

基于分数阶谱相减的语音增强法

王振力 张雄伟

(解放军理工大学通信工程学院 南京 210007)

摘要: 该文提出了基于分数阶谱相减的语音增强法(FSS)。该方法通过对带噪声语音信号作分数阶傅里叶变换(FRFT),将得到的分数阶语噪混合谱与估计的分数阶噪声谱相减,最后利用分数阶 Fourier 反变换获得去噪后的语音信号。理论分析表明,所提方法存在一个最佳分数阶阶数,使得语噪混合信号能在分数阶变换域得到最好的分离,从而有效地提高了增强语音的性能。计算机仿真表明,对于混有加性白噪声的男/女声发音信号,所提方法在信噪比提高量和 Itakura 距离减少量两个方面都优于传统的谱相减法(SS),并且增强语音中的音乐噪声得到了明显抑制。

关键词: 分数阶傅里叶变换; 语音增强; 谱相减; 音乐噪声

中图分类号: TN912.3

文献标识码: A

文章编号: 1009-5896(2007)05-1096-05

A Method Based on Fractional Spectral Subtraction for Speech Enhancement

Wang Zhen-li Zhang Xiong-wei

(Institute of Communications Engineering, PLAUST, Nanjing 210007, China)

Abstract: A method based on fractional spectral subtraction for speech enhancement (FSS) is proposed. It applies FRFT (FRactional Fourier Transform) to noisy speech, the estimated fractional noise spectrum is then subtracted from the derived fractional speech-noise spectrum. Finally, the denoised speech is obtained by inverse fractional Fourier transform. Theory analysis indicates that the new method can find an optimal fractional order, which can best separate speech for noisy data in the fractional Fourier domain. The result is that the performance of the enhanced speech is effectively improved. Computer simulation shows that the SNR improvement amount and Itakura distance decrease amount of the proposed method are superior to those of Spectral Subtraction (SS) for male/female speech corrupted by additive white Gaussian noise. In addition, the music noise of the enhanced speech is remarkably suppressed.

Key words: FRFT (FRactional Fourier Transform); Speech enhancement; Spectral Subtraction (SS); Music noise

1 引言

在信号处理领域中,传统的Fourier变换是一个研究最为成熟、应用最为广泛的数学工具。然而Fourier变换只能用来处理确定性的平稳信号,对于时变的非平稳信号则无能为力。为了解决这一问题,学者们提出了一种广义的Fourier变换,一般称之为分数阶 Fourier 变换——FRFT (FRactional Fourier Transform)^[1-6]。近年来,FRFT已经在许多领域得到了广泛应用^[1-4, 7-9]。

语音增强是指为了提高受噪声污染的语音信号的质量而对含噪语音所作的处理,其主要目标是从带噪语音信号中提取尽可能纯净的原始语音。目前具有代表性的语音增强方法主要有:谱相减法^[10-13]、噪声对消法^[14, 15]、基于语音参数模型的方法^[16, 17]、基于状态转移的HMM方法^[18, 19]、神经网络

络法^[20]、小波去噪法^[21-23]和信号子空间法^[24-26]等。以上这些方法从本质上讲都是在时域或者在频域对带噪语音进行增强处理的,其中谱相减法由于其实施简单得到了广泛的应用。我们知道传统的谱相减法是在Fourier域进行语噪分离的,事实上这个Fourier域并不一定是最好的语噪分离域,结果必然导致不必要的语音信号损失,但是在分数阶Fourier变换域则可较好地完成语噪分离。本文通过对带噪语音信号作不同阶数的FRFT,提出了基于分数阶谱相减的语音增强法。理论分析和仿真结果表明了本文方法的有效性。

2 FRFT 的定义及其快速离散算法

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 Fourier 域上的表示方法。

信号 $s(x)$ 的 α 阶 ($0 < \alpha < 2$) FRFT 定义为

$$S_{\alpha}(u) = F^{\alpha}[s(x)] = \int_{-\infty}^{+\infty} K_{\alpha}(u, x) s(x) dx \quad (1)$$

其中变换核函数

2005-08-15 收到, 2006-04-12 改回

江苏省自然科学基金(BK2006001)和江苏省图像处理与图像通信实验室(KJS03036)资助课题

$$K_{\alpha}(u, x) = \begin{cases} A_{\phi} \exp\{j\pi[u^2 \cot(\phi) - 2ux \csc(\phi) + x^2 \cot(\phi)]\}, & \alpha \neq 0, \pm 2 \\ \delta(u - x), & \alpha = 0 \\ \delta(u + x), & \alpha = \pm 2 \end{cases} \quad (2)$$

而 $K_{\alpha}(u, x)$ 的幅值函数

$$A_{\phi} = \frac{\exp\{-j\pi \operatorname{sgn}[\sin(\phi)]/4 + i\phi/2\}}{|\sin(\phi)|^{1/2}}, \quad \phi = \frac{\alpha\pi}{2} \quad (3)$$

对于任意整数 l , 通过等式 $F^{\alpha+4l}[s(x)] = F^{\alpha}[s(x)]$ 可将此 FRFT 定义扩展到区间 $[-2, 2]$ 之外。该 FRFT 定义满足:

(1) 唯一性, $(F^{\alpha})^{-1} = F^{-\alpha}$; (2) 指数可加性: $F^{\alpha_1} F^{\alpha_2} = F^{\alpha_2} F^{\alpha_1} = F^{\alpha_1 + \alpha_2}$; (3) 当 $\alpha = 1$ 时退化为普通 Fourier 变换。

在 FRFT 的实际工程应用中, 必需采用其离散形式。目前国内外学者提出了多种离散 FRFT 的定义及其快速算法^[5,6,27-29]。本文采用文献[5]提出的 FRFT 快速算法。我们将 FRFT 的定义式变形为

$$S_{\alpha}(u) = A_{\phi} \exp[j\pi \cot(\phi)u^2] \int_{-\infty}^{+\infty} \exp[-j2\pi \csc(\phi)ux] \cdot \{\exp[j\pi \alpha x^2] s(x)\} dx \quad (4)$$

假定信号 $s(x)$ 的 Wigner 分布限定在以原点为中心, 直径为 Δx 的圆内。在此假设条件下, 我们将阶数 α 限制在 $0.5 \leq \alpha \leq 1.5$ 。于是, 线性调制函数 $\exp[j\pi \cot(\phi)x^2]s(x)$ 在频域具有带宽 Δx 。因此, $\exp[j\pi \cot(\phi)x^2]s(x)$ 可以用 Shannon 插值公式表示为

$$\exp[j\pi \cot(\phi)x^2]s(x) = \sum_{n=-N}^N \exp[j\pi \cot(\phi)(n/2\Delta x)^2] \cdot s\left(\frac{n}{2\Delta x}\right) \operatorname{sinc}\left[2\Delta x\left(x - \frac{n}{2\Delta x}\right)\right] \quad (5)$$

其中 $N = (\Delta x)^2$ 。由于假定 $s(x)$ 在 $[-\Delta x/2, \Delta x/2]$ 以外为零, 所以求和区间是从 $-N$ 到 N 。将式(5)代入式(4)并交换积分与求和的顺序得到

$$S_{\alpha}(u) = A_{\phi} \exp[j\pi \cot(\phi)] \sum_{n=-N}^N \exp[j\pi \cot(\phi)(n/2\Delta x)^2] \cdot s\left(\frac{n}{2\Delta x}\right) \int_{-\infty}^{+\infty} \exp[-j2\pi \csc(\phi)ux] \cdot \operatorname{sinc}\left[2\Delta x\left(x - \frac{n}{2\Delta x}\right)\right] dx \quad (6)$$

上式中的积分项等于 $\exp\left\{-\frac{j2\pi[\csc(\phi)]un}{2\Delta x}\right\} \frac{1}{2\Delta x} \cdot \operatorname{rect}\left\{\frac{[\csc(\phi)]x}{2\Delta x}\right\}$ 。在 $0.5 \leq \alpha \leq 1.5$ 的范围内, 矩形函数 $\operatorname{rect}\left\{\frac{[\csc(\phi)]x}{2\Delta x}\right\}$ 在变换函数的支撑区将总是等于 1。因此可以得到

$$S_{\alpha}(u) = \frac{A_{\phi}}{2\Delta x} \sum_{n=-N}^N \exp[j\pi \cot(\phi)u^2] \exp\left\{-\frac{j2\pi[\csc(\phi)]un}{2\Delta x}\right\} \cdot \exp\left\{\frac{j\pi[\cot(\phi)]n^2}{(2\Delta x)^2}\right\} s\left(\frac{n}{2\Delta x}\right) \quad (7)$$

这样, 变换后函数的样本值为

$$S_{\alpha}\left(\frac{m}{2\Delta x}\right) = \frac{A_{\phi}}{2\Delta x} \sum_{n=-N}^N \exp\left\{\frac{j\pi[\cot(\phi)]m^2}{(2\Delta x)^2} - \frac{j\pi[\csc(\phi)]mn}{(2\Delta x)^2} + \frac{j\pi[\cot(\phi)]n^2}{(2\Delta x)^2}\right\} s\left(\frac{n}{2\Delta x}\right) \quad (8)$$

这是一有限求和, 使得我们可以利用原函数 $s(x)$ 的离散采样值求出其分数阶 Fourier 变换 $S_{\alpha}(u)$ 的离散采样值。直接计算式(8)的计算量为 $o(N^2)$, 为了减少计算量, 我们将上式改写为

$$S_{\alpha}\left(\frac{m}{2\Delta x}\right) = \frac{A_{\phi}}{2\Delta x} \exp\left\{\frac{j\pi[\cot(\alpha) - \csc(\alpha)]m^2}{(2\Delta x)^2}\right\} \cdot \sum_{n=-N}^N \exp\left\{\frac{j\pi[\csc(\phi)](m-n)^2}{(2\Delta x)^2}\right\} \cdot \exp\left\{\frac{j\pi[\cot(\phi) - \csc(\phi)]n^2}{(2\Delta x)^2}\right\} s\left(\frac{n}{2\Delta x}\right) \quad (9)$$

可以看出式(9)的求和实际上是信号 $\exp[j\pi \csc(\phi)n^2/(2\Delta x)^2]$ 和线性调频调制函数 $\exp\left\{\frac{j\pi[\cot(\phi) - \csc(\phi)]n^2}{(2\Delta x)^2}\right\} s\left(\frac{n}{2\Delta x}\right)$ 的卷积运算。该卷积可用

FFT 计算, 计算复杂度为 $o(N \log(N))$ ^[5]。

3 分数阶谱相减法

分数阶谱相减法通过对带噪声语音信号作分数阶 Fourier 变换, 最终实现语音和噪声在分数阶域的分隔。

设加窗分帧带噪声语音信号

$$x(n) = s(n) + d(n) \quad (10)$$

其中 $s(n)$ 表示纯净语音信号, $d(n)$ 表示加性高斯白噪声, 且 $s(n)$ 和 $d(n)$ 相互独立。对式(10)两边分别作离散 FRFT 得

$$X_{\alpha}(u_k) = S_{\alpha}(u_k) + D_{\alpha}(u_k) \quad (11)$$

其中 $X_{\alpha}(u_k)$, $S_{\alpha}(u_k)$ 和 $D_{\alpha}(u_k)$ 分别表示离散信号 $x(n)$, $s(n)$ 和 $d(n)$ 的离散分数阶 Fourier 变换, α 表示 FRFT 的阶数。当 $\alpha = 1$ 时式(11)退化为普通的傅里叶变换。设 $X_{\alpha}(u_k) = X_{\alpha,k} \exp(j\varphi_{\alpha,x,k})$, $S_{\alpha}(u_k) = S_{\alpha,k} \exp(j\varphi_{\alpha,s,k})$, $D_{\alpha}(u_k) = D_{\alpha,k} \exp(j\varphi_{\alpha,d,k})$, 将它们代入式(11)可得

$$X_{\alpha,k} \exp(j\varphi_{\alpha,x,k}) = S_{\alpha,k} \exp(j\varphi_{\alpha,s,k}) + D_{\alpha,k} \exp(j\varphi_{\alpha,d,k}) \quad (12)$$

其中 $X_{\alpha,k}$, $S_{\alpha,k}$ 和 $D_{\alpha,k}$ 分别表示 $X_{\alpha}(u_k)$, $S_{\alpha}(u_k)$ 和 $D_{\alpha}(u_k)$ 的第 k 个分数阶 (α 阶) 谱分量的幅值, $\varphi_{\alpha,x,k}$, $\varphi_{\alpha,s,k}$ 和 $\varphi_{\alpha,d,k}$ 分别表示 $X_{\alpha}(u_k)$, $S_{\alpha}(u_k)$ 和 $D_{\alpha}(u_k)$ 的第 k 个分数阶 (α 阶) 谱分量的相位。根据理想谱相减模型, 我们知道带噪声语音谱幅度是纯净语音谱幅度和噪声谱幅度之和, 即

$$X_{\alpha,k} = S_{\alpha,k} + D_{\alpha,k} \quad (13)$$

分数阶谱相减法的任务就是利用已知的噪声谱信息 $D_{\alpha}(u_k)$, 从 $X_{\alpha}(u_k)$ 中估计出 $S_{\alpha}(u_k)$ 。由于人耳对相位不敏感^[30], 所以只需估计出 $S_{\alpha,k}$, 然后利用带噪声语音相位 $\varphi_{\alpha,x,k}$, 进行反分数阶 Fourier 变换就可得到增强语音。即

$$\hat{s}(n) = \text{IFRFT}[\hat{S}_{\alpha,k} \exp(j\varphi_{\alpha,x,k})] \quad (14)$$

其中 IFRFT 表示分数阶 Fourier 反变换。当 $\alpha = 1$ 时分数阶谱

相减法退化为传统的谱相减法^[10-13]。图 1 为分数阶谱相减法流程图。

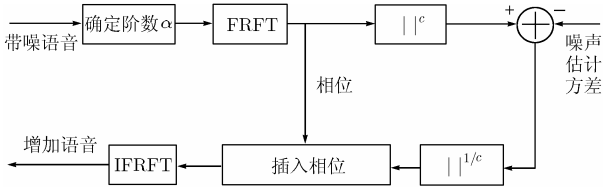


图 1 分数阶谱相减法流程图

根据文献[12,13], 可以得到

$$\hat{S}_{\alpha,k} = X_{\alpha,k} - \gamma E\{D_{\alpha,k}\} \quad (15)$$

其中 γ 表示噪声相减因子, $\gamma \geq 1$; $E\{D_{\alpha,k}\}$ 表示无音期噪声统计平均谱。根据 MMSE 准则, 我们定义归一化误差函数

$$\hat{\varepsilon}_\alpha = (S_{\alpha,k} - \hat{S}_{\alpha,k}) / S_{\alpha,k} \quad (16)$$

将式(13)和式(15)代入式(16)得

$$\hat{\varepsilon}_\alpha = [D_{\alpha,k} - \gamma E\{D_{\alpha,k}\}] / S_{\alpha,k} \quad (17)$$

将式(17)两边平方并取 $E\{\cdot\}$ 得

$$E\{\hat{\varepsilon}_\alpha^2\} = \frac{E\{D_{\alpha,k}^2\} - 2\gamma E\{D_{\alpha,k}\}E\{D_{\alpha,k}\} + \gamma^2 E\{D_{\alpha,k}\}^2}{E\{S_{\alpha,k}^2\}} \quad (18)$$

假设噪声谱分量是具有时变方差的零均值复高斯随机变量^[11], 那么噪声谱服从瑞利分布^[31], 即

$$f(D_{\alpha,k}) = \frac{2D_{\alpha,k}}{E\{D_{\alpha,k}^2\}} \exp[-D_{\alpha,k}^2 / E\{D_{\alpha,k}^2\}], \quad D_{\alpha,k} > 0 \quad (19)$$

根据式(19)和均值定义^[32]可知

$$E\{D_{\alpha,k}^q\} = \int_0^{+\infty} D_{\alpha,k}^q f(D_{\alpha,k}) dD_{\alpha,k} = \Gamma\left(\frac{q}{2} + 1\right) E\{D_{\alpha,k}^2\}^{q/2} \quad (20)$$

其中 q 是能量指数, $\Gamma(\cdot)$ 表示 Gamma 函数。在式(20)中, 取 $q = 1$ 得

$$E\{D_{\alpha,k}\} = \Gamma(1.5) E\{D_{\alpha,k}^2\}^{1/2} \quad (21)$$

将式(21)代入式(18)并整理得

$$E\{\hat{\varepsilon}_\alpha^2\} = [(\gamma^2 - 2\gamma)\Gamma^2(1.5) + 1] \frac{E\{D_{\alpha,k}^2\}}{E\{S_{\alpha,k}^2\}} \quad (22)$$

将式(21)代入式(15)得

$$\hat{S}_{\alpha,k} = X_{\alpha,k} - \gamma \Gamma(1.5) E\{D_{\alpha,k}^2\}^{1/2} \quad (23)$$

图 2-图 4 给出了 $E\{\hat{\varepsilon}_\alpha^2\}$ 和 $E\{D_{\alpha,k}^2\}$ 与 γ , $E\{S_{\alpha,k}^2\}$ 和 α 的理想关系曲线图。由图 4 可见, 存在一个最佳阶数 $\alpha \in (0.5, 1)$ 或者 $\alpha \in (1, 1.5)$, 可以使得 $E\{\hat{\varepsilon}_\alpha^2\}$ 最小。因此, 根据不同说话人的语音及其与噪声的混合程度, 从理论上我们可以按照

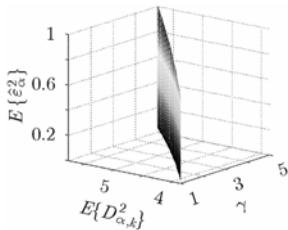


图 2 $E\{\hat{\varepsilon}_\alpha^2\}$ 、 $E\{D_{\alpha,k}^2\}$ 和 γ 关系曲线

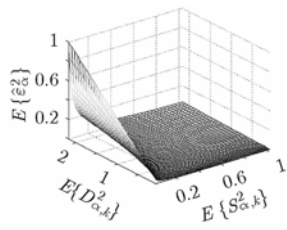


图 3 $E\{\hat{\varepsilon}_\alpha^2\}$ 、 $E\{D_{\alpha,k}^2\}$ 和 $E\{S_{\alpha,k}^2\}$ 关系曲线

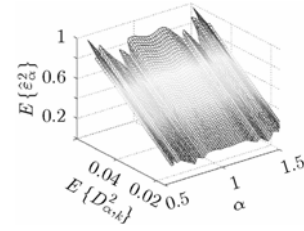


图 4 $E\{\hat{\varepsilon}_\alpha^2\}$ 、 $E\{D_{\alpha,k}^2\}$ 和 α 关系曲线

$E\{\hat{\varepsilon}_\alpha^2\}$ 最小的原则来确定最优阶数 α_{opt} , 但是实际应用中, 我们必须结合人的主观听觉感受来确定 α_{opt} 。

4 仿真实验与分析

4.1 客观评价

为了研究分数阶谱相减法的有效性, 我们采用男/女声标准发音语句“他去无锡市, 我到黑龙江”, 其采样频率为 8000Hz; 混杂在纯净语音中的噪声取加性高斯白噪声; 噪声的估计是在讲话人没有说话前的无音期进行的。我们用输入信噪比(SNR_{in})和输出信噪比(SNR_{out})分别表示输入带噪语音的信噪比和输出增强语音的信噪比, 它们分别定义如下:

$$SNR_{in} = 10 \lg \frac{\sum_{m=1}^M \sum_{n=1}^N s_m(n)}{\sum_{m=1}^M \sum_{n=1}^N d_m(n)} \quad (24)$$

$$SNR_{out} = 10 \lg \frac{\sum_{m=1}^M \sum_{n=1}^N s_m(n)}{\sum_{m=1}^M \sum_{n=1}^N [\hat{s}_m(n) - s_m(n)]} \quad (25)$$

式中 N 表示帧长, M 表示帧长的总数。实验中我们用汉明窗对带噪语音进行加窗, 帧长取 $N = 180$ (相当于每帧时长 22.5ms), 帧与帧之间的重叠度为 50%; 根据旋转角度的对称性(如图 4 所示), FRFT 的阶数 α 取值范围为 1~1.5; 谱相减因子取 $\gamma = 1$ 。图 5 和图 6 分别给出了在 $SNR_{in} = 0$ dB 时, 本文方法的 SNR_{out} 和 Itakura 距离与 α 变化关系图(当 $\alpha = 1$ 时对应传统的谱相减法)。由图 5 和图 6 可知, 当 α 由 1 增大到 1.5 时, 本文方法的 SNR_{out} 和 Itakura 距离均存在着全局最优值, 分别比传统的谱相减法提高 3.09dB(女声)、1.16dB(男声)和降低 0.22(女声)、0.27(男声)。由于 SNR_{out} 和 Itakura 距离取最优值时的 α 不相等, 我们可以均方误差 $E\{\hat{\varepsilon}_\alpha^2\}$ 为依据, 同时结合人耳的听觉感受(如非常接近人主观听觉质量的 Itakura 距离)来判定最优阶数 α_{opt} 。

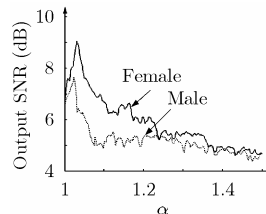


图 5 输出信噪比与 α 关系曲线

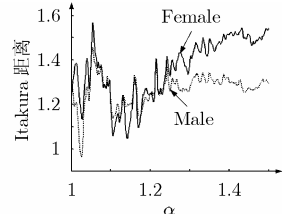


图 6 Itakura 距离与 α 关系曲线

表 1 和表 2 分别给出了 $SNR_{in} \in (-10, 10)$ dB 时女/男声的 SNR_{in} 与 SNR_{out} 和 Itakura 距离的对应关系(括号中的数字表示 SNR_{in} 对应的 Itakura 距离值), 表 1 对应的

表 1 SNR_{in} 与 SNR_{out} 和 Itakura 距离的对应关系(女声)

| SNR_{in} (dB) | SS | | FSS | |
|-----------------|------------------|------------|------------------|------------|
| | SNR_{out} (dB) | Itakura 距离 | SNR_{out} (dB) | Itakura 距离 |
| -9.46(2.15) | -5.07 | 1.84 | -1.29 | 1.62 |
| -6.57(1.99) | -1.94 | 1.63 | 1.45 | 1.38 |
| -2.13(1.73) | 3.06 | 1.34 | 5.56 | 1.21 |
| -0.55(1.63) | 4.81 | 1.26 | 7.10 | 1.04 |
| 1.95(1.49) | 7.23 | 1.13 | 9.06 | 0.93 |
| 4.72(1.32) | 9.26 | 1.02 | 10.83 | 0.84 |
| 7.11(1.18) | 10.97 | 0.95 | 12.14 | 0.78 |
| 9.47(1.04) | 12.23 | 0.87 | 13.45 | 0.74 |

表 2 SNR_{in} 与 SNR_{out} 和 Itakura 距离的对应关系(男声)

| SNR_{in} (dB) | SS | | FSS | |
|-----------------|------------------|------------|------------------|------------|
| | SNR_{out} (dB) | Itakura 距离 | SNR_{out} (dB) | Itakura 距离 |
| -9.78(1.87) | -4.25 | 1.76 | -2.45 | 1.55 |
| -6.82(1.73) | 0.41 | 1.60 | 2.42 | 1.38 |
| -2.98(1.53) | 4.38 | 1.42 | 5.97 | 1.16 |
| -0.80(1.41) | 6.01 | 1.32 | 7.27 | 1.04 |
| 2.40(1.23) | 7.99 | 1.15 | 9.01 | 0.91 |
| 4.20(1.14) | 8.96 | 1.06 | 10.07 | 0.83 |
| 6.56(1.01) | 10.30 | 0.95 | 11.37 | 0.73 |
| 9.66(0.86) | 11.98 | 0.80 | 12.92 | 0.61 |

$\alpha_{opt} = 1.141$, 表 2 对应的 $\alpha_{opt} = 1.024$ 。由表 1 和表 2 可知, 当发音为女声时, 本文方法(FSS)的 SNR_{out} 和 Itakura 距离对应的比谱相减法(SS)平均提高 2.22dB 和降低 0.19; 当发音为男声时, 对应的 SNR_{out} 提高量和 Itakura 距离减少量分别为 1.35dB 和 0.23。

4.2 非正式主观测试

对于表 1 和表 2 混有加性白噪声并且 $SNR_{in} \in (-10, 10)$ dB 的带噪语音, 经传统的谱相减法和本文方法增强后, 由听力正常的 3 男 3 女进行听力测试, 测试结果为: 6 人均认可本文方法增强语音的质量高于谱相减法, 并且增强语音明显抑制了传统的谱相减法无法解决的音乐噪声。

5 结束语

本文提出了一种基于分数阶谱相减的语音增强法。该方法通过最优分数阶阶数的确定, 可以使得语音和噪声在分数阶变换域获得最好的分离, 从而提高了增强语音的性能。与传统的谱相减法相比, 无论对于男声还是女声发音带噪信号, 本文方法均可以获得更好的语音增强效果, 并且增强语音中的音乐噪声得到了明显抑制。当 FRFT 阶数取 1 时, 本文方法退化为传统的谱相减法。可以这样说, 本文方法是传统谱相减法的一种推广, 具有更灵活的适用性和更广阔的推广使用价值。

参 考 文 献

- [1] Ozaktas H M, Barshan B, Mendlovic D, and Onural L. Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms. *J. Opt. Soc. Amer. A*, 1994, 11(2): 547-559.
- [2] Namias V. The fractional Fourier transform and its application in quantum mechanics. *J. Inst. Maths. Application*, 1980, 25(1): 241-265.
- [3] McBride A C and Kerr F H. On Namias's fractional Fourier transform. *IMAJ. Applied Math.*, 1987, 39(2): 159-175.
- [4] Almeida L B. The fractional Fourier transform and time-frequency representations. *IEEE Trans. on Signal Processing*, 1994, 42(11): 3084-3091.
- [5] Ozaktas H M, Arikan O, Kutay M A, and Bozdagi G. Digital computation of The fractional Fourier transform. *IEEE Trans. on Signal Processing*, 1996, 44(9): 2141-2150.
- [6] Candan C, Kutay M A, and Ozaktas H M. The discrete fractional Fourier transform. *IEEE Trans. on Signal Processing*, 2000, 48(5): 1329-1337.
- [7] Vogel K and Risken H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 1989, 40(5): 2847-2849.
- [8] Mendlovic D, Ozaktas H M, and Lohmann A W. Fractional correlation. *Appl. Opt.*, 1995, 34(2): 303-309.
- [9] Fonollosa J R and Nikias C L. A new positive time-frequency distribution. in Pro. IEEE Int. Conf. Acoust., Speech, Signal Processing, New Jersey, 1994: 301-304.
- [10] Boll S F. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans. on Acoust, Speech, Signal*

- Processing*, 1979, ASSP-27(2): 113–120.
- [11] Sim B L, Tong Y C, Chang J S, and Tan C T. A parametric formulation of the generalized spectral subtraction method. *IEEE Trans. on Speech and Audio Processing*, 1998, 6(4): 328–337.
- [12] Berouti M, Schwartz R, and Makhoul J. Enhancement of speech corrupted by acoustic noise. in Proc. ICASSP, Washington DC, 1979: 208–211.
- [13] Kushner W M, Vladimir G, Wu C, Nguyen V, and Damoulakis J N. The effects of subtractive-type speech enhancement/noise reduction algorithms on parameter estimation for improved recognition and coding in high noise environments. In Proc. ICASSP, Glasgow Scotland, 1989: 211–214.
- [14] Widrow B, *et al.* Adaptive noise cancelling, principles and applications. *IEEE Proc.*, 1975, 63(12): 1692–1716.
- [15] Frazier R H, Samsam S, Braida L D, and Oppenheim A V. Enhancement of speech by adaptive filtering. ICASSP'76, Philadelphia, 1976: 251–253.
- [16] Gibson J D, Koo B, and Gray S D. Filtering of colored noise for speech enhancement and coding. *IEEE Trans. on Acoust., Speech, Signal Processing*, 1991, ASSP-39(8): 1732–1742.
- [17] Lim J S and Oppenheim A V. All-pole modeling of degraded speech. *IEEE Trans. on Acoust., Speech, Signal Processing*, 1978, ASSP-26(3): 197–210.
- [18] Ephraim Y. A Bayesian estimation approach for speech enhancement using Hidden Markov models. *IEEE Trans. on Signal Processing*, 1992, 40(4): 725–735.
- [19] Ephraim Y, Malah D, and Juang H. On the application of Hidden Markov models for enhancing noise speech. *IEEE Trans. on Acoust., Speech, Signal Processing*, 1989, 37(12): 1846–1856.
- [20] Knecht W G, Schenkel M E, and Moschytz G S. Neural network filters for speech enhancement. *IEEE Trans. on Speech. Audio Processing*, 1995, 3(6): 433–438.
- [21] Donoho D L. De-Noising by Soft-Thresholding. *IEEE Trans. on Inform. Theory*, 1995, 41(3): 613–627.
- [22] Donoho D L and Johnstone I M. Adapting to Unknown Smoothness via Wavelet Shrinkage. *Journal of the American Statistical Association*, 1995, 90(432): 1200–1224.
- [23] Donoho D L, Johnstone I M, Kerkyacharian G and Picard D. Wavelet Shrink: Asymptopia. *Journal of the Royal Statistical Society, Series B*, 1995, 57(2): 301–369.
- [24] Ephraim Y and Van Trees H L. A signal subspace approach for speech enhancement. *IEEE Trans. on Speech Audio Processing*, 1995, 3(4): 251–266.
- [25] Mittal U and Phamdo N. Signal/noise KLT based approach for enhancing speech degraded by colored noise. *IEEE Trans. on Speech Audio Processing*, 2000, 8(2): 159–167.
- [26] Rezayee A and Gazor S. An adaptive KLT approach for speech enhancement. *IEEE Trans. on Speech Audio Processing*, 2001, 9(2): 87–95.
- [27] Santhanam B and McClellan J H. The discrete rotational Fourier transform. *IEEE Trans. on Signal Processing*, 1996, 42(4): 994–998.
- [28] Pei Soo-Chang, Yeh Min-Huang, and Tseng Chien-Cheng. Discrete fractional Fourier Transforms based on orthogonal projections. *IEEE Trans. on Signal Processing*, 1999, 47(5): 1335–1347.
- [29] 平先军, 陶然, 周思永等. 一种新的分数阶 Fourier 变换快速算法. *电子学报*, 2001, 29(3): 406–408.
- [30] Wang D L and Lim J S. The unimportance of phase in speech enhancement. *IEEE Trans. on Acoust., Speech, Signal Processing*, 1982, ASSP-30(8): 679–681.
- [31] Picinbono B. Random Signals and Systems. Englewood Cliffs, NJ: Prentice Hall, 1993: 136–138.
- [32] Spiegel M R. Schaum's Outline of Theory and Problems of Mathematical Handbook of Formulas and Tables. Int.ed. New York: McGrawHill, 1990.
- 王振力: 男, 1977 年生, 博士生, 感兴趣方向为阵列信号处理、多媒体信息处理等。
- 张雄伟: 男, 1965 年生, 教授, 博士生导师, 主要研究方向为数字通信、多媒体信息处理等。