

基于短时分数阶傅里叶变换的语音增强算法

杨桂芹, 徐红莉, 唐铭卓

(兰州交通大学 电子与信息工程学院, 甘肃 兰州 730070)

摘要: 提出了一种基于短时分数阶傅里叶变换 (STFRFT) 的语音增强新方法。该方法首先将带噪语音信号进行短时分数阶傅里叶变换, 然后在分数阶傅里叶域 (FRFD) 对信号进行滤波, 最后对滤波后的信号进行短时分数阶傅里叶逆变换, 得到增强后的语音信号。实验表明在选定最佳的分数阶阶数时, 可使噪声得到最大限度的滤除, 大大提高了语音增强效果。

关键词: 短时分数傅里叶变换; 滤波; 语音增强

中图分类号: TN 912 **文献标识码:** A **文章编号:** 1000-8829(2011)01-0042-03

Method for Speech Enhancement Based on Short-Time Fractional Fourier Transform

YANG Guì-qín, XU Hong-li, TANG Míng-zhuo

(School of Electronics & Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract A new method for speech enhancement is proposed which is based on short-time fractional Fourier transform (STFRFT). Firstly noisy signal is processed by STFRFT. Then, the noisy signal is filtered in fractional Fourier domain (FRFD). Finally, the enhanced signal is obtained by using inverse short-time fractional Fourier transform (ISTFRFT) to the filtered signal. Simulation experiments indicate that when it is given an optimal fractional order, noise can be denoised maximumly which greatly improves the quality of speech enhancement.

Key words STFRFT; filtering; speech enhancement

在实际的环境中, 语音信号不可避免地会受到来自周围环境和传输媒介的噪声、通信设备内部的电噪声和周围讲话者的干扰。这些干扰会使语音处理系统的性能急剧恶化并且会影响语音的质量, 严重的情况下会将语音完全淹没在噪声中, 无法分辨出语音。为了提高系统性能并接收到较清晰的语音信号, 必须对带噪语音进行增强处理。语音增强就是从带噪语音中提取尽可能纯净的原始语音信号。

从时频分析的观点来看, 经典的滤波方法大都只限于在频域或时域进行加窗或遮隔运算, 但是由于语音信号与噪声之间存在着较强的时频耦合, 这使得经典的滤波方法难以实现有效的信噪分离^[1]。随着时频分析方法及应用的研究不断深入, 针对这类问题, 出现了一类基于时频平面旋转的滤波方法, 其核心思想

是将混有噪声的语音信号在时频平面上旋转特定的角度, 使得信号在新的时频平面上变为单频正弦信号; 再对变换后的信号分帧处理; 在此基础上, 通过简单的一维遮隔 (加窗) 处理^[2]或信号子空间的分解^[3], 将信号与噪声、干扰相分离; 然后, 通过对时频平面的反向旋转, 恢复出不含噪声与干扰的信号。笔者对一段带噪语音信号采用了基于短时谱估计的分数阶傅里叶变换进行降噪处理, 恢复出了较纯净的原始语音信号, 并根据实验结果分析了其降噪效果。结果表明基于短时分数阶傅里叶变换的技术有很明显的语音增强效果。

1 短时分数阶傅里叶变换

短时分数阶傅里叶变换 (STFRFT) 是一种新的信号分析和处理工具。短时处理时要将信号分帧处理, 分帧时时域信号首先被分成相叠的帧信号, 再分别对每帧信号进行分数阶傅里叶变换 (FRFT)。作为傅里叶变换的一种推广形式, FRFT 可以解释为信号在时频平面内坐标轴绕原点逆时针旋转任意角度后构成的分数阶傅里叶域上的表示, 如图 1 所示。如果信号的傅里叶变换可看成信号在时间轴上逆时针旋转 $\pi/2$

收稿日期: 2010-04-22; 修回日期: 2010-06-04

基金项目: 甘肃省自然科学基金资助项目 (3ZS061-A25-056)

作者简介: 杨桂芹 (1970-), 女, 河南许昌人, 副教授, 硕士, 主要从事现代数字通信理论与现代数字信号处理教学和科研工作。

到频率轴上的表示,那么 FRFT 可以看成将信号在时间轴上逆时针旋转角度 α 到 u 轴上的表示 (u 轴被称为分数阶傅里叶域)。从本质上讲,信号在分数阶傅里叶域上的表示,同时融合了信号在时域和频域的信息。因此,FRFT 被认为是一种新的时频分析方法^[4]。

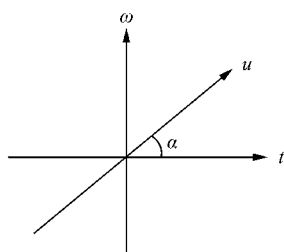


图 1 FRFT 变换原理图

信号 $x(t)$ 的 FRFT 定义为

$$X_p(u) = F^p\{x(t)\}(u) = \int_{-\infty}^{\infty} x(t) K_p(\tau, u) dt \quad (1)$$

式中, $K_p(\tau, u)$ 为核函数^[11]。

而短时分数阶傅里叶变换可表示为

$$STFRFT_{x_p}(t, u) = \int_{-\infty}^{\infty} x(\tau) g(\tau - t) K_p(\tau, u) d\tau \quad (2)$$

信号的短时分数阶傅里叶变换可以视为在做 FRFT 时对信号乘上一个窗函数。为了和那些有着相同窗函数但不同阶数的 FRFT 相区别,此式称为 P 阶 STFRFT。在做变换时,以窗函数 t 为中心占取一部分信号,这种部分的 FRFT 可以被看作信号在 t 时刻的瞬时分数阶傅里叶域谱;沿 t 轴移动窗函数,在每一个时刻就可以得到这样的一个谱。再对滤波后的信号进行逆 FRFT,便恢复出抑制噪声后的语音信号。

2 短时分数阶傅里叶域滤波

设带噪信号表示为

$$x(t) = s(t) + n(t) \quad (3)$$

式中, $s(t)$ 为原始信号, $n(t)$ 为干扰噪声。根据分数阶傅里叶变换的叠加性可知:两个独立信号和的分数阶傅里叶变换等于各自分数阶傅里叶变换的和。对式 (3) 两边分别做分数阶傅里叶变换得到

$$X_p(u) = S_p(u) + N_p(u) \quad (4)$$

式中, $X_p(u)$ 、 $S_p(u)$ 、 $N_p(u)$ 分别为信号 $x(t)$ 、 $s(t)$ 、 $n(t)$ 的离散分数阶傅里叶变换, p 为阶数。虽然语音信号和噪声在时域 (式 (3) 表示) 存在较强的时频耦合,但是在进行短时分数阶傅里叶变换之后,分数阶傅里叶域 (式 (4) 表示) 语音信号和噪声就会在不同的区域发生能量聚集。所以根据此原理可知,在做分数阶傅里叶变换时,选取合适的阶数,是可以将语音和噪声分离的。

文献 [5] 研究表明,分数阶傅里叶变换对语音信号中的浊音和清音均有一定的能量聚集性,不同点在

于能量聚集的区域不同:对浊音的能量聚集在分数阶变换域上体现为波形的中央区域,对清音的能量聚集则体现为波形的两端区域。由于分数阶傅里叶变换对白噪声没有能量聚集的特性,这一特性通过在信噪比 $-5 \sim 20$ dB 范围内做的一个 200 次蒙特卡洛实验证实,其仿真结果^[6]如图 2 所示。

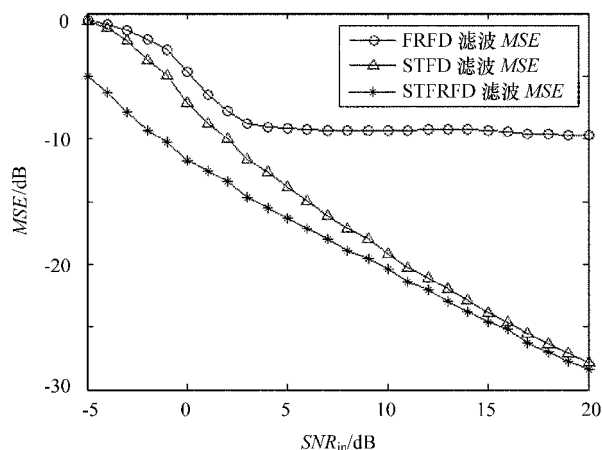


图 2 蒙特卡洛实验

这个仿真结果证明,在广范围的信噪比条件下,短时分数阶傅里叶域 (STFRFD) 滤波优于分数阶傅里叶域 (FRFD) 滤波和短时傅里叶域 (STFD) 滤波。这是因为:FRFD 滤波不能消除干扰,即使在较高的信噪比条件下其均方误差 (MSE) 也很大;在 STFD 滤波时,在较低的信噪比条件下,信号不能被很好地恢复,原因在于短时傅里叶变换对信号有较低的能量聚集性;短时分数阶傅里叶变换既能把干扰和信号分离,又能使信号很密集,所以 STFRFD 滤波效果比另外两种滤波效果都好,尤其是在较低的信噪比情况下。并且局部的最优窗和 STFRFD 滤波器都可以根据不同的信噪比和均方误差的情况设计出来。

为了更多地去去除噪声,根据式 (4) 中的分数阶傅里叶变换,变换后必须对 $X_p(u)$ 进行低通滤波处理。选用在通带内幅值特性比较平坦的巴特沃斯滤波器,再对滤波后的信号进行分数阶傅里叶逆变换即得到去噪后的语音信号。

3 仿真实例分析

为了验证此方法的语音增强效果,下面给出一仿真实例。实验中采用的语音材料为一段在安静的实验室环境下录制的干净语音。噪声选自 NOISEX-92 数据库,噪声类型为 F16 战斗机噪声。实验首先通过观察仿真图谱和靠听觉主观感受恢复后的语音信号,比较后确定短时分数阶傅里叶变换的最佳阶数。图 3 为做出的最佳语音增强的效果图,图中分别显示了纯净信号、带噪语音信号和增强后的输出信号。比较三者

可知, 输出信号与带噪信号相比, 明显滤除了大部分的噪声, 更接近于纯净语音信号。听觉上的主观感受也表明输出的信号与纯净语音信号已经相差无几, 不影响听觉效果。

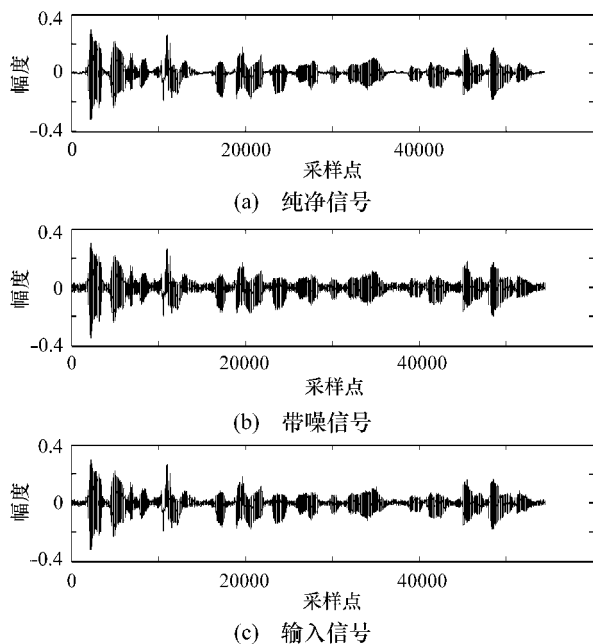


图3 仿真示例图

4 结束语

文中对带噪语音信号使用了短时分数阶傅里叶变换, 比较了分数阶傅里叶域、短时傅里叶域、短时分数阶傅里叶变换域的信号滤波效果。与传统的傅里叶变换去噪法相比, 文中方法将信号和噪声变换到短时分数阶傅里叶域, 使信号和噪声尽可能多地分离, 从而达到更好的去噪效果。研究结果表明, 选择一个最佳分数阶阶数, 可以使得去噪效果最佳, 优于传统的傅里叶变换去噪方法。

参考文献:

- [1] 陶然, 齐林, 王越. 分数阶 Fourier 变换的原理与其应用 [M]. 北京: 清华大学出版社, 2004.
- [2] Wood J C, Barry D T. Linear signal synthesis using the Radon-Wigner transform [J]. IEEE Transactions on Signal Processing 1994, 42(8): 2105-2111.
- [3] 邹红星, 周小波, 李衍达. 基于 Radon-STFT 变换的含噪 LFM 信号子空间分解 [J]. 电子学报, 1999, 27(12): 4-8.
- [4] Almeida L B. The fractional Fourier transform and time-frequency representations [J]. IEEE Transactions on Signal Processing 1994, 42(11): 3084-3091.
- [5] 王振力. 单/双通道语音增强方法的研究 [D]. 南京: 解放军理工大学, 2006.
- [6] Tao R, Li Y L, Wang Y. Short time fractional Fourier transform and its applications IEEE Transactions on Signal Processing 2010, 58(5): 2568-2580.
- [7] 李永强, 王振力. 基于分数阶 Fourier 变换的信号去噪方法 [J]. 弹箭与制导学报, 2006, 26(3): 235-238, 241.
- [8] 王振力, 张雄伟, 白志强. 语音增强新方法的研究 [J]. 南京邮电大学学报 (自然科学版), 2007, 27(2).
- [9] Soo C P, Wen L H, Ding J J. Discrete fractional Fourier transform based on new nearly triangular commuting matrices [J]. IEEE Transactions on Signal Processing 2006, 54(10): 3815-3828. □

(上接第 41 页)

4 结束语

系统采用基于 VXI 总线的数据采集技术, 显著地提高了系统的稳定性和可靠性。系统自投入运行以来, 工作稳定可靠, 测试精度高, 扩展性强, 操作维护简单, 大大降低了试车人员的工作强度, 提高了试车效率。

参考文献:

- [1] 张宝诚. 航空发动机试验和测试技术 [M]. 北京: 北京航空航天大学出版社, 2005.
- [2] 樊尚春, 吕俊芳, 张庆荣, 等. 航空测试系统 [M]. 北京: 北京航空航天大学出版社, 2005.
- [3] 杨训, 陈桂英, 李长征, 等. 某型涡扇发动机试车计算机辅助试验系统 [J]. 计算机测量与控制, 2006, 14(8): 981-999.
- [4] 卢艳娟, 雷勇, 王旺元. 某型涡轴航空发动机试车测控系统的设计 [J]. 计算机测量与控制, 2008, 16(2): 168-170.
- [5] 杨占星, 张沛超, 刘之尧, 等. 基于多线程技术的数字保护实时仿真系统 [J]. 电力系统保护与控制, 2009, 37(19): 73-76. □

免费阅读, 电子版期刊
全方位检索, 测控技术文章

请登录 www.mct.com.cn (测控在线)!

投稿、查稿,
浏览国内外测控动态,
发布企业资讯,
了解行业产品,
讨论测控话题,

.....