

# Noise Compression Based on One-Sided Autocorrelation Linear Prediction

XU Jingbo, YU Hongtao, RAN Chongsen

Information and Engineering Institute, PLA, Zhengzhou, China

**Abstract:** A new method proposed, which reduces white noise using correlation functions of one-sided autocorrelation and linear prediction residual in the autocorrelation domain. When One-sided autocorrelation sequence is obtained, speech sequence is replaced by One-sided autocorrelation sequence. Then linear prediction coefficients and linear prediction residual are calculated. Based on it, a method on speech denoising is proposed that the noise residual is subtracted from the noisy speech signal in the proportion of subtracting coefficient  $\mu$ , which is gained from proportion of residual and signal energy. Experimental results showed that the structure of the speech was retained well and quality of the speech was not reduced in lower SNR.

**Keywords:** one-sided autocorrelation; prediction residual; residual energy; subtracting coefficient

## 基于单边自相关线性预测的语音噪声衰减

徐静波, 于洪涛, 冉崇森

解放军信息工程大学信息工程学院, 郑州, 中国, 450002

**摘要:** 提出了一种在自相关域上, 以相关函数值为参数, 利用单边自相关序列的线性预测误差去除语音中加性噪声的方法。方法首先, 对含噪语音进行单边自相关处理, 以语音信号的单边自相关序列替代语音信号序列, 进而对该序列进行线性预测分析后, 获得线性预测分析系数, 并求得线性预测误差。根据误差能量与信号能量的比例关系, 确定减因子  $\mu$ , 从含噪语音中依减因子  $\mu$  的大小减去预测误差, 即可抑制噪声误差能量。实验表明: 上述方法在低信噪比时, 仍能较好地保留语音信号的频谱结构, 使语音音质不至于下降。

**关键词:** 单边自相关; 预测误差; 误差能量; 减因子

### 1 引言

目前的一些去除语音中白噪声的方法, 如基于短时谱估计的谱相减法, wiener 滤波法, Kalman 滤波法等。这些方法在去除语音中噪声成分的同时, 对语音成分本身也造成了损伤, 降低了语音的质量, 原因就在于在处理过程中, 改变了语音信号的频谱结构, 特别是改变了语音共振峰的位置和宽度。因此, 为了弥补算法的这一缺憾, 本文提出一种以语音的单边自相关序列替代语音数据参数, 对该序列进行线性预测分析, 获得线性预测分析系数, 抑制噪声残差能量的方法。该方法对语音共振峰影响很小, 较好地保留了语音信号的频谱。

#### 1.1 单边自相关序列

一个离散随机实序列  $x(n)$  的自相关函数定义为

$$R(m) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x(n)x(n+m) \quad (1-1)$$

在语音处理中, 语音信号的分析一般是按帧进行

的。设一帧语音为  $x(n)$ ,  $n = 0, 1, \dots, N-1$ , 其中  $N$  是一帧语音的长度。根据式 (1-1) 可以得到该帧语音自相关序列的估计式为

$$\hat{R}(m) = \frac{1}{N} \sum_{n=-N}^{N-m-1} x(n)x(n+m), \quad (1-2)$$

$$m = 0, 1, \dots, N-1$$

(1-2)式是自相关序列的有偏估计式。自相关序列的无偏估计式为

$$\hat{R}(m) = \frac{1}{N-m} \sum_{n=-N}^{N-m-1} x(n)x(n+m), \quad (1-3)$$

$$m = 0, 1, \dots, N-1$$

(1-2)、(1-3)中虽然限定  $m \geq 0$ , 但  $\hat{R}(m)$  在  $m = -1, -2, \dots, -N+1$  时也有定义, 取值规则为  $\hat{R}(m) = \hat{R}(-m)$ 。

单边自相关序列  $R^+(n)$  是自相关序列  $R(n)$  的因果部分<sup>[1,2]</sup>, 即

$$R^+(n) = \begin{cases} R(n) & n > 0 \\ R(0)/2 & n = 0 \\ 0 & n < 0 \end{cases} \quad (1-4)$$

式(1-4)定义的单边自相关序列  $R^+(n)$  满足:

$$R^+(n) + R^+(-n) = R(n), -\infty \leq n \leq \infty。$$

单边自相关序列  $R^+(n)$  有两个特性: (1) 单边自相关序列  $R^+(n)$  和信号  $x(n)$  本身有着相同的极点<sup>[3]</sup>, (2) 抗噪声。与自相关序列  $R(n)$  的频谱相比, 单边自相关序列  $R^+(n)$  的增强了能量分布最高的那一部分。不在此频段范围内的噪声分量会有较大的衰减。

设语音信号  $s(n)$ , 在加性噪声环境下的观测值是

$$x(n) = s(n) + d(n) \quad (1-5)$$

$d(n)$  为稳定的噪声信号, 均值为 0, 且与语音信号无关。

$x(n)$  的单边自相关函数为:

$$\begin{aligned} R^+(n) &= E[x(m)x(m+n)] \\ &= E[s(m)s(m+n)] + E[s(m)d(m+n)] \\ &\quad + E[d(m)s(m+n)] + E[d(m)d(m+n)] \quad n \geq 0 \end{aligned} \quad (1-6)$$

由于噪声与语音无关, 所以有

$$E[s(m)d(m+n)] = E[s(m)]E[d(m+n)] \quad (1-7)$$

$$E[s(m+n)d(m)] = E[s(m+n)]E[d(m)] \quad (1-8)$$

由于噪声均值为 0, 即有

$$E[d(m)] = E[d(m+n)] = 0 \quad (1-9)$$

设  $s(n)$  和  $d(n)$  的单边自相关函数分别为  $R_s^+(n)$ 、 $R_d^+(n)$ , 即

$$R_s^+(n) = E[s(m)s(m+n)], \quad n \geq 0$$

$$R_d^+(n) = E[d(m)d(m+n)], \quad n \geq 0$$

式(1-6)可写为:

$$R^+(n) = R_s^+(n) + R_d^+(n), \quad (1-10)$$

式(1-6)说明, 带噪语音信号的自相关函数等于清晰语音信号的自相关函数与噪声信号自相关函数之和。由于语音是按帧处理的, 式(1-10)又可以表示为

$$\begin{aligned} R^+(k, n) &= R_s^+(k, n) + R_d^+(k, n), \\ 0 \leq k \leq K-1, \quad 0 \leq n \leq N-1 \end{aligned} \quad (1-11)$$

其中  $k$  是帧序号,  $N$  和  $K$  分别是一段语音所有帧的数目和每帧的长度。如果噪声是稳定的, 因  $R_d^+(k, n)$  不随  $k$  的变化而变化, 可用  $R_d^+(n)$  代替, (1-11)式可表示

为:

$$\begin{aligned} R^+(k, n) &= R_s^+(k, n) + R_d^+(n), \\ 0 \leq k \leq K-1, \quad 0 \leq n \leq N-1 \end{aligned} \quad (1-12)$$

根据式(1-12)写出清晰语音的自相关序列为

$$\begin{aligned} R_s^+(k, n) &= R^+(k, n) - R_d^+(n), \\ 0 \leq k \leq K-1, \quad 0 \leq n \leq N-1 \end{aligned} \quad (1-13)$$

因此, 通过下面的步骤从带噪声语音中估算出每一帧清晰语音的单边自相关序列: 在静音段, 选取几帧信号, 分别计算它们的单边自相关序列, 然后取它们的平均值作为噪声单边自相关序列  $\hat{R}_d^+(n)$  的估计值。

1) 在有声段, 分别计算每帧带噪语音的自相关序列  $R^+(k, n)$ 。

2) 已知噪声的单边自相关序列的估计值  $\hat{R}_d^+(n)$  和每帧带噪语音的单边自相关序列  $R(n)$ , 可求得清晰语音自相关序列的估计值:

$$\begin{aligned} \hat{R}_s^+(k, n) &= R^+(k, n) - \hat{R}_d^+(n), \\ 0 \leq k \leq K-1, \quad 0 \leq n \leq N-1 \end{aligned} \quad (1-14)$$

以上是在自相关域中, 从带噪语音中估算清晰语音自相关序列的方法。

## 2 单边自相关线性预测分析

采用全极点模型, 辐射、声道以及声门激励的组合谱效应的传输函数为:

$$H(Z) = \frac{S(z)}{E(z)} = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}}$$

$p$  是预测器阶数,  $G$  是声道滤波器增益 (可以取  $G=1$ , 不影响后续的分析), 由此, 语音信号  $s(n)$  和  $e(n)$  之间的关系可以如下表示:

$$s(n) = e(n) + \sum_{i=1}^p a_i s(n-i) \quad (2-1)$$

线性预测分析的基本思想是: 用过去  $P$  个样点值来预测现在或未来的样点值:

$$\hat{s}(n) = \sum_{i=1}^p a_i s(n-i) \quad (2-2)$$

预测误差  $e(n)$  为:

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{i=1}^p a_i s(n-i) \quad (2-3)$$

这样在均方误差最小准则下, 确定唯一的一组线性预测系数  $a_i$  ( $i=1, 2, \dots, p$ )。

线性预测系数的求解, 可以用自相关法转化为线性

方程组的求解问题:

$$\begin{bmatrix} R^+(0) & R^+(1) & \cdots & R^+(p-1) \\ R^+(1) & R^+(0) & \cdots & R^+(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R^+(p-1) & R^+(p-2) & \cdots & R^+(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} R^+(1) \\ R^+(2) \\ \vdots \\ R^+(p) \end{bmatrix} \quad (2-4)$$

式(2-4)简写为:  $R^+ a = r$  (2-5)

其中:  $a = [a_1, a_2, \dots, a_p]^T$

$r = [R^+(1), R^+(2), \dots, R^+(p)]^T$

$$R^+ = \begin{bmatrix} R^+(0) & R^+(1) & \cdots & R^+(p-1) \\ R^+(1) & R^+(0) & \cdots & R^+(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R^+(p-1) & R^+(p-2) & \cdots & R^+(0) \end{bmatrix}$$

则  $a = (R^+)^{-1} r$

令  $s_n = [s(n-1), s(n-2), \dots, s(n-p)]^T$ , 则由式(2-1)

可得纯净语音信号的线性预测误差为:

$$e(n) = s(n) - a^T s_n \quad (2-6)$$

可以用相对误差能量  $g$  来确定预测误差相对于因信号的大小:

$$g = E_e / E \quad (2-7)$$

式中,  $E_e$  和  $E$  分别为  $e(n)$ 、 $s(n)$  在分析帧内的能量分贝值:

$$E_e = \sum_{n=0}^{N-1} \lg e^2(n), E = \sum_{n=0}^{N-1} \lg s^2(n) \quad (2-8)$$

$N$  为语音帧长度,  $g$  小于 1, 并随着采样信号之间的线性相关程度增大而减小。

### 3 噪声语音信号的单边自相关线性预测误差分析

参见文献[4]。

设式(1-5)中  $d(n)$  为白噪声信号, 且  $s(n)$  与  $d(n)$  不相关, 则

$$a_x = R_x^+{}^{-1} r_x \quad (3-1)$$

式中,  $R_x$ 、 $r_x$  分别是  $x(n)$  的短时单边自相关矩阵和短时自相关矢量。令  $R_s$ 、 $r_s$  分别是  $s(n)$  的短时自相关矩阵和短时自相关矢量;  $R_d$ 、 $r_d$  分别是  $d(n)$  的短时自相关矩阵和短时自相关矢量。由于  $s(n)$  与  $d(n)$  不相关, 因此:

$$a_x = (R_s + R_d)^{-1} (r_s + r_d) \quad (3-2)$$

令:  $x_n = [x(n-1), x(n-2), \dots, x(n-p)]^T$ ,

$d_n = [d(n-1), d(n-2), \dots, d(n-p)]^T$

信号  $x(n)$  的线性预测误差为

$$\begin{aligned} e_x(n) &= x(n) - a_x^T x_n \\ &= [s(n) - a_x^T s_n] + [d(n) - a_x^T d_n] \end{aligned} \quad (3-3)$$

设

$$e_x(n) = e_s(n) + e_d(n)$$

此时

$$e_s(n) = s(n) - a_x^T s_n,$$

$$e_d(n) = d(n) - a_x^T d_n$$

令  $e_x(n)$ ,  $e_s(n)$ ,  $e_d(n)$  的  $\alpha$  次方的和式分别为:  $SU_{ex}$ ,  $SU_{es}$ ,  $SU_{ed}$ 。那么,  $SU_{ex}$  的期望值为:

$$\begin{aligned} E[SU_{ex}] &= E\left[\sum_{n=0}^{N-1} e_x^\alpha(n)\right] \\ &= E\left[\sum_{n=0}^{N-1} (e_s(n) + e_d(n))^\alpha\right] \\ &= \left\{ \sum_{n=0}^{N-1} E[e_s^\alpha(n) + e_d^\alpha(n) + \sum_{i=1}^{\alpha-1} C_\alpha^i e_s^i(n) e_d^{\alpha-i}(n)] \right\} \\ &= \sum_{n=0}^{N-1} E[e_s^\alpha(n) + e_d^\alpha(n)] \\ &= E[SU_{es} + SU_{ed}] \end{aligned}$$

所以,  $SU_{ex} \approx SU_{es} + SU_{ed}$

根据  $d(n)$  的短时单边自相关函数, 可以对  $SU_{ed}$  进行计算。设  $a_{x0} = -1$ , 则有:

$$\begin{aligned} SU_{ex} &= \sum_{n=0}^{N-1} e_x^\alpha(n) \approx NE[e_x^\alpha(n)] \\ &= NE\left\{ \left[ \sum_{i=0}^p (-a_{xi} d(n-i)) \sum_{j=0}^p (-a_{yj} d(n-j)) \cdots \right] \right\} \\ &= NE\left\{ \left[ \sum_{i=0}^p \sum_{j=0}^p (-a_{xi} d(n-i)) (-a_{yj} d(n-j)) \cdots \right] \right\} \\ &= N \left\{ \sum_{i=0}^p \sum_{j=0}^p \cdots \sum E[(-a_{yj} d(n-j)) (-a_{xi} d(n-i)) \cdots] \right\} \\ &= N \left\{ \sum_{i=0}^p \sum_{j=0}^p \cdots \sum E[a_{yj} a_{xi} \cdots d(n-j) d(n-i) \cdots] \right\} \\ &\approx N \left\{ \sum_{i=0}^p \sum_{j=0}^p \cdots \sum a_{yj} a_{xi} \cdots R_d(|i-j| \cdots) \right\} \\ &= (1 + \sum_{i=1}^p a_{xi}^\alpha) R_d(0) \end{aligned}$$

由上可以导出  $x(n)$ 、 $s(n)$ 、 $d(n)$  帧内满足:

$$SU_x \approx SU_s + SU_d$$

$$l_d = \frac{SU_{ed}}{SU_d} = 1 + \sum_{i=1}^p a_{xi}^\alpha; \quad l_s \approx \frac{SU_{ex} - SU_{ed}}{SU_x - SU_d}$$

则  $l_d > l_s$  [4]。

### 4 单边自相关线性预测误差算法基本原理与框图

用单边自相关序列来替代语音，需要采用适当的模型来得到该序列的线性预测系数，以避免采用低阶的 AR 模型计算含噪声较强的语音自相关函数所引起的严重失配。采用高阶的 AR 模型能够降低模型失配而引起的偏差，只是随着阶数的增加，假共振峰出现的可能性也增大。为了抑制假共振峰的出现，阶数最大可能取为  $N/2$ ， $N$  为每帧语音的样点数[5]。

对具有一定信噪比的含噪声语音，得到其单边自相关序列的线性预测系数后，由于  $l_d > l_s$ ，因此，缩小含噪声语音的预测残差，抑制其能量，可以提高语音信噪比，从而在一定程度上达到去噪的目的。同时，经过这样的处理，语音能量的损失相对其本身而言是有限的。

因此，按一定比例，从原始含噪声信号中减去预测误差部分，就可提高信噪比，即

$$\begin{aligned} \hat{x}(n) &= x(n) - ue_x(n) \\ &= \mathbf{a}_x^T \mathbf{x}_n + e_x(n) - ue_x(n) \\ &= \mathbf{a}_x^T \mathbf{x}_n + (1-u)e_x(n) \end{aligned} \quad (4-1)$$

式 (4-1) 中  $u$  的取值应适当，若取值太大，语音信息损失会很大；若取值太小，去噪效果又不明显。我们选择  $u$  的原则是，使噪声误差  $\alpha$  次方的减小量  $SU_{ex}$  近似等于噪声  $\alpha$  次方  $SU_d$ ，如下式示：

$$\begin{aligned} \Delta SU_{ex} &= \left\{ \sum [e_d^\alpha(n)] - \sum [(1-u)e_d(n)]^\alpha \right\} \\ &= \sum [1 - (1-u)^\alpha] e_d^\alpha(n) \\ &= [1 - (1-u)^\alpha] l_d SU_d \end{aligned}$$

解方程可得： $u = 1 - (1 - \frac{1}{l_d})^{\frac{1}{\alpha}}$ ， $u$ ， $l_d$  及  $\alpha$  的

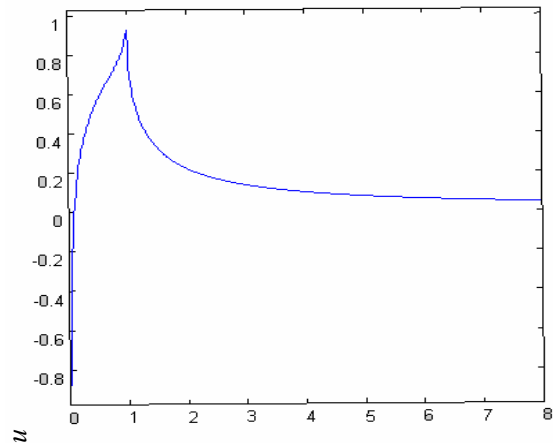
关系曲线如图 1 所示

### 5 试验结果与讨论

试验中采用的数据包括 10 个男声和 10 个女声，噪声为白噪声，含噪声程度分为 4 种，分别是信噪比 >25、20、10、0 dB。采样率 8 kHz，数据帧长 256，采用汉明窗，LPC 阶数为 18。

去噪前后，语音各帧信噪比如表 1 所示。

由表 1 可见，经过去噪处理后，含噪声语音的信噪比有了较大的提高。将平稳白噪声（峰值为 1V）和平



单边自相关线性预测误差算法框图如图 2 所示。

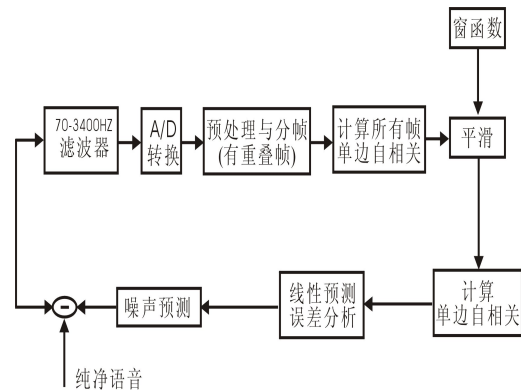


图 2 单边自相关线性预测误差算法框图

表 1 去噪前后语音各帧信噪比

帧	6	7	8	9	10	20	30	40
去噪前	-29	-5	-1	0.1	0.2	51	-10	-29
去噪后	-9	0	18	19	15	70	0	-10

稳高斯白噪声（均方根值为 1V）作为干扰信号，考察处理前后的平均功率变化，考察结果表明噪声明显降低。

表 2 平稳白噪声和 平稳高斯白噪声 考察处理前后的平均功率变化

噪声类型	平稳白噪声	平稳高斯白噪声
平均输入功率/W	0.2362	1.0034
平均输出功率/W	0.0098	0.0132

$\alpha = 2.8$  时的去噪语音与纯净语音共振峰短时频谱差值的比较（纵轴为共振峰短时频谱差值的相对变化量），如图 3 所示。由图 3 可见。

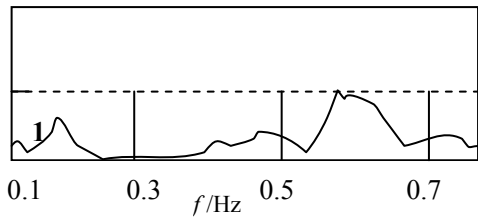


图3 去噪语音与纯净语音共振峰短时频谱差值的比较

去噪后的语音与纯净语音相比较,变化很小。测听试验表明,去噪后的语音比去噪前的明显清晰,并且语音音质和音色保持良好。

## 6 结论

本文在分析语音单边自相关的基础上,提出了单边自相关线性预测语音去噪的方法。这种方法是基于单边自相关抗噪声的特性,以信号的单边自相关序列替代信号本身,并获得相关的线性预测系数,通过抑制预测误

差中的噪声成分来控制语音中噪声的能量,以达到去噪的目的。这样的处理算法,对语音的频谱结构影响很小,从而保证了语音的质量和可懂度。

## References (参考文献)

- [1] Hermando J, Nadeu C. Speech recognition in noisy car environment based on OSALPC representation and robust similarity measuring technique [J]. Proc IEEE Internet Conf Acoust Speech Signal Processing, 1994, 2:69-72.
- [2] Hermando J, Nadeu C. Linear prediction of the one-sided autocorrelation sequence for noisy speech recognition [J]. IEEE Transactions on Speech Audio Processing, 1997, 5(1): 80-84.
- [3] McGinn D P, Johnson D H. Reduction of all-pole parameter estimation bias by successive autocorrection [J]. Proc IEEE Internet Conf Acoust Speech Signal Processing, 1983, 2: 1088-1091.
- [4] Yegnanarayana B, Carlos Avendano, Hynek Herma- nsky, et al. Speech enhancement using linear pred- iction residual [J]. Speech communication, 1999, 28: 25-42.
- [5] SM.key, Jianguo Huang ,Yanxiang Wu,Shixing. Yang, Spectral estimation Principle and Application, Science Press, Beijing. 1994 estimation Principle and Application, Science Press, Beijing. 1994.