

改进的小波阈值去噪在语音识别中的应用

李克粉,王 直

(江苏科技大学 计算机科学与工程学院,江苏 镇江 212003)

摘要:目前消除环境噪声提高语音信号质量是语音识别技术研究的重要方向之一。小波阈值去噪法是一种常见的得到广泛应用的语音增强方法,然而传统的小波阈值去噪法存在许多不足,如小波系数经过硬阈值函数处理后在正负阈值处不连续,重构后的信号会产生振荡;经过软阈值函数去噪后的语音过于平滑而丢失语音的突变部分影响语音信号的质量。文中将小波去噪应用于语音识别系统的前端处理中,提出一种新的阈值函数,对小波阈值去噪方法中传统阈值函数的缺点进行改进,将带噪语音信号去噪后再进行识别。实验结果表明,新提出的方法使带背景噪声的语音识别系统的识别性能得到显著的提高。

关键词:语音识别;噪声;小波变换;阈值函数

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2013)05-0231-04

doi:10.3969/j.issn.1673-629X.2013.05.060

Application of Improved Wavelet Threshold Denoising in Speech Recognition

LI Ke-fen, WANG Zhi

(College of Computer Science and Engineering, Jiangsu University of Science and Technology,
Zhenjiang 212003, China)

Abstract: At present, removal of environmental noise to improve the quality of voice signal has become an important direction for voice recognition technology. Wavelet threshold denoising method is a method widely applied to speech enhancement. However, the traditional wavelet threshold denoising method has many deficiencies, like wavelet coefficients processed by the hard threshold function is discontinuous in the positive and negative threshold values, resulting in the reconstructed signal oscillation; denoised voice processed by soft threshold function is too smooth and lose voice mutation part, which would affect the quality of the voice signal. It applied wavelet denoising to the front-end processing of speech recognition system, and proposed a new threshold function to improve the shortcomings of the traditional threshold function of the wavelet threshold denoising method. It denoised speech signal with noise, and then identified. The experimental results show that the new proposed method makes speech recognition performance of the system with background noise significantly improved.

Key words: speech recognition; noise; wavelet transform; threshold function

0 引言

目前,在纯净语音环境下,语音识别系统的识别率已经达到很高的水平,但在现实中,噪声无处不在,语音受到噪声干扰会导致待识别语音的特征矢量与训练模版不匹配,系统识别率降低。因此,抗噪问题成为语音识别迈向实用化的关键。

小波分析,又称为多分辨率分析,在时域、频域都有良好的局部化特性,特别适用于非平稳信号(如语

音信号)的研究。Witkin 首先提出利用小波分解中不同尺度信号的空间相关性去除外界噪声的思想。近年来,国内外很多学者已经开始注意小波阈值去噪方法, D. L. Donoho^[1]和 I. M. Johnstone 对此方法进行了较深入的理论研究^[2]。德克萨斯大学的 X. P. Zhang 和 M. D. Desai^[3]在此方法基础上提出了一种新的阈值函数,把原阈值函数由间断函数变为连续函数。在国内西北工业大学的潘泉、戴冠中等^[4]对该方法在阈值选择问题上提出了一种新方案,拓宽了此方法的应用前景。

1999年,上海交通大学的李冲泥、胡光锐^[5]把小波阈值去噪方法应用于语音去噪的过程中,但由于此方法所采用的硬、软阈值函数都存在缺陷和不足,小波

收稿日期:2012-08-16;修回日期:2012-11-20

基金项目:国家自然科学基金资助项目(51008143)

作者简介:李克粉(1986-),女,硕士研究生,主要研究方向为模式识别、语音信号处理;王 直,教授,硕士生导师,主要研究方向为模式识别、船舶控制。

系数用硬阈值函数处理后在正负阈值处不连续,重构后的信号会产生振荡。软阈值函数在正负阈值处连续,但它的导数不连续,而且估计小波系数与带噪小波系数之间存在恒定偏差,会影响语音信号的质量。因此,文中提出一种新的阈值函数,与原阈值函数相比,它连续且可导,弥补了原函数分段性和不连续性的缺点和不足,使语音识别率得到显著提高。

1 小波分析

小波就是在较短时间区间上有振荡的波,满足

$$\int_{-\infty}^{\infty} \psi(t) dt = 0.$$

由于重构信号的需要,小波还必须满足容许性条件:

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty$$

在应用中会将小波进行平移和伸缩,得到一族小波。函数 f 在 $L^2(R)$ 上的积分小波变换定义为:

$$W_f(a, b) = \int_{-\infty}^{\infty} \hat{\psi}_{a,b}(x) f(x) dx$$

积分小波反变换为:

$$f(x) = \frac{1}{C_{\psi}} \iint W_f(a, b) \hat{\psi}_{a,b}(x) \frac{da db}{a^2}$$

小波的分解可以看成是信号通过一组滤波器后得到的结果,通过低通滤波器得到其轮廓信号,通过高通滤波器得到其细节信号,小波的分解结果框图^[6]如图 1 所示。

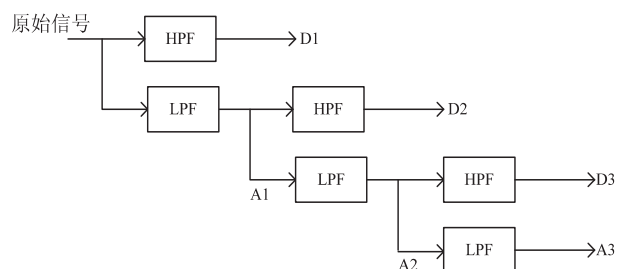


图 1 小波的分解结果图

2 基于清浊音分离的小波阈值去噪原理

设带噪语音信号模型表示为:

$$f(k) = s(k) + n(k) \quad k = 1, 2, \dots, N$$

其中, $f(k)$ 为带噪语音信号, $s(k)$ 为原始纯净语音信号, $n(k)$ 为方差为 σ^2 的高斯白噪声,服从 $N(0, \sigma^2)$ 。 N 为采样序列的长度。将信号进行 J 层小波分解, $d(j, k)$ ($1 \leq j \leq J, 1 \leq k \leq N$) 为分解得到的各尺度上的小波系数。

考虑到语音信号本身的特殊性,使用阈值对语音信号进行处理时,不要影响语音的清音部分,因为清音段包含了许多类似噪声的高频有用成分,若误除会降低语音识别率。因此,首先需要从带噪信号中把清音

段识别并分离出来,然后对不同部分采用不同的阈值进行处理。区分清音与浊音的方法:首先,对语音信号分帧并进行离散小波变换,将小波系数分为 $M+1$ 个子带(M 为小波分解层数),计算出每个子带的平均能量,如果最高频率带的能量比其他频率带的能量大,而且最低频率带的能量和最高频率带的能量比小于某个阈值(如 0.9),则将这段语音信号视为清音,否则视为浊音。如果为清音,则只对最高频率带的小波系数用阈值进行处理,如果为浊音,对所有的小波系数都要用阈值进行处理。

利用小波进行语音去噪^[7]的具体方法:首先选择合适的小波基和小波分解层数 J ,将带噪语音信号分帧,然后对每帧信号加窗,做离散小波变换,计算出各个小波系数 $d(j, k)$;再将小波系数分成若干个子带,利用上述方法对清浊音进行判断。然后分别用文中提出的阈值去噪方法和传统的阈值去噪方法对清音段和浊音段进行不同的阈值处理。若为清音,则只对最高频率带的小波系数进行阈值处理;若为浊音,则需要对所有小波系数进行处理,处理后得到小波系数估计值 $D(j, k)$,最后将处理过的小波系数进行离散反变换、重构,得到去噪后的原始语音信号。其基本流程如图 2 所示。

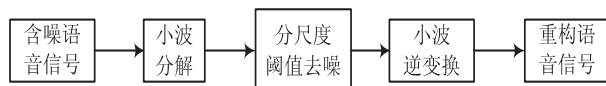


图 2 小波去噪流程图

2.1 选取阈值函数

对小波系数进行估计必须对阈值函数和阈值^[8]进行选取,目前常用的阈值函数的选取方式主要有硬阈值和软阈值两种。

硬阈值函数:

$$D(j, k) = \begin{cases} d(j, k), & |d(j, k)| \geq \lambda \\ 0, & |d(j, k)| < \lambda \end{cases}$$

软阈值函数:

$$D(j, k) = \text{sgn}(d(j, k)) \max(0, |d(j, k)| - \lambda)$$

其中, $d(j, k)$ 为带噪语音信号分解后得到的小波系数, $D(j, k)$ 为经阈值函数处理后的小波系数, λ 为阈值。

虽然通过硬阈值函数或软阈值函数消噪后的效果不错,但仍存在一些不足:

(1) 小波系数经过硬阈值函数处理后在正负阈值处不连续,它可以很好地保留信号边缘信息,但是重构后的信号会产生振荡。

(2) 软阈值函数,在正负阈值处连续,但它的导数是不连续的,小波系数经软阈值函数处理后,能量会有

一部分损失。并且,估计小波系数与带噪小波系数之间存在恒定偏差,去噪后的语音过于平滑而丢失语音的突变部分,从而影响语音信号的质量。

由此,文中提出了介于软、硬阈值函数之间的一种改进的阈值函数^[9],如下:

$$D(j,k) = \begin{cases} \operatorname{sgn}(d(j,k)) \left\{ \frac{|d(j,k)| - a\lambda_j}{e^{b(\frac{|d(j,k)| - \lambda_j}{\lambda_j})}} \right\} & |d(j,k)| \geq \lambda_j \\ 0 & |d(j,k)| < \lambda_j \end{cases}$$

式中, λ_j 为阈值,参数 a, b 为调节因子。其中 $a \in [0, 1], b > 0$ 。当 $a = 0$ 时,此函数成为硬阈值函数。当 $a = 1, b = 0$ 时,此函数便成为软阈值函数。恰当地选择参数可以达到阈值去噪的最佳效果。文中选用 $a = 1, b = 1$ 。此新函数在小波域内既具有连续性,又具有高阶可导性。它具有连续性,不会导致重构后的信号产生振荡,保证了有用信号能量不丢失,有效减少了估计小波系数与带噪小波系数之间存在的恒定偏差;它具有高阶可导性,能进行各种数学处理。该阈值函数在保存了两种传统阈值函数优点的同时弥补了它们在处理小波系数时的一些缺点和不足,使用起来更加灵活方便,去噪效果更好。另外,由于噪声小波变换模值上界的衰减是符合指数规律的,文中提出的函数正好符合噪声的衰减规律。

2.2 阈值选取方法

如何确定阈值 λ 是小波变换去噪的关键^[9], λ 值可以用下式来确定: $\lambda = \sqrt{2\log N}\sigma$,其中 N 为带噪数据的序列长度, σ^2 为噪声的估计方差,高斯白噪声的标准差可用下式来计算:

$$\sigma = \sqrt{\frac{\operatorname{median}(|d(j,k)|)}{0.6745}}$$

文中采用的阈值公式^[10]:

$$\lambda_j = \sigma \sqrt{2\ln N} / \ln(e + 2^{j-1} - 1)$$

当 $j = 1$ 时,即为原阈值函数。当 $j > 1$ 时,随着尺度 j 的增加,阈值随着减小,正好与噪声的小波变换在不同尺度上的衰减规律相一致。表1即为改进的阈值随 j 值变化的情况($\sigma = 1, N = 100$)。

表1 阈值随 j 值变化情况

j 的值	改进前的阈值	改进后的阈值
1	2	2
2	2	1.289
4	2	0.844
8	2	0.410
16	2	0.192
32	2	0.093

3 实验结果及分析

文中语音采集通过 Cooledit 声音处理软件完成,采集过程中,以0~9这10个语音数字作为实验对象,通过对这10个数字进行多遍录音,生成30组(共600个样本),分别来自20个男生,10个女生,用11025Hz的采样率,单声道录音,量化为16bits,提取每帧信号的24阶MFCC作为语音特征参数提取,噪声为高斯白噪声。在小波变换中,文中选用 Daubiches 小波阶数为8阶,小波分解层次为4层,语音分帧时帧长采用128样值点。

一般采用信噪比^[11](SNR)作为评定去噪结果的客观评价标准。其定义形式如下:

$$p_s = \sum ((x - \operatorname{mean}(x))^2)$$
$$p_n = \sum ((x - x_d)^2)$$
$$\operatorname{SNR} = 10\log_{10}(p_s/p_n)$$

x 代表原始纯净语音信号, x_d 代表含噪语音信号。

测试所用的含噪语音信号是由纯净语音信号叠加不同强度的高斯白噪声得到,图3是数字“4”的发音及在经过各不同阈值方法去噪后得到的语音波形(信噪比为5dB)。

三种阈值去噪方法的去噪性能比较与三种去噪方法去噪后的语音识别率比较分别如表2,表3所示。

表2 三种阈值去噪方法的去噪性能比较

输入 SNR (dB)	硬阈值	软阈值	文中方法
0	5.0522	6.092	8.407
5	12.1297	12.0846	16.529
10	18.385	19.173	22.748
15	22.829	24.839	27.275
20	29.375	28.935	30.0747

表3 三种去噪方法去噪后的语音识别率比较

输入 SNR (dB)	未去噪语音识别率	去噪后语音识别率		
		硬阈值	软阈值	文中方法
0	18.937	22.763	23.936	29.197
5	24.674	26.878	26.819	34.359
10	31.689	37.943	40.865	45.683
15	48.239	52.578	53.693	68.204
20	61.378	70.784	70.583	78.183
25	70.679	81.268	83.573	92.372
30	86.538	92.268	93.965	97.267

表2,表3比较了传统的小波阈值及文中所述方法的去噪性能及对语音识别率的影响。仿真结果表明文中提出的阈值函数对于提高信噪比及语音识别率有更好的效果,从表中数据可以看出该方法的有效性,去噪后语音更清晰,语音识别率也得到很大的提高。

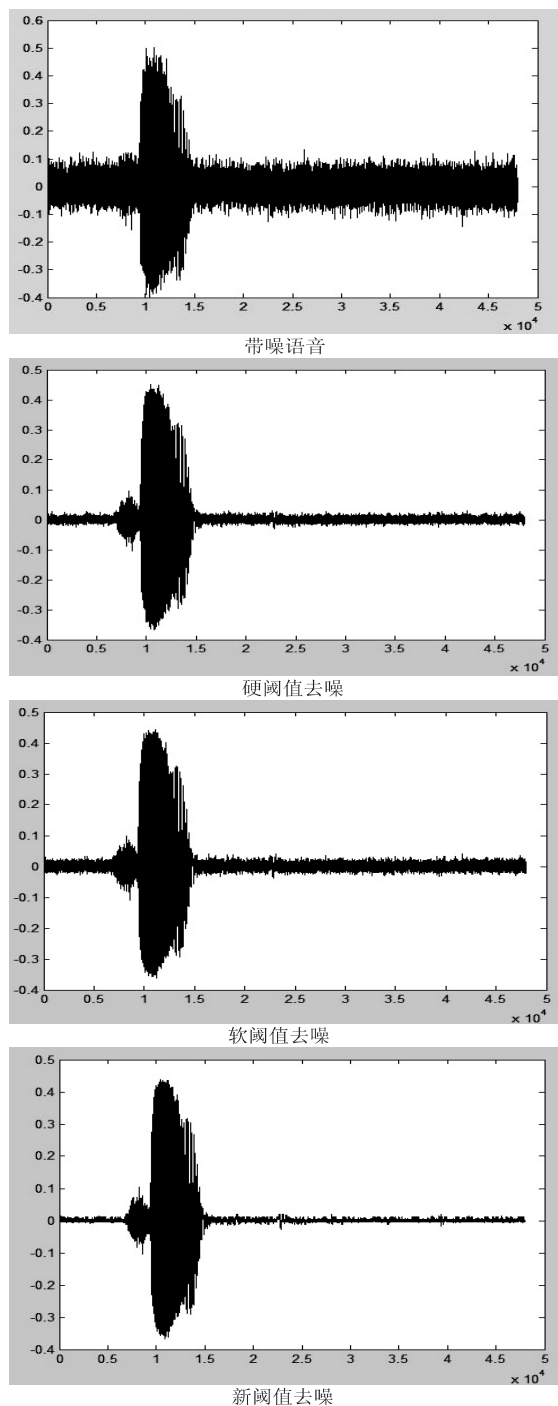


图 3 各种不同方法处理后的“4”语音波形比较

4 结束语

小波阈值去噪法是语音信号处理中一种重要的语音增强技术,可以对语音的背景噪声进行有效地去除。传统的软、硬阈值去噪法由于阈值选取和阈值函数本身存在的问题而使得去噪效果存在很多的缺陷和不足。

文中对传统阈值去噪的不足进行改进,提出一种新的阈值函数。仿真结果表明,文中提出的新的阈值去噪方法比起传统方法能够更好地消除噪声,语音识别率也得到显著提高。

参考文献:

- [1] Donoho D L. De-noising by soft-thresholding[J]. IEEE Trans. on Inform Theory, 1995, 41(3): 612-627.
- [2] Donoho D L, Johnstone I M. Adapting to Unknown Smoothness Via Wavelet Shrinkage[J]. Journal of American Stat. Assoc., 1995, 90(432): 1200-1224.
- [3] Zhang Xiaoping, Desai M D. Adaptive Denoising Based on SURE Risk [J]. IEEE Signal Processing Letters, 1998, 5(10): 265-267.
- [4] 潘 泉, 戴冠中, 张洪才, 等. 基于阈值决策的子波阈去噪方法[J]. 电子学报, 1998, 26(1): 115-117.
- [5] 李冲泥, 胡光锐. 一种改进的子波阈语音增强方法[J]. 通信学报, 1999, 20(4): 88-91.
- [6] 李海东, 李 青. 基于阈值法的小波去噪算法研究[J]. 计算机技术与发展, 2009, 19(7): 56-58.
- [7] 李如玮, 鲍长春, 窦慧晶. 基于小波变换的语音增强算法综述[J]. 数据采集与处理, 2009, 24(3): 362-368.
- [8] 林 劼. 抗噪连续语音识别研究[D]. 成都: 电子科技大学, 2006.
- [9] 陈小曦, 王延杰, 刘 恋. 小波阈值去噪法的深入研究[J]. 激光与红外, 2012, 42(1): 3-5.
- [10] Huang Zhenghong, Fang Bin, He Xiping, et al. Image denoising based on the dyadic wavelet transform and improved threshold [J]. Multiresolution and Information Processing, 2009, 7(3): 4-7.
- [11] 张翠芳. 小波阈值降噪效果影响因素的研究[J]. 西安邮电学院学报, 2008, 13(5): 13-15.

(上接第 230 页)

- 的设计实现[J]. 计算机测量与控制, 2008, 16(3): 327-329.
- [8] 何益波, 金 瓯, 贺建彪. 基于 GSM 短消息的自动售货机监控终端开发[J]. 计算机测量与控制, 2009, 17(4): 679-681.
 - [9] 陈小刚, 吴志国, 朱成健, 等. 自动售货远程监控系统关键技术的研究与设计[J]. 计算机测量与控制, 2010, 18(6): 1301-1302.
 - [10] 张捍东, 朱 林. 物联网中的 RFID 技术及物联网的构建

- [J]. 计算机技术与发展, 2011, 21(5): 56-59.
- [11] 闫 涛, 吕丽民. 物联网技术在企业安全生产中的应用[J]. 计算机技术与发展, 2012, 22(2): 226-228.
 - [12] Kim D K, Tok C K. Mobile multicasting in wireless ATM networks[J]. Mobile Networks and Applications, 2000, 5(2): 55-64.
 - [13] Kalden R, Meirick I, Meyer M. Wireless Internet Access Based on GPRS[J]. IEEE Personal Communications, 2000, 7(2): 8-18.

改进的小波阈值去噪在语音识别中的应用

作者: [李克粉, 王直](#)
作者单位: [江苏科技大学 计算机科学与工程学院, 江苏 镇江 212003](#)
刊名: [计算机技术与发展](#)
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2013(5)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjfz201305062.aspx