语音质量评估方法研究

路 萌 工业和信息化部电信研究院泰尔实验室工程师

摘 要 对各标准化组织中针对语音质量评估方法制定的标准进行了介绍，分析了国际电信联盟中 PSQM，PESQ，POLQA，欧洲电信标准协会 ETSIEG 202-396，TOSQA 等评估方法的模型和算法。

关键词 语音质量 绝对种类评定值 声学模型 多媒体传输质量

1 引言

从移动终端通信业务发展开始，语音质量就是移动用户体验的最重要的一个衡量标准。语音质量受到个人区别、可理解性、语音特征、周围环境、背景噪声传输、网络状况和人的期望等复杂的因素影响，难以得到一个统一且直观的评估结果。各网络运营商、终端厂商和各国的电信联盟组织也都致力于研究能够直观反映用户通话体验的评估方案。本文就业界主流的语音质量评估方法进行介绍和分析。

2 主观评估方法

被广泛认同的语音质量评估方法是人们凭主观的判断、通过实际通话，由人耳来感知通话质量的好坏。人类的听觉和感知语音的行为被量化后，从而得到语音质量的级别。在 20 世纪 90 年代由国际电信联盟 ITU-T 发布了 P.800，使用 MOS（平均主观评分）以规范化主观的评估语音质量的方法。利用人本身的主观感受为听到的语音的满意度进行评分，评分的基本指标是可理解程度。其中，定义了绝对种类评定值（ACR）测量方法。该方法是要求 20～50 人分别听完一段语音片段后根据自己的感受打出一个主观分值，分值分布如表 1 所示。

测试完成后，对所有评分取平均值，这个平均主观值 MOS（Mean Opinion Score）是被广泛应用的语音质量量化标准；因此，本文中所提出的所有客观测量方法的评估结果都会对应到平均主观值上。

由于 P.800 中所规定的主观评估方法对试验样本数量、试验者背景、试验环境、语音样本等试验程序都有严格的要求，要完成该项测试要消耗大量的人力物力，成本高，耗时长。最重要的是受到试验环境的影响，该测试可重复性差，仅能反应某特定语言环境的语音质量主观评估，并不能在多个语言环境中进行比较。

表 1 ACR 评分指标

ACR 指标评分	
分值	语音质量
5	非常好
4	好
3	一般
2	较差
1	差

3 客观评估方法

3.1 PSQM

ITU-T 在 1998 年发布了 P.861 感知通话质量测量（PSQM），此方法使用计算机产生的波型文件，通过比较其通过通信网络传输前后的变化计算出 PSQM 中相对应的级别及好坏程度。测量方法需要发送一个语音参考信号通过电话网络，在网络的另一端采用数字信号处理的方式比较样本信号和接收到的信号，进而估算网络的语音质量。

流程示意图如图 1 所示，其中 PSQM 模型的核心是听觉变换，模拟了人的听觉系统的主要心理和物理上的处理过程。并引入了认知模型来描述劣化信号和

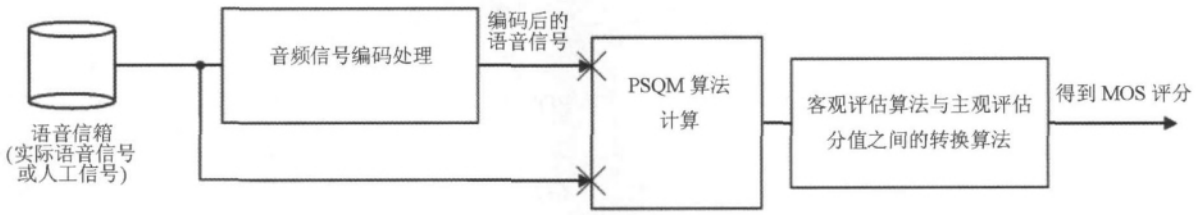

图 1 PSQM 模型流程图

原始输入信号在听觉变换过程中产生的差距。

但是对某些类型的编解码、背景噪声和端到端的影响，比如滤波和时延变化，PSQM 不能给出精确的预测值。它仅对网络及终端设备的包括编码器语音输入电平、传输速率、编码转换等编码质量进行评估，且只局限于窄带编码测量。

3.2 PESQ

ITU-T 在 2001 年 2 月发布了 P.862 来取代 PSQM。P.862 感知评估通话质量测量 PESQA (Perceptual Evaluation of Speech QualityAnalysis) 是用来计算语音样本的 MOS-LQO (Mean Opinion Score-Listening Quality Objective) 值，把在信号传输通过设备时提取的输出信号与参照信号进行比较计算出差异值。在 PEQM 的基础上，改进了时间延时和线性滤波这两个方面。

对于所有模拟网络来说，为了计算统计指标，已经使用了按条件平均的方法，至少要 4 位谈话人，2 位女性和 2 位男性。对真实网络的数据库，已经用按样本的客观和主观评分来计算统计指标。

如图 2 所示，PESQA 的听觉质量模型包括了人类感受声音所应用的感知模型和认知模型。通过将原始输入信号和劣化信号切分为多个时域帧信号，评估二

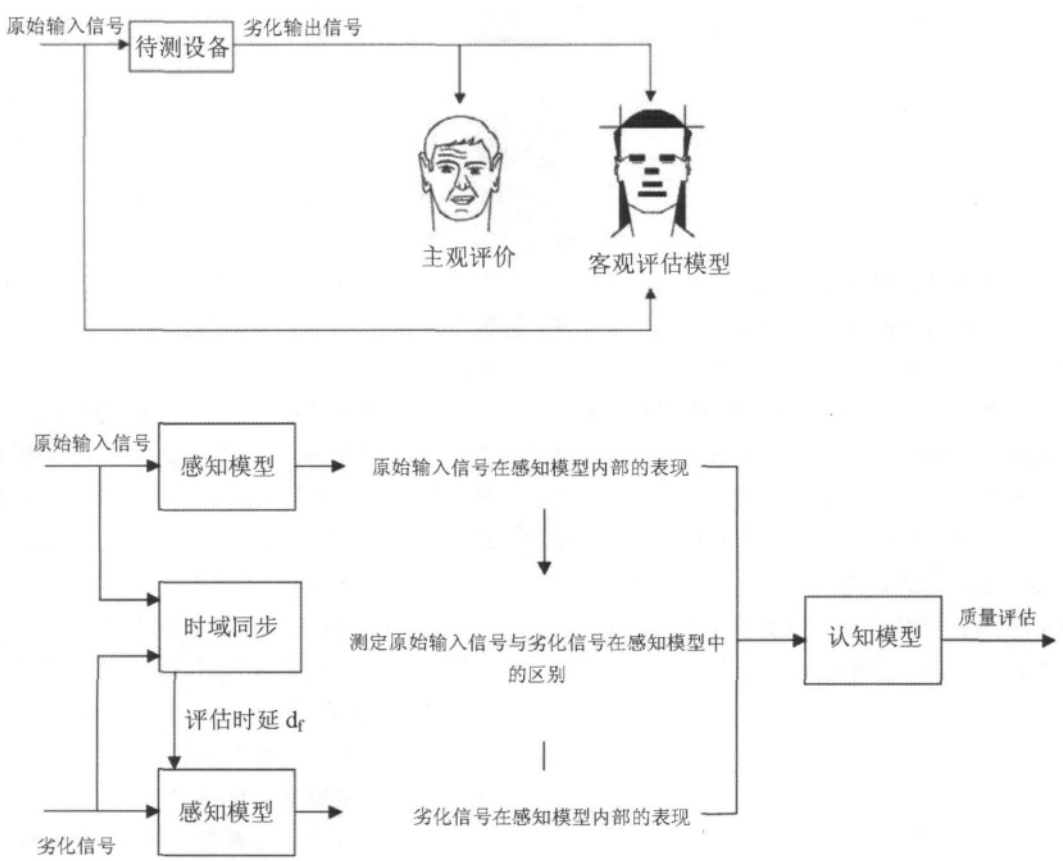

图 2 PESQ 模型示意图

者的时域延时，消除延时后将二者同步，通过感知模型中的算法模拟人类真实听觉对测试信号的反应。

简言之，通过比较原始音信号和劣化语音信号，这些客观质量模型使用规定的认知模型，通过规定的算法计算得到客观质量分值。它综合考虑了感知中的编码器语音输入电平，误码率、编码速率、编码传输、多速率编码传输下的编解码匹配、噪声等因素的影响来客观地评价语音信号的质量，从而提供可以完全量化的语音质量衡量方法。但只适用于窄带网络通信，即300Hz～3.4kHz的语音信号编解码网络。在随后颁布的ITU-T P.862.2建议书中，完善了PESQ，将带宽扩展至50Hz～7kHz，实现了对宽带网络通信的评估。

但对于线性失真的评估缺失使得其并不能完全适用于声对声接口的通信场景，也就是说该评估结果并不能将通信终端对语音处理所带来的影响考虑在内。

3.3 POLQA

对于现代移动通信终端的测试要求考虑到从嘴到耳朵包括终端在内的整个传输通路，数字终端做了大量处理语音编码信号的工作。因此，仅考虑电接口不足以评估端到端的通信业务。由于PESQ测试标准的局限性，ITU-T于2011年开发了P.863 POLQA（Perceptual Objective Listening Quality Analysis）作为下一代语音质量测试技术，是对P.862的改进。可用于固定电话网络包括LTE在内的移动网络及IP电话网络，还将高频带宽扩展至14kHz，支持超宽带语音业务。

在POLQA的声学模型中，是ITU-T P.86x系列中第一次参考了以下影响语音质量的因素：

- EVRC编码。
- 降噪和语音增强技术。
- UCC和VoIP时间规整。
- 非最优呈现等级。
- 滤波及频谱整合。
- 人工耳录制语音信号。

P.863模型示意图参见图3。在P.862的基础上，POLQA计算模型完成了时域同步、幅值同步、频率规整、响度压缩等步骤，并考虑了人耳响度感受对评估结果的影响。在认知模型输出中，频率指标、噪声指标、室内混响指标，以及时间、响度和声调在听觉中的差别指标综合评估得出MOS分值。

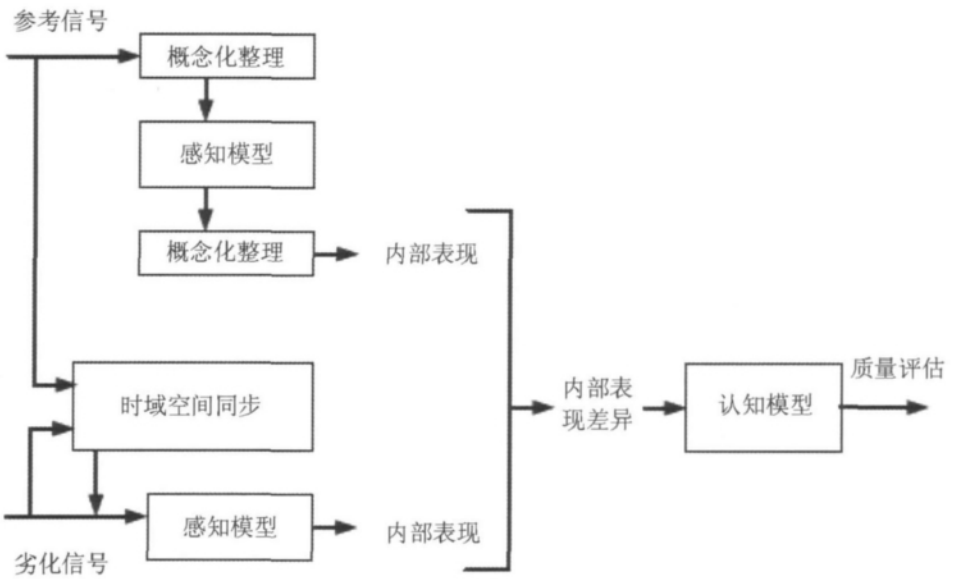

图3 P.863模型示意图

但是P.863同样存在其局限性，它只考虑了收听场合下的语音质量评估，如单向语音失真、噪音和响度等对语音传输质量的影响，并没有包括双向通话情况下的指标。

3.4 TOSQA

电信客观语音质量评价TOSQA是一种基于心理声学的测试方法，可用于评估窄带或宽带网络以及终端的语音传输质量。该方法虽未被任何标准化组织正式颁布，但却因为其成熟的测试系统和运营商的采用而被业界广泛采用。

TOSQA算法考虑了包括频率响应、失真、空闲信道噪声、测音、回声抑制、双讲性能等音频效果在内的多种因素，是根据大量的主观测试，将处理的语音信号和原始语音信号相对比，收敛训练序列得到的。算法收敛的流程如图4所示。

如图5所示，大量验证测试显示TOSQA算法的语音质量预测结果与人的主观评价结果的相关性大

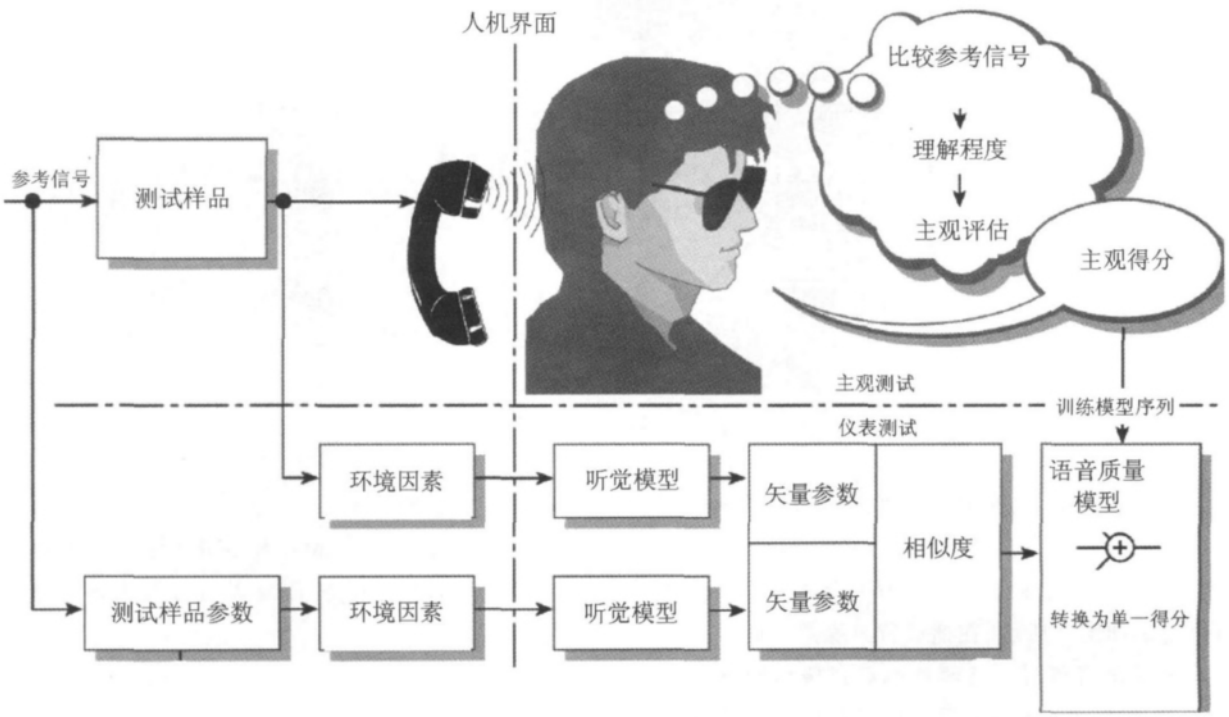

图 4 TOSQA 模型示意图

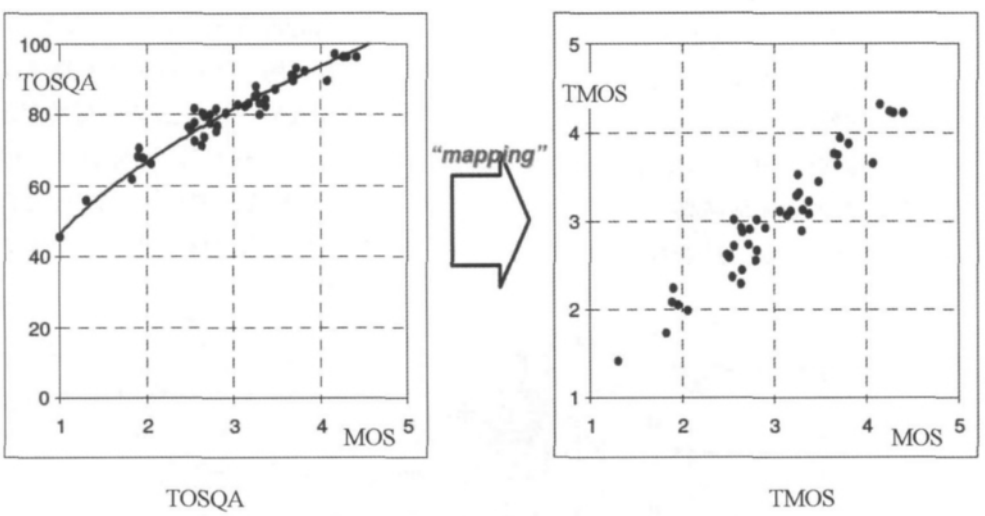

图 5 TOSQA 算法与 MOS 分值相关性验证结果

于 90%。

采用 TOSQA 算法得到语音质量评估分，再对应到平均主观值 MOS 所得结果即为 TMOS。该指标得到了许多运营商及终端厂家的认可，并适用于世界多个运营商内部标准。

3.5 ETSI 202 396

欧洲电信标准协会 ETSI EG 202 396 在语音和多媒体传输质量(STQ)中定义了在背景噪音下的语音质量性能。其中，第一部分规范了多种环境噪声音源数据库，第三部分规范化了背景噪音传输客观测试方法。

图 6 表示在这种测试方法中，与上述评估方法相同，原始语音信号经过主观测试采样，对多种语言的八个句子在不同测试条件下进行评价，在相同的采样个数条件下，收敛出听觉模型的算法。不同的是，在终

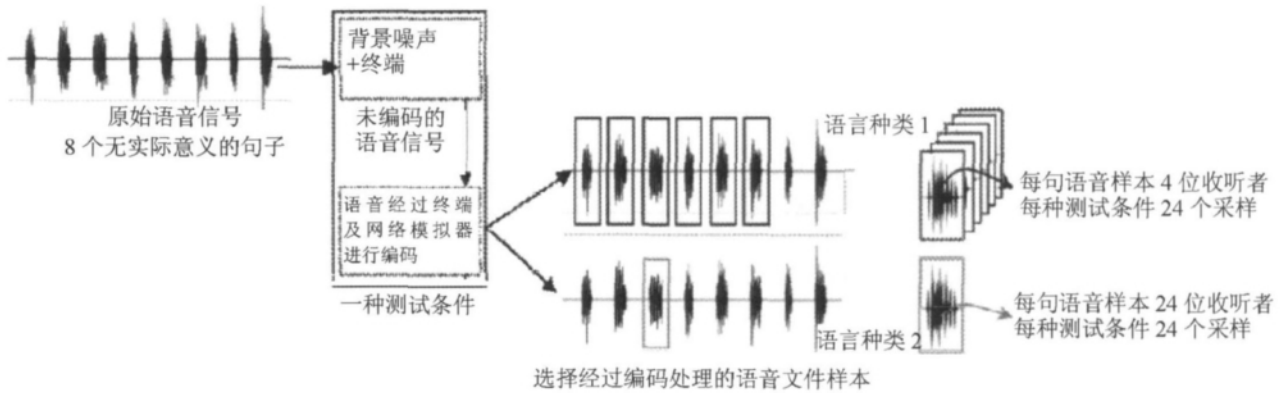

图6 采样流程示意图

端播放语音的同时,在声接口层面引入了背景噪声作为测试参数。

第一部分中还规定了背景噪声模拟系统的具体测试环境。如图7所示,在播放背景噪声的同时,使用人工头发送语音信号。待测终端经过模拟网络通话,对噪声信号和背景噪声信号共同进行内部算法处理,得到处理后的信号p(k)。待测终端麦克风旁放置参考麦克风,接收语音及噪声信号,得到未处理信号u(k)。原始语音信号c(k)与前二者形成三组不同数据参数。将这三者通过对应算法相比较,得出三个反应被测设备在背景噪声表现优劣的评估结果:

- S-MOS:在噪声环境下语音信号传输评估结果。
- N-MOS:噪声信号的传输评估结果。
- G-MOS:S-MOS和N-MOS经过线性二次回归得到的概括性评估结果。

此评估方法被GSMA和3GPP等标准组织所采用,ITU-T SG12也正在开发基于202 396-3的标准P.ONRA。

4 结束语

语音质量评估方法不仅局限于本文中所介绍的,还有例如PSQM99和PAMS等测试方法,也被不同的电信运营商所采用。用户对于语音质量的需求使得各方面都对语音质量评估方法有更高的要求。上文所述的多种评估方法各有利弊,业界也没有停止优化评估模型、开发新的客观模型算法。以满足直观了解网络条件,终端软硬件能力的要求,达到优化用户直观听觉感受的目的。

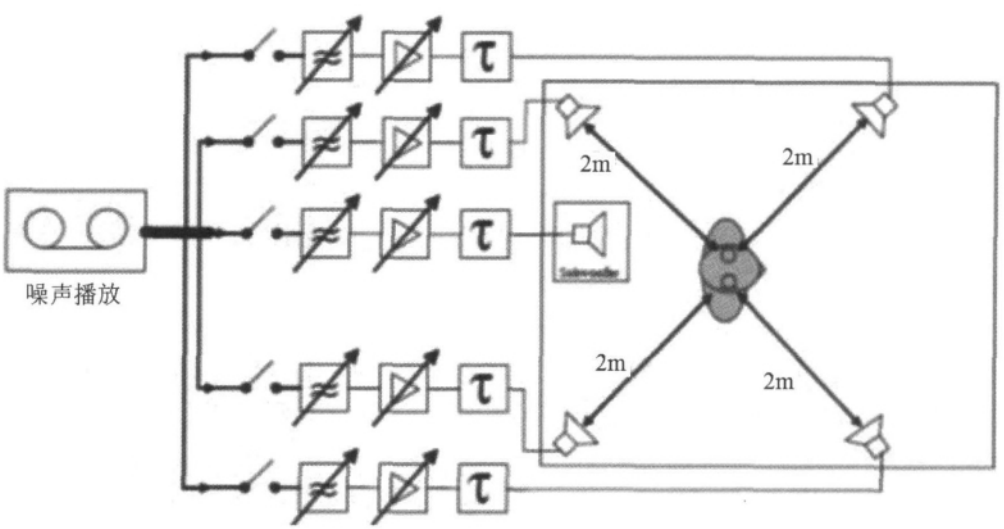

图7 背景噪声回放系统示意图

(收稿日期:2012-07-20)