

# AMR-WB+中帧错误隐藏算法的优化

周凤<sup>1</sup>, 吴云<sup>2</sup>, 赵勇<sup>2</sup>

(1. 贵州大学计算机科学与技术学院, 贵阳 550025; 2. 北京大学深圳研究生院, 深圳 518055)

**摘要:** AMR-WB+是应用于第3代移动通信系统的最新音频压缩编解码标准。分析 AMR-WB+中基音延迟参数的帧错误隐藏算法, 给出算法优化方案。实验结果表明, 优化后的基音延迟参数的帧错误隐藏算法在平均情况下, 客观测试得分高于原算法, 主观测试得分与原算法相当。

**关键词:** AMR-WB+标准; 音频编码; 错误隐藏; 算法优化

## Optimization of Frame Error Concealment Algorithm in AMR-WB+

ZHOU Feng<sup>1</sup>, WU Yun<sup>2</sup>, ZHAO Yong<sup>2</sup>

(1. College of Computer Science and Technology, Guizhou University, Guiyang 550025;

2. Shenzhen Graduate School, Peking University, Shenzhen 518055)

**【Abstract】** AMR-WB+ is the newest audio compression standard which is used in 3G mobile communication systems. This paper analyzes the frame error concealment algorithm for pitch delay parameter in AMR-WB+, and gives the scheme of optimization. Experimental results show that, on average, the objective test score of the optimized algorithm is higher than the original algorithm, and the subjective test score of the optimized algorithm is equal to the original algorithm.

**【Key words】** AMR-WB+ standard; audio encoder; error concealment; algorithm optimization

### 1 概述

近年来随着移动多媒体技术及无线网络接入技术的日益成熟, 移动音频业务迅速增长。但由于移动无线传输信道的传输环境比较恶劣、误码率较高, 语音数据包在传输过程中难免会发生丢失。一旦数据包丢失, 解码端将无法正确解码, 就会产生噪声。

AMR-WB+<sup>[1]</sup>是对 AMR-WB<sup>[2]</sup>标准的升级, 它是目前适用于低码率下最优秀的音频编码标准。AMR-WB+混合了 ACELP/TCX 模型, 其中, ACELP(代数码本激励线性预测)适合对语音信号编码; TCX(变换编码激励)适合对音乐信号进行编码。为了提高编码质量, 避免或减少因语音包丢失而产生的噪声, AMR-WB+应用了错误隐藏技术。基音延迟参数的帧错误隐藏技术是 AMR-WB+中关于 ACELP 的错误隐藏技术, 这种技术延用了 AMR-WB 中的方法, 选择合适的参数替代那些丢失的数据包, 其目的就是解决上述的噪声问题。但是这种延用 AMR-WB 的技术在 AMR-WB+中已经不能完全适用了。

### 2 AMR-WB+编解码原理

AMR-WB+采用新型的编码技术框架, 融合当前语音和音频编码领域内最新的技术, 在编码时, AMR-WB+根据音频信号的内容, 自适应地选择选用 ACELP 编码或 TCX<sup>[3]</sup>编码, 并能在 2 种编码方式之间实现无缝切换。这种模型非常适合对语音、音乐和混合音频等复杂的音频信号编码, 满足了对移动音频中较为复杂音频信号编码的要求。AMR-WB+标准是 AMR-WB 的升级版本, 完全兼容 AMR-WB 标准。相对于 AMR-WB, AMR-WB+主要增加了 TCX 变换编码技术、频带扩展技术(BWE)以及立体声编码, 支持更高的采样速率

(立体声采样速率 8 kb/s~48 kb/s, 非立体声采样速率 6 kb/s~36 kb/s)。本文只针对单声道信号(mono signal)编码进行分析。

AMR-WB+语音编码器由多速率语音编码器, 包括端点检测、舒适噪声生成系统的源控制速率方案和对付传输错误和丢包后的错误隐藏机制组成。编码器整体框架如图 1 所示。

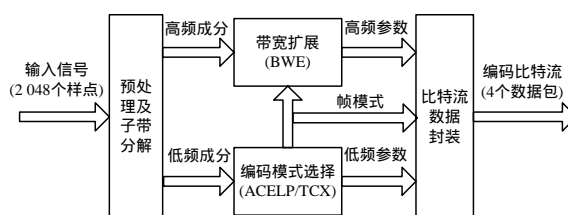


图1 编码器整体框架

ACELP 编码与 TCX 编码都是利用线性预测的方法(linear prediction), 对待编码信号利用 Levinson-Durbin 算法得到预测系数, 从而得到预测误差滤波器即分析滤波器, 对输入信号进行线性预测分析, 得到残差信号(residual signal)。对残差信号提取参数, 协同预测系数一起进行量化编码并传输或保存。在解码端由接收到的参数恢复出编码端的残差信号, 将此残差信号作为激励信号通过由预测系数构成的全极

**基金项目:** 贵州省科学技术基金资助项目“基于 DCT 变换域参考图像的频谱分级视频编码方法”(黔科合 J 字[2009]2124)

**作者简介:** 周凤(1976-), 女, 讲师、硕士, 主研方向: 音视频压缩, 数字图像处理; 吴云, 博士; 赵勇, 副教授、博士、博士生导师

**收稿日期:** 2009-05-04 **E-mail:** zhoufeng77@21cn.com

点滤波器即合成滤波器即可得到重构的语音或音乐信号。ACELP 与 TCX 的不同之处在于对残差信号编码的方法不一样,而之前的预处理、线性预测分析和感知加权基本一致。

解码基本是编码的逆过程,但为获得更佳的重构语音或乐音质量,解码时还采用了一些特殊的处理技术,如增益平滑、后置处理等。

### 3 AMR-WB+中基音延迟参数的帧错误隐藏算法

AMR-WB+中的错误隐藏技术<sup>[4]</sup>基本沿用了 AMR-WB 中的错误隐藏技术。但由于 AMR-WB+中结合了 ACELP 和 TCX 2 种编码模型,比较复杂,先前的错误隐藏技术不能完全适用,在 AMR-WB+中作了简单化处理<sup>[1]</sup>。

基音延迟参数的帧错误隐藏技术是 ACELP 中的错误隐藏技术。在标准 3GPP TS 26.290<sup>[5]</sup>中没有规定基音延迟参数的帧错误隐藏算法。但是在 AMR-WB+浮点参考代码 3GPP TS 26.304 中 AMR-WB+扩展出来的几种模式下,使用了如下简单算法<sup>[4-5]</sup>：

```

if (bfi) /* 发生帧错误 */
{
    old_T0_frac = old_T0_frac + 1; /* 将上一帧基音延迟增加
1/4, 即分数部分增加 1 */
    if (old_T0_frac > 3)
    {
        old_T0_frac = old_T0_frac - 4;
        old_T0 = old_T0 + 1;
    }
    if (old_T0 >= PIT_MAX)
    {
        old_T0 = PIT_MAX - 5;
    }
    T0 = old_T0;
    T0_frac = old_T0_frac;
}

```

其中, bfi 是帧错误标志; old\_T0\_frac 是上一帧的基音延迟参数的分数部分,精度为 1/4 样点,取值为{1, 2, 3}; PIT\_MAX 是基音延迟参数的上限,取值为 231。

AMR-WB+中的原算法是让基音延迟参数每次增加 1/4 个样点即最小精度。其基本思想基于平稳的语音信号基音延迟参数变化很缓慢的事实,并为了避免过度周期性,因此,每次只增加 1/4 个样点。另外,增加基音延迟参数意味着降低语音频率,因此,只是单调递增基音延迟参数也是出于避免或减少尖锐噪声的考虑。但该算法存在如下缺点：

(1)实际上音基音延迟参数并不是单调变化的,尤其是在清浊音转换时具有随机性,因此,每次增加 1/4 不能很好地模拟真实情况。

(2)WB+支持 ACELP/TCX 混合模型,它是在编码端分别用这 2 种编码方式编码后计算信噪比而选择合适的编码方式。即使是语音,也会有很多帧是采用 TCX 方式编码的。而该算法没有考虑到 ACELP/TCX 切换的影响,它所用到的只是先前的 ACELP 的基音延迟参数信息,如果当前帧与上一 ACELP 帧存在若干帧 TCX 帧,则基音延迟参数信息的相关性大大减弱。

基于对 AMR-WB+中基音延迟参数的帧错误隐藏算法的分析,下文提出一种优化的基音延迟参数的帧错误隐藏算法。

## 4 算法的优化

### 4.1 算法优化思想

(1)保存 TCX 帧中的周期信息。尽管传输的 TCX 参数中不含有周期信息,但是在解码端对每个 TCX 帧都会对周期做简略的估计,这主要是用来恢复丢失的 20 ms TCX 帧。

(2)在改变基音延迟参数前首先做方向变化的判断,即判断增加还是减小。

### 4.2 算法优化方案

(1)增加存储的基音周期的个数,由 1 个增至 7 个。用 2 个长度为 7 的一维数组分别保存基音延迟参数的整数部分和小数部分即可,最后一位为最新更新的值。初始化时值为 64。

(2)正常解码时,将数组里的基音延迟参数移位并将正确的基音延迟参数保存在最后一位。

(3)发生帧错误时,对保存的相邻的两两基音延迟参数值进行比较,若变化值为负(即变化趋势为减小)则计数器减 1;否则,计数器加 1。

(4)若计数器大于 0(增加的趋势较多),则由上一基音延迟参数值加上 1/4;否则,减去 1/4。

(5)对由(4)计算出的值进行上下限控制,若超过上限值(231)则减去 5;若超过下限值(34)则加上 5。

### 4.3 客观测试及结果

客观评测采用的评测软件为 ITU-T 的 P.862 客观评测标准 WB-PESQ。测试序列为 AVS\_M 提供的 12 个测试序列及 WB+的 12 个测试序列。测试序列均编解码为单声道 wav 波形文件,采样率为 16 kHz。帧错误序列是通过 ITU-T G.191 STL 工具包中的 gen-patt.exe 软件生成的二进制文件(0x6b21 代表正确帧,0x6b20 代表错误帧),测试中模拟的帧错误率有 2%, 3%, 4%, 6%和 10%。测试的时候,需要对 AMR-WB+参考代码加入少量代码,使其能够读取 gen-patt.exe 生成的帧错误序列。

由于同时存在 ACELP 和 TCX 编码,不同的模拟丢包序列可能会产生不同的实验结果,因此需要对大量不同的模拟丢包序列进行测试才能体现出统计意义,这里限于篇幅,只给出其中测试序列的平均测试值。如表 1 所示,在平均情况下,在帧错误率为 2%, 3%, 4%, 6%和 10%时,优化算法客观测试平均得分均高于 AMR-WB+原算法。

表 1 2 种编码速率下的平均值客观测试结果

丢包率 /(%)	编码速率为 12 Kb/s 情况下		编码速率为 24 Kb/s 情况下	
	WB+原算法客观 测试平均分	优化算法客观 测试平均分	WB+原算法客观 测试平均分	优化算法客观 测试平均分
2	2.905 167	2.906 042	3.218 024	3.219 542
3	2.806 042	2.806 708	3.067 125	3.067 917
4	2.887 917	2.890 510	3.189 625	3.191 208
6	2.699 792	2.702 417	2.944 625	2.947 542
10	2.466 125	2.467 208	2.634 417	2.634 958
总平均分	2.753 009	2.753 975	3.010 763	3.011 833

#### 4.4 主观测试及结果

本文使用CMOS(Comparison Mean Opinion Score)方法对优化算法和原算法进行比较。每次主观听力测试材料由Ref/A/B组成,Ref是原始的未编码的信号,如果A是优化算法的解码结果,那么B就是原算法的解码结果,反之亦然。

具体测试方案如下:每个测试项目中Ref信号的位置是固定的,对A,B位置的分配是随机的,且对测听者来说是未知的。在高质量测听打分中使用7级比较等级,如表2所示。

表2 等级比较评分准则

比较结果	等级
A 远好于 B	+3.0
A 好于 B	+2.0
A 稍好于 B	+1.0
A 与 B 一致	0.0
B 稍好于 A	-1.0
B 好于 A	-2.0
B 远好于 A	-3.0

测听者的打分必须为表中所列的整数;打分的原则是较为接近原始信号的被测方案得分应较高,根据接近原始信号的程度来决定具体的分值;测听结果由平均得分和95%置信区间范围组成,在统计方差时计入所有测听者的测听结果;测试序列为AVS\_M的12个序列及WB+的12个序列,采样率为16kHz,用4%模拟帧错误率进行解码。编码速率为12Kb/s时的测试结果如图2所示。编码速率为24Kb/s时的测试结果如图3所示。

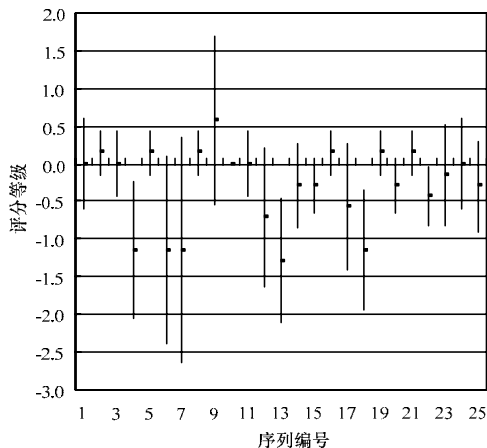


图2 编码速率为12Kb/s时的测试结果

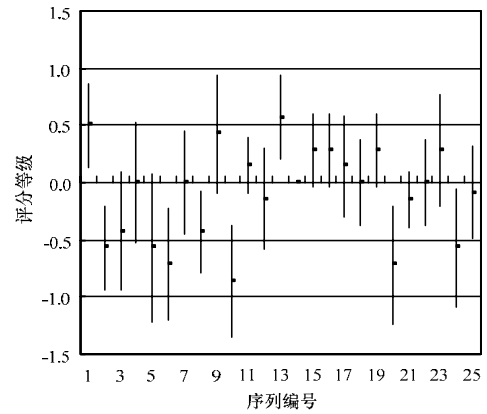


图3 编码速率为24Kb/s时的测试结果

在图2、图3中,横坐标表示测试序列的编号1~24,第25条直线是24个测试序列的平均;纵坐标表示评分等级3.0~-3.0,每一条直线上的点代表均值,顶点是均值加,底点是均值减。可以看出,在12Kb/s和24Kb/s的编码码率下,所采用的优化基音延迟参数算法和AMR-WB+原算法的主观质量相当。

#### 5 结束语

本文介绍了AMR-WB+音频编解码原理,分析了AMR-WB+中关于ACELP的用于中基音延迟参数的帧错误隐藏算法,在此基础上提出对原基音延迟参数的帧错误隐藏算法的优化方案。AMR-WB+标准是目前用于移动多媒体的最先进的音频编码标准,具有较高的实用性,但目前国内的相关研究还不多,尤其是TCX编码模块,因此,下一步将对其进行全面分析和深入研究。

#### 参考文献

- [1] Makien J, Bessette B, Bruhn S, et al. AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services[C]// Proc. of the International Conference on Acoustics, Speech, and Signal Processing. Philadelphia, USA: [s. n.], 2005.
- [2] 焦传斌, 于保华, 李治柱. 新型宽带语音编解码器 AMR-WB 的研究[J]. 计算机仿真, 2005, 22(1): 150-152.
- [3] Lefebvre R, Salami R, Laflamme C, et al. High Quality Coding of Wideband Audio Signals Using Transform Coded Excitation (TCX)[C]//Proc. of the International Conference on Acoustics, Speech, and Signal Processing. Montreal, Canada: [s. n.], 1994.
- [4] 3GPP TS 26.191-2007 Adaptive Multi-Rate-Wideband(AMR-WB) Speech Codec: Error Concealment of Erroneous or Lost Frames[S]. 2007.
- [5] 3GPP TS 26.290-2007 Extended AMR Wideband Codec: Transcoding Functions[S]. 2007.

编辑 顾姣健