

水声语音通信体验质量的实时测量方法

袁飞¹ 陈炜玲¹ 李晔² 程恩¹

(1. 厦门大学水声通信与海洋信息技术教育部重点实验室, 厦门, 361005; 2. 山东省计算中心山东省计算机网络重点实验室, 济南, 250014)

摘要: 水声语音通信质量的实时测量是保障通信质量的重要环节, 利用实时测量结果可及时调整语音业务的调制参数, 提高链路的自适应能力。本文提出了一种基于参数表示的语音通信质量实时估测模型, 该模型提取语音 3 个特征参数, 即: Mel 频率倒谱系数 (Mel-frequency cepstrum coefficient, MFCC)、线性预测倒谱系数 (Linear predictive cepstrum coefficient, LPCC) 及加权对数谱 (Log spectral deviation, LSD), 构建 3 种特征参数的权重谱失真测度。利用失真测度与接收语音质量 (Perceptual evaluation of speech quality-mean opinion score, PESQ-MOS) 之间的映射关系, 建立语音质量估测模型。引入动态 MFCC (Dynamic Mel-frequency cepstrum coefficient, DMFCC) 的谱失真测度作为质量估测模型的调节因子, 使估测系统具有更好的适应性。实验及海测结果显示, 利用本模型估测的语音 MOS 值与人主观感受误差较小, 具有一定的实用性。

关键词: 语音质量评价; 水声信道特性; 参数提取; 语音感知质量评测

中图分类号: TP393 **文献标识码:** A

Real-Time Measurement for Experience Quality of Underwater Acoustic Voice Communication

Yuan Fei¹, Chen Weiling¹, Li Ye², Cheng En¹

(1. Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of Ministry of Education, Xiamen University, Xiamen, 361005, China; 2. Shandong Provincial Key Laboratory of Computer Network, Jinan, 250014, China)

Abstract: The real-time measurement of quality of the underwater acoustic voice communication is crucial to communication quality. Real-time measurements can help to adjust voice modulation parameters timely and improve the adaptive ability of link. An objective assessment model for voice quality is proposed based on parameter extraction, based on the underwater acoustic channel characteristics. Three feature parameters of voice are extracted: Mel-frequency cepstrum coefficient (MFCC), linear predictive cepstrum coefficient (LPCC) and log spectral deviation (LSD). The three parameters form the weight spectral distortion evaluation, the mapping relationship of distortion evaluation and receive voice quality of perceptual evaluation of speech quality - mean opinion score (PESQ-MOS) is used to conduct quantization. Dynamic Mel-frequency cepstrum coefficient (DMFCC) spectral distortion evaluation is also introduced as regulatory factor, which improves the adapt-ability. The results of simulation and sea test show that the measured MOS via the assessment model is close to PESQ-MOS, which indicated the model has practical value.

Key words: voice quality assessment; characteristics of underwater acoustic channel; parameter extraction; perceptual evaluation of speech quality

引言

在能见度较低的水下环境中,语音通信是人与人之间最直接的交流手段^[1,2]。水下模拟语音通信由于无需声码器对语音信号进行复杂的编解码,具有低延迟、低复杂度、高性价比和宽信道容忍性等优点,至今仍是商用化水声语音通信系统的主流形式。模拟语音通信由于无法通过纠错编码机制提高可靠性,故及时测量终端语音质量并据此调整调制参数是维持其可用性,提高系统性能的重要手段。当前终端语音质量的测量方法分为主观和客观两种。常用的主观评价方法有平均意见得分(Mean opinion score, MOS)、诊断韵字测试(Diagnostic rhyme test, DRT)和诊断满意度测量(Diagnostic acceptability measure, DAM)等,其优点是符合人耳听觉效果,缺点是受人的主观影响大、耗费大和灵活性低等。客观评价方法在现代语音评测领域有着广泛的运用,这方面的研究也有许多^[3-13],语音质量客观评价方法大多数用数值距离或者描述听觉系统如何来感知质量的模型来量化语音质量。在现有的语音质量客观评价方法中,主观语音质量(Perceptual evaluation of speech quality, PESQ)的评估值与实际主观评估值的相关度为 0.97,达到最高^[14],但 PESQ 的测量方法需要先发送经过网络的参考信号,接收以数字信号处理的方式比较样本信号与接收到的信号估算网络的语音质量,由于在实际的通信过程中,参考信号往往无法获取,所以该方法并不适合直接应用在针对通信终端的实时测量中。E-Model 考虑窄带电话的通话质量影响因素,并将其归纳为相关传输参数,这些传输参数在终端与应用环境正常时可以直接使用 ITU-T G. 107 提供的默认值,有效避免了对参考信号的需求,但 E-Model 只是针对数字 IP 类语音,并不适用模拟语音评测应用。

本文提出 FM 模拟水声语音通信质量的估测模型,提取原始语音数据的特征参数 Mel 频率倒谱系数(Mel-frequency cepstrum coefficient, MFCC)、线性预测倒谱系数(Linear predictive cepstrum coefficient, LPCC)和加权对数谱(Log spectral deviation, LSD),并通过对其进行更好地保护后传输,解调端提取接收语音特征参数以及受更好保护,质量几乎没有受损的参考特征参数,计算获取特征参数谱失真测度。由于 PESQ 与 MOS 相关度很高,本文采用 PESQ-MOS 代替 MOS 与特征参数谱失真测度建立映射函数,避免了复杂费时的人工测评过程的同时,也提高了测评结果与主观感受的相似度。将特征参数谱失真测度作为映射函数的输入,映射出数的输出为估测的 PESQ-MOS,解调语音的质量等级依据该输出的 PESQ-MOS 来进行评估。本文综合 3 种特征参数与 PESQ-MOS 的拟合曲线情况,提出了权重谱失真测度与 PESQ-MOS 的估测模型,同时引入调节因子动态 MFCC 谱失真测度,从而依据不同情况选择模型中不同拟合曲线,使估测系统有更好的适应性。无失真的特征和语音同时进行传输,且评测所用的时间较短,故可以实现实时语音评测,该模型适用于对实时性要求高的通信系统中。

1 模拟水声语音通信质量估测模型

模拟水声语音通信质量的估测模型主要在于形成 FM 解调端提取的语音特征参数与 FM 解调信号语音质量 PESQ-MOS 得分的映射函数,解调端对信号进行特征参数 MFCC, LPCC 和 LSD 的提取,再依据特征参数谱失真测度与 PESQ-MOS 得分的函数映射来评估解调语音的质量等级,估测模型的系统框图如图 1 所示。

1.1 语音信号的特征参数

特征提取采用少量参数来有效表征语音信号,在语音编码、合成、识别以及语音质量客观评价等运用中,特征提取是非常重要的环节。在本文提出的语音质量评测模型中,语音信号的特征参数将作为评价语音信号质量的依据,选用的 3 种特征参数分别为:MFCC, LPCC 以及 LSD。

MFCC 利用听觉原理与倒谱的解相关性,并结合语音的产生机制与人耳的听觉感知特征,是目前大

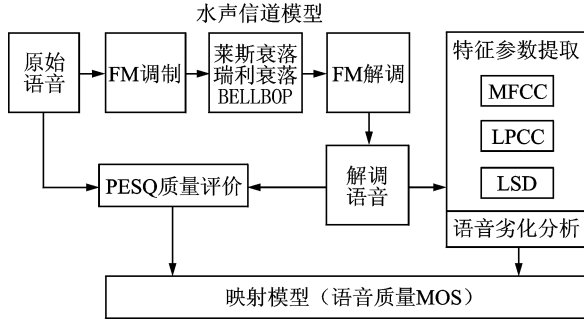


图1 模拟水声语音通信系统质量估测模型的系统

Fig.1 Voice quality assessment model for analogue underwater voice communication system

多数语音识别系统中广泛使用的语音特征参数^[15,16]。LPCC特征参数的主要优点是语音产生过程中的激励信息得到彻底地去除,能更好地反映声道响应特性,一般通过几个倒谱系数的分析,语音的共振峰特性就可以被很好地描述^[17]。LSD参数是在LPCC分析的基础上进一步加权,它是最早采用的一种谱失真距离测度,具有代表性的是运用人耳对频率的非线性感知特性的听觉频率加权对数谱。

1.2 权重谱失真测度

客观语音质量评价方法一般是建立在语音信号特征参数及其谱失真测度的基础上,在逐帧计算输入语音信号 $X(n)$ 与输出失真信号 $Y(n)$ 的特征参数后,分析这些原信号与失真信号矢量对之间失真值的加权和,得到语音通信系统的输出语音质量评测值,本文通过建立特征参数的谱失真测度与语音质量PESQ-MOS分的关系进行语音质量评价。谱失真测度算法中倒谱距离谱失真测度的定义为

$$d(C, C') = \sum_{i=1}^L (C_i - C'_i)^2 \quad (1)$$

式中: $C = \{c_1, c_2, \dots, c_L\}$, $C' = \{c'_1, c'_2, \dots, c'_L\}$ 分别为两组倒谱系数; L 为倒谱系数的个数,在谱失真测度定义中通常不用反映频谱能量的0阶倒谱系数。

本文中,MFCC特征参数、LPCC特征参数和LSD特征参数的谱失真测度 D_{MFCC} , D_{LPCC} 和 D_{LSD} 分别为

$$D_{\text{MFCC}} = E[(\text{MFCC}, \nabla \text{MFCC}) - (\text{MFCC}, \nabla \text{MFCC})']^2 \quad (2)$$

$$D_{\text{LPCC}} = \sum [(\text{LPCC} - \text{LPCC}')^2] \quad (3)$$

$$D_{\text{LSD}} = \text{LSD} \quad (4)$$

特征参数MFCC, LPCC和LSD通过添加更强的保护机制后与语音信号一起进行传输,接收端提取保护完好、质量几乎没有受损的参考特征参数,以及接收语音信号的特征参数,将两者的谱失真测度与接收端解调信号语音质量PESQ-MOS进行拟合,拟合函数为

$$\text{PESQ-MOS} = f(D) = f(D_{\text{MFCC}}, D_{\text{LPCC}}, D_{\text{LSD}}) \quad (5)$$

式中: D 为权重谱失真测度。模型中的函数拟合考虑MFCC, LPCC和LSD特征参数谱失真测度对MOS的不同影响程度,本文引入权重谱失真测度

$$D = \sum K_1 \cdot f_1(D_{\text{MFCC}}) + K_2 \cdot f_2(D_{\text{LPCC}}) + K_3 \cdot f_3(D_{\text{LSD}}) \quad (6)$$

式中: $K_1 + K_2 + K_3 = 1$ 。权重值 K_1 , K_2 和 K_3 需要依据不同谱失真测度对MOS的影响等级而定,通常MFCC谱失真测度所占比重要高一些。另外,考虑到MFCC特征参数相对复杂,且动态MFCC谱失真测度 D_{DMFCC} 可以很好地体现语音质量劣化情况,故为了提高拟合模型的准确度,本文引入调节因子-动态MFCC谱失真测度 D_{DMFCC} 作为判定条件,从而选择模型中两条质量估测拟合曲线中的一条曲线,即把调节因子 D_{DMFCC} 与预先设置的阈值进行比较,依据比较结果,判定选择语音质量评测时采用的估测拟合

曲线。

2 模拟水声语音通信质量估测模型

2.1 权重谱失真测度与 PESQ-MOS 关系曲线拟合

本文选取不同类型(包括单词、词组和语句 3 种类型)以及对应的不同说话速率(包括快、中和慢)的语音样本,选择通过不同水声信道模型(包括模拟近程水声信道的瑞利模型、模拟中远程水声信道的莱斯模型以及基于射线声学模型的 Bellhop 模型 3 种信道模型^[18]),同时改变信道条件,从而得到语音样本所对应的 MFCC 特征参数谱失真测度 D_{MFCC} ,LPCC 特征参数谱失真测度 D_{LPCC} ,LSD 特征参数谱失真测度 D_{LSD} 以及接收端语音质量 PESQ-MOS 得分,在 D_{MFCC} , D_{LPCC} , D_{LSD} 与 PESQ-MOS 关系拟合的基础上,分析得到权重谱失真测度与 PESQ-MOS 关系的拟合曲线。通过不同语音样本的多次仿真,分别得到 D_{MFCC} , D_{LPCC} , D_{LSD} 与 PESQ-MOS 的拟合函数,拟合函数式为

$$\text{PESQ-MOS} = \frac{2.689}{1 + (D_{\text{MFCC}}/1.829)^{1.980}} + 1.876 \quad (7)$$

$$\text{PESQ-MOS} = \frac{4.485}{1 + 1.123 \cdot e^{-0.173 \cdot D_{\text{LPCC}}}} \quad (8)$$

$$\text{PESQ-MOS} = \frac{2.804}{1 + (D_{\text{LSD}}/1.523)^{4.221}} + 1.993 \quad (9)$$

通过分别分析以上 MFCC,LPCC 和 LSD 特征参数谱失真测度与 PESQ-MOS 的拟合曲线情况,本文确定权重谱失真测度的最终表达式为

$$D = \sum K_1 \cdot D_{\text{MFCC}} + K_2 \cdot \left(\frac{1}{D_{\text{LPCC}}} \right) + K_3 \cdot D_{\text{LSD}} \quad (10)$$

式中: $K_1 = 0.6$, $K_2 = 0.1$, $K_3 = 0.3$ 。通过不同语音样本多次仿真,得到权重谱失真测度 D 与接收端语音质量 PESQ-MOS 关系的拟合函数为

$$\text{PESQ-MOS} = \frac{2.601}{1 + (D/1.794)^{2.058}} + 1.953 \quad (11)$$

拟合曲线如图 2 所示。

2.2 改进权重谱失真测度与 PESQ-MOS 关系曲线拟合

动态 MFCC 谱失真测度 D_{DMFCC} 可以很好体现语音质量劣化情况,仿真实验表明,当 $D_{\text{DMFCC}} > 1.39$ 时,语音质量较差,语音质量恶化程度急剧增大。为了提高拟合模型的准确度,本文将原权重谱失真测度与 PESQ-MOS 的拟合模型中的样点分为原拟合曲线以上和原拟合曲线以下分别进行拟合,如图 2 所示。分别将两类样点标识为红色和黑色。同时,引入调节因子动态 MFCC 谱失真测度 D_{DMFCC} 作为判定条件,当调节因子大于阈值 1.39 时,说明语音劣化情况较严重,对语音进行质量评测时采用原拟合曲线以下的样点拟合模型,否则采用原拟合曲线以上的样点拟合模型。通过拟合原曲线以下的样点,得到改进的权重谱失真测度 D 与接收端语音质量 PESQ-MOS1 关系的拟合函数为

$$\text{PESQ-MOS1} = \frac{2.421}{1 + (D/1.927)^{2.241}} + 2.099 \quad (12)$$

通过原拟合曲线以上的样点,得到改进的权重谱失真测度 D 与接收端语音质量 PESQ-MOS2 关系的拟合函数为

$$\text{PESQ-MOS2} = \frac{2.98733}{1 + (D/1.661)^{1.634}} + 1.706 \quad (13)$$

通过以上方法,本文将改进拟合模型分为较差环境下和较优环境下两种,将能反映语音质量劣化状况的调节因子 D_{DMFCC} 作为改进拟合模型中环境优劣的标识,从而依据不同情况选择不同的拟合曲线。结合调节因子 D_{DMFCC} 与权重谱失真测度 D 来估测接收端的语音质量 PESQ-MOS 得分,这样使估测系统

具有更好的适应性,原拟合曲线与改进拟合曲线的对比效果如图 3 所示。

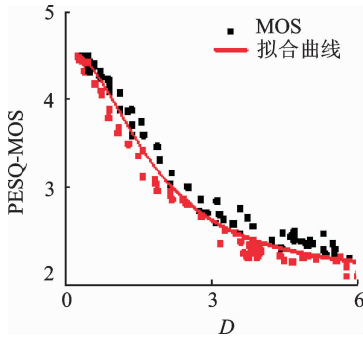


图 2 权重谱失真测度与 PESQ-MOS 关系拟合曲线

Fig. 2 Fitting curve of weighted spectrum distortion measure vs. PESQ-MOS

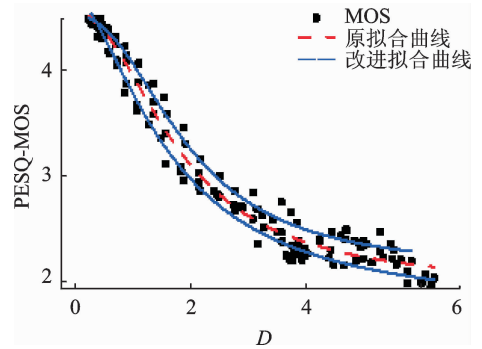


图 3 权重谱失真测度与 PESQ-MOS 拟合曲线(原拟合与改进拟合曲线)

Fig. 3 Fitting curve of weighted spectrum distortion measure vs. PESQ-MOS (original fitting curve and modified fitting curve)

3 实验结果分析

3.1 模型性能仿真验证

本文通过另外的语音样本库来验证估测模型的相关性能,仿真验证中选取不同类型(包括单词、词组和语句 3 种类型)的语音样本,选择通过不同水声信道模型(包括瑞利信道、莱斯信道、Bellhop 信道 3 种信道),同时改变信道条件,得到语音样本的模型估测 MOS, PESQ-MOS 以及两者的差值如表 1 所示。分析表 1 数据可知,语句和词组语音样本的估测误差不超过 0.1, 单词语音样本的估测误差不超过 0.5, 相对于相邻语音质量等级之间的 MOS 基本跨度 1.0, 模型估测 MOS 与 PESQ-MOS 的估测误差不超过 0.5, 所以大多情况下通过估测 MOS 得到的语音质量等级与 PESQ-MOS 得到的语音质量等级相同。

表 1 估测模型的仿真验证参数

Tab. 1 Simulation parameter of evaluation model

语音类型	质量状况	调节因子	PESQ-MOS	估测 MOS	估测差值
语句	好	0.148 790	4.432 20	4.442 76	0.010 56
	中	1.248 510	2.560 19	2.573 52	0.013 33
	差	2.918 540	1.214 73	1.272 37	0.057 64
词组	好	0.121 778	4.480 82	4.420 48	-0.060 34
	中	1.253 520	2.604 67	2.677 81	0.073 14
	差	2.741 410	1.581 88	1.581 88	0.000 00
单词	好	0.172 716	4.463 14	4.421 29	-0.041 85
	中	1.156 220	2.484 420	2.544 3	0.059 88
	差	2.629 840	1.659 54	1.888 42	0.228 88

3.2 模型性能海测验证

本文通过厦大水库、厦门五缘湾海域的测试实验来验证并完善估测模型,实验地点坐标图和实验流程图分别如图 4, 5 所示。测试实验信息如表 2 所示。

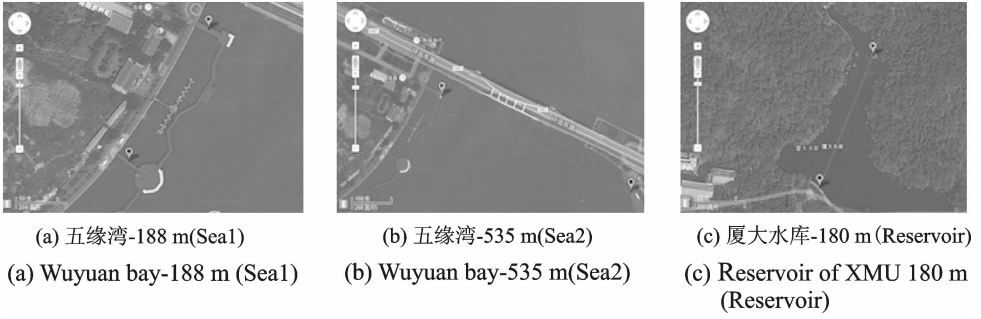


图 4 实验测试地点坐标图

Fig. 4 Coordinate graph of test site

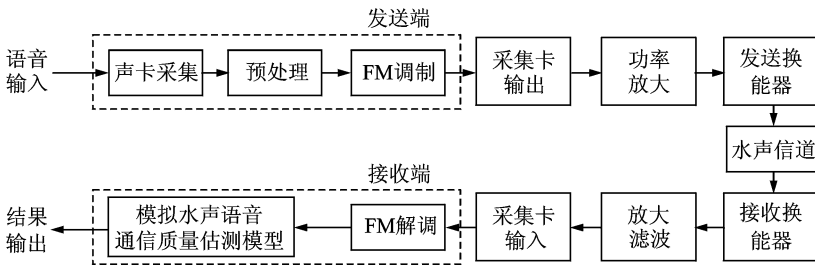


图 5 厦大水库与厦门五缘湾海域实验流程

Fig. 5 Flow diagram of experiment in reservoir of XMU and Wuyuan bay waters in Xiamen

表 2 测试实验相关信息

Tab. 2 Test information

地点	环境标记	收发距离/m	发送数据
厦大水库	Reservoir	180	FM 调制信号 4 组
五缘湾	Sea1	188	载波频率均为 25 kHz
五缘湾	Sea2	535	采样率均为 96 kHz

实验中采用采集速率最高为 1.25 MS/s 的 NI-6259 采集卡、带宽为 20~30 kHz 的水声换能器以及 100 W 功率放大器,换能器距离水面 1.5 m,选取具有一定代表性的 4 组中文男声语音样本进行测试,语音内容分别为“厦门大学通信工程”、“前进后退上升下降”、“右转”和“左转”。依据解调语音的人耳听觉效果,将解调语音的实际可懂度分为水声环境下的质量等级:差、较差、中等、良好和好。海测样本的模型估测 MOS, PESQ-MOS 与两者的差值等参数如表 3 所示。

海测实验中“左转”原始与解调语音的时域波形图如图 6 所示,图 6(a~c)中(I, II)分别表示不同环境下的原始语音和解调语音。由图 6 可以看出,海洋环境下语音的失真情况比较严重。综合分析可知,实际海洋环境信道特性复杂,其语音 MOS 要比依据水声仿真信道建立的模型估测 MOS 低,且降低的程度与调节因子 D_{DMFCC} 存在一定关系。分析此次海测实验结果,本文认为水声仿真信道建立的 FM 水声语音通信质量估测模型可以做出如下改进:在原估测 MOS 的基础上,降低 CE 值,则在改进后的估测模型中,估测 $MOS-G = MOS - CE$ 。其中 CE 值与调节因子 D_{DMFCC} 的关系如表 4 所示。

4 结束语

本文主要研究模拟水声语音通信质量的估测模型,归纳了常用的语音特征参数表示,给出了提取特征参数的水印方法,并通过分别分析特征参数 MFCC, LPCC 和 LSD 的谱失真测度与语通信音质量

PESQ-MOS 的拟合关系,改进得到权重谱失真测度与 PESQ-MOS 的估测模型,通过语音样本库的仿真实验对估测模型的相关性能进行验证。海测实验表明,本文提出的语音质量评估模型与人的主观感受相近,具有一定的实用价值。

表 3 估测模型的海测实验参数表

Tab. 3 Sea test parameter table of evaluation model

语音	环境	调节因子	PESQ-MOS	估测 MOS	估测差值	实际可懂度
厦门大学 通信工程	Sea1	2.900 2	1.690 7	2.170 5	0.479 8	良好
	Sea2	2.716 9	1.539 8	2.261 6	0.721 8	良好
	Reservoir	3.709 5	1.500 3	2.263 5	0.763 2	良好
前进后退 上升下降	Sea1	4.103 5	0.967 8	2.159 2	1.191 4	较差
	Sea2	4.029 0	0.972 2	2.231 2	1.259 0	较差
	Reservoir	4.394 1	0.941 0	2.167 8	1.226 8	较差
右转	Sea1	3.221 8	0.961 1	2.032 1	1.071 0	中等
	Sea2	2.867 5	1.201 3	2.187 2	0.985 9	良好
	Reservoir	3.465 3	1.164 5	2.217 1	1.052 6	良好
左转	Sea1	3.115 4	1.791 8	2.131 3	0.339 5	良好
	Sea2	2.657 4	1.696 0	2.285 7	0.589 7	良好
	Reservoir	3.530 0	1.510 6	2.196 2	0.685 6	良好

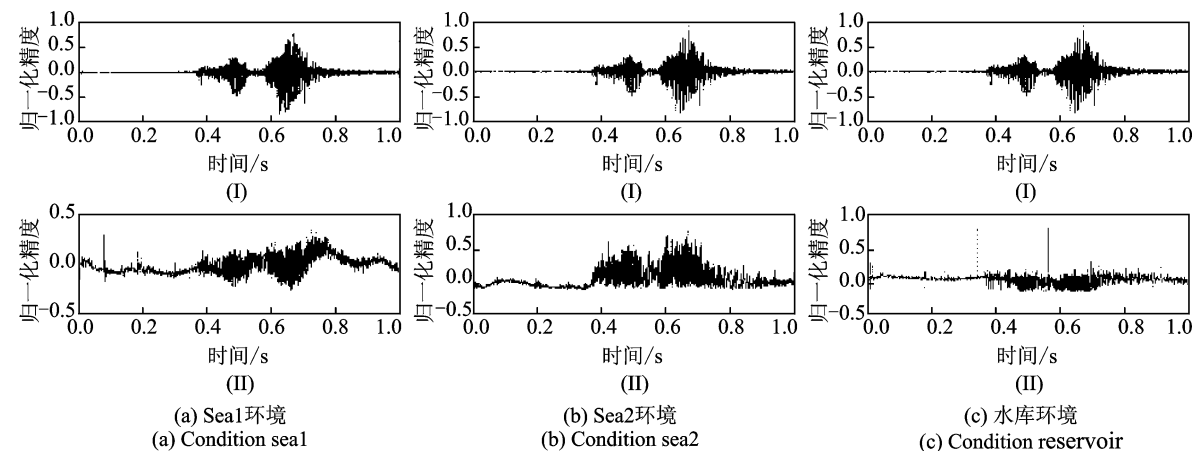


图 6 “左转”原始语音与解调语音时域波形

Fig. 6 Time domain waveform of original voice and demodulation voice of "turn left"

表 4 CE 值与 D_{DMFCC} 的关系表Tab. 4 Correlation chart of CE value and D_{DMFCC}

D_{DMFCC} 值范围	CE 值	D_{DMFCC} 值范围	CE 值
[0, 1.5)	0.0	[3, 3.5)	0.8
[1.5, 2)	0.3	[3.5, 4.5)	1.0
[2, 3)	0.5	≥ 4.5	0.3

参考文献:

- [1] 易克初. 语音信号处理[M]. 北京: 机械工业出版社, 2003: 223-224.
Yi Kechu. Voice signal processing[M]. Beijing: China Machine Press, 2003: 223-224.
- [2] 张雄伟, 陈亮, 杨吉斌. 现代语音处理技术及应用[M]. 北京: 机械工业出版社, 2003: 13-14.
Zhang Xiongwei, Chen Liang, Yang Jibin. Modern speech processing technology and application[M]. Beijing: China Ma-

chine Press, 2003: 13-14.

- [3] 戴礼荣, 张仕良. 深度语音信号与信息处理: 研究进展与展望[J]. 数据采集与处理, 2014, 29(2): 171-169.
Dai Lirong, Zhang Shiliang. Deep speech signal and information processing: Research progress and prospect [J]. Journal of Data Acquisition and Processing, 2014, 29(2): 171-179.
- [4] Voran S. Estimation of perceived speech quality using measuring normalizing blocks[C]//IEEE Workshop on Speech Coding for Telecommunications Proceedings. Pocono Manor PA, USA; IEEE, 1997: 83-84.
- [5] Voran S. Objective estimation of perceived speech quality part-I; Development of the measuring normalizing block technique [J]. IEEE Transactions on Speech and Audio Processing, 1999, 7(4): 371-382.
- [6] Voran S. Objective estimation of perceived speech quality part-II; Evaluation of the measuring normalizing block technique [J]. IEEE Transactions on Speech and Audio Processing, 1999, 7(4): 383-390.
- [7] 秦基伟. 语音质量客观评价系统的研究及实现[M]. 重庆: 重庆大学出版社, 2013.
Qin Jiwei. The research and implementation of objective quality evaluation of voice[M]. Chongqing: Chongqing University Press, 2013.
- [8] Liang J, Kubichek R. Output-based objective speech quality [C]//1994 IEEE 44th, Vehicular Technology Conference. [S. l.]; IEEE, 1994: 1719-1723.
- [9] Jin C, Kubichek R. Output-based objective speech quality using vector quantization techniques[C]// ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing. Atlanta Piscataway, NJ, United States; IEEE, 1995: 1291-1294.
- [10] Yu Ping, Wang Zhijian, Liu Shanshan, et al. Multidimensional acoustic analysis for voice quality assessment based on the GRBAS scale[C]// Proceedings of the 9th International Symposium on Chinese Spoken Language Processing, ISCSLP 2014. Singapore: Institute of Electrical and Electronics Engineers Inc, 2014: 321-325.
- [11] Lai Yinghui, Tang Tsang, Tsai Kuenshian, et al. Objective speech quality measurement[J]. Journal of Medical and Biological Engineering, 2013, 33(6): 576-582.
- [12] Siddique M A R, Kamruzzaman J, Hossain M J. An analytical approach for voice capacity estimation over WIFI network using ITU-T e-model[J]. Multimedia IEEE Transactions on, 2014, 16(2): 360-372.
- [13] 成卫青, 龚俭, 丁伟. 基于流特性和真值程度的 VoIP 语音质量单端客观评价[J]. 通信学报, 2008, 29(4): 30-39.
Cheng Weiqing, Gong Jian, Ding Wei. Flow features and truth scale based single-ended objective assessment of perceived quality of VoIP service[J]. Journal of Communication, 2008, 29(4): 30-39.
- [14] 付俊. 通信系统中语音质量客观评测技术研究[D]. 成都: 电子科技大学, 2011.
Fu Jun. Research on objective quality evaluation of voice of communication system[D]. Chengdu: University of Electronic Science and Technology of China, 2011.
- [15] Zhang Qiuyu, Liu Yangwei, Di Yanjun, et al. MFCC-based perceptual hashing for compressed domain of speech content identification [J]. Journal of Chemical and Pharmaceutical Research, 2014, 6(7): 379-386.
- [16] 钟巍, 孔祥维, 尤新刚, 等. 基于分数倒谱变换的取证语音拼接特征提取与分析[J]. 数据采集与处理, 2014, 29(2): 248-253.
Zhong Wei, Kong Xiangwei, You Xingang, et al. Splicing feature extraction and analysis based on fractional cepstrum transform in voice forensics[J]. Journal of Data Acquisition and Processing, 2014, 29(2): 248-253.
- [17] Mohammed E M, Sayed M S, Moselhy A M, et al. LPC and MFCC performance evaluation with artificial neural network for spoken language identification[J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2013, 6(3): 55-66.
- [18] 段乐峥. 基于 BELLHOP 的水声信道时变模型[J]. 电子世界, 2014, 9: 105-112.
Duan Lezheng. The time-varying multipath underwater acoustic channel model based on the BELLHOP [J]. Electronic Worlds, 2014, 9: 105-112.

作者简介:



袁飞 (1979-), 男, 副教授, 博士, 研究方向: 水声通信、水下多媒体信号传输及处理, E-mail: yuanfei@xmu.edu.cn.



陈炜玲 (1991-), 女, 博士研究生, 研究方向: 水声通信, 信号处理。



李晔 (1982-), 男, 副研究员, 研究方向: 信号处理和语音编码。



程恩 (1965-), 男, 教授, 博士生导师, 研究方向: 水下通信和数字信号处理。

