

文章编号:1002-8684(2004)10-0031-05

自适应多速率宽带编码标准
AMR-WB 及应用

· 论文 ·

张 韬, 裴正定

(北方交通大学 信息所, 北京 100044)

【摘 要】 比较了现有的宽带语音编码标准, 重点介绍了自适应多速率的宽带语音压缩编码标准 AMR-WB, 分析了它的语音质量, 以及在未来的广泛应用。

【关键词】 AMR-WB 语音编码标准; 编解码器; MOS 值

【中图分类号】 TN912

【文献标识码】 A

Brief Introduction of Adaptive Multi-rate Wideband Standards AMR-WB

ZHANG Tao, QIU Zheng-ding

(Institute of Information Science, Northern Jiaotong University, Beijing 100044, China)

【Abstract】 In this paper, comparison among current standards of wideband speech compression coding are briefly introduced. New generation standards of wideband speech compression coding (AMR-WB) is emphasized and its speech quality is analyzed. Finally, its application in the future is expected.

【Key words】 AMR-WB speech coding standards; CODEC; MOS score

1 引言

公共电话交换网(PSTN)都有基带带宽的限制, 这对于增强语音质量是很大的障碍。因此, 目前大多数语音编码在窄带语音的前提下获得良好的性能, 即带宽一般在 200~3 400 Hz, 用 8 kHz 采样。但音频带宽若扩展为 50~7 000 Hz, 用 16 kHz 采样的宽带语音编码后,

就可提供更优质的语音质量, 尤其在语音的自然性方面, 宽带编码已超越了其它有线音频传输的质量。图 1 给出了窄带和宽带通话语音信号的声频能量图。图 2 是类似的未通话的声频图。在未通话期间, 超过 4 kHz 的能量被窄带语音过滤掉, 这就影响语音的清晰度, 如“s”和“f”就不易区分。在通话期间, 多数信号是低频的, 低于 200 Hz 的被过滤掉, 这会影响语音的自然性。

小, 纹波频率为原来的 3 倍左右, 于是把功放的输入及反馈回路中的电容加以改变提高功放整体频响, 在中小功率下, 该机的反应速度有所提高; 在大动态运行下, 发现同样的音乐负荷下, 改后多相电源的温升明显低于未改造的电源。

由此可见, 多相电源的优点是显而易见的。实际上多相电源的控制主从芯片在计算机上已有成熟的产品, 如 HIP6302CB + HIP6620BCB, HIP6301CB + HIP6602BCB, RT9327, IRU3007 等。这些芯片在当前计算机主板的多项电源中应用广泛。计算机中多相电源是低压大电流应用, 在大功率音频功放中只需改进采样、推动和末级电路提高输出电压电流就能为音乐系统所用。

参考文献

- [1] George Shuellin. 可扩展的灵活多相解决方案. 电子产品世界, 2003, (5): 45-47.
- [2] 严家耀. 宽调节开关电源可选择的一种电路结构. 电子设计与应用, 2003, (6): 81-82.
- [3] 阿凡. 主板供电技术面面观. 微型计算机, 2003, (15): 110-112.
- [4] 俞阿龙. 高效率可调复合开关电源. 电声技术, 2003, (5): 43-44.

作者简介

周平, 讲师, 主要从事物理学、电子技术以及音视频技术与器材方面应用与研究。

[收稿日期] 2004-05-15

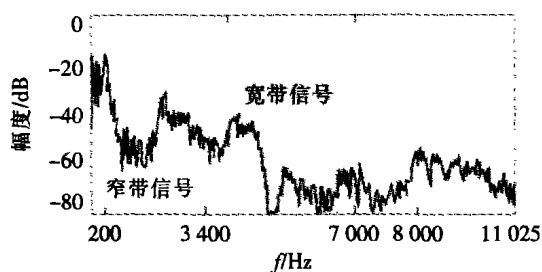


图1 窄带和宽带通话语音信号的声频能量图

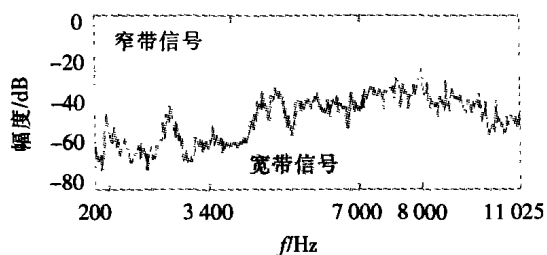


图2 未通话的声频图

为使数字电话达到自然语音的质量,人们在语音编码的标准化方面做了很多努力。宽带语音编解码器的标准化目标是提供一种高质量、低比特率的宽带语音编解码器,以适应无线和包交换网的需求。丰富的语音带宽可为使用者创造出亲近的效果。表1为宽带和窄带语音编码的MOS值比较。

表1 宽带和窄带语音编码的MOS值比较

编码	MOS值
GSM EFR(窄带;3.4 kHz)	3.30
G.722, 48 Kb/s(宽带;7 kHz)	4.06
G.722, 56 Kb/s(宽带;7 kHz)	4.57

宽带语音编码标准主要有以下几种:

(1) 16 bit linear PCM(16 bit 线性 PCM)

取样率为 16 kHz 的宽带语音,不经压缩的 16 bit 线性 PCM 编码比特率为 256 kb/s。这种编码提供了最高的质量和动态范围以及最低的计算的复杂度。

(2) G.711

G.711 是基于取样的编解码器,因此 G.711 PCM μ 律和 A 律压缩扩展可以和宽带 16 kHz 的语音取样共同使用。有效的比特率是 128 Kb/s。G.711 编解码器计算的复杂度很低,少于 2 个 DSP MIPS。算法的时延是 1 个取样周期(0.065 ms)。

(3) G.722

G.722 编解码器使用子带自适应微分脉冲编码调制(SB-ADPCM)将宽带语音编码为 64 Kb/s, 56 Kb/s 或 48 Kb/s。具体方法是将频带分成 2 个子带,每个子带用 ADPCM(与 G.726 类似)来编码。这种算法的计算复杂度适中,需要 10~12 MIPS。算法延时为 2 个取样周期(0.125 ms)。

(4) G.722.1

G.722.1 是一种低速率宽带编码,编码速率为 24 Kb/s 或 32 Kb/s。在 32 Kb/s 时的话音质量与 G.722 SB-ADPCM 在 64 Kb/s 时相同。G.722.1 使用了一种变换编码方案(Modulated Lapped Transform, MLT),帧长为 20 ms。这种算法的计算复杂度适中,需要 14 MIPS。算法时延为 40 ms (20 ms frame size+20 ms look-ahead)。

(5) AMR-WB

AMR-WB 是一种多速率编解码器,产生的比特率为 6.6~23.85 Kb/s。在更高的比特率时质量会更高。它使用多速率 ACELP 语音编码,帧长 20 ms。算法复杂度很高,约 38 MIPS。算法的时延 25 ms。它已经被 3GPP 标准化,是 16 Kb/s G.722 标准的候选标准。

(6) MPEG-4 Audio

MPEG-4 规范包括了多速率宽带 CELP 语音编码标准,速率范围为 13.2~24 Kb/s。MPEG-4 标准制定了广泛多样的技巧和模式,如可配置工具箱。而且译码表也可根据应用来设置。取样率为 16 kHz 时可使用 10 ms 或 20 ms 的帧长。这种算法的复杂度很高,时延随模式和比特率不同而从 18.75~41.75 ms 变化。MPEG-4 的 CELP 算法提供不同的比特率,以及比特率、带宽、复杂度的可测量性。

最后 3GPP 和 ITU-T 都选择了 AMR-WB 在比特率 6.6~23.85 Kb/s 作为新的宽带语音编码。对于 AMR-WB 来说,同时被 ITU-T 和 ETSI/3GPP 采用具有很大意义,因为这是第一次有一种编解码标准同时被选作有线和无线通用的标准。

笔者将具体介绍这种宽带语音编码标准 AMR-WB,语音质量及其特点,并给出它的广泛应用。

2 AMR-WB 标准

2.1 AMR-WB 语音编码的基本描述

AMR-WB 语音编码使用 ACELP 技术,这个技术用于 AMR-NB/FR 语音编码,也用于 ITU-TG.729 和 G.723.1 的 5.3 Kb/s 模式的语音编码中。AMR-WB 语音

编码包括 9 种比特率, 分别是 23.85 Kb/s, 23.05 Kb/s, 19.85 Kb/s, 18.25 Kb/s, 15.85 Kb/s, 14.25 Kb/s, 12.65 Kb/s, 8.85 Kb/s, 6.6 Kb/s。编码还包括一个背景噪声模式, 用于非连续转换 (DTX) 操作和一个低比特率资源模式, 用于在其它系统中编码背景噪声。在 GSM 中, 这个模式的比特率是 1.75 Kb/s。在 AMR-WB 语音编码中, 12.65 Kb/s 模式以及上面提到的模式可提供高质量的宽带语音。8.85 Kb/s 和 6.6 Kb/s 的 2 个模式只在网络繁忙或收音信道忙碌时使用。

AMR-WB 在 16 kHz 取样率的运作, 2 个频率带 50~6 400 Hz 和 6 400~7 000 Hz 进行编码, 用来降低复杂度, 将位算法集中到更重要的频率区。要注意的是低频带已超出了窄带电话学的范围。

低频带使用 ACELP 算法进行编码。添加几个特征来达到一个高的主观质量。线性预测(LP)算法是在每隔 20 ms 的帧要进行一次线性预测算法, 每 5 ms 搜索一次自适应码本。这个过程是在 12.8 Kbs 速率下进行的。

高频带是在解码器端使用低带和随机激励的参数重建的, 目的是调整与在声音基础上的低频有关的高频带。高频带的声频通过使用由低带 LP 过滤器产生的 LP 过滤器进行重建。

AMR-WB 语音编码的计算复杂度总共是 38.9 WMOps(加权百万次操作每秒)。这相当于最坏情况的复杂度, 复杂度包括 VAD, DTX 和 CGN 功能。AMR-WB 和 AMR-NB 计算复杂度示于表 2。

表 2 AMR-WB 和 AMR-NB 的计算复杂度

	计算复杂度	
	AMR-WB	AMR-NB
语音编码	31.1	14.2
语音解码	7.8	2.6
数据 RAM/K 字	6.5	5.3
数据 ROM/K 字	9.9	14.6

2.2 AMR-WB 的语音质量

AMR-WB 提供多种比特率, 使它在 2G 和 3G 系统中能适合多种应用。这种高语音质量也使它适合在宽带语音的应用, 因此, ITU-T 采用 AMR-WB。

图 3 是在 GSM 全速信道下的 AMR-WB 与 AMR-NB 比较的说明图。在特定环境下 ($C/I > 10$ dB), AMR-

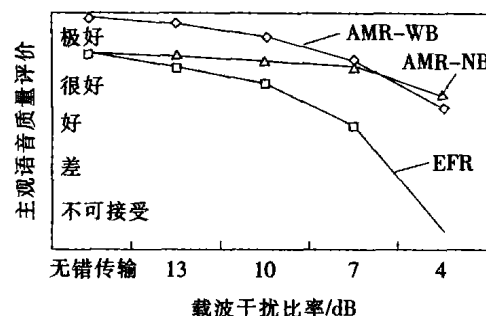


图 3 在 GSM 全速信道下 AMR-WB 与 AMR-NB 的比较

WB 比其它 GSM 编码质量都优越, 即使在较坏的情况下 ($C/I < 7$ dB), AMR-WB 仍比 AMR-NB 优越并远远超过 GSM 的质量。

在 AMR-WB 标准化形成的过程中, 编码要经历选择、验证和描述阶段。2000 年春季, 这个编码加入到 ITU-T Q.7/16 的选择阶段。

在 3GPP 的选择阶段, AMR-WB 编码经历 6 个独立实验, 有 5 种语言(英语、法语、日语、普通话和西班牙语)。测试内容包括输入级别、前后、背景噪声和 VAD/DTX 的性能。另外, 编码要在 GSM 和 WCDMA 3G 中通过不同的错误测试。AMR-WB 编码显示出良好的性能, 在所有的实验中都达到了要求。

在验证阶段, 对 AMR-WB 编码的不同情况进行测试来验证其性能, 包括 DTMF 音调和和其它特殊输入信号、过载性能、传送延迟、频率反映、复杂度分析和噪音总结。该编码在所有的实验中显示了良好的性能。

描述阶段进行包含 AMR-WB 的 9 种模式的测试, 使用 6 种语言(英语、芬兰语、法语、德语、日语和西班牙语)。实验包括输入级别、现实世界宽带和窄带的交叉操作、VAD/DTX、纯音、有背景噪声的语音、在 GSM 和 WCDMA 3G 的渠道错误以及分组交换应用的测试。

另外, AMR-WB 编码还在 ITU-T 选择阶段测试, 包括输入级别、前后、4 种背景噪音、帧磨损测试和窄带语音信号测试。使用语言是英语、法语、德语和日语。在 ITU-T 测试中, 只测试 12.65 Kb/s, 15.85 Kb/s, 19.85 Kb/s, 23.85 Kb/s 这一组模式。

AMR-WB 编码的质量在下面不同的测试阶段进行描述。

2.2.1 基本质量

AMR-WB 模式提供的在 23.85~14.25 Kb/s 的纯音

质量相当于或超过 ITU-T 宽带编码 G.722 在 64 Kb/s 的质量;12.65 Kb/s 模式至少相当于 G.722 在 56 Kb/s 的质量;85 Kb/s 模式相当于 G.722 在 48 Kb/s 的质量。

图 4 说明了 AMR-WB 编码的纯音质量。这是从 ITU-T 选择测试中选录出来的。图中说明了正常的信号级别-26 dB,单一信号的纯音编码(1T)和自我阻塞(2T)。12.65 Kb/s 和 23.85 Kb/s 模式是用法语进行的。ITU-T 宽带编码 G.722 中 48 Kb/s,56 Kb/s,64 Kb/s 被用做参考编码。结果显示 AMR-WB 12.65 Kb/s 模式超出 G.722 在 48 Kb/s 的质量,并与 G.722 在 56 Kb/s 的质量相媲美。AMR-WB 在 23.85 Kb/s 的性能相当于 G.722 在 64 Kb/s 的性能。上述观察结果在单一信号和自我阻塞中都有效。

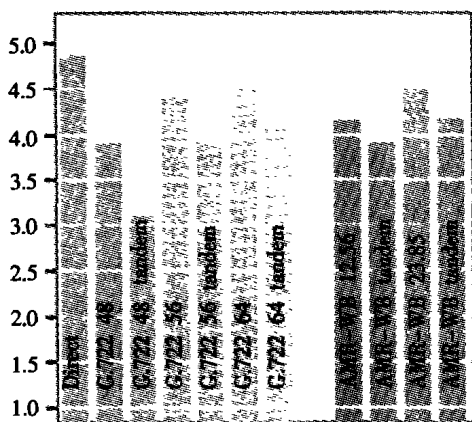


图 4 AMR-WB 编码的纯音质量

图 5 给出了 AMR-WB 编码的背景噪声性能。这是从 ITU-T 用英语进行的实验中抽取出来的。实验中包含 AMR-WB 模式 12.6 Kb/s,15.85 Kb/s,19.85 Kb/s,

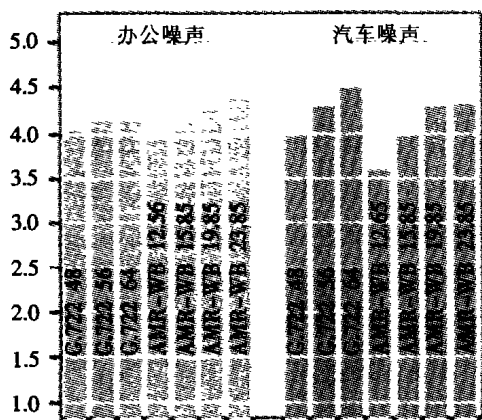


图 5 AMR-WB 编码的背景噪声性能

23.85 Kb/s。办公和汽车噪声都在 SNR15 dB。对于办公噪声来说,在 AMR-WB 最低模式 12.65 Kb/s 下性能相当于 G.722 在 48 Kb/s 的性能。而其它模式相当于或好于 G.722 在 64 Kb/s 的性能。对于汽车噪声来说,2 个最高模式的性能相当于 G.722 在 56 Kb/s 的性能。15.85 Kb/s 模式可与 G.722 在 48 Kb/s 的性能相媲美。

2.2.2 单元环境

AMR-WB 编码开发用于移动系统中,包括信道错误和背景噪声。在这些条件下,提供了良好的性能。在好的操作环境条件下,C/I 在 13 dB 以上。C/I 在 6 dB 以下是不好的操作环境的反映。这种情况下收音信道的位错误率较高。

在 GSM 全速信道纯音条件下,AMR-WB 提供了相当于或超过 G.722 在 64 Kb/s 的质量,11 dB C/I 以上。有一个相当于 G.722 在 56 Kb/s 的质量,10 dB C/I 以上。在背景噪声(15 dB 汽车噪声和 20 dB 办公噪声)下,AMR-WB 在 GSM 全速信道提供的质量相当于或超过 G.722 在 56 Kb/s 的质量,C/I 率在 10 dB 以上。

图 6 显示了 AMR-WB 在 GSM 全速信道下的性能,15 dB 的汽车背景噪声。这个实验是在 3GPP 描述阶段用英语进行的。图中显示了每个模式的性能。因为不是 GSM 的全速应用的目的,所以 AMR-WB 的 2 个最高位率不包含在实验中。在 3GPP 的描述测试中,VAD/DTX/CNG 的操作对听众来说是透明的。

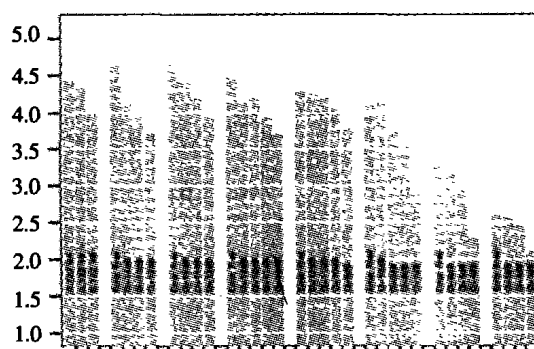


图 6 AMR-WB 在 GSM 全速信道下的性能

2.3 AMR-WB 编解码器的特点

(1)一组固定速率编码模式,可根据网络拥塞程度进行调整,这就确保增强了 QoS。

(2)由于带宽范围为 50~7 000 Hz,AMR-WB 不仅使语音更加易懂和自然,而且提供了一种透明通信的

感觉使通话者更易于识别。

(3) 话音自适应检测功能。

(4) 在 GSM 中使用的不连续传送 (DTX) 功能和在 3G 中使用的信源控制速率 (SCR) 功能。

(5) 带内信令 In-band signaling for codec mode transmission。

(6) 有线和无线应用的桥梁, 不同的有线和无线应用之间具有协作性。

(7) 尽管是在现存的窄带标准上规定的高质量语音, 但因为 AMR-WB 有多速率处理和自适应性, 所以对于传输错误还是有很强的鲁棒性。

(8) 对于 AMR-WB 来说, 同时被 ITU-T 和 ETSI/3GPP 采用具有很大意义, 因为这是第一次有一种编解码标准同时被选做有线和无线通用的标准。

3 AMR-WB 标准的应用

宽带语音编码的自然性对于电话的高保真度和远程电子通信有非常重要的意义, 因此具有广泛的应用。

(1) 第三代移动通信系统。3G 的一个主要目标就是传递多媒体信息, 并且更加强调语音的质量。即使在 3G 声音电话学应用领域中, 宽带语音编码的选择也是无线业务提供商作出超过传统的 PSTN 语音质量的重要一步。3GPP 认识到宽带语音的重要性, 最终选择了 AMR-WB。

(2) 视频和声频电子会议。宽带语音编码改进视频和声频电子会议的音质, 通过 ISDN 和基于分组网络提供更加现实的声音背景。

(3) 因特网应用。宽带电话能够加强好几种因特网应用, 如广播、聊天、虚拟现实环境、多媒体实时协作工具、叙述文的归档和分配以及网络语言学习程序。

(4) 数字化收音广播。宽带语音可应用于 AM 收音广播、新闻的视听广播、TV 节目和封闭循环演讲。

(5) 基于分组网络的宽带电话业务。基于分组网络 (VoPN) 的语音应用也引起了注意。不久的将来, 多数声音交流将通过分组网络如 IP, ATM, 帧延迟和 DSL 进行传输。VoPN 很容易引进宽带电话, 呈现出比传统的窄带电话网络领先的优势。

在分组交换网络中, 编码的健壮性十分重要。数字网络世界仍然将编码看作黑盒子。几年来, 虽然一直许诺 VoIP 的到来, 但是到现在仍未实现。服务供应商不愿接纳 VoIP 的主要原因就是服务质量。为使 VoIP 腾飞, VoIP 必须更加便宜, 而且有相当于或超过 PSTN 的

质量。网络专家试着通过改进路由器和网关的智能性来解决这个问题, 但这使解决方案更加昂贵, 资源管理效果更差。解决方法是对编码黑盒子智能化, 保证良好的 QoS 的关键主要是解决包组丢失和激励问题。在分组网络中可使用宽带语音, 就可以超过 PSTN 的质量。

宽带语音编码可以有效进行错误封闭和机动管理, 这是 VoIP 服务的竞争法码。随着远程通信市场的发展, 新的服务供应商需要这个解决方法来与 PSTN 竞争。服务供应商会以低成本、高速数据服务提供宽带电话。

4 结束语

分析了宽带语音编码相对窄带语音编码在语音自然性等方面的优势。宽带语音编码主要在主观评价上提高了语音的质量, 与窄带电话语音相比, 低频增加的 50~200 Hz 使语音更自然、更舒适、更有现场感。高频增加的 3 400~7 000 Hz 使摩擦音更易于区分, 从而增加了可懂度和清晰度。比较了现有的宽带语音编码标准, 重点介绍了新一代的宽带语音编码标准 AMR-WB。对于 AMR-WB 来说, 同时被 ITU-T 和 ETSI/3GPP 采用具有很大意义, 因为这是第一次有一种编解码标准同时被选作有线和无线通用的标准, 在未来一定会有广泛应用。

参考文献

- [1] (美) Ken C. Pohlmann 著. 数字音频原理与应用. 苏菲译. 北京: 电子工业出版社, 2002.
- [2] (美) Marcus Goncalves 著. IP 网络语音技术, 王文东, 白刚等译. 北京: 机械工业出版社, 1999.
- [3] On Wideband Codec Performance. Source: Nokia, ETSI SMG11#10, 1999-05.
- [4] AMR Wideband Speech Coding Development - Proposal for Design Constraints. Source: Nokia, ETSI SMG11#9, 1999-01.
- [5] ITU-T Subjective Selection Test Plan for the ITU-T Wideband (7 kHz) Speech Coding Algorithm 24 kb/s and 32 kb/s. Version 2.1. 1998-12.
- [6] G-series Recommendations. G.700 to G.799. <http://www.itu.int/itudoc/itu-t/rec/g/g700-799/index.html>.

作者简介

张韬, 北方交通大学信息所硕士生研究生。

袁正定, 北方交通大学信息所教授, 博士生导师。

【收稿日期】2004-05-12