

基于 EEMD 数据预处理和 DNN 的 语音增强算法研究

陈建明 梁志成

(陆军装甲兵学院 信息通信系, 北京 100072)

摘要: 提出了基于总体平均经验模态分解(EEMD)预处理和深度神经网络的语音增强算法,首先将带噪语音信号和纯净语音信号进行EEMD分解,获得一组频率从高到低的本征模态函数IMF分量,然后从各IMF中提取时域的信号特征,组成特征向量,输入神经网络中进行训练。实验表明:该算法与经典无监督算法相比,无需任何假设条件,可以较好地学习带噪语音和纯净语音之间复杂的非线性关系,在语音质量和可懂度方面优势明显,显示了深度神经网络在语音增强方面的独特作用。

关键词: 语音增强; EEMD分解; 语音信号特征提取; 深度神经网络; 语音质量; 可懂度

本文引用格式: 陈建明, 梁志成. 基于 EEMD 数据预处理和 DNN 的语音增强算法研究[J]. 兵器装备工程学报, 2019, 40(6): 96-103.

Citation format: CHEN Jianming, LIANG Zhicheng. Research on Speech Enhancement Algorithm Based on EEMD Data Preprocessing and DNN[J]. Journal of Ordnance Equipment Engineering 2019, 40(6): 96-103.

中图分类号: TN912.35

文献标识码: A

文章编号: 2096-2304(2019)06-0096-08

Research on Speech Enhancement Algorithm Based on EEMD Data Preprocessing and DNN

CHEN Jianming, LIANG Zhicheng

(Department of Information and Communication,
Academy of Armored Force for Land Army, Beijing 100072, China)

Abstract: A speech enhancement algorithm based on EEMD(Ensemble Empirical Mode Decomposition) preprocessing and deep neural network was proposed. The noisy speech signal and the pure speech signal were decomposed by EEMD to get a set of IMF (Intrinsic Mode Function) components from high to low frequency. Then, the signal features in the time domain are extracted from each IMF, and the feature vectors were constructed, they are input into the neural network for training. Experiments show that the algorithm does not require any assumptions when compared with the classical unsupervised algorithm. It can better learn the complex nonlinear relationship between noisy speech and pure speech. It has obvious advantages in speech quality and intelligibility, showing the unique role of deep neural network in speech enhancement.

Key words: speech enhancement; EEMD decomposition; speech signal feature; deep neural network; speech quality; intelligibility

语音增强是语音信号处理领域的一个重要分支,主要目标是从带噪声的语音信号中提取尽可能纯净的原始语音信

号,提高语音信号的质量、清晰度和可懂度。语音增强算法主要分为两大类:有监督算法和无监督算法。比较典型的无

收稿日期: 2018-11-03; 修回日期: 2018-12-20

作者简介: 陈建明(1964—),男,博士,教授,主要从事信号与信息处理研究。

监督算法有:谱减法、自适应滤波法、小波变换法、EEMD (Ensemble Empirical Mode Decomposition) 法等^[1],基本原理是通过先估计噪声的谱信息,然后从带噪语音谱中将估计的噪声谱减去以获得干净的语音谱。其优点是运算量比较小,可实现实时语音增强,便于实际应用。缺点是各个无监督算法都会不同程度的产生“音乐噪声”^[2]。其中谱减法中的“音乐噪声”最明显;自适应滤波法主要是利用滤波器估计发音器官的声道参数和激励源的参数,但在信噪比较低的情况下,难以估计出正确的参数,会产生严重的语音失真;小波法可以灵活选择小波基,由于语音信号的时变性,如何有效地选择小波基是一大难点^[3]。EEMD 法是由于剔除不同频率分量造成高频语音损失,语音失真的问题在低信噪比下更加严重。再者,传统无监督算法都基于平稳噪声的假设,对于非平稳噪声,尤其是极端非平稳噪声(比如机关枪射击声)难以起到有效的增强作用。最后,传统的无监督算法中,存在一些为了计算而采取的假设,例如噪声是加性的、噪声和语音互相独立,这些假设和近似,对传统语音增强算法性能的进一步提高形成制约^[4]。就目前有关语音增强算法的研究来说,无监督算法的研究已趋于成熟,由于算法原理本身的局限性,无法满足非线性、非平稳语音信号的去噪要求,也不可能通用的去噪算法。有监督算法主要分为:浅层模型和深层模型。浅层模型有浅层神经网络、隐马尔科夫模型 HMM (Hidden Markov Model) 和非负矩阵分解 NMF (Nonnegative Matrix Factorization) 等。基于 HMM 和 NMF 的语音增强算法,都是基于噪声和语音之间是独立的假设,限制了语音增强性能的上限;对于浅层神经网络来说,网络层数少,隐含层节点数少,无法精确地学习到带噪语音和纯净语音间复杂的非线性关系,而且浅层神经网络缺乏有效的初始化方法,随机初始化造成每次训练结果不稳定,容易陷入局部最优解^[5]。深层模型起于 2006 年 Hinton 等人提出深度信念网络^[6] (DBN),使用受限玻尔兹曼机 (RBM) 逐层训练神经网络的初始化权值^[7],使得深度神经网络 (DNN) 开始受到国内外研究者的广泛关注,并在图像、语音、视频和自然语言处理等领域取得巨大进步。DNN 应用于语音增强是近几年研究的热点,通过网络学习带噪语音和纯净语音之间复杂的非线性映射关系,并不断调整网络参数以找到最优的函数关系,从而去除噪声。文献[7]提出了一种基于理想二元时频掩蔽估计的语音增强方法,该方法把语音增强问题转化成用 DNN 估计理想二元时频掩蔽估计的分类问题,该方法对于低信噪比非平稳语音增强可得到高可懂度的增强语音,语音音质损失严重。文献[8]提出了一种基于 DNN 的最小均方误差回归拟合语音增强方案,该方法基于对数功率谱最小均方误差准则,通过 DNN 对带噪语音和干净语音间的复杂关系进行回归拟合建模。文献[9]提出了采用一种堆叠式去噪自编码器来进行语音增强的方法。目前,监督性语音分离的框架基本成熟,即利用深度模型学习一个从带噪特征到分离目标的映射函数,很难在框架层面进行重大的改进。DNN 用于语音增强处理研究主要集中在如何进一步提升对不包含在

训练集噪声环境下的语音增强性能,即提升系统的泛化能力;研究语音增强 DNN 模型在噪声环境的自适应问题和进一步将深度学习方法应用到多通道语音增强中等^[5]。本文针对典型装甲车辆运动噪声背景下语音传输质量差问题,采用基于 EEMD 预处理和 DNN 结合的方法,研究 DNN 模型在该噪声环境下的自适应性,达到提高语音质量和可懂度的目的。

1 算法整体框图

本算法分三部分:数据预处理阶段、模型训练阶段和增强处理阶段。算法整体框图如图 1。

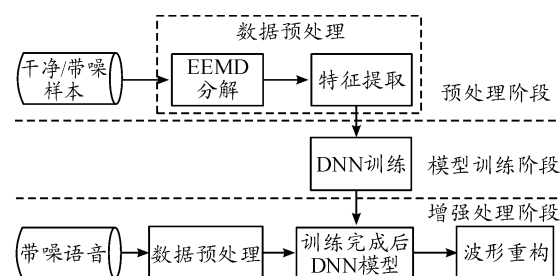


图1 算法整体框图

① 数据预处理阶段。包括两个部分: EEMD 分解和特征提取。首先将带噪语音信号和纯净语音信号进行 EEMD 分解,获得一组频率从高到低的本征模态函数 IMF (Intrinsic Mode Function) 分量,原一维时域信号分解成多维时域信号,更便于 DNN 学习带噪语音与纯净语音间的关系。提取时域的信号特征,组成特征向量,输入神经网络中进行训练,网络的训练目标为纯净语音信号的特征。

② 模型训练阶段。经过预处理后,带噪语音和纯净语音分别被分为相对应的特征样本,在训练阶段,将特征样本和对应的训练目标按批次输入到 DNN 中进行训练,经过信号的正向传播和误差的反向传播,得到最优的网络模型并进行保存。DNN 训练模型如图 2 所示。

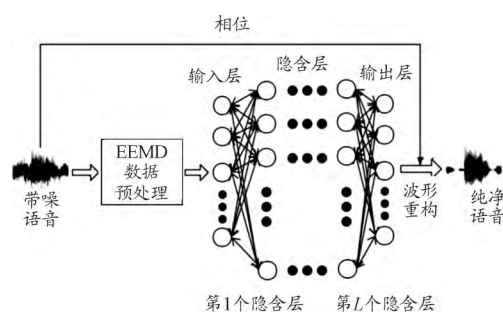


图2 DNN 训练模型

③ 增强阶段。增强阶段是对真实带噪语音进行处理。首先将待测语音按照①中步骤完成预处理,得到特征向量。然后输入到②中获得的最优网络模型中进行增强处理,最后

将输出的特征向量结合带噪语音的相位一起合成可主观测听的语音波形文件。

2 数据预处理

2.1 EEMD 分解

EEMD 分解^[10]的实质是对信号进行平稳化处理,不借助任何基函数,只需根据信号自身的时间尺度特征进行分解,使整个“筛分”过程具有直接性和自适应性,突出了信号的局部特征,能更直观地观察分析信号,对于非线性、非平稳信号的处理有很大帮助^[11]。

选取含有 -5 dB 装甲车运动噪声的带噪语音“四百米障碍” $y(t)$ 进行 EEMD 分解,波形图如图 3 所示,分解后各 IMF 分量如图 4 所示。从图 4 可以直观的看到每个 IMF 分量的

时域波形,不同级数的 IMF,其包含的信号频率、能量等均不相同。各 IMF 占原信号的能量比如表 1 所示,由表 1 可知信号的能量主要集中在级数较低的 IMF 中,前 7 个 IMF 大约占总能量的 85%,其余的只占总能量的 15%。能量较低的低频虚假分量对语音分析没有意义,因此在 EEMD 分解后,本文只保留信号的前 7 个分量进行分析。

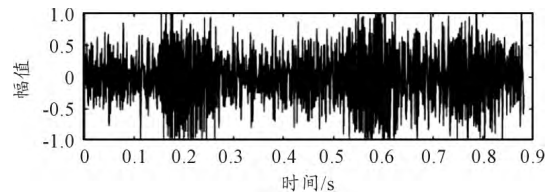


图 3 带噪语音“四百米障碍”波形

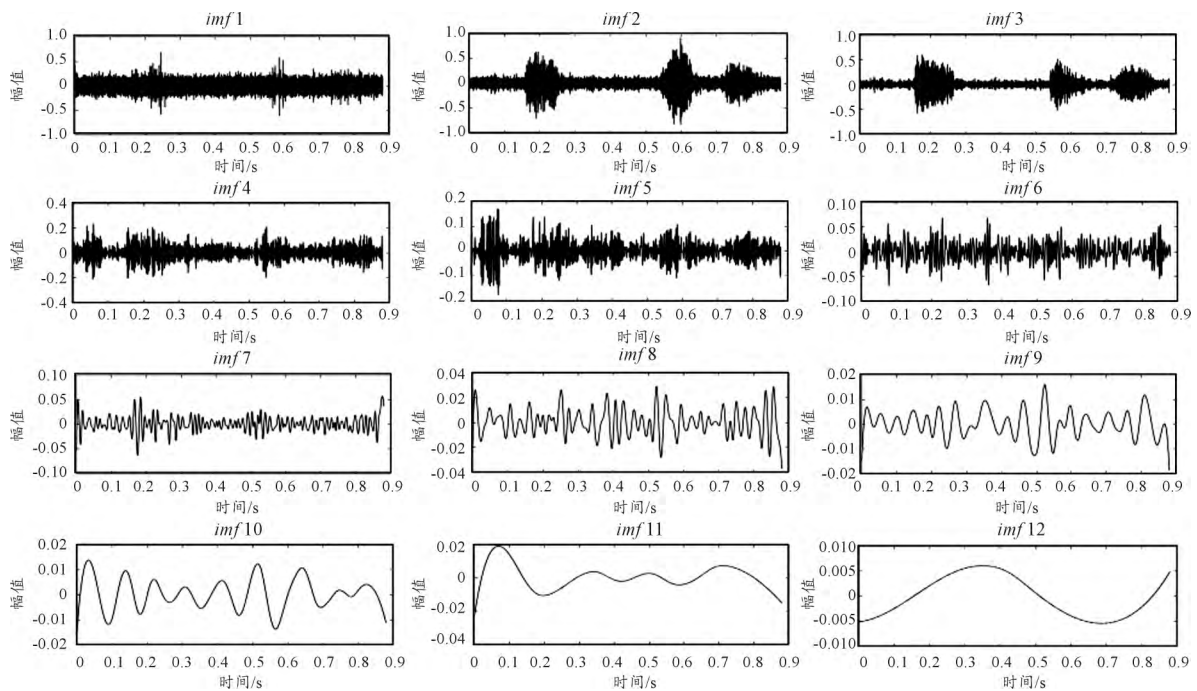


图 4 EEMD 分解后 IMF(1-12) 的波形

表 1 各 IMF 占原信号的能量比

IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11	IMF12
0.09	0.22	0.24	0.10	0.09	0.07	0.05	0.03	0.02	0.02	0.01	0.06

2.2 特征提取

步骤如下:

① 经 EEMD 分解后获得 IMF 分量。语音信号经 EEMD 分解后,能初步区分噪声分量和语音分量,这对于 DNN 来说是非常重要的信息,可以排除许多干扰因素。

② 选择有效的 IMF 分量,并对各 IMF 分量分帧。

③ 短时傅立叶变换。对各 IMF 每一帧信号进行 FFT 变

换,从时域数据变为频域数据。

④ 计算各 IMF 的谱线能量。

⑤ 计算对数功率谱。对各频点能量取对数。

⑥ 组成子带特征向量。每一帧 N 个频点的短时能量组成一个子带的特征向量: $P_i = [E_i(1) E_i(2) \cdots E_i(N)]$

对所有特征数据进行高斯归一化处理,即所有训练数据的均值归一化为 0,方差规整为 1。

3 DNN 训练

深度神经网络的训练过程由信号的正向传播与误差的反向传播两个部分组成。模型框架如图 2 所示。信号正向传播与误差反向传播的过程,是循环往复且权值不断修改的过程,也就是网络的训练过程,这个过程一直进行到输出层的输出误差逐步减小到可以接受的范围或者达到预先设定的学习次数为止。

3.1 信号正向传播

正向传播时,样本特征作用于输入层,经隐含层逐层处理之后,传到输出层,具体过程分以下。

DNN 训练之前首先要进行初始化。据早期研究^[6]可知,DNN 的网络初始化对于找到全局最优的结果非常重要,DNN 大部分利用受限玻尔兹曼机 RBM 组成的 DBN,通过对比散度算法进行无监督的预训练来对网络进行初始化。最近的研究显示,数据量的增加和 Dropout 技术^[12]的使用,同样能使网络获得较好的性能,因此本文采用随机初始化并利用 Dropout 来获得全局最优解。

本算法采用经验值 $[-\sqrt{2}, \sqrt{2}]$ 进行随机初始化,其中 n_{inputs} 为每一层输入节点数, n_{outputs} 为每一层输出节点数。

第 l 个隐含层输出为:

$$G_j^l = f\left(\sum_{i=1}^J w_{ij}^{l-1} H_i^{l-1}\right), j = 1, 2, \dots, J, l = 1, 2, \dots, L-1 \quad (1)$$

式(1)中 w_{ij}^{l-1} 为第 $l-1$ 层第 i 个节点与第 l 层第 j 个节点之间的权重; H_i^{l-1} 为第 $l-1$ 层第 i 个节点的输出值; J 为各层节点数;当 $l=1$ 时 w_{ij}^{l-1} 为输入层与第一隐含层权重 $H_i^{l-1} = P_i$, J 为输入层节点数 n ; f 为隐含层的激活函数,隐含层的激活函数选择 ReLU,即 $f(x) = \max(0, x)$ 。

输出层输出为:

$$O_i = f\left(\sum_{j=1}^J w_{ji}^L G_j^L\right), i = 1, 2, \dots, n \quad (2)$$

$$O = [O_1, O_2, \dots, O_n] \quad (3)$$

式(2)中 J 为第 l 层隐含节点数; f 为输出层激活函数,输出层的激活函数选择 sigmoid 函数,即 $f(x) = \frac{1}{1 + e^{-x}}$ 。

如果输出层未能够得到期望的输出,则转入到误差的反向传播阶段,将输出误差按照某种方式,通过隐含层向输入层返回,并把它“分摊”给各层的所有单元,从而获得各层单元的参考误差,用来修改各单元权值的依据。采用纯净语音和增强后语音对应特征之间的最小均方误差作为损失函数,反向传播算法允许来自损失函数的信息通过网络向后流动,被用于计算梯度,利用小批量随机梯度下降法(Mini-batch gradient descent algorithm)进行学习,从而更新整个 DNN 的参数。

将网络的实际输出 O 与期望输出 T 带入损失函数,计算网络误差 e 。由于语音增强是一个回归问题,因此选用均方误差 MSE(Mean-Square Error) 作为损失函数^[1]。为了避免模型出现过拟合问题,在损失函数中加入了 $L2$ 正则化,限制模型的复杂程度。

$$e = \text{MSE}(T, O) = \frac{\sum_1^n (T - O)^2}{n} \quad (4)$$

$$R(W) = \|W\|_2^2 = \sum |w^2| \quad (5)$$

$$E(W) = e + \frac{\lambda}{2} R(W), \lambda = 0.5 \quad (6)$$

式(4)~(6)中, N 为样本数; T 为期望输出; $R(W)$ 为体现模型复杂度的函数; W 为网络中所有权重参数; $E(W)$ 为网络优化目标函数; λ 为模型复杂度损失在总损失中的比例。

3.2 权重更新

首先调整第 l 隐含层与输出层之间的权重,沿着梯度最快下降的方向调整。然后误差向前传播,对前面各层权重进行调整。

第 l 个隐含层权重调整为:

$$w_{ij}^{l-1} = (1 - \eta\lambda) w_{ij}^{l-1} + \eta \frac{\partial e}{\partial G_j^l} G_j^l, j = 1, 2, \dots, J \quad (7)$$

式(7)中 η 为权重更新的学习速率(Learning Rate),来定义每次参数更新的幅度(如果 η 太大,权重更新跳转太快,容易训练坏掉,如果 η 太小,收敛速度太慢,增加运算时间); G_j^l 为第 l 层第 j 个节点的输出值,当 $l=L+1$ 时 $G_j^l = O_j, J=n$ 。

至此,一轮权重调整结束。调整规则归纳为:权重调整量 = 学习率 \times 局部梯度 \times 上一层输出信号。

4 实验及数据分析

为了检验本算法的有效性,通过 DNN 模型训练后与语音增强经典算法中的谱减法(算法1)、自适应滤波法(算法2)以及神经网络中的基于短时 FFT 变换的对数功率谱特征提取方法(算法3)进行比较。选择一段语音“四百米障碍”进行分析对比研究,纯净语音波形及语谱图如图 5 所示,噪声为装甲车运动噪声如图 6 所示。

4.1 实验参数设置

特征位数:896 维(7×128),即 IMF 分量有效个数为 7,每帧为 128 个点。因而输入输出节点数分别为 896。

特殊函数:隐含层激活函数选择 ReLU,输出层激活函数选择 Sigmoid,损失函数选择 MSE,采用随机初始化,训练算法采用小批量随机梯度下降法,采用 Dropout 对模型进行优化,损失比例设置为 0.2,学习率设置为 0.01,并按照 0.004 的差值逐渐减小。

数据源:本文实验纯净语音全部来自于清华大学语音与语言技术中心(CSLT)出版的开放式中文语音数据库,抽取 500 条纯净语音被用来与装甲车运动噪声相加在一起,相加的信噪比为 15 dB、10 dB、5 dB、0 dB 和 -5 dB,来构建一个多种情况的平行训练数据。另外,从语音数据库里抽取另外

100 条语音与上述装甲车运动噪声相加构成测试数据集。

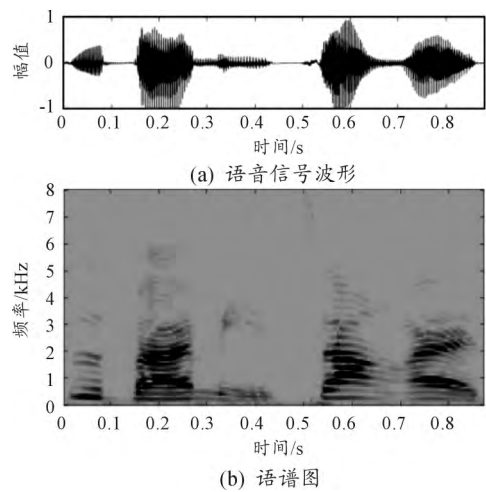


图5 纯净语音

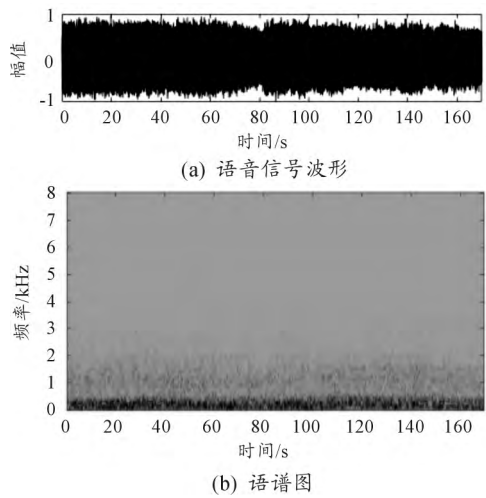


图6 装甲车运动噪声

4.2 DNN 隐含层及节点数的确定

一般来说,深度神经网络要求隐含层大于3层、每层节点数大于1 000。首先假定隐含层数为4,实验结果如表2所示,每个隐含层的节点数为1 024时效果较好,隐含层数为5时亦如此,确定隐含层节点数为1 024。然后节点数为1 024时实验数据如表3所示,确定隐含层为4层。

表2 隐含层不同节点数的 DNN 增强后 PESQ

SNR/dB	节点数			
	512	800	1 024	1 600
-10	1.493 1	1.557 4	1.624 3	1.419 4
-5	1.875 4	1.893 0	1.773 6	1.835 5
0	2.343 5	2.395 1	2.436 7	2.259 3
5	2.673 2	2.704 1	2.736 3	2.581 7

表3 不同隐含层数的 DNN 增强后 PESQ

SNR/dB	层数			
	2	3	4	5
-10	0.708 0	1.334 4	1.419 4	1.453 2
-5	1.108 0	1.750 5	1.835 5	1.863 3
0	1.653 6	2.174 3	2.259 3	2.276 0
5	2.163 0	2.517 7	2.581 7	2.597 1

4.3 实验数据及比较

4.3.1 模拟带噪语音下各算法比较结果及分析

选取纯净语音“四百米障碍”和装甲车运动噪声在-5 dB、0 dB和5 dB下混合构成带噪语音信号,检验指标为信噪比 SNR、感知语音质量评估 PESQ 和短时客观可懂度 STOI,数据如表4、表5、表6所示。-5 dB(限于篇幅只以-5 dB为例)带噪语音(如图7所示)经各算法增强后得到的波形与语谱图如图8、图9、图10和图11所示。

表4 不同算法在-5 dB 装甲运动噪增强后的信噪比

指标	算法1	算法2	算法3	本文算法
SNR	0.700 0	3.780 0	8.571 5	8.978 1
PESQ	0.822 2	1.435 4	2.132 1	2.360 2
STOI	0.553 8	0.886 3	0.833 0	0.876 7

表5 不同算法在0 dB 装甲运动噪增强后的信噪比

指标	算法1	算法2	算法3	本文算法
SNR	2.659 9	3.608 8	10.224 4	11.299 4
PESQ	1.339 1	1.553 8	2.403 3	2.578 9
STOI	0.583 5	0.891 1	0.843 1	0.890 5

表6 不同算法在5 dB 装甲运动噪增强后的信噪比

指标	算法1	算法2	算法3	本文算法
SNR	3.365 2	3.569 4	13.509 1	13.937 3
PESQ	1.725 0	1.777 3	2.660 1	2.857 2
STOI	0.614 2	0.887 9	0.908 0	0.921 8

由表4、表5和表6可看出:算法3和本文算法在语音增强处理后信噪比均提高2倍以上,PESQ和STOI与经典算法比均显著提高,本文算法优势更明显。

比较各算法增强后的语音与纯净语音的波形图可知,在-5 dB噪声环境下,谱减法去除的噪声最多,语音波形幅度较纯净语音有所减小,甚至有的语音波形已经消失不见,且在语音间隙期,新产生了几小的波形凸起;自适应滤波法去除了大部分噪声,但还有小部分残留噪声,语音波形的幅度没有明显减小;算法3和本文算法很好地滤除了背景噪声,波形更接近于原始纯净语音,去噪效果都强于经典算法。

本文算法较算法 3 的 PESQ 和 STOI 值分别提高 0.228 1 和 0.013 8,说明本文算法在可懂度和语音质量方面均有所提高,这是因为信号经 EEMD 分解后,首先能初步区分噪声分量和语音分量,更能反映信号的时频特性。

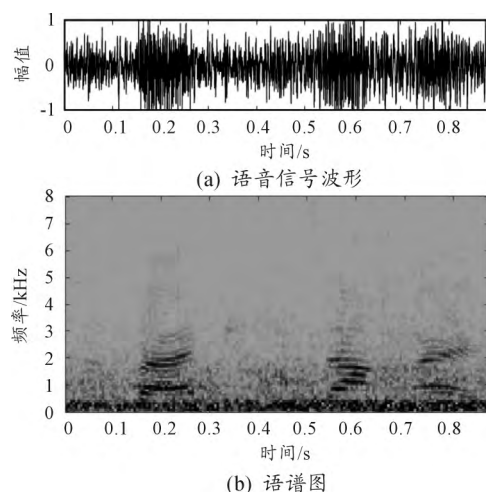


图 7 -5 dB 合成带噪声语音

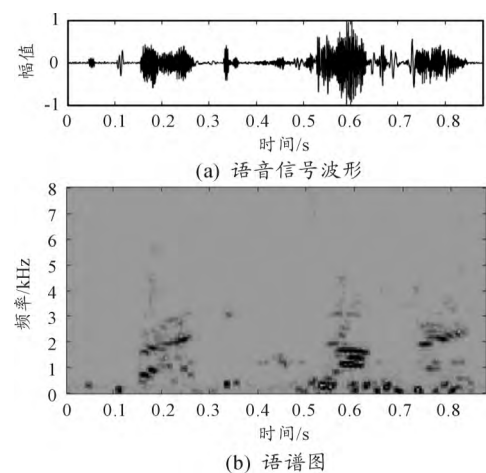


图 8 算法 1 结果

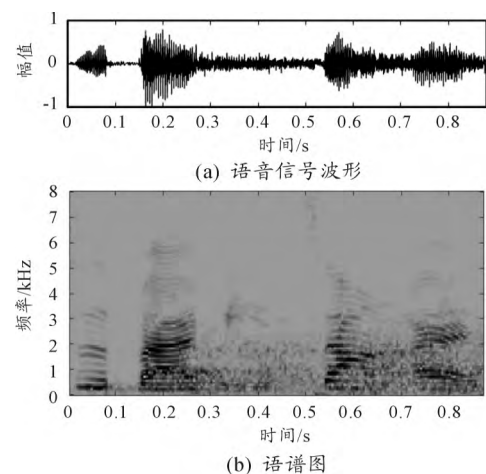


图 9 算法 2 结果

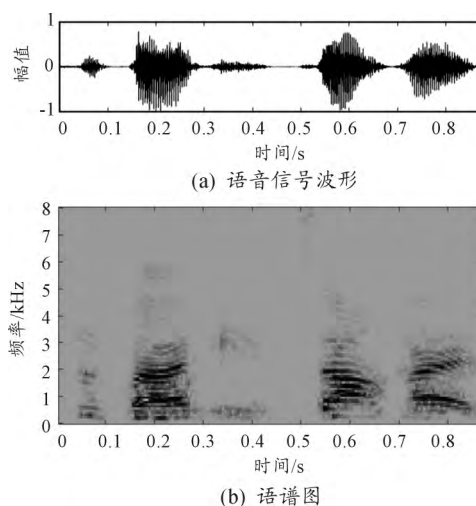


图 10 算法 3 结果

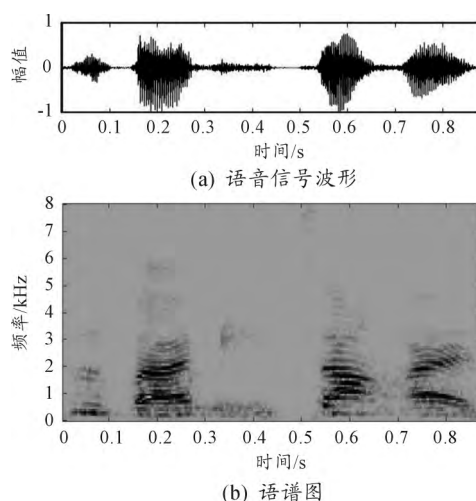


图 11 本文算法结果

比较各算法增强后的语谱图,谱减法处理后不仅噪点消失不见,语音的谱线也几乎被消除,没有明显的语音变化趋势。自适应滤波法处理后的语谱图较清晰,高频噪声分量几乎完全去除,语音趋势清晰可见;算法 3 和本文算法处理后语谱图清晰,与原始纯净语音的语谱图一致,语音趋势清晰,本文算法较算法 3 比,能保留有效的低频信息。

4.3.2 真实带噪声语音下各算法比较结果及分析

对实际采集的带噪声语音“刀光枪影”进行语音增强实验,原始带噪声语音及语谱图如图 12 所示,图 13、图 14、图 15 和图 16 分别是算法 1、算法 2、算法 3 和本文算法增强后语音的波形图和语谱图。

图 12 是从 59D 装甲车运动现场采集的语音信号,语音“刀光枪影”淹没在噪声中。从图 13 可看出,经谱减法处理后,噪声波形几乎消失不见,但是语谱图中只有几个高能量的点随机分布,观察不到语音的趋势,说明谱减法对语音造成了严重失真。从图 14 看出,经自适应滤波法处理后的增强语音语谱图虽然有噪点,但是能基本看清语音分量的趋势。从图 15 看出,算法 3 处理后,噪声几乎被完全去除,语

谱图比较干净且语音趋势清晰明显,但是低频信号去除太多,会造成部分语音失真。观察图 16,本文算法处理后,波形十分清晰,从语谱图中可以发现,背景几乎没有噪点,语音趋势清晰明显,低频部分的噪声被去除,而有效的低频成分保留,由经验可知,这去除的恰是装甲车的运动噪声。

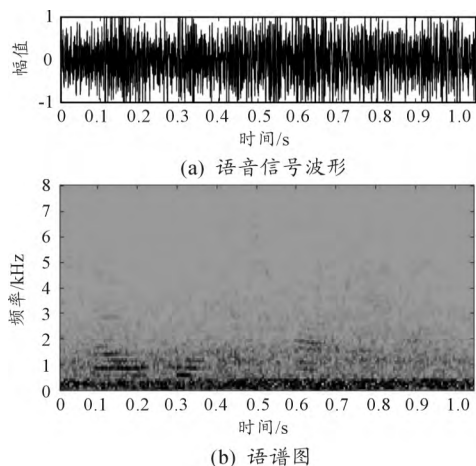


图 12 实际采集带噪语音

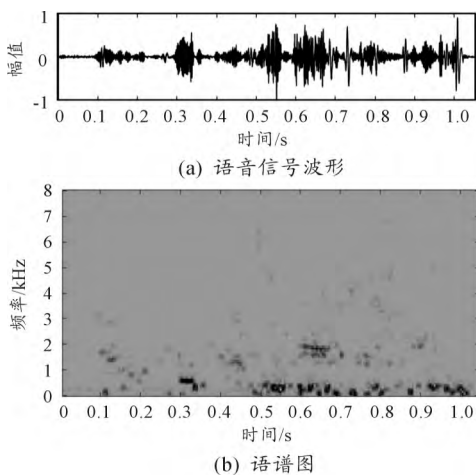


图 13 算法 1 结果

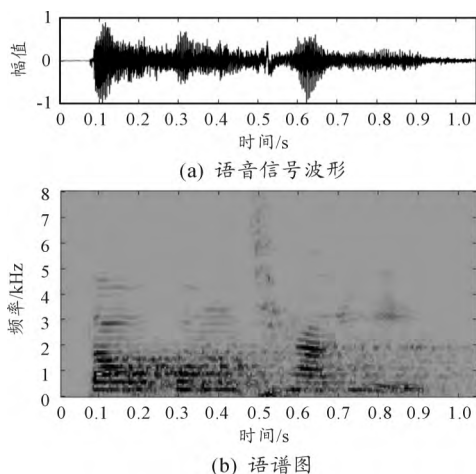


图 14 算法 2 结果

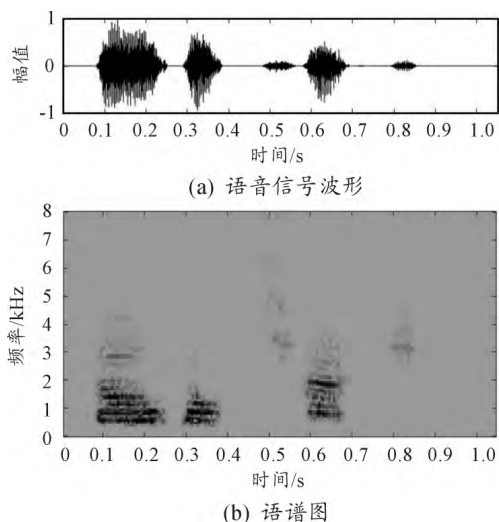


图 15 算法 3 结果

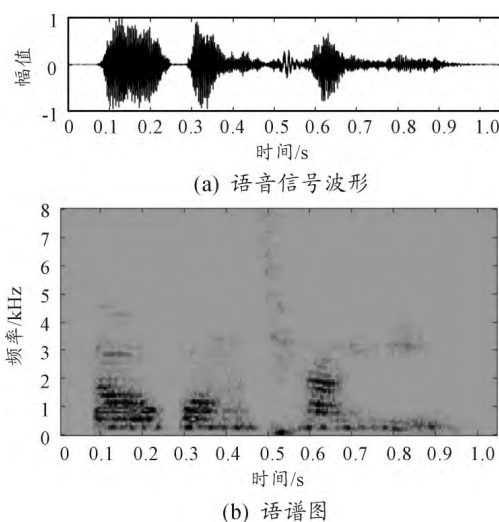


图 16 本文算法结果

由于实验采集的是真实带噪语音,没有纯净语音做参考源,无法直接得到 PESQ 和 STOI 的分值,因此以各算法增强后的语音做参考源与原始带噪语音计算,获得相对分值进行分析,分值越低,表示效果越好,结果如表 7 所示。

表 7 不同算法增强后相对 PESQ 和 STOI 分值

指标	算法 1	算法 2	算法 3	本文算法
相对 PESQ	2.973 0	1.842 2	0.512 2	0.182 9
相对 STOI	0.662 3	0.316 6	0.675 8	0.396 8

从表 7 可看出:本文算法处理后的增强语音质量和可懂度都比较高,算法 2 处理后的音频能够听懂,但质量稍差,算法 1 音频更多的是“音乐噪声”。结合表中分值可知,基于深度神经网络的语音增强算法对装甲车运动时的声场环境有很好的增强作用,与算法 3 相比,本文算法的预处理对增强

效果有更进一步的改善。

5 结论

本文算法提出在深度神经网络训练之前,先对数据进行预处理,包括 EEMD 分解和特征提取两步。本文预处理方法能有助于语音增强性能的提高,并在实际语音环境中进行检验。

本文算法主要针对装甲车运动噪声背景的语音增强,对其他噪声(如:餐厅噪声、极端噪声等)还需进一步研究,以充分发挥深度神经网络的独特作用。

参考文献:

- [1] 徐岩,王春丽. 语音信号增强技术及其应用[M]. 北京: 科学出版社, 2014.
- [2] ESCH T, VARY P. Efficient musical noise suppression for speech enhancement system[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP 2009). 2009: 4409 – 4412.
- [3] 张彦霞,肖清泰,徐建新,等. 基于经验模态分解的小波神经网络预测模型[J]. 计算机应用与软件, 2016, 33(10): 284 – 287.
- [4] 徐勇. 基于深度神经网络的语音增强方法研究[D]. 合肥: 中国科学技术大学, 2015.
- [5] 时文华,张雄伟,张瑞昕,等. 深度学习方法在语音增强中的应用[J]. 军事通信技术, 2016, 37(3): 98 – 104.
- [6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural computation, 2006, 18(7): 1527 – 1554.
- [7] HINTON G. E. A practical guide to training restricted boltzmann machines[J]. Momentum, 2010, 9(1): 926.
- [8] XU Y, DU J, DAI L R, et al. An experimental study on speech enhancement based on deep neural networks[J]. IEEE Signal Processing Letters, 2014, 21(1): 65 – 68.
- [9] LU X G, TSAO Y, MATSUDA S, et al. Speech enhancement based on deep denoising auto-encoder [C]//Proc. Inter speech. Lyon: IEEE, 2013: 436 – 440.
- [10] 陈建明,杨龙. 基于总体平均经验模态分解的语音增强算法研究[J]. 计算机应用与软件, 2017, 34(9): 328 – 333.
- [11] HUANG N E, WU M L C, LONG S R, et al. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis[J]. Proceedings Mathematical Physical & Engineering Sciences, 2003, 459(2037): 2317 – 2345.
- [12] SRIVASTAVA N, HINTON G E, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from over-fitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929 – 1958.

(责任编辑 杨继森)