

脉冲噪声环境下基于卡尔曼滤波的语音增强

何志勇¹,朱忠奎²

(1. 苏州大学 机电工程学院,江苏 苏州 215021; 2. 苏州大学 城市轨道交通学院,江苏 苏州 215021)

(hezhiyong@suda.edu.cn)

摘要:语音增强的目标在于从含噪信号中提取纯净语音,纯净语音在某些环境下会被脉冲噪声所污染,但脉冲噪声的时域分布特征却给语音增强带来困难,使传统方法在脉冲噪声环境下难以取得满意效果。为在平稳脉冲噪声环境下进行语音增强,提出了一种新方法。该方法通过计算确定脉冲噪声样本的能量与含噪信号样本的能量之比最大的频段,利用该频段能量分布情况逐帧判别语音信号是否被脉冲噪声所污染。进一步地,该方法只在被脉冲噪声污染的帧应用卡尔曼滤波算法去噪,并改进了传统算法执行时的自回归(AR)模型参数估计过程。实验中,采用白色脉冲噪声以及有色脉冲噪声污染语音信号,并对低输入信噪比的信号进行语音增强,结果表明所提出的算法能显著地改善信噪比和抑制脉冲噪声。

关键词:脉冲噪声;语音增强;卡尔曼滤波;自回归模型;滤波器组

中图分类号: TN912.3 **文献标志码:** A

Removal of impulsive noise based on Kalman filtering for speech enhancement

HE Zhi-yong¹, ZHU Zhong-kui²

(1. School of Mechanical and Electrical Engineering, Soochow University, Suzhou Jiangsu 215021, China;

2. School of Urban Rail Transportation, Soochow University, Suzhou Jiangsu 215021, China)

Abstract: Speech enhancement aims at extracting clean speech from noisy speech. Traditional speech enhancement algorithms proposed for continuous noise can not work well when clean speech is polluted by impulsive noise. The feature of the distribution of impulsive noise in time domain has to be concerned for speech enhancement. This paper provided a new speech enhancement method for the removal of noise when clean speech is polluted by stationary impulsive noise. By analyzing the samples of impulsive noise and noisy speech, the method found the frequency band in which the energy of impulsive noise was most evident compared with the energy of noisy speech. Furthermore, filter banks were used in recognizing whether the speech frame was polluted. The method based on Kalman filtering removed impulsive noise from the polluted signal frame. Compared with the traditional algorithms based on Kalman filtering, the method also improved the process of the parameters estimation of Auto Regressive (AR) model. In some speech enhancement tests, speech was polluted by white impulsive noise or colored impulsive noise, the results show that the method has achieved an improved Signal-to-Noise Ratio (SNR) and removes impulsive noise even when the SNR of noisy speech is low.

Key words: impulsive noise; speech enhancement; Kalman filter; autoregressive model; filter banks

0 引言

语音增强的目标在于从被污染的信号中提取纯净语音信号。对于某些环境下的脉冲噪声,如高频电磁干扰造成的通信系统噪声和扩音器噪声等,当它们污染语音时,传统的适于连续噪声的语音增强算法并不能取得好的效果。对脉冲噪声环境下的语音增强,为适应脉冲噪声在时域中的非连续和间歇特征,应当在时域中较准确地识别脉冲噪声时域分布,以使噪声消除算法只在脉冲噪声存在的时段运行,而对非脉冲噪声段保留原信号以避免在该段因运算引起语音信号失真。

文献[1]中基于一维语音信号的自回归(Auto Regressive, AR)模型提出了脉冲噪声的检测方法,该方法受模型估计的准确性影响较大。文献[2-4]也描述了一些对脉冲噪声检测和降噪的算法,这些算法都没有在时频域中对信号进行分析。文献[5]中引入了基于离散小波变换的方法

消除脉冲噪声,该方法的优点是在时频域对脉冲噪声进行滤波,但对于高频段信息的分析并没有足够的分辨率。

卡尔曼滤波在白色及有色噪声环境下的语音增强中得到了成功的应用^[6-9],这些算法逐帧估计语音信号的AR模型,并对每帧信号用卡尔曼滤波器滤波,取得了好的效果。此外,在传统基于卡尔曼滤波器的语音增强算法基础上,文献[10-11]将卡尔曼滤波与小波及人耳听觉掩蔽特性等相结合发展了一些新的有效的算法。但上述基于卡尔曼滤波的算法均不能直接用于脉冲噪声,因为并不是每帧信号都被噪声污染。针对平稳脉冲噪声环境,本文提出了如下的基于卡尔曼滤波的语音增强方法:在比较脉冲噪声样本与含噪信号样本的频域能量分布特征的基础上,用滤波器组对含噪信号逐帧进行分解并判别信号帧是否被噪声所污染,进而可以采用传统卡尔曼滤波语音增强算法^[6-7]对时域内被脉冲噪声污染的帧滤波,并改进了卡尔曼滤波算法中对信号模型参数估计的方法。

收稿日期:2011-06-21;修回日期:2011-07-31。

基金项目:国家自然科学基金资助项目(50905121);江苏省高校自然科学基金资助项目(10KJD510008)。

作者简介:何志勇(1976-),男,内蒙古包头人,讲师,硕士,主要研究方向:信号检测与处理、机器视觉;朱忠奎(1974-),男,山东济宁人,副教授,博士,主要研究方向:机械故障诊断、信号检测与分析。

仿真实验中对一些被白色及有色脉冲噪声污染的语音信号进行脉冲噪声污染帧的判别和滤波,都获得了较好的语音增强效果。

1 信号模型

含噪语音信号的数学模型可表示为:

$$y(k) = s(k) + n(k); k = 1, 2, \dots, L \quad (1)$$

其中: $s(k)$ 和 $n(k)$ 分别表示纯净语音信号和脉冲噪声, $y(k)$ 表示含噪语音信号, L 为信号长度。式(1)亦可以被改写为向量形式:

$$\mathbf{Y} = \mathbf{S} + \mathbf{N} \quad (2)$$

其中: \mathbf{Y} 表示含噪语音信号序列, \mathbf{S} 表示纯净语音信号序列, \mathbf{N} 表示脉冲噪声序列。

语音信号可以被看作短时平稳过程,对语音信号分帧,取帧长为 20 ~ 40 ms,则每帧内的信号可被表述为由白噪声驱动的、全极点线性系统输出自递归的过程^[12]。将纯净语音信号序列 $\{s(k)\}$ 用 p 阶 AR 模型表示为:

$$s(k) = \sum_{i=1}^p a_i(t)x(k-i) + w(k) \quad (3)$$

其中: $w(k)$ 为零均值的方差为 σ_w^2 的高斯白噪声序列, p 的取值范围通常为 8 ~ 12, $a_i(t)$ 为第 t 帧 AR 模型的参数。

$n(k)$ 假定为加性脉冲噪声,且脉冲噪声应满足:时域中每段噪声在频域内的能量分布特征基本一致。对于 $n(k)$ 的每段噪声,若每段噪声都是平稳的,则可用自回归滑动平均 (Autoregressive Moving Average, ARMA) 模型来逼近该过程,而 ARMA 模型可采用有足够阶数的全极点模型来拟合。因此对于每段噪声可视为由一个均值为 0, 方差为 σ_η^2 的白噪声序列 $\eta(k)$ 驱动一个 AR 过程,该段内的噪声序列可表示如下:

$$n(k) = \sum_{j=1}^q b_j n(k-j) + \eta(k) \quad (4)$$

特殊地,若每段噪声为白噪声序列,即对每段噪声有 $n(k) = \eta(k)$,因此每段噪声均值为 0, 方差为 σ_η^2 。

2 算法设计

对于含噪语音信号的语音增强算法是对 \mathbf{Y} 逐帧处理的,算法的流程如图 1 所示。

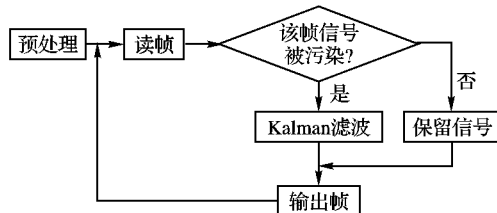


图1 算法流程

处理过程中对经判别未被脉冲噪声污染的帧数据,算法保留原信号,以避免在该帧内运行滤波算法引起的信号失真。

2.1 预处理

假设脉冲噪声的样本数据也可以采集到,而且对含脉冲噪声的语音信号也采集一段作为样本。设脉冲噪声样本信号序列为 \mathbf{NS} , 含噪语音信号样本序列为 \mathbf{YS} , 二者的采样频率相同。

从含噪信号中识别出某帧信号是否被脉冲噪声所污染,需要明确被污染的含噪信号与纯净信号之间的差异特征。由于纯净信号是不可获的,考虑到式(1),该差异特征在识别过

程中可由 \mathbf{NS} 和 \mathbf{YS} 之间的差异特征代替。

由离散小波变换理论可知,尺度函数 $\varphi(t)$ 和小波函数 $\psi(t)$ 存在如下关系:

$$\varphi(t) = \sum_n h_0(n) \sqrt{2} \varphi(2t - n)$$

$$\psi(t) = \sum_n h_1(n) \sqrt{2} \varphi(2t - n)$$

其中 $h_0(n)$ 、 $h_1(n)$ 为由系数序列组成的滤波器,且:

$$h_1(n) = (-1)^n h_0(1 - n)$$

一个函数基于尺度函数和小波函数的展开可以表示为:

$$f(t) = \sum_k c_j(k) 2^{j/2} \varphi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k)$$

展开系数 $c_j(k)$ 和 $d_j(k)$ 均可由内积运算求得:

$$c_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle$$

$$d_j(k) = \langle f(t), \psi_{j,k}(t) \rangle$$

上述计算展开系数的过程亦可采用迭代方法计算:

$$c_j(k) = \sum_m h_0(m - 2k) c_{j+1}(m)$$

$$d_j(k) = \sum_m h_1(m - 2k) c_{j+1}(m)$$

在分析 \mathbf{NS} 和 \mathbf{YS} 之间的差异特征时,首先由离散小波分解中所用的滤波器组 h_0 和 h_1 分别将 \mathbf{NS} 和 \mathbf{YS} 的每一频段不断分解为高频部分和低频部分,低频部分为对输入信号的粗略逼近,高频部分为输入信号的细节部分。如图 2 所示,滤波器组由低通滤波器 h_0 和高通滤波器 h_1 组成。滤波器组的选择有不同类型,可以是正交镜像的,或是双正交的。在图 2 中整个频域被分解为 2^j 个频段, j 为分解层数,因此 \mathbf{NS} 和 \mathbf{YS} 也在频域被二义完全分解。

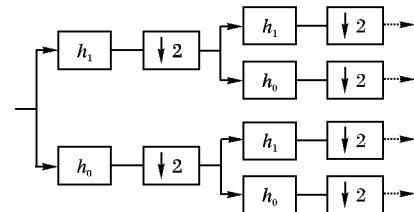


图2 滤波器组 (h_0 : 低通滤波器; h_1 : 高通滤波器)

从 Parseval 能量定理可知,信号在时域与频域的能量应该相等,因此,在用滤波器组对信号分解后,信号在每个频段上的能量即等于信号的时域序列在该频段上的映射系数的平方和。

在经滤波器组分解所得的 2^j 个频段中,算法选取 \mathbf{NS} 的能量与 \mathbf{YS} 的能量之比最大的一个频段作为判别帧信号是否被噪声污染时所考查的频段。设该频段为 j 层第 m 个频段, $m \in (0, 1, \dots, 2^j - 1)$, 且 m 值越大,所对应频段频率越高, m 值按能使式(5)成立的原则选取:

$$\frac{\sum |C_{NS}(j, m)|^2}{\sum |C_{YS}(j, m)|^2} = \max \left\{ \frac{\sum |C_{NS}(j, k)|^2}{\sum |C_{YS}(j, k)|^2} \right\} \quad (5)$$

其中, $k, m \in (0, 1, \dots, 2^j - 1)$, $C_{NS}(j, m)$ 和 $C_{NS}(j, k)$ 分别为 \mathbf{NS} 经 j 层分解后在第 m 和 k 个频段映射的系数序列, $C_{YS}(j, m)$ 和 $C_{YS}(j, k)$ 分别为 \mathbf{YS} 经 j 层分解后在第 m 和 k 个频段映射的系数序列。

对 \mathbf{Y} 进行分帧,帧长应使语音信号满足短时平稳特征,并对帧数据加窗,以减少频谱泄露。

2.2 帧信号是否被噪声污染的判别

对 \mathbf{Y} 的任意帧第 t 帧,采用 h_0 和 h_1 滤波器组按照预处理阶段获得第 j 层第 m 个频段的路径将帧信号进行分解,获得

该帧信号在第 j 层第 m 个频段的映射系数序列 $C_Y(t, j, m)$, 该序列的能量为 $\sum |C_Y(t, j, m)|^2$ 。

图3展示了一段被脉冲噪声污染了的语音信号各帧信号在用前述方法分解和计算所获得的各帧在第 m 个频段的能量; 该例中, 选用 Db4 小波分解中的滤波器组, 帧长为 20 ms, $j=3, m=6$, 语音信号的采样频率为 16 KHz, 数据长度为 101 888 个点。从图中可以看出, 噪声污染帧在第 m 个频段的能量与其余帧相比十分突出, 有利于帧信号是否被噪声污染的判别。

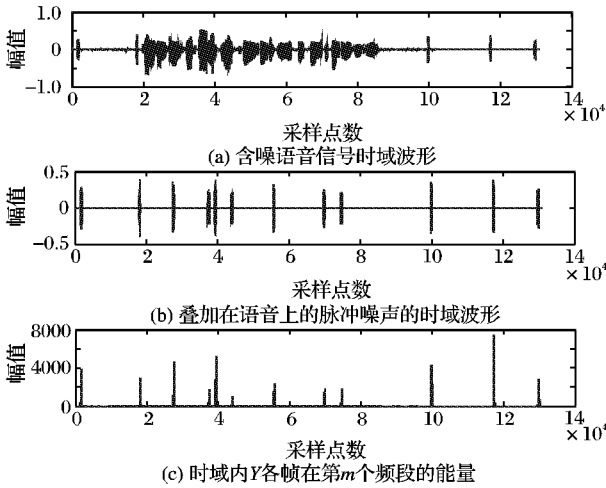


图3 Y 各帧第 m 频段能量与 Y, N 在时域分布特征上的比较

设 Y 的第 $1, 2, \dots, t-1$ 帧已经完成滤波并获得语音信号的估计值, 在处理第 t 帧时需判定第 t 帧信号是否被噪声污染。将语音信号估计值序列 \hat{S} 的第 $1, 2, \dots, t-1$ 帧在第 m 个频段的能量以序列形式表示为 $E_{\hat{S}}(i, m)$ ($i=1, 2, \dots, t-1$)。该序列中每帧能量的计算方法与计算 Y 中帧在第 j 层第 m 个频段的能量的方法相同。假设语音信号中第 m 个频段的能量在相邻帧间是缓慢变化的, 考虑到泰勒级数展开式并且忽略余项, 第 t 帧语音信号在第 m 个频段的能量可近似估计为:

$$\hat{E}_S(t, m) = E_{\hat{S}}(t-1, m) + \left. \frac{\partial [E_{\hat{S}}(i, m)]}{\partial i} \right|_{i=t-1} + \frac{1}{2} \left. \frac{\partial^2 [E_{\hat{S}}(i, m)]}{\partial i^2} \right|_{i=t-1} \quad (6)$$

式(6)中, 对于 $E_{\hat{S}}(i, m)$ 的计算采用后向差分方法, 即:

$$\left. \frac{\partial [E_{\hat{S}}(i, m)]}{\partial i} \right|_{i=t-1} = E_{\hat{S}}(t-1, m) - E_{\hat{S}}(t-2, m)$$

进一步地, 判定第 t 帧信号被噪声污染的条件为同时满足式(7)和式(8):

$$\sum |C_Y(t, j, m)|^2 > l * \hat{E}_S(t, m) \quad (7)$$

$$\sum |C_Y(t, j, m)|^2 > \frac{1}{t-1} \sum_{i=1}^{t-1} E_Y(i, m) \quad (8)$$

式(7)中, 系数 l 为阈值倍数, 其恰当值与 NS 和 YS 在第 m 个频段的帧平均能量之比呈正相关, 且与 $\hat{E}_S(t, m)$ 的估计误差有关, 误差越大, 该值越大。式(8)中要求第 t 帧在第 m 频段的能量大于 Y 过去 $t-1$ 帧在第 m 频段的能量的均值, 该条件的设置有助于消除单纯由式(7)判断时将语音信号帧判定为脉冲噪声污染帧的错误, 以避免对纯净语音帧滤波而导致的信号失真。

2.3 噪声污染帧的卡尔曼滤波

对于已确定被噪声污染的语音帧, 即可以如在连续噪声

环境下基于卡尔曼滤波的语音增强算法^[7]一样在该帧内滤波。为应用卡尔曼滤波器, 将式(1)、(3)、(4)改写为状态方程形式, 可得:

$$\begin{aligned} y(k) &= H_s S(k) + n(k) \\ S(k) &= F_s S(k-1) + G_s w(k) \\ N(k) &= F_n N(k-1) + G_n \eta(k) \end{aligned}$$

其中:

$$\begin{aligned} S(k) &= [s(k-p+1) \quad s(k-p+2) \quad \cdots \quad s(k)]^T \\ N(k) &= [n(k-q+1) \quad n(k-q+2) \quad \cdots \quad n(k)]^T \end{aligned}$$

$$F_s = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$F_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$G_s = H_s^T = [0 \quad 0 \quad \cdots \quad 0 \quad 1]^T$$

$$G_n = H_n^T = [0 \quad 0 \quad \cdots \quad 0 \quad 1]^T$$

其中, H_s 为 $1 \times p$ 阶矩阵, H_n 为 $1 \times q$ 阶矩阵。于是有:

$$\begin{bmatrix} S(k) \\ N(k) \end{bmatrix} = \begin{bmatrix} F_s & 0 \\ 0 & F_n \end{bmatrix} \begin{bmatrix} S(k-1) \\ N(k-1) \end{bmatrix} + \begin{bmatrix} G_s & 0 \\ 0 & G_n \end{bmatrix} \begin{bmatrix} w(k) \\ \eta(k) \end{bmatrix}$$

$$y(k) = [H_s \quad H_n] \begin{bmatrix} S(k) \\ N(k) \end{bmatrix}$$

令:

$$X(k) = \begin{bmatrix} S(k) \\ N(k) \end{bmatrix}, F = \begin{bmatrix} F_s & 0 \\ 0 & F_n \end{bmatrix}, G = \begin{bmatrix} G_s & 0 \\ 0 & G_n \end{bmatrix},$$

$$W(k) = \begin{bmatrix} w(k) \\ \eta(k) \end{bmatrix}, H = [H_s \quad H_n]$$

可得离散卡尔曼滤波器的时间更新方程为:

$$\hat{X}^-(k) = F \hat{X}^-(k-1)$$

$$P^-(k) = F P^-(k-1) F^T + G Q G^T$$

其中, $Q = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix}$, 对应的状态更新方程为:

$$K(k) = P^-(k) H^T [H P^-(k) H^T]^{-1}$$

$$\hat{X}(k) = \hat{X}^-(k) + K(k) [y(k) - H \hat{X}^-(k)]$$

$$P(k) = [I - K(k) H] P^-(k)$$

最终, k 时刻信号估计值为: $\hat{s}(k) = [H_s \quad 0] \hat{X}(k)$ 。

上述算法对于可用 AR 模型来拟合的平稳噪声均适用, 若 $\{n(k)\}$ 每段噪声均为方差是 σ_η^2 的白噪声, 可以采用计算量更小的卡尔曼滤波算法。首先, 令:

$$X(k) = S(k), W(k) = w(k), Q = \sigma_w^2 I, R = \sigma_\eta^2 I$$

于是, 可得离散卡尔曼滤波时间更新方程为:

$$\hat{X}^-(k) = F_s \hat{X}^-(k-1)$$

$$P^-(k) = F_s P^-(k-1) F_s^T + G_s Q G_s^T$$

对应的状态更新方程则为:

$$K(k) = P^-(k) H_s^T [R + H_s P^-(k) H_s^T]^{-1}$$

$$\hat{X}(k) = \hat{X}^-(k) + K(k) [y(k) - \hat{X}_s^-(k)]$$

$$P(k) = [I - K(k) H_s] P^-(k)$$

最终, k 时刻信号估计值为: $\hat{s}(k) = H_s \hat{X}(k)$ 。

2.4 滤波算法中 AR 模型参数的估计

滤波算法在执行时,需要明确 σ_w^2 和 σ_η^2 以及 F_s 和 F_n 。其中, σ_η^2 和 F_n 中的 b_q, b_{q-1}, \dots, b_1 可由 NS 采用线性预测系数 (Linear Predictor Coefficient, LPC) 算法求得。特殊地,若 $\{n(k)\}$ 每段噪声为白噪声,则对于 σ_η^2 可由 NS 采用统计方法求得。而对于 σ_w^2 以及 F_s 中的 a_p, a_{p-1}, \dots, a_1 , 可采用如下迭代的方法逐步逼近:对于含噪语音帧,用 LPC 算法求出 $\hat{\sigma}_w^2$ 和 $\hat{a}_p, \hat{a}_{p-1}, \dots, \hat{a}_1$, 并将这些参数代入算法中进行滤波,对于滤波后的语音估计帧,再用 LPC 算法求语音参数的估计值,并将新参数代入算法中对信号语音估计帧滤波,反复迭代。

关于迭代结束的条件,传统的卡尔曼滤波语音增强算法^[7]是考虑在 $\hat{\sigma}_w^2$ 小于某个阈值时结束迭代,并且认为一般应迭代 3~6 次。该条件涉及到阈值的合理选择,但文献^[7]中并未给出有普遍意义的阈值选择方法,并且由于不同环境下含噪语音的信噪比差异可能较大,实际的最佳迭代执行次数并不一定是 3~6 次。针对上述情况,将迭代结束的条件改进为:每次滤波后对该帧原始含噪语音序列与该次滤波获得的语音信号估计序列求差,并用 LPC 算法对差值序列计算 $\hat{\sigma}_\eta^2$, 对于白噪声则是用统计方法对差值序列计算 $\hat{\sigma}_\eta^2$, 若 $\sigma_\eta^2 - \hat{\sigma}_\eta^2$ 的值已趋于 0 或基本稳定,则可结束迭代,给出该帧最终的语音信号估计。

3 语音增强实验

实验中,语音信号的采集由麦克风、声卡及 PC 组成的采集系统完成,采样频率为 16 kHz,采集的语音信号作为纯净语音信号,由数据生成含噪语音及去噪语音时,频率亦为 16 kHz。语音采集及将数据输出为语音文件的过程,均用 Matlab 完成。

对于污染纯净信号的脉冲噪声,分别选取有色脉冲噪声信号和白色脉冲噪声信号为实验对象。首先在 Matlab 中生成连续噪声,然后利用连续噪声在 Matlab 中生成一系列在时域中起始位置、宽度都是随机的脉冲噪声段。实验中在用脉冲噪声污染纯净信号时,使用的脉冲噪声段数也是随机的。采用上述方法获得脉冲噪声序列及含噪信号,是为了实验能较好地检验算法的适应性和有效性。处理数据时,选择帧长为 20 ms,数据窗为汉明窗,帧间重叠率为 40%;预处理中选用 Db4 小波分解的滤波器组,对数据进行 3 层完全二叉分解;在卡尔曼滤波算法中取纯净语音 AR 模型阶数为 $p = 10$ 。

图 4 展示了本文的脉冲噪声消除算法在消除有色噪声方面的效果,该例中,所采用的语音数据为一段女声德语语音,内容为:“Kann ich am Samstagabend nach 22:00 von Frankfurt nach Muenchen kommen”,数据长度为 101 888 个点。在图 4 所示实验中生成连续有色噪声时采用全极点 AR 过程,极点分别为 $0.9e^{-\frac{3}{8}j\omega}, 0.9e^{\frac{3}{8}j\omega}, 0.9e^{-\frac{7}{8}j\omega}, 0.9e^{\frac{7}{8}j\omega}$ 。逼近有色噪声的 AR 模型阶数为 $q = 6$ 。从图 4 中可以看出脉冲噪声被较好地抑制,从图 4 中(c)与(b)的波形比较亦可以看出,算法降噪时并未引起纯净信号的明显失真。

当算法应用于对被白色脉冲噪声污染的语音进行增强时,算法的效果如图 5 中的语谱图所示。该例中,语音数据为一段男声德语语音,内容为:“Ich brauche den naechsten Zug Muenchen”,数据长度为 42 496 个点。

在图 5(a)中,可以清晰地看到时域中存在若干段白色脉

冲噪声,从与图 5(b)的比较中可以看出,一部分在时域中与语音重叠,还有一些则处于无语音信号的时域区间。从图 5(c)、(b)和(a)的比较中,可以看出算法较好地抑制了脉冲噪声,同时语音信号在时频域的能量分布保留得较好,没有大的能量损失。

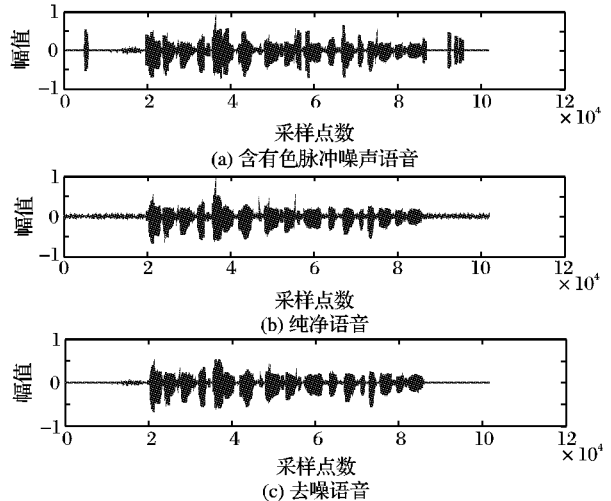


图4 信号时域波形

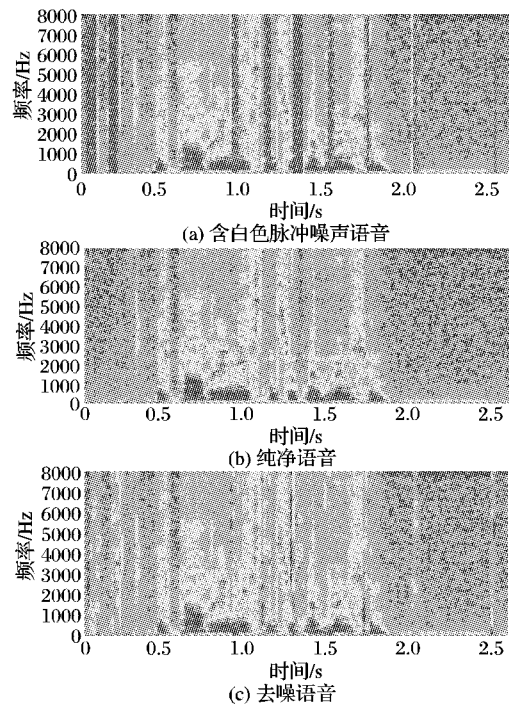


图5 信号语谱图

比较算法的信噪比(SNR)有助于客观地评价算法的有效性。在此,定义输入信噪比为:

$$SNR_{in} = 10 \lg \frac{\sum s(k)^2}{\sum [y(k) - s(k)]^2}$$

而输出信噪比定义为:

$$SNR_{out} = 10 \lg \frac{\sum s(k)^2}{\sum [\hat{s}(k) - s(k)]^2}$$

表 1 中分别列出了当信号被有色脉冲噪声或白色脉冲噪声污染,且输入信噪比较低的情况下,不同去噪算法所获得的输出信噪比。在表中,算法 2 为文献^[5]中所描述算法,该算法对信号进行离散小波分解,在小波域内检测脉冲噪声,并随不同的脉冲噪声映射区间使用 Donoho 阈值法^[13]和软阈值策

略分别计算滤波阈值;算法 3 为在整个时域未进行语音帧是否被噪声污染的判别,而直接对每一帧都使用卡尔曼滤波的算法。从表中所列实验结果可以看出,本文的新算法可以使输出信噪比比输入信噪比显著提高,且提高幅度在几种算法

中是最大的。对算法去噪后的语音信号,依然采用 16 kHz 的频率生成语音文件,并组织听力测试,结果表明去噪后的语音信号虽较原纯净语音信号有轻微变化,但语音可懂度保留较好,且噪声被明显抑制。

表 1 脉冲噪声环境下语音增强算法的输出信噪比 dB

SNR _{in}	白色脉冲噪声			SNR _{in}	有色脉冲噪声		
	本文算法	算法 2	算法 3		本文算法	算法 2	算法 3
-2.1992	14.9306	6.5640	0.1563	-1.4920	16.8432	10.7491	5.3509
-1.8609	11.3629	6.9827	0.9847	1.6628	17.4650	11.8369	6.8087
3.7350	17.2579	10.8689	5.2106	2.5427	19.9721	16.4852	9.3625
5.5784	20.3964	13.0124	6.3148	5.1864	20.4591	17.3470	12.3846

4 结语

针对脉冲噪声环境下的语音增强,本文提出一种新方法。首先,新方法利用滤波器组对噪声和含噪语音样本进行分解,确定一个脉冲噪声样本能量与含噪语音样本能量之比最大的频段,然后对含噪语音逐帧考查该频段的能量状况,并对每帧是否被脉冲噪声污染进行逐帧判别;然后在降噪时,新方法只对被噪声污染帧应用卡尔曼滤波器进行去噪,并改进了传统卡尔曼滤波语音增强算法在估计信号模型参数时的迭代结束条件。语音增强实验表明,新方法对被平稳脉冲噪声污染的语音能有效地改善信噪比,对脉冲噪声抑制效果明显,并且不会引起语音信号的明显失真。

参考文献:

- [1] ESQUEF P, KARJALAINEN M, VALIMAKI V. Detection of clicks in audio signals using warped linear prediction[C]// The 14th International Conference on Digital Signal Processing. New York: IEEE, 2002, 2: 1085 - 1088.
- [2] KERPEZ K J. Minimum mean squared error impulse noise estimation and cancellation[J]. IEEE Transactions on Signal Processing, 1995, 43(7): 1651 - 1662.
- [3] LABEAU F, CHIANG J C, KIEFFER M, *et al.* Oversampled filter banks as error correcting codes: theory and impulse noise correction[J]. IEEE Transactions on Signal Processing, 2005, 53(12): 4619 - 4630.
- [4] ZAHEDPOUR S, FEIZI S, AMINI A, *et al.* Impulsive noise cancellation based on soft decision and recursion[J]. IEEE Transactions on Instrumentation and Measurement, 2009, 58(8): 2780 - 2790.
- [5] NONGPIUR R C. Impulse noise removal in speech using wavelets

- [C]// IEEE International Conference on Acoustics, Speech, and Signal Processing. New York: IEEE, 2008: 1593 - 1596.
- [6] PALIWAL K K, BA A. A speech enhancement method based on Kalman filtering[C]// IEEE International Conference on Acoustics, Speech, and Signal Processing. New York: IEEE, 1987: 177 - 180.
- [7] 沈亚强,程仲文. 建立在卡尔曼滤波基础上的语音增强方法[J]. 声学学报, 1994, 19(3): 227 - 233.
- [8] SO S, PALIWAL K K. A long state vector Kalman filter for speech enhancement[EB/OL]. [2011 - 02 - 20]. <http://www.interspeech2008.org/>.
- [9] SO S, PALIWAL K K. Fast converging iterative Kalman filtering for speech enhancement using long and overlapped tapered windows with large side lobe attenuation[EB/OL]. [2011 - 02 - 21]. http://www.isca-speech.org/archive/interspeech_2010/i10_1081.html.
- [10] MA NING, BOUCHARD M, GOUBRAN R A. Speech enhancement using a masking threshold constrained Kalman filter and its heuristic implementations[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2006, 14(1): 19 - 32.
- [11] MA NING, BOUCHARD M, GOUBRAN R A. A wavelet Kalman filter with perceptual masking for speech enhancement in colored noise[C]// IEEE International Conference on Acoustics, Speech, and Signal Processing. New York: IEEE, 2005: 149 - 152.
- [12] QUATIERI T F. Discrete-time speech signal processing: principles and practice[M]. New Jersey: Pearson Education, 2002: 176 - 177.
- [13] DONOHO D L. De-noising by soft-thresholding[J]. IEEE Transactions on Information Theory, 1995, 41(3): 613 - 627.

(上接第 3440 页)

5 结语

本研究实现了基于模糊综合评价的可穿戴心电信号质量评估算法,为确定心电信号质量提供了一个客观的定量指标,可有效反映信号的干扰水平,在严重干扰存在时仍能准确评价。进一步研究基于信号质量评估的可穿戴智能监护系统,可望有效改善失真带来的误判。

参考文献:

- [1] 王远干,喻洪麟,黄良明. 基于小波分析的光栅信号质量评估方法[J]. 中国激光, 2005, 32(3): 336 - 340.
- [2] 王继祥,刘贵忠,刘广健,等. 基于小波谱失真测度的语音质量评估[J]. 兵工学报, 2008, 29(1): 33 - 36.
- [3] WANG J Y. A new method for evaluating ECG signal quality for multi-lead arrhythmia analysis[J]. Computers in Cardiology, 2002, 29(11): 85 - 88.
- [4] ZIGEL Y, COHEN A, KATZ A. ECG signal compression using a-

- nalysys by synthesis coding[J]. IEEE Transactions on Biomedical Engineering, 2000, 47(10): 308 - 316.
- [5] 李桥,俞梦孙. 重症监护病人心电导联信号质量评估[J]. 山东大学学报, 2007, 45(9): 869 - 870.
- [6] 孟兆辉,张永红,白净. 模糊算法在心律失常病类判别中的应用[J]. 电子学报, 2001, 29(9): 1246 - 1248.
- [7] FRIESEN G M, JANNETT T C, JADALLAH M A, *et al.* A comparison of the noise sensitivity of nine QRS detection algorithms[J]. IEEE Transactions on Biomedical Engineering, 1990, 37(8): 85 - 98.
- [8] HE T, CLIFFORD G D, TARASSENKO L. Application of independent component analysis in removing artifacts from the melectrocardiogram[J]. Neural Computing and Applications, 2006, 15(6): 105 - 116.
- [9] LIFFORD G D, AZUAJE F, MCSHARRY P E. Advanced methods and tools for ECG data analysis[M]. Norwood: Artech House, 2006: 69 - 70.