

## 基于小波变换的自适应多分辨率语音增强算法<sup>1</sup>

郑元谨 李乐民 闻懋生\*

(电子科技大学光纤通信国家重点实验室 成都 610054)

\*(西安交通大学信息与通信工程系 西安 710049)

**摘要** 本文提出了一种基于小波变换的自适应多分辨率语音增强算法,它在尺度上和尺度间同时对受噪声污染的语音信号作自适应滤波处理,从而使得对听觉影响最严重的频段上的噪声被有效地滤除掉,滤波后的语音信噪比和主观听觉质量都得到了很大的改善。

**关键词** 小波变换,多分辨分析,自适应滤波,语音增强

**中图分类号** TN912.3

### 1 引言

从有噪信号中提取出有用信号是一个传统的问题,其根本原理都是利用信号与噪声的不同统计特性来分离信噪,但它本身就是一个病态问题,很多算法都是基于对信号加噪声作带通滤波以提取原始信号,这主要是考虑到噪声一般占据很宽的频带,而信号的带宽一般是有限的,主要集中在低频段。一般自适应滤波的 LMS 算法是针对平稳信号,而语音信号是准周期的非平稳信号,可以预计采用 LMS 算法滤波效果不会很好。文献 [1] 中采用适当调整合成滤波器柜,也就是内插算子的传递函数,以最小均方误差作为判则,通过解一组线性方程组得一组最佳合成滤波器。与文献 [1] 不同,本文提出一种自适应多分辨去噪算法,它先把原信号在多个尺度上分解,每个尺度上的信号是原信号在不同频带上的带通信号,对不同尺度上和尺度间的信号作线性组合,使组合以后的信号与原信号之间的均方误差为最小。算法采用迭代修正权系数的方法,使得滤波以后的信号与原信号之间的均方误差逐渐减小,从而达到自适应滤波的效果。

### 2 小波变换自适应滤波去噪的原理

#### 2.1 多分辨分析和小波变换等效滤波器

Mallat 最早建立了多分辨分析框架与小波分析的关系<sup>[2]</sup>。把一个信号在有限尺度上作小波分解,小尺度对应着信号的高频分量,时域分辨率高;大尺度对应着信号的低频分量,频域分辨率高。

实际的物理信号分辨率是有限的,设分解的最大尺度是  $J$ , 定义一个平滑算子  $S_{2^j} f(t)$ :

$$S_{2^j} f(t) = f(t) * \varphi_{2^j} f(t). \quad (1)$$

同时定义一个补偿算子  $W_{2^j} f(t)$

$$W_{2^j} f(t) = f(t) * \psi_{2^j} f(t). \quad (2)$$

<sup>1</sup> 1996-07-08 收到, 1997-08-11 定稿

其中“\*”表示卷积，即： $y(n) = f(n) * g(n) = \sum_{k=-\infty}^{+\infty} f(k)g(n-k) = \sum_{k=-\infty}^{+\infty} g(k)g(n-k)$ ，以下表示同。

经平滑算子作用后所得信号  $S_{2^j} f(t)$  表示在尺度  $2^j$  上信号基本变化趋势，而  $W_{2^j} f(t)$  表示在尺度  $2^j$  上信号的细节部分， $S_{2^j} f(t)$  虽然丢失原信号  $f(t)$  中的众多高频成份，但完全可由  $\{W_{2^j} f(t)\}$  来补偿。

Mallat 给出了离散二进小波变换的快速算法<sup>[3]</sup>，在本文中，我们取的小波  $\psi(x)$  是紧支集二次样条小波，它是三次样条函数的一阶导数，其等效滤波器系数，请参见文献 [3]。

如果设

$$d_n = f * \phi(n), \quad H(\omega) = \sum_{n=0}^{+\infty} h_n e^{-in\omega}, \quad G(\omega) = \sum_{n=0}^{+\infty} g_n e^{-in\omega}, \quad K(\omega) = \sum_{n=0}^{+\infty} k_n e^{-in\omega}, \quad (3)$$

则有：分解算法：

$$\left. \begin{aligned} f * \phi_{2^j}(x) &\Rightarrow \bar{f}(\omega)\bar{\phi}(2^j\omega) = \bar{f}(\omega)\bar{\phi}(2^{j-1}\omega)H(2^{j-1}\omega) \Rightarrow S_{2^j} f(x) = \sum_n h_n S_{2^{j-1}} f(x - 2^{j-1}n), \\ f * \psi_{2^j}(x) &\Rightarrow \bar{f}(\omega)\bar{\psi}(2^j\omega) = \bar{f}(\omega)\bar{\phi}(2^{j-1}\omega)G(2^{j-1}\omega) \Rightarrow W_{2^j} f(x) = \sum_n g_n S_{2^{j-1}} f(x - 2^{j-1}n), \\ f * \chi_{2^j}(x) &\Rightarrow \bar{f}(\omega)\bar{\chi}(2^j\omega) = \bar{f}(\omega)\bar{\phi}(2^{j-1}\omega)K(2^{j-1}\omega) \Rightarrow T_{2^j} f(x) = \sum_n k_n S_{2^{j-1}} f(x - 2^{j-1}n). \end{aligned} \right\} \quad (4)$$

其中  $\bar{f}(\omega)$  表示  $f(x)$  的傅里叶变换， $\bar{\phi}(\omega)$  表示  $\phi(x)$  的傅里叶变换。重构算法：

$$\begin{aligned} \overline{S_{2^j} f}(\omega) &= \bar{f}(\omega)\bar{\phi}(2^{j-1}\omega) = \bar{f}(\omega)\bar{\phi}(2^{j-1}\omega)[|H(2^{j-1}\omega)|^2 + G(2^{j-1}\omega)K(2^{j-1}\omega)] \\ &= \overline{S_{2^j} f}(\omega)\bar{H}(2^{j-1}\omega) + \overline{W_{2^j} f}(\omega)K(2^{j-1}\omega), \end{aligned}$$

即

$$S_{2^{j-1}} f(x) = \sum_n [\bar{h}_{-n} S_{2^j} f(x - 2^{j-1}n) + k_n W_{2^j} f(x - 2^{j-1}n)]. \quad (5)$$

上面介绍的二进小波变换的快速算法可采用等效滤波器的结构，如图 1 所示。

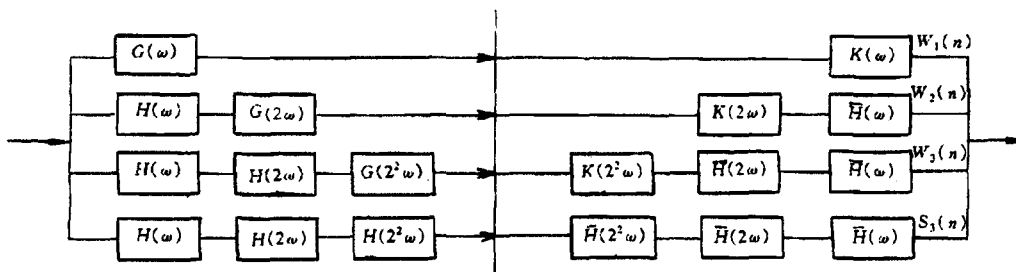


图 1 二进小波变换一种等效滤波器结构

### 2.2 自适应滤波算法

设等效滤波器的输出分别为  $W_i(n)$ ,  $i = 1, 2, \dots, J$  和  $S_J(n)$ . 设第  $i$  个尺度上小波分解序列的滤波系数为  $\{aa(i, k), k = 1, 2, \dots, NN\}$ , 在尺度  $j = J$  上平滑序列的滤波系数为  $\{bb(k), k = 1, 2, \dots, NN\}$ , 尺度间的滤波系数为  $\{a(k)_{k=1,2,\dots,J}, b\}$ . 这样如果在各个尺度上滤波可得

$$\left. \begin{aligned} W_i(n) &= \sum_{k=1}^{NN} W_i(n-k)aa(i, k), \\ S_J(n) &= \sum_{k=1}^{NN} S_J(n-k)bb(k). \end{aligned} \right\} \quad (6)$$

再对各个尺度间的小波系数作线性组合可得  $f(n)$  的一个估计  $\hat{f}(n)$ :

$$\hat{f}(n) = \sum_{i=1}^J a(i)W_i(n) + bS_J(n) = \sum_{i=1}^J \sum_{k=1}^{NN} a(i)aa(i, k)W_i(n-k) + b \sum_{k=1}^{NN} bb(k)S_J(n-k). \quad (7)$$

要使  $f(n)$  与  $\hat{f}(n)$  的均方误差最小, 即使  $e(kk) = \sum_{n=1}^u [f(n) - \hat{f}(n)]^2$  ( $e(kk)$  表示迭代到第  $kk$  步时的均方误差) 最小, 我们可以采用 LMS 算法来修正系数.

## 3 算法设计和实验结果

### 3.1 算法优化设计

在实际的算法设计中, 考虑到运算复杂性和存储量的大小, 我们对上述算法作一修正, 设计一种二维的自适应滤波器来实现语音信号的去噪处理. 首先对原信号作  $N$  个尺度上的小波变换, 然后用等效滤波器恢复可得  $N$  个小波分解序列  $W_i(n)$ ,  $i = 1, 2, \dots, N$  和低通平滑序列  $S_N(n)$ , 假设每个尺度上的滤波系数为固定长度  $NN$ , 即  $SS = [ss(1), ss(2), \dots, ss(NN)]$ , 而尺度间的滤波系数设其长度为  $N$ , 即  $S = [s(0), s(1), s(2), \dots, s(N)]$ , 再定义矩阵  $WW(k)$  为

$$WW(k) = \begin{bmatrix} S_J(k-1) & S_J(k-2) & S_J(k-3) & \dots & S_J(k-NN) \\ W_1(k-1) & W_1(k-2) & W_1(k-3) & \dots & W_1(k-NN) \\ W_2(k-1) & W_2(k-2) & W_2(k-3) & \dots & W_2(k-NN) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ W_N(k-1) & W_N(k-2) & W_N(k-3) & \dots & W_N(k-NN) \end{bmatrix} \quad (8)$$

显然, 原信号  $f(k)$  的一个逼近  $\hat{f}(k)$  可以表示为

$$\hat{f}(k) = S \cdot WW(k) \cdot SS^T. \quad (9)$$

这样要使均方误差:

$$e(kk) = \sum_{n=1}^u [f(n) - \hat{f}(n)]^2 \quad (e(kk) \text{ 表示迭代到第 } kk \text{ 步时的均方误差}) \quad (10)$$

为最小, 可采用以下修正 LMS 算法:

$$\partial e / \partial S = -2 \sum_{k=1}^u [f(k) - \hat{f}(k)] S S \cdot W W(k)^T, \quad (11)$$

$$\partial e / (\partial S S) = -2 \sum_{k=1}^u [f(k) - \hat{f}(k)] S \cdot W W(k),$$

$$\left. \begin{aligned} S S_{kk+1} &= S S_{kk} - \mu \partial e / (\partial S S), \\ S_{kk+1} &= S_{kk} - \mu \partial e / \partial S, \end{aligned} \right\} \quad (12)$$

其中  $\mu$  为步长;  $\mu$  比较大时, 算法收敛速度快, 但容易发散;  $\mu$  较小时, 收敛速度慢, 但收敛精度高。

当相对误差  $\delta = \frac{e(kk-1) - e(kk)}{e(kk)} \leq T$ ,  $T$  为门限值时, 停止迭代, 输出去噪的语音信号。

### 3.2 实验结果和分析

设输入干净语音信号为  $s(k)$ ,  $k = 1, 2, \dots, 1$ , 均值为零; 受污染语音信号为  $\tilde{s}(k) = s(k) + n(k)$ ; 其中  $n(k) = \tilde{s}(k) - s(k)$  是干扰白噪声, 均值为零; 滤波去噪后输出语音信号为  $\hat{s}(k)$ 。

图 2 是对一小段语音信号 (采样率为 8kHz, 段长 512 点) 作降噪处理的结果。从均方误差曲线可见算法是收敛的, 且收敛速度很快 (大约迭代 8 次就收敛到稳态值了)。语音输入信噪比为  $\text{SNR}_b = 5.237\text{dB}$ , 迭代 9 次后输出信噪比为  $\text{SNR}_a = 20.573\text{dB}$ , 提高了 15dB 多。

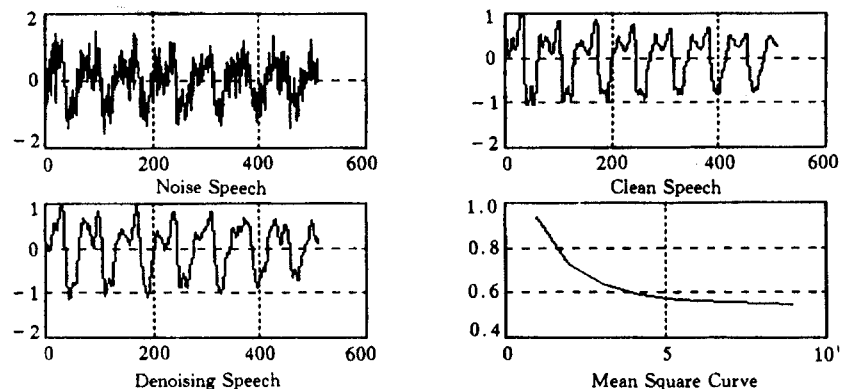


图 2 语音信号自适应多分辨降噪处理

上面指出过,  $NN$  是尺度上滤波器系数长度,  $N$  是尺度间滤波器系数长度, 在不同的输入信噪比前提下, 它们对滤波性能和收敛速度有不同的影响, 参看表 1 和表 2。

表 1 较大输入信噪比时滤波器长度  $N$  和  $NN$  对滤波性能影响 ( $SNR_b=8.733\text{dB}$ )

NN	N	
	6	4
4	$SNR_a=15.764\text{dB}$ itst=2	$SNR_a=16.153\text{dB}$ itst=3
2	$SNR_a=19.470\text{dB}$ itst=7	$SNR_a=21.056\text{dB}$ itst=9

$SNR_b$  表示输入信噪比,  $SNR_a$  表示滤波后信噪比, itst 表示到收敛时的迭代次数

表 2 低输入信噪比时滤波器长度  $N$  和  $NN$  对滤波性能影响 ( $SNR_b=0.375\text{dB}$ )

NN	N	
	6	4
4	$SNR_a=16.582\text{dB}$ itst=2 $N*NN=24$	$SNR_a=15.546\text{dB}$ itst=3 $N*NN=16$
2	$SNR_a=13.982\text{dB}$ itst=12 $N*NN=12$	$SNR_a=12.563\text{dB}$ itst=11 $N*NN=8$

$SNR_b$  表示输入信噪比,  $SNR_a$  表示滤波后信噪比, itst 表示到收敛时的迭代次数

仔细分析上面两个表, 我们可得结论: (1) 收敛速度主要由  $NN$  决定, 而与输入信噪比和  $N$  无关。当  $NN$  小时, 收敛速度慢, 当  $NN$  大时, 收敛速度快。 $NN$  越大, 尺度上的滤波进行的越快, 而收敛速度主要由各尺度上滤波速度决定, 因而收敛速度快。

(2) 滤波后的输出信噪比由输入信噪比和  $N \times NN$  之积共同决定。当输入信噪比较大时 ( $SNR_b \geq 5\text{dB}$ ),  $N \times NN$  值越大, 滤波后输出信噪比越低, 增益越小; 反之,  $N \times NN$  值越小, 滤波后输出信噪比越高, 增益越大。这说明在噪声干扰较小时, 我们应降低二维滤波器的阶数, 这样系统滤波效果反而较好。另一方面, 当输入信噪比较小时 ( $SNR_b < 5\text{dB}$ ),  $N \times NN$  值越大, 滤波后输出信噪比越高, 增益越大; 反之,  $N \times NN$  值越小, 滤波后输出信噪比越低, 增益越小。这说明在噪声干扰较大时, 提高二维滤波器的阶数, 有利于系统的滤波降噪效果。

(3) 不论输入信噪比是哪种情况, 与其相应的最佳  $N, NN$  组合滤波器都能带来  $12\text{dB}$  以上的增益, 这对改善语音质量是非常可观的。

在实际应用中, 我们可先估计输入信噪比的大小, 再根据以上的规则作综合考虑, 选定最佳的滤波器系数长度。在本章的最后, 我们对一段长为  $612.5\text{ms}$  (采样率为  $8\text{kHz}$ ) 的有噪语音信号用上面提出的滤波算法做处理, 结果如下 (图 3):

噪声语音输入信噪比为:  $SNR_b=1.433\text{dB}$ , 取滤波器参数: 尺度上滤波器长度  $NN=4$ , 尺度间滤波器长度  $N=6$ , 滤波去噪后语音输出信噪比:  $SNR_a=17.946\text{dB}$ , 滤波增益:  $GN=SNR_a - SNR_b=16.513\text{dB}$ 。

算法达到这个增益只用了 4 次迭代, 信噪比改观达  $16\text{dB}$  多, 并且听觉实验表明, 恢复语音主观听觉质量良好, 不仅听上去很清晰, 而且能区分男女声。应该指出, 通常的 LMS 算法因为是基于过去序列的线性组合来预测当前值, 因而往往假定被滤波的一段信号是平稳的。语音信号可区分为浊音和清音, 即使是浊音, 随着基音周期的变化, 语音信号各段的统计性能剧烈变化, 因而是非平稳信号, 采用 LMS 算法不能直接对语音信号滤波 (算法不收敛), 除非先对语音信号分段, 再采用 LMS 算法。大量仿真结果表明, 用 LMS 算法对语音滤波, 收敛速度慢, 信噪比增益在  $5\text{--}10\text{dB}$  之间。而多分辨率滤波算法的突出优点在于它将非平稳的语音信号分解成各尺度上的平稳信号, 然后只需对平稳信号加以处理; 它的另一个突出优点在于充分利用了尺度间的相似信息来去除噪声。

图 3 表示了有噪波形, 无噪波形和去噪波形的对比关系。

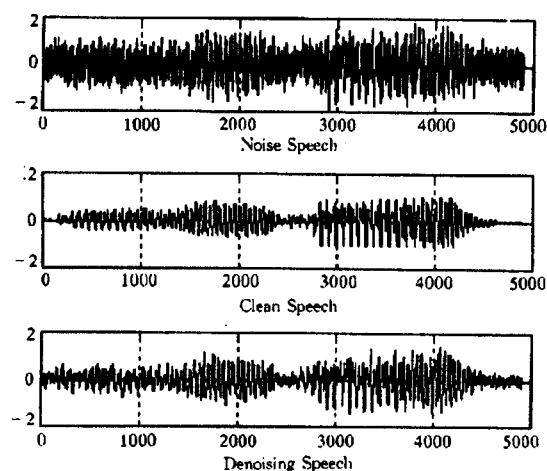


图 3 一段受扰语音去噪结果

## 4 结 论

本文提出了一种多分辨分析自适应滤波去噪的算法，并成功地应用于受污染语音信号的增强处理，信噪比和主观听觉质量的改善都是十分可观的。该算法充分利用了信号多尺度的信息，在尺度上和尺度间同时做滤波，有利于将信号和噪声作多分辨率的分离，达到最佳消噪效果。算法的设计体现了小波变换与神经网络结合的思想<sup>[4-7]</sup>，不仅仅能应用于语音信号去噪，对其它有噪信号一样能进行处理。

## 参 考 文 献

- [1] Delopoulos A, Kollias S. Optimal filter banks for signal reconstruction from noisy subband components. *IEEE Trans. on SP*, 1996, SP-44(2):212-224.
- [2] Mallat S. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. on PAMI*, 1989, PAMI-11(7): 674-691.
- [3] Mallat S, Liang Hwang Wen. Singularity detection and processing with wavelet. *IEEE Trans. on IT*, 1992, IT-38(2): 617-643.
- [4] Qinghua Zhang, Benveniste Albert. Wavelet networks. *IEEE Trans. on NN*, 1992, NN-3(6): 889-898.
- [5] Sahiner B, Yagle A E. Image reconstruction from projection under wavelet constraints. *IEEE Trans. on SP*, 1993, SP-41(12): 3579-3583.
- [6] Gilloire A, Vetterli M. Adaptive filtering in subband. in *Proc. IEEE Int. Conf. on ASSP*, New York: Apr., 1988, 1572-1575.
- [7] Zhang J, Walter G G, Lee W N W. Wavelet neural networks for function learning. *IEEE Trans. on SP*, 1995, SP-43(6): 1489-1496.

## AN ADAPTIVE MULTIREOLUTION SPEECH ENHANCEMENT ALGORITHM BASED ON WAVELET TRANSFORM

Zheng Yuanjin    Li Lemin    Wen Maosheng\*

*(National Key Lab. of Optical Fiber Comm.,*

*University of Electron. Science and Tech. of China, Chengdu 610054)*

*\*(Inform. and Comm. Department, Xi'an Jiaotong University, Xi'an 710049)*

**Abstract** In this paper, an adaptive multiresolution speech enhancement algorithm based on wavelet transform is put forward. It can make adaptive filtering to noise speech both at scales and among scales. So that the noise parts during the frequency intervals which decrease hearing quality mostly are reduced efficiently. Both the SNR and subject hearing quality of denoised speech are high and good.

**Key words** Wavelet transform, Multiresolution analysis, Adaptive filtering, Speech enhancement

郑元谨: 男, 1972 年生, 博士生, 主要从事神经网络在通信系统中的运用, 信号自适应多分辨分析处理, 小波分析理论及应用, 语音压缩编码, ATM 建模及性能分析。

李乐民: 男, 1932 年生, 教授, 博士生导师, 中国工程院院士, 《通信学报》编委, 光纤通信国家重点实验室主任, 主要研究领域为 ATM 建模及复用技术, 多媒体通信技术, 计算机通信网络及无线 CDMA。

闻懋生: 男, 1932 年生, 教授, 博士生导师, 西安交通大学信息与通信技术研究所所长, 主要研究领域为语音信号处理, 计算机通信网络, 信道优化与均衡, 宽带 ISDN 和交换技术等。