

国 际 电 信 联 盟

ITU-T

国际电信联盟
电信标准化部门

P.800.1
(07/2006)

P系列：电话传输质量、电话装置和本地线路网络
质量的客观和主观评定方法

平均评价计分（MOS）的术语

ITU-T P.800.1建议书



ITU-T P-系列建议书
电话传输质量、电话装置和本地线路网络

名词术语和传输参数对用户传输质量意见的影响	系列	P.10
用户线和话机	系列	P.30
		P.300
传输标准	系列	P.40
客观测量装置	系列	P.50
		P.500
客观电声测量	系列	P.60
与话音响度有关的测量	系列	P.70
质量的客观和主观评定方法	系列	P.80
		P.800
多媒体业务的音视频质量	系列	P.900
IP 端点的传输性能和业务质量问题	系列	P.1000

欲了解更详细信息，请查阅 ITU-T 建议书目录。

平均评价计分（MOS）的术语

摘要

本建议书所列术语应在描述语音质量的平均评价计分（MOS）时使用。制定这些新术语的动因是为了避免在下述情况下产生误解，即某特定MOS值是否与收听、通话或对话质量相关，以及它们究竟是源于主观测试、客观模型还是网络规划的模型。此外，此次修订为带宽和接口类型（电气或声音）提供了标识符。

来源

ITU-T第12研究组（2005-2008年）按照ITU-T A.8建议书规定的程序，于2006年7月14日批准了ITU-T P.800.1建议书。

前 言

国际电信联盟（ITU）是从事电信领域工作的联合国专门机构。ITU-T（国际电信联盟电信标准化部门）是国际电信联盟的常设机构，负责研究技术、操作和资费问题，并且为在世界范围内实现电信标准化，发表有关上述研究项目的建议书。

每四年一届的世界电信标准化全会（WTSA）确定 ITU-T 各研究组的研究课题，再由各研究组制定有关这些课题的建议书。

WTSA 第 1 号决议规定了批准建议书须遵循的程序。

属 ITU-T 研究范围的某些信息技术领域的必要标准，是与国际标准化组织（ISO）和国际电工技术委员会（IEC）合作制定的。

注

本建议书为简明扼要起见而使用的“主管部门”一词，既指电信主管部门，又指经认可的运营机构。

遵守本建议书的规定是以自愿为基础的，但建议书可能包含某些强制性条款（以确保例如互操作性或适用性等），只有满足所有强制性条款的规定，才能达到遵守建议书的目的。“应该”或“必须”等其它一些强制性用语及其否定形式被用于表达特定要求。使用此类用语不表示要求任何一方遵守本建议书。

知识产权

国际电联提请注意：本建议书的应用或实施可能涉及使用已申报的知识产权。国际电联对无论是其成员还是建议书制定程序之外的其它机构提出的有关已申报的知识产权的证据、有效性或适用性不表示意见。

至本建议书批准之日止，国际电联尚未收到实施本建议书可能需要的受专利保护的知识产权的通知。但需要提醒实施者注意的是，这可能并非最新信息，因此特大力提倡他们通过下列网址查询电信标准化局（TSB）的专利数据库：<http://www.itu.int/ITU-T/ipr/>。

© 国际电联 2007

版权所有。未经国际电联事先书面许可，不得以任何手段复制本出版物的任何部分。

目录

	页码
1 参考文献	1
2 建议使用的MOS的术语	1
2.1 与只收情况相关的MOS	2
2.2 与对话情况相关的MOS	3
2.3 与讲话情况相关的MOS	3
3 某些 MOS 限定词之间的关系	4

平均评价计分（MOS）的术语

1 参考文献

下列ITU-T建议书和其他参考文献的条款，在本建议书中的引用而构成本建议书的条款。在出版时，所指出的版本是有效的。所有的建议书和其它参考文献均会得到修订，本建议书的使用者应查证是否有可能使用下列建议书或其它参考文献的最新版本。当前有效的ITU-T建议书清单定期出版。本建议书引用的文件自成一体时不具备建议书的地位。

- [1] ITU-T Recommendation P.10/G.100 (2006), *Vocabulary for performance and quality of service*.
- [2] ITU-T Recommendation G.107 (2005), *The E-model, a computational model for use in transmission planning*.
- [3] ITU-T Recommendation G.113 Appendix I (2002), *Provisional planning values for the equipment impairment factor I_e and packet-loss robustness factor B_{pl}* .
- [4] ITU-T Recommendation P.562 (2004), *Analysis and interpretation of INMD voice-service measurements*.
- [5] ITU-T Recommendation P.800 (1996), *Methods for subjective determination of transmission quality*.
- [6] ITU-T Recommendation P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.
- [7] ITU-T Recommendation P.833 (2001), *Methodology for the derivation of equipment impairment factors from subjective listening-only tests*.
- [8] ITU-T Recommendation P.834 (2002), *Methodology for the derivation of equipment impairment factors from instrumental models*.
- [9] ITU-T Recommendation P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*.

2 建议使用的MOS术语

ITU-T P.10/G.100建议书对MOS（平均评价得分）这一术语做出了下述定义：

- 评价得分的平均值，即在一个预定的取值范围内，主观评价者对用于对话或收听讲话素材的电话传输系统的性能给出评价。

除主观评价外，MOS这一缩略语亦用于源自客观模型或网络规划模型的评分。为区分应用的领域，建议在使用MOS这一缩略语的同时使用下述标识符，其中N代表窄带、W代表宽带、LQ代表收听质量、CQ代表通话质量、S表示主观、O表示客观，E表示估计。

	只听	对话	讲话
主观	MOS-LQSy	MOS-CQSy	MOS-TQSy
客观	MOS-LQOy	MOS-CQOy	MOS-TQOy
估计	MOS-LQEy	MOS-CQEy	MOS-TQEy

注 –上述首字缩写后面的字母"y"是为相关音频带宽描述符预留的位置，请见下述临时说明。

- N表示窄带（300-3 400 Hz）语音相对于窄带高质量基准的MOS得分。此值可用于窄带专项主观测试或P.862.1评分等情况。
- W表示宽带（50-7 000 Hz）语音相对于宽带高质量基准的MOS得分。此值可用于宽带专项主观测试或P.862.2评分等情况。
- 在混合带宽的情况下，M表示窄带或宽带语音相对于宽带高质量基准的MOS得分。此值可用于混合带宽主观测试等情况。

目前，ITU-T第12研究组正在对音频带宽对MOS评分的影响进行研究。如果带宽分母N、W或M不能适当地反映实际情况，则建议用适当的符号来临时替代"y"这一位置预留符。

2.1 与只收听情况相关的MOS

这些MOS评分适用于只听的情况。对此，有必要区分三种不同情况。

2.1.1 MOS-LQS

评分已通过实验室测试采集，其计算方式是依照ITU-T P.800建议书中的定义，使用5分制的ACR质量评分标准来计算出主观评价的算术平均值。依照ITU-T P.830、P.835和P.840建议书进行的主观测试给出了MOS-LQS的测试结果。

2.1.2 MOS-LQO

此项评分的计算是通过一种预测只听测试情况质量的客观模型。采用ITU-T P.862.1和P.862.2建议书中所述模式进行的客观测量，给出了MOS-LQO的测量结果。

应当注意的是，ITU-T在P.862.1和P.862.2建议书中推荐的方法仅适用于电接口。目前，国际电联有关收听质量测量的建议书中还未包括针对声音接口的建议书。旨在将声音接口也包括在内的新ITU-T P.OLQA建议书正在制定之中。

对此，应当注意，有必要对收听质量的介入式端对端测量进行区分，以使：

- **MOS-LQO（电）**

此类测量仅在电接口进行。为预测用户感受到的收听质量，对终端的IRS或经更正的IRS频率响应作出了假设；而这便隐含了这样一种假设，即电话接收器与用户耳朵之间处于密闭状态。ITU-T P.862建议书属于这一范围。

- **MOS-LQO（音频）**

此类测量在声音接口进行。为预测用户感受到的收听质量，此测量中包括了由制造商或供应商提供的电话机。结合实验室测试中选择的语音接收机（“仿真耳”）看，在电话接收机与仿真耳间将或多或少存在缝隙。因此，对于更为真实的测试方案，测出的MOS值可能会下降，而对于仿真程度更高的测试方案，则可能会存在可忽略的差异。

2.1.3 MOS-LQE

此项评分是通过网络规划模型计算得出，旨在预测只听应用的质量。

2.2 与对话情况相关的MOS

此类MOS评分适用于对话情况。对此，有必要区分三种不同情况。

2.2.1 MOS-CQS

评分已通过实验室测试采集，其计算方式是依照ITU-T P.800建议书中的定义，使用5分制的ACR质量评分标准来计算出主观评价的算术平均值。依照ITU-T P.800、P.831和P.832建议书进行的对话测试给出了MOS-CQS的测试结果。

2.2.2 MOS-CQO

此项评分是通过一种旨在预测对话测试情况质量的客观模型算出的。采用ITU-T P.562和P.563建议书中所述模式进行的客观测量，给出了MOS-CQO的测量结果。

2.2.3 MOS-CQE

此项评分是通过一种旨在预测对话应用质量的网络规划模型算出的。根据ITU-T G.107建议书对对话质量进行的预计，在转换为平均评价评分后，给出了MOS-CQE的结果。

2.3 与讲话情况相关的MOS

讲话质量仅描述了讲话一方感受到的电话质量。讲话质量主要受到回声信号和背景噪声倒换效应以及通话重叠的影响。

2.3.1 MOS-TQS

评分已通过实验室测试采集，其计算方式是依照ITU-T P.800建议书中的定义，使用5分制的ACR质量评分标准来计算出主观评价的算术平均值。

2.3.2 MOS-TQO

此项评分是通过一种旨在预测只讲测试情况质量的客观模型算出的。生成MOS-TQO的方法目前仍在开发阶段且并未标准化。

2.3.3 MOS-TQE

此项评分是通过一种旨在预测只讲应用测试情况质量的客观模型算出的。生成MOS-TQO的方法目前尚未标准化。

3 某些MOS限定词之间的关系

图1所示为MOS-LQS、MOS-LQO和MOS-LQE间的关系概述。

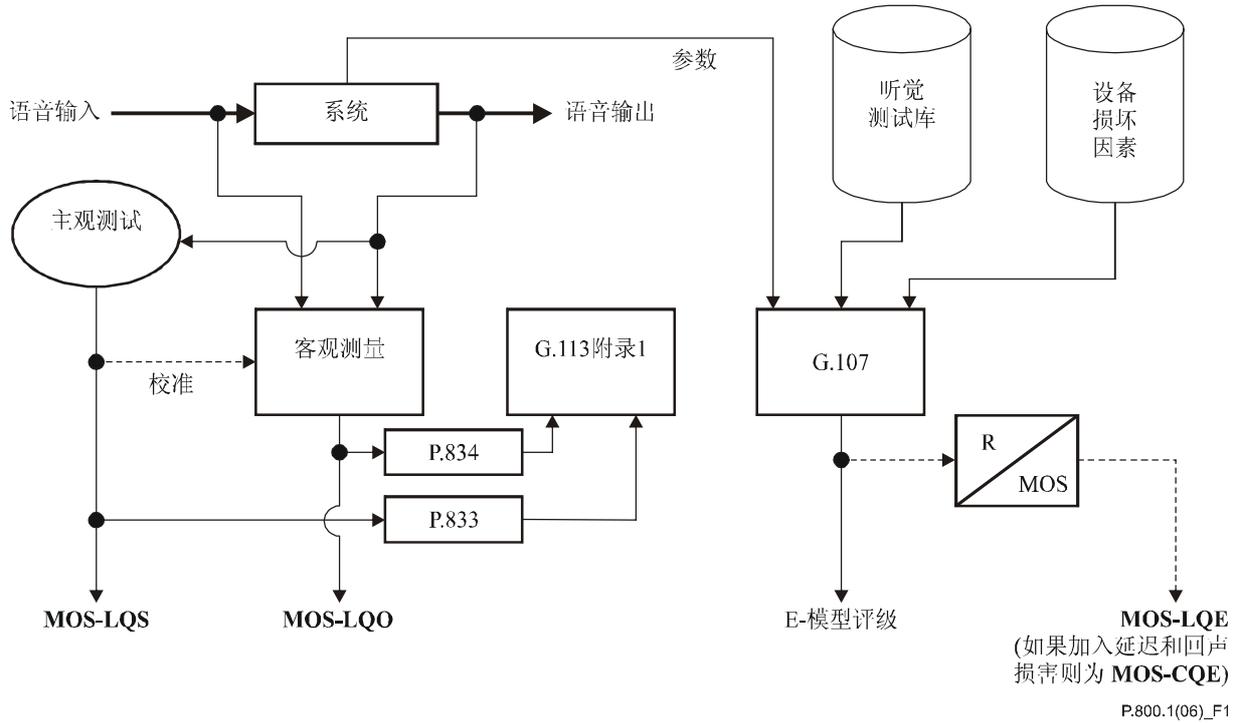


图1/P.800.1 – 某些MOS限定词间的关系

ITU-T 系列建议书

A系列	ITU-T工作的组织
D系列	一般资费原则
E系列	综合网络运行、电话业务、业务运行和人为因素
F系列	非话电信业务
G系列	传输系统和媒质、数字系统和网络
H系列	视听和多媒体系统
I系列	综合业务数字网
J系列	有线网和电视、声音节目和其他多媒体信号的传输
K系列	干扰的防护
L系列	线缆的构成、安装和保护及外部设备的其他组件
M系列	电信管理，包括TMN和网络维护
N系列	维护：国际声音节目和电视传输电路
O系列	测量设备技术规程
P系列	电话传输质量、电话装置、本地线路网络
Q系列	交换和信令
R系列	电报传输
S系列	电报业务终端设备
T系列	远程信息处理业务的终端设备
U系列	电报交换
V系列	电话网上的数据通信
X系列	数据网和开放系统通信及安全
Y系列	全球信息基础设施、互联网的协议问题和下一代网络
Z系列	用于电信系统的语言和一般软件问题