2 外设总线

- 以太网 DMA 总线

- USB OTG HS DMA 总线

- 七条被控总线:

- 内部 Flash ICode 总线

- 内部 Flash DCode 总线

- 主要内部 SRAM1 (112 KB)

- 辅助内部 SRAM2 (16 KB)

- 辅助内部 SRAM3 (64 KB)

- AHB1 外设(包括 AHB-APB 总线桥和 APB 外设)

- AHB2 外设

— FSMC

借助总线矩阵, 可以实现主控总线到被控总线的访问, 这样即使在多个高速外设同时运行期间, 系统也可以实现并发访问和高效运行。

2.2 TSNR 硬件实现方案

嵌入式语音增强处理过程如图 1 所示, 经过模拟解调后的语音信号通过 I2C 进入 STM32F4 存储器中, 在 STM32F4 中完成二步噪声消除和时间反转算法的处理过程, 最后将数据输出。

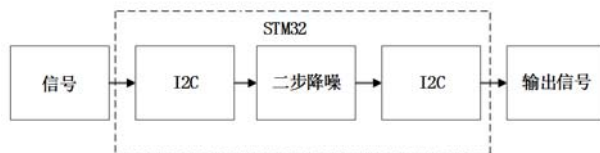


图 1 TSNR 硬件处理流程

接收端具体程序流程如下:

系统上电以后, STM32 首先进行初始化, 初始化功能包括时钟、GPIO、I2C、定时器、中断控制器、DMA 控制器。初始化完成后进入接收模式, 开启 I2C 口、中断和 DMA 传输模块。STM32F4 开始读取 I2C 的数据并通过 DMA 直接存储在 BUFF1 中, 读取 1024 位数据后结束 BUFF1 读取, DMA 中断, DMA 中断后开始向 BUFF2 传输数据。同时 BUFF1 进入数据处理程序, 进行二步噪声消除和被动时间反转, 数据处理完毕后输出, 等待 BUFF2 传输结束进入中断。BUFF2 传输结束后, 同理转向 BUFF1, BUFF2 进入数据处理程序。程序循环至设置的计时器结束停止。

3 实验结果

为了验证本文嵌入式二步噪声消除算法的性能, 在实验室内进行了噪声背景下的语音增强实验。实验环境为尺寸 10m*10m 的室内办公环境, 噪声源为电机噪声。语音信号采样率为 16kHz, 通过麦克风采集了该背景噪声下的短语音语句, 采集到的语音信号时域波形如图 2(a)所示, 时频图如图 2(b)所示, 可看到明显的背景噪声污染。

将采集到的语音信号输入 Matlab 中进行降噪处理, 二步噪声消除算法设置单窗长度为 1024 点, 经过降噪的语音波形如图 3(a)所示, 时频图如图 3(b)所示。经过本文 STM32 嵌入式系统进行语音增强处理后, 语音采样率 16kHz, 降噪的语音波形如图 4(a)所示, 时频图如图 4(b)所示, 可以发现, 虽然仍然有明显的减噪效果, 但比软件处理明显下降。

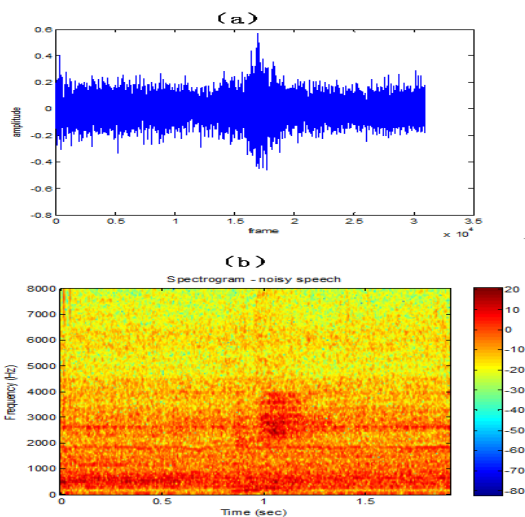


图2 带噪语音波形及声谱图

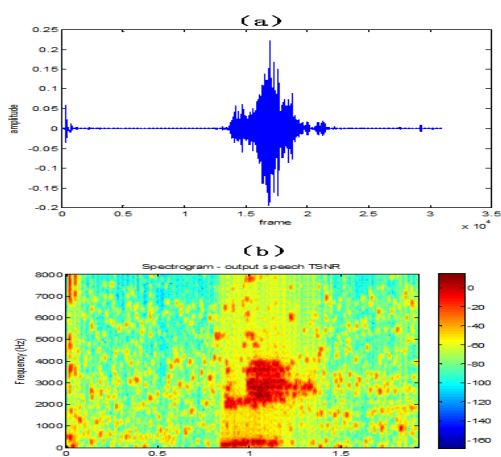


图3 消噪语音波形及声谱图

通过与图1对比可以发现，经过二步噪声消除算法处理的语音信号的噪声基本被消除并且语音部分信息保留完整，语音质量较好。这说明二步噪声消除算法对稳定的噪声影响有着明显的抑制效果，利用该算法可以极大地提升语音质量。进一步

分析算法性能，如图4所示，可以看到在语音部分段，算法对语音有较大的增强，而在噪音段部分则没有，说明对于稳定频率段的噪声，二步噪声消除算法的性能良好。

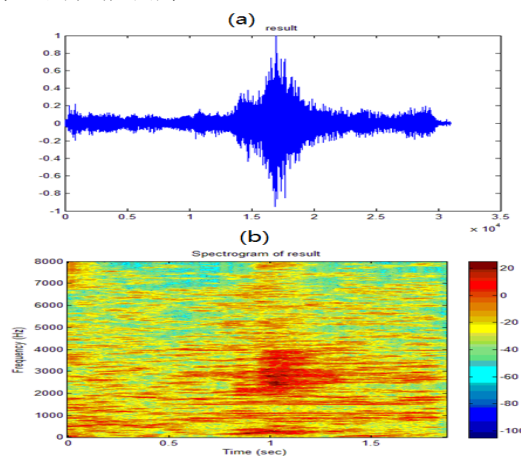


图4 消噪语音波形及声谱图

4 结论

针对实际应用中背景噪声对语音信号处理的影响，本文设计了基于嵌入式系统的二步噪声消除系统并进行实验测试。

参考文献

- [1] 丁伟, 吴小培.基于改进谱减方法的语音增强研究[J].计算机技术与发展,2008, 18(9):98-100
- [2] 李晔, 崔慧娟, 唐昆.基于谱减的语音增强算法的改进[J].清华大学报, 2006, 46(10):1685-1686
- [3] 叶顺舟,付仕明.基于维纳滤波的改进语音增强算法研究[J].广东通信技术.2011.12.63-65
- [4] 欧世峰, 王显云, 高颖, 赵晓晖.基于两步噪声消除技术与高斯统计模型的语音增强算法[J].信号处理,2011,27(8),1172-1178

Embedded Speech Enhancement System based on STM32

Jiang Wei-hua¹, Tong Feng¹, Hong qin-yang², Li Lin²

1. UAC MOE Key Lab, Xiamen University, Xiamen 361005, China

2. School of Information Science and Technology, Xiamen university, xiamen,361005, China

Abstract: Background noise pose significant difficulties to the signal processing systems such as speech recognition and voiceprint recognition. TWO-STEP NOISE REDUCTION (TSNR) method has been extensively applied for speech enhancement. In this paper, STM32 microprocessor is adopted to develop an embedded speech enhancement system. Experimental results are provided to validity the effectiveness of the proposed scheme.

Key words: TSNR; speech enhancement; embedded system