



(12)发明专利

(10)授权公告号 CN 103337243 B

(45)授权公告日 2017.02.08

(21)申请号 201310272820.1

(22)申请日 2013.06.28

(65)同一申请的已公布的文献号
申请公布号 CN 103337243 A

(43)申请公布日 2013.10.02

(73)专利权人 大连理工大学
地址 116024 辽宁省大连市高新园区凌工
路2号

(72)发明人 陈喆 殷福亮 李文月

(74)专利代理机构 大连东方专利代理有限责任
公司 21212
代理人 曲永祚 李洪福

(51)Int.Cl.
G10L 19/24(2013.01)
G10L 19/12(2013.01)

(56)对比文件

JP 特开2001-100773 A,2001.04.13,
WO 02/056295 A2,2002.07.18,
CN 1514559 A,2004.07.21,
CN 101359474 A,2009.02.04,
段盼爽 等.人工语音带宽扩展研究.《中国
优秀硕士学位论文全文数据库》.2009,(第4期),

审查员 赵云峰

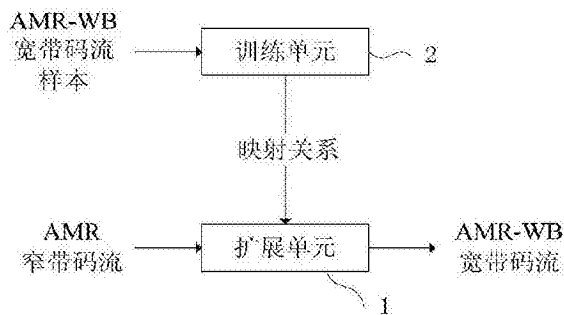
权利要求书4页 说明书23页 附图8页

(54)发明名称

一种AMR码流转换成AMR-WB码流的方法

(57)摘要

本发明公开了一种AMR码流转换成AMR-WB码流的方法,属于编码技术领域。一种AMR码流转换成AMR-WB码流的方法;AMR窄带码进入扩展单元后转换成AMR-WB码流,扩展单元和训练单元,所述的训练单元为扩展单元提供参数扩展过程所需的映射关系。



1. 一种AMR码流转换成AMR-WB码流的方法,其特征在于:AMR窄带码进入扩展单元后转换成AMR-WB码流,扩展单元和训练单元,所述的训练单元为扩展单元提供参数扩展过程所需的映射关系;

所述的扩展单元包括AMR解码单元、参数提取单元、窄带能量计算单元、SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B、升采样单元以及AMR-WB部分编码单元,所述的AMR解码单元的输入端输入AMR的窄带码流,输出端与参数提取单元、窄带能量计算单元以及升采样单元的输入端相连,所述的参数提取单元的输入端与AMR解码单元的输出端相连,其输出端与SVR预测单元、函数映射单元A、码本映射单元以及AMR-WB部分编码单元的输入端相连;所述的窄带能量计算单元的输入端与AMR解码单元的输出端相连,其输出端与函数映射单元B的输入端相连,所述的SVR预测单元、函数映射单元A以及码本映射单元的输入端与参数提取单元的输出端相连并接收训练单元提供的映射关系,它们的输出端都与AMR-WB部分编码单元的输入端相连,所述的函数映射单元B的输入端与窄带能量计算单元的输出端相连接接收训练单元提供的映射函数,其输出端与AMR-WB部分编码单元的输入端相连,所述的升采样单元的输入端与AMR解码单元的输出端相连,其输出端与AMR-WB部分编码单元的输入端相连,AMR-WB部分编码单元的输入端与SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B以及升采样单元的输出端相连,其输出端输出AMR-WB宽带码流。

2. 根据权利要求1所述的一种AMR码流转换成AMR-WB码流的方法,其特征在于:所述的AMR解码单元包括窄带码流分离单元、LSP解码单元、自适应码本解码单元、增益解码单元、固定码本解码单元、4子帧插值单元、激励重建单元、LSP到 $A(z)$ 转换单元、合成滤波器单元以及后置滤波器单元,所述的窄带码流分离单元的输入端输入AMR窄带码流,其输出端分别与LSP解码单元、自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端相连;所述的LSP解码单元的输入端与码流分离单元的输出端相连、其输出端与4子帧插值单元的输入端相连;所述的自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端都与码流分离单元的输出端相连、他们的输出端都与激励重建单元的输入端相连,所述的4子帧LSP插值单元的输入端与LSP解码单元的输出端相连,其输出端与LSP到 $A(z)$ 转换单元的输入端相连,所述的激励重建单元的输入端分别与自适应码本解码单元、增益解码单元以及固定码本解码单元的输出端相连;所述的LSP到 $A(z)$ 转换单元的输入端与4子帧LSP插值单元的输出端相连,其输出端与合成滤波器单元的输入端相连,所述的合成滤波器单元的输入端分别与激励重建单元和LSP到 $A(z)$ 转换单元的输出端相连,其输出端与后置滤波器单元的输入端相连,所述的后置滤波器单元的输入端与合成滤波器单元的输出端相连,其输出单元输出合成语音。

3. 根据权利要求1所述的一种AMR码流转换成AMR-WB码流的方法,其特征在于:所述的参数提取单元包括VAD提取单元、LSP提取单元、开环基音周期及固定码本提取单元,所述的VAD提取单元的输入端与AMR解码单元的输出端相连,其输出端与AMR-WB部分编码的输入端相连,所述的LSP提取单元的输入端与AMR解码单元的输出端相连,其输出端与SVR预测单元的输入端相连,所述的开环基音提取单元的输入端与AMR解码单元的输出端相连,其输出端与映射单元A的输入端相连,所述的固定码本单元的输入端与AMR解码单元的输出端相连,其输出端与码本映射单元的输入端相连。

4. 根据权利要求1所述的一种AMR码流转换成AMR-WB码流的方法,其特征在于:所述的

AMR-WB部分编码单元包括加权语音计算单元、4子帧差值单元A、ISP到ISF转换单元、开环基音搜索单元、闭环基音搜索单元、自适应码本计算单元、4子帧差值单元B、ISF量化单元、自适应码本贡献计算单元、自适应滤波器选择单元、固定码本目标信号计算单元、固定码本搜索单元、增益矢量量化单元、脉冲响应计算单元以及AMR-WB码流生成单元；所述的加权语音计算单元的输入端输入升采样之后的AMR合成语音和VAD并与4子帧插值单元A的输出端相连，其输出端与开环基音搜索单元的输入端相连；所述的4子帧插值单元A的输入端与输入16维的ISP，其输出端分别与加权语音计算单元、自适应码本计算单元以及脉冲响应计算单元的输入端相连；所述的ISP到ISF转换单元的输入端输入16维的ISP，其输出端与ISF量化单元的输入端相连；所述的ISF量化单元的输入端与ISP到ISF转换单元的输出端相连，其输出端分别与4子帧插值单元B和AMR-WB码流生成单元的输入端相连；所述的开环基音搜索单元的输入端接收扩展得到的开环基音并与加权语音的输出端相连，其输出端与闭环基音搜索单元的输入端相连；所述的4子帧差值单元B的输入端与ISF量化单元的输出端相连，其输出端分别与自适应码本信号计算单元和脉冲响应计算单元的输入端相连；所述的自适应码本计算单元的输入端输入升采样后的AMR合成语音并与4子帧插值单元A的输出端相连，其输出端与固定码本目标信号计算单元的输入端相连；所述的闭环基音搜索单元的输入端与自适应码本计算单元的输出端相连，其输出端分别与自适应码本贡献计算单元和AMR-WB码流生成单元的输入端相连；所述的自适应码本贡献计算单元的输入端与闭环基音搜索单元的输出端相连，其输出端分别与自适应滤波器选择单元和增益矢量量化单元的输入端相连；所述的增益矢量量化单元的输入端分别与自适应码本贡献计算单元和固定码本搜索单元的输出端相连，其输出端与AMR-WB码流生成单元的输入端相连；所述的自适应滤波器选择单元的输入端自适应码本贡献计算单元的输出端相连，其输出端分别与固定码本目标信号计算单元和AMR-WB码流生成单元的输入端相连；所述的固定码本计算单元的输入端输入扩展得到宽带固定码本并分别与自适应码本目标信号计算单元和自适应滤波器选择单元的输出端相连，其输出端与固定码本搜索单元的输入端相连；所述的固定码本搜索单元的输入端分别与固定码本目标信号计算单元和脉冲响应计算单元的输出端相连，其输出端分别与增益矢量量化单元和AMR-WB码流生成单元的输入端相连；所述的AMR-WB码流生成单元的输入端接收扩展得到高频增益索引并分别与固定码本搜索单元、自适应滤波器选择单元、增益矢量量化单元、闭环基音搜索单元以及ISF量化单元的输出端相连，其输出端输出AMR-WB宽带码流。

5. 根据权利要求1所述的一种AMR码流转换成AMR-WB码流的方法，其特征在于：所述的训练单元包括窄带码流分离单元、窄带码流解析单元、AMR-WB编码单元、SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元；所述的窄带码流分离单元的输入端输入窄带码流，其输出端与窄带码流解析单元的输入端相连；所述的窄带码流解析单元的输入端与窄带码流分离单元的输出端相连，其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连；所述的AMR-WB编码单元的输入端输入宽带语音，其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连；所述的SVR训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输出端相连，其输出端输出SVR映射模型；所述的开环基音映射函数

训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输出端相连,其输出端输出开环基音映射函数;所述的固定码本映射码本训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出映射码本;所述的高频增益映射函数训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出高频增益映射函数。

6. 根据权利要求5所述的一种AMR码流转换成AMR-WB码流的方法,其特征在于:所述的AMR-WB编码单元包括预处理单元、线性预测分析单元、ISP量化单元、4子帧ISP插值单元A、加权语音计算单元、4子帧ISP插值单元B、开环基音搜索单元、目标信号计算单元、最佳基音延迟和增益搜索单元、自适应码本分量计算单元、自适应码本滤波器选择单元、冲激响应计算单元、高频增益索引计算单元、固定码本搜索单元、滤波器更新单元、激励计算单元以及增益量化单元;所述的预处理单元的输入端输入采样率为16KHz的宽带语音,其输出端分别与线性预测分析单元、加权语音计算单元以及目标信号计算单元的输入端相连;所述的线性预测分析单元的输入端与预处理单元的输出端相连、其输出端分别与ISP量化单元和4子帧ISP插值单元B的输入端相连;所述的ISP量化单元的输入端与线性预测分析单元的输出端相连,其输出端与4子帧ISP差值单元A的输入端相连;所述的4子帧插值单元A的输入端与ISP量化单元的输出端相连,其输出端与冲激响应计算单元的输入端相连;所述的加权语音计算单元的输入端分别与预处理单元和四子帧ISP插值单元B的输出端相连,其输出端与开环基音搜索单元的输入端相连;所述的4子帧插值单元B的输入端与线性预测分析单元的输出端相连,其输出端分别与目标信号计算单元、加权语音计算单元以及冲激响应计算单元的输入端相连;所述的开环基音搜索单元的输入端与加权语音计算单元的输出端相连,其输出端与最佳基音延迟和增益搜索单元的输入端相连;所述的目标信号计算单元的输入端分别与预处理单元、4子帧ISP插值单元B以及4子帧ISP插值单元A的输出端相连,其输出端分别与固定码本搜索单元和最佳基音延迟和增益搜索单元的输入端相连;所述的最佳基音延迟和增益搜索单元的输入端分别与目标信号计算单元、开环基音搜索以及冲激响应计算单元的输出端相连,其输出端输出基音索引并与自适应码本贡献计算单元的输入端相连;所述的自适应码本贡献计算单元的输入端与最佳基因延迟和增益上搜索单元的输出端相连,其输出端分别与自适应码本滤波器选择单元和增益量化单元的输入端相连;所述的自适应码本滤波器选择单元的输入端与自适应码本贡献计算单元的输出端相连,其输出端输出滤波器索引并与脉冲响应计算单元的输入端相连;所述的冲激响应计算单元的输入端分别与自适应码本滤波器选择单元、4子帧ISP插值单元A以及4子帧ISP插值单元B的输出端相连,其输出端分别与最佳基音延迟和增益搜索单元以及固定码本搜索单元的输入端相连;所述的固定码本搜索单元的输入端分别与目标信号计算单元、自适应码本滤波器选择单元以及冲激响应计算单元的输出端相连,其输出端输出固定码本增益索引并与增益量化单元的输入端相连;所述的增益量化单元的输入端分别与固定码本搜索单元以及自适应码本贡献计算单元的输出端相连,其输出端输出增益索引并与激励计算单元的输入端相连;所述的激励计算单元的输入端与增益量化单元的输出端相连,其输出端分别与滤波器状态更新单元和高频增益索引计算单元的输入端相连;所述的滤波器状态更新单元的输入端与激励计算单元的输出端相连;所述的高频增益索引计算单元的输入端输入采样率为16KHz的宽带语音并分别与4子帧ISP插值单元和激励计算单元的输出端相连,其输出端输出高频增益

索引。

一种AMR码流转换成AMR-WB码流的方法

技术领域

[0001] 本发明涉及一种AMR码流转换成AMR-WB码流的方法,属于编码技术领域。

背景技术

[0002] 在很多通信系统中,如公共电话交换网(PSTN)和全球移动通信系统(GSM),系统传输的语音带宽被限制在4KHz以内。虽然4KHz窄带语音可满足基本的通信要求,但在一些对音质要求较高的场合,如会议电视系统等,由于4KHz窄带语音缺乏高频成分,其听起来“闷闷”的,自然度和可懂度均较低,已不能满足音质方面的需求。这些应用上的需求促使人们开始关注宽带语音编码技术,并相继提出了一些宽带编码标准,如AMR-WB[和G.729.1等。但这些宽带编码标准未考虑对已有网络通信协议的兼容,即对编码速率和码流格式改变都较大,很难直接应用于既有网络中。长时间建设而成的既有通信网络极其庞杂,因而网络的升级必然是一个复杂、渐进的过程,因此,短时间内对既有通信网络进行全面升级是不现实的,如何在现有通信网络条件下,获得宽带语音质量就成为一个亟待解决的问题。为此,人们提出了人工语音带宽技术,所谓的人工带宽扩展就是借助语音信号处理方法,扩展出窄带语音的其它频带分量,进而合成出宽带语音。早在1933年,人们就提出了语音带宽扩展概念并尝试通过线性运算实现该技术。后来在上世纪70年代的早期,开始有公司试图通过数字信号处理技术来重建宽带语音信号。但当时没考虑声音特性和人耳听觉特点,早期的尝试都以失败告终。直到70年代末期,学者们提出语音的线性预测模型,才使语音带宽扩展技术取得突破性的进展,相继提出了多种带宽扩展算法。

发明内容

[0003] 本发明针对以上问题的提出,而研制一种AMR码流转换成AMR-WB码流的方法。

[0004] 本发明的技术手段如下:

[0005] 一种AMR码流转换成AMR-WB码流的方法;AMR窄带码进入扩展单元后转换成AMR-WB码流,扩展单元和训练单元,所述的训练单元为扩展单元提供参数扩展过程所需的映射关系。

[0006] 所述的扩展单元包括AMR解码单元、参数提取单元、窄带能量计算单元、SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B、升采样单元以及AMR-WB部分编码单元,所述的AMR解码单元的输入端输入AMR的窄带码流,输出端与参数提取单元、窄带能量计算单元以及升采样单元的输入端相连,所述的参数提取单元的输入端与AMR解码单元的输出端相连,其输出端与SVR预测单元、函数映射单元A、码本映射单元以及AMR-WB部分编码单元的输入端相连。所述的窄带能量计算单元的输入端与AMR解码单元的输出端相连,其输出端与函数映射单元B的输入端相连,所述的SVR预测单元、函数映射单元A以及码本映射单元的输入端与参数提取单元的输出端相连并接收训练单元提供的映射关系,它们的输出端都与AMR-WB部分编码单元的输入端相连,所述的函数映射单元B的输入端与窄带能量计算单元的输出端相连接接收训练单元提供的映射函数,其输出端与AMR-WB部分编码单元的输入端相

连,所述的升采样单元的输入端与AMR解码单元的输出端相连,其输出端与AMR-WB部分编码单元的输入端相连,AMR-WB部分编码单元的输入端与SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B以及升采样单元的输出端相连,其输出端输出AMR-WB宽带码流。

[0007] 所述的AMR解码单元包括窄带码流分离单元、LSP解码单元、自适应码本解码单元、增益解码单元、固定码本解码单元、4子帧插值单元、激励重建单元、LSP到 $A(z)$ 转换单元、合成滤波器单元以及后置滤波器单元,所述的窄带码流分离单元的输入端输入AMR窄带码流,其输出端分别约LSP解码单元、自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端相连。所述的LSP解码单元的输入端与码流分离单元的输出端相连、其输出端与4子帧插值单元的输入端相连。所述的自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端都与码流分离单元的输出端相连、他们的输出端都与激励重建单元的输入端相连,所述的4子帧LSP插值单元的输入端与LSP解码单元的输出端相连,其输出端与LSP到 $A(z)$ 转换单元的输入端相连,所述的激励重建单元的输入端分别与自适应码本解码单元、增益解码单元以及固定码本解码单元的输出端相连。所述的LSP到 $A(z)$ 转换单元的输入端与4子帧LSP插值单元的输出端相连,其输出端与合成滤波器单元的输入端相连,所述的合成滤波器单元的输入端分别与激励重建单元和LSP到 $A(z)$ 转换单元的输出端相连,其输出端与后置滤波器单元的输入端相连,所述的后置滤波器单元的输入端与合成滤波器单元的输出端相连,其输出单元输出合成语音。

[0008] 所述的参数提取单元包括VAD提取单元、LSP提取单元、开环基音周期及固定码本提取单元,所述的VAD提取单元的输入端与AMR解码单元的输出端相连,其输出端与AMR-WB部分编码的输入端相连,所述的LSP提取单元的输入端与AMR解码单元的输出端相连,其输出端与SVR预测单元的输入端相连,所述的开环基音提取单元的输入端与AMR解码单元的输出端相连,其输出端与映射单元A的输入端相连,所述的固定码本单元的输入端与AMR解码单元的输出端相连,其输出端与码本映射单元的输入端相连。

[0009] 所述的AMR-WB部分编码单元包括加权语音计算单元、4子帧差值单元A、ISP到ISF转换单元、开环基音搜索单元、闭环基音搜索单元、自适应码本计算单元、4子帧差值单元B、ISF量化单元、自适应码本贡献计算单元、自适应滤波器选择单元、固定码本目标信号计算单元、固定码本搜索单元、增益矢量量化单元、脉冲响应计算单元以及AMR-WB码流生成单元。所述的加权语音计算单元的输入端输入升采样之后的AMR合成语音和VAD并与4子帧插值单元A的输出端相连,其输出端与开环基音搜索单元的输入端相连。所述的4子帧插值单元A的输入端与输入16维的ISP,其输出端分别与加权语音计算单元、自适应码本计算单元以及脉冲响应计算单元的输入端相连。所述的ISP到ISF转换单元的输入端输入16维的ISP,其输出端与ISF量化单元的输入端相连。所述的ISF量化单元的输入端与ISP到ISF转换单元的输出端相连,其输出端分别与4子帧插值单元B和AMR-WB码流生成单元的输入端相连。所述的开环基音搜索单元的输入端接收扩展得到的开环基音并与加权语音的输出端相连,其输出端与闭环基音搜索单元的输入端相连。所述的4子帧差值单元B的输入端与ISF量化单元的输出端相连,其输出端分别与自适应码本信号计算单元和脉冲响应计算单元的输入端相连。所述的自适应码本计算单元的输入端输入升采样后的AMR合成语音并与4子帧插值单元A的输出端相连,其输出端与固定码本目标信号计算单元的输入端相连。所述的闭环基音搜索单元的输入端与自适应码本计算单元的输出端相连,其输出端分别与自适应码本贡献

计算单元和AMR-WB码流生成单元的输入端相连。所述的自适应码本贡献计算单元的输入端与闭环基音搜索单元的输出端相连,其输出端分别与自适应滤波器选择单元和增益矢量量化单元的输入端相连。所述的增益矢量量化单元的输入端分别与自适应码本贡献计算单元和固定码本搜索单元的输出端相连,其输出端与AMR-WB码流生成单元的输入端相连。所述的自适应滤波器选择单元的输入端与自适应码本贡献计算单元的输入端相连,其输出端分别与固定码本目标信号计算单元和AMR-WB码流生成单元的输入端相连。所述的固定码本计算单元的输入端输入扩展得到宽带固定码本并分别与自适应码本目标信号计算单元和自适应滤波器选择单元的输入端相连,其输出端与固定码本搜索单元的输入端相连。所述的固定码本搜索单元的输入端分别与固定码本目标信号计算单元和脉冲响应计算单元的输入端相连,其输出端分别与增益矢量化单元和AMR-WB码流生成单元的输入端相连。所述的AMR-WB码流生成单元的输入端接收扩展得到高频增益索引并分别与固定码本搜索单元、自适应滤波器选择单元、增益矢量量化单元、闭环基音搜索单元以及ISF量化单元的输入端相连,其输出端输出AMR-WB宽带码流。

[0010] 所述的训练单元包括窄带码流分离单元、窄带码流解析单元、AMR-WB编码单元、SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元。所述的窄带码流分离单元的输入端输入窄带码流,其输出端与窄带码流解析单元的输入端相连;所述的窄带码流解析单元的输入端与窄带码流分离单元的输入端相连,其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连;所述的AMR-WB编码单元的输入端输入宽带语音,其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连;所述的SVR训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出SVR映射模型;所述的开环基音映射函数训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出开环基音映射函数;所述的固定码本映射码本训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出映射码本;所述的高频增益映射函数训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连,其输出端输出高频增益映射函数。

[0011] 所述的AMR-WB编码单元包括预处理单元、线性预测分析单元、ISP量化单元、4子帧ISP插值单元A、加权语音计算单元、4子帧ISP插值单元B、开环基音搜索单元、目标信号计算单元、最佳基音延迟和增益搜索单元、自适应码本分量计算单元、自适应码本滤波器选择单元、冲激响应计算单元、高频增益索引计算单元、固定码本搜索单元、滤波器更新单元、激励计算单元以及增益量化单元。所述的预处理单元的输入端输入采样率为16KHz的宽带语音,其输出端分别与线性预测分析单元、加权语音计算单元以及目标信号计算单元的输入端相连;所述的线性预测分析单元的输入端与预处理单元的输入端相连、其输出端分别与ISP量化单元和4子帧ISP插值单元B的输入端相连;所述的ISP量化单元的输入端与线性预测分析单元的输入端相连,其输出端与4子帧ISP差值单元A的输入端相连;所述的4子帧插值单元A的输入端与ISP量化单元的输入端相连,其输出端与冲激响应计算单元的输入端相连;所述的加权语音计算单元的输入端分别与预处理单元和四子帧ISP插值单元B的输入端相连,其输出端与开环基音搜索单元的输入端相连;所述的4子帧插值单元B的输入端与线性预测分

析单元的输出端相连,其输出端分别与目标信号计算单元、加权语音计算单元以及冲激响应计算单元的输入端相连;所述的开环基音搜索单元的输入端与加权语音计算单元的输出端相连,其输出端与最佳基音延迟和增益搜索单元的输入端相连;所述的目标信号计算单元的输入端分别与预处理单元、4子帧ISP插值单元B以及4子帧ISP插值单元A的输出端相连,其输出端分别与固定码本搜索单元和最佳基音延迟和增益搜索单元的输入端相连;所述的最佳基音延迟和增益搜索单元的输入端分别与目标信号计算单元、开环基音搜索以及冲激响应计算单元的输出端相连,其输出端输出基音索引并与自适应码本贡献计算单元的输入端相连;所述的自适应码本贡献计算单元的输入端与最佳基音延迟和增益搜索单元的输入端相连,其输出端分别与自适应码本滤波器选择单元和增益量化单元的输入端相连;所述的自适应码本滤波器选择单元的输入端与自适应码本贡献计算单元的输出端相连,其输出端输出滤波器索引并与脉冲响应计算单元的输入端相连;所述的冲激响应计算单元的输入端分别与自适应码本滤波器选择单元、4子帧ISP插值单元A以及4子帧ISP插值单元B的输出端相连,其输出端分别与最佳基音延迟和增益搜索单元以及固定码本搜索单元的输入端相连;所述的固定码本搜索单元的输入端分别与目标信号计算单元、自适应码本滤波器选择单元以及冲激响应计算单元的输出端相连,其输出端输出固定码本增益索引并与增益量化单元的输入端相连;所述的增益量化单元的输入端分别与固定码本搜索单元以及自适应码本贡献计算单元的输出端相连,其输出端输出增益索引并与激励计算单元的输入端相连;所述的激励计算单元的输入端与增益量化单元的输出端相连,其输出端分别与滤波器状态更新单元和低频增益索引计算单元的输入端相连;所述的滤波器状态更新单元的输入端与激励计算单元的输出端相连;所述的低频增益索引计算单元的输入端输入采样率为16KHz的宽带语音并分别与4子帧ISP插值单元和激励计算单元的输出端相连,其输出端输出低频增益索引。

[0012] 本发明的有益效果:

[0013] (1)本发明可以准确地恢复出窄带信号所对应的高频部分,因此可以实现AMR窄带码流到AMR-WB宽带码流的转换。

[0014] (2)AMR窄带码流解码得到窄带语音相比,扩展后的AMR-WB宽带码流解码得到的宽带语音音质有明显提高。

[0015] (3)与AMR到AMR-WB的时域带宽扩展方法相比,本发明提出的码流域带宽扩展方法编解码部分计算量大幅度降低,可以降低30%左右。

附图说明

[0016] 图1一种将AMR窄带码流转换成AMR-WB宽带码流的转换装置。

[0017] 图2本发明扩展单元结构简图。

[0018] 图3本发明AMR解码单元结构简图。

[0019] 图4本发明参数提取单元结构简图。

[0020] 图5本发明AMR-WB部分编码单元。

[0021] 图6本发明训练单元结构简图。

[0022] 图7本发明AMR-WB编码单元。

[0023] 图8本发明AMR编码器速率表。

- [0024] 图9本发明AMR-WB编码器速率表。
- [0025] 图10本发明AMR在10.20kbps编码速率下的比特分配表。
- [0026] 图11本发明确定轨道最大最小位置的算法流程图。
- [0027] 图12本发明AMR-WB固定码本搜索流程图。
- [0028] 图13本发明AMR-WB在23.85kbps编码模式下的参数索引比特分配。
- [0029] 图14本发明SVR参数设置。

具体实施方式

[0030] 下面结合附图对本发明做进一步说明：

[0031] 本发明是根据AMR窄带码流按照一定的方法生成AMR-WB宽带码流，本发明的技术方案如下：

[0032] 一种将AMR窄带码流转换成AMR-WB宽带码流的转换装置如图1所示：包括扩展单元和训练单元，所述的训练单元为扩展单元提供参数扩展过程所需的映射关系，只在扩展单元工作之前“离线”地运行一次。

[0033] 所述的扩展单元如图2所示：包括AMR解码单元、参数提取单元、窄带能量计算单元、SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B、升采样单元以及AMR-WB部分编码单元。所述的AMR解码单元的输入端输入AMR的窄带码流，输出端与参数提取单元、窄带能量计算单元以及升采样单元的输入端相连。所述的参数提取单元的输入端与AMR解码单元的输入端相连，其输出端与SVR预测单元、函数映射单元A、码本映射单元以及AMR-WB部分编码单元的输入端相连。所述的窄带能量计算单元的输入端与AMR解码单元的输入端相连，其输出端与函数映射单元B的输入端相连。所述的SVR预测单元、函数映射单元A以及码本映射单元的输入端与参数提取单元的输入端相连并接收训练单元提供的映射关系，它们的输出端都与AMR-WB部分编码单元的输入端相连。所述的函数映射单元B的输入端与窄带能量计算单元的输入端相连接接收训练单元提供的映射函数，其输出端与AMR-WB部分编码单元的输入端相连。所述的升采样单元的输入端与AMR解码单元的输入端相连，其输出端与AMR-WB部分编码单元的输入端相连。AMR-WB部分编码单元的输入端与SVR预测单元、函数映射单元A、码本映射单元、函数映射单元B以及升采样单元的输入端相连，其输出端输出AMR-WB宽带码流。

[0034] 所述的AMR解码单元如图3所示：包括窄带码流分离单元、LSP解码单元、自适应码本解码单元、增益解码单元、固定码本解码单元、4子帧插值单元、激励重建单元、LSP到A(z)转换单元、合成滤波器单元以及后置滤波器单元。所述的窄带码流分离单元的输入端输入AMR窄带码流，其输出端分别约LSP解码单元、自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端相连。所述的LSP解码单元的输入端与码流分离单元的输入端相连，其输出端与4子帧插值单元的输入端相连。所述的自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端都与码流分离单元的输入端相连、他们的输出端都与激励重建单元的输入端相连。所述的4子帧LSP插值单元的输入端与LSP解码单元的输入端相连，其输出端与LSP到A(z)转换单元的输入端相连。所述的激励重建单元的输入端分别与自适应码本解码单元、增益解码单元以及固定码本解码单元的输入端相连。所述的LSP到A(z)转换单元的输入端与4子帧LSP插值单元的输入端相连，其输出端与合成滤波器单元的输入端相

连。所述的合成滤波器单元的输入端分别与激励重建单元和LSP到 $A(z)$ 转换单元的输出端相连,其输出端与后置滤波器单元的输入端相连。所述的后置滤波器单元的输入端与合成滤波器单元的输入端相连,其输出单元输出合成语音。

[0035] 所述的参数提取单元如图4所示:包括VAD提取单元、LSP提取单元、开环基音周期及固定码本提取单元。所述的VAD提取单元的输入端与AMR解码单元的输入端相连,其输出端与AMR-WB部分编码的输入端相连。所述的LSP提取单元的输入端与AMR解码单元的输入端相连,其输出端与SVR预测单元的输入端相连。所述的开环基音提取单元的输入端与AMR解码单元的输入端相连,其输出端与映射单元A的输入端相连。所述的固定码本单元的输入端与AMR解码单元的输入端相连,其输出端与码本映射单元的输入端相连。

[0036] 所述的AMR-WB部分编码单元如图5所示:包括加权语音计算单元、4子帧差值单元A、ISP到ISF转换单元、开环基音搜索单元、闭环基音搜索单元、自适应码本计算单元、4子帧差值单元B、ISF量化单元、自适应码本贡献计算单元、自适应滤波器选择单元、固定码本目标信号计算单元、固定码本搜索单元、增益矢量量化单元、脉冲响应计算单元以及AMR-WB码流生成单元。所述的加权语音计算单元的输入端输入升采样之后的AMR合成语音和VAD并与4子帧插值单元A的输出端相连,其输出端与开环基音搜索单元的输入端相连。所述的4子帧插值单元A的输入端与输入16维的ISP,其输出端分别与加权语音计算单元、自适应码本计算单元以及脉冲响应计算单元的输入端相连。所述的ISP到ISF转换单元的输入端输入16维的ISP,其输出端与ISF量化单元的输入端相连。所述的ISF量化单元的输入端与ISP到ISF转换单元的输入端相连,其输出端分别与4子帧插值单元B和AMR-WB码流生成单元的输入端相连。所述的开环基音搜索单元的输入端接收扩展得到的开环基音并与加权语音的输出端相连,其输出端与闭环基音搜索单元的输入端相连。所述的4子帧差值单元B的输入端与ISF量化单元的输入端相连,其输出端分别与自适应码本信号计算单元和脉冲响应计算单元的输入端相连。所述的自适应码本计算单元的输入端输入升采样后的AMR合成语音并与4子帧插值单元A的输出端相连,其输出端与固定码本目标信号计算单元的输入端相连。所述的闭环基音搜索单元的输入端与自适应码本计算单元的输入端相连,其输出端分别与自适应码本贡献计算单元和AMR-WB码流生成单元的输入端相连。所述的自适应码本贡献计算单元的输入端与闭环基音搜索单元的输入端相连,其输出端分别与自适应滤波器选择单元和增益矢量量化单元的输入端相连。所述的增益矢量量化单元的输入端分别与自适应码本贡献计算单元和固定码本搜索单元的输入端相连,其输出端与AMR-WB码流生成单元的输入端相连。所述的自适应滤波器选择单元的输入端与自适应码本贡献计算单元的输入端相连,其输出端分别与固定码本目标信号计算单元和AMR-WB码流生成单元的输入端相连。所述的固定码本计算单元的输入端输入扩展得到宽带固定码本并分别与自适应码本目标信号计算单元和自适应滤波器选择单元的输入端相连,其输出端与固定码本搜索单元的输入端相连。所述的固定码本搜索单元的输入端分别与固定码本目标信号计算单元和脉冲响应计算单元的输入端相连,其输出端分别与增益矢量量化单元和AMR-WB码流生成单元的输入端相连。所述的AMR-WB码流生成单元的输入端接收扩展得到高频增益索引并分别与固定码本搜索单元、自适应滤波器选择单元、增益矢量量化单元、闭环基音搜索单元以及ISF量化单元的输入端相连,其输出端输出AMR-WB宽带码流。

[0037] 所述的训练单元如图6所示,包括窄带码流分离单元、窄带码流解析单元、AMR-WB

编码单元、SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元。所述的窄带码流分离单元的输入端输入窄带码流，其输出端与窄带码流解析单元的输入端相连；所述的窄带码流解析单元的输入端与窄带码流分离单元的输出端相连，其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连；所述的AMR-WB编码单元的输入端输入宽带语音，其输出端分别与SVR训练单元、开环基音映射函数训练单元、固定码本映射码本训练单元以及高频增益映射函数训练单元的输入端相连；所述的SVR训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输出端相连，其输出端输出SVR映射模型；所述的开环基音映射函数训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输出端相连，其输出端输出开环基音映射函数；所述的固定码本映射码本训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连，其输出端输出映射码本；所述的高频增益映射函数训练单元的输入端分别与窄带码流解析单元和AMR-WB编码单元的输入端相连，其输出端输出高频增益映射函数。

[0038] 所述的AMR-WB编码单元如图7所示，包括预处理单元、线性预测分析单元、ISP量化单元、4子帧ISP插值单元A、加权语音计算单元、4子帧ISP插值单元B、开环基音搜索单元、目标信号计算单元、最佳基音延迟和增益搜索单元、自适应码本分量计算单元、自适应码本滤波器选择单元、冲激响应计算单元、高频增益索引计算单元、固定码本搜索单元、滤波器更新单元、激励计算单元以及增益量化单元。所述的预处理单元的输入端输入采样率为16KHz的宽带语音，其输出端分别与线性预测分析单元、加权语音计算单元以及目标信号计算单元的输入端相连；所述的线性预测分析单元的输入端与预处理单元的输出端相连、其输出端分别与ISP量化单元和4子帧ISP插值单元B的输入端相连；所述的ISP量化单元的输入端与线性预测分析单元的输出端相连，其输出端与4子帧ISP差值单元A的输入端相连；所述的4子帧插值单元A的输入端与ISP量化单元的输出端相连，其输出端与冲激响应计算单元的输入端相连；所述的加权语音计算单元的输入端分别与预处理单元和四子帧ISP插值单元B的输出端相连，其输出端与开环基音搜索单元的输入端相连；所述的4子帧插值单元B的输入端与线性预测分析单元的输出端相连，其输出端分别与目标信号计算单元、加权语音计算单元以及冲激响应计算单元的输入端相连；所述的开环基音搜索单元的输入端与加权语音计算单元的输出端相连，其输出端与最佳基音延迟和增益搜索单元的输入端相连；所述的目标信号计算单元的输入端分别与预处理单元、4子帧ISP插值单元B以及4子帧ISP插值单元A的输出端相连，其输出端分别与固定码本搜索单元和最佳基音延迟和增益搜索单元的输入端相连；所述的最佳基音延迟和增益搜索单元的输入端分别与目标信号计算单元、开环基音搜索以及冲激响应计算单元的输出端相连，其输出端输出基音索引并与自适应码本贡献计算单元的输入端相连；所述的自适应码本贡献计算单元的输入端与最佳基因延迟和增益上搜索单元的输出端相连，其输出端分别与自适应码本滤波器选择单元和增益量化单元的输入端相连；所述的自适应码本滤波器选择单元的输入端与自适应码本贡献计算单元的输出端相连，其输出端输出滤波器索引并与脉冲响应计算单元的输入端相连；所述的冲激响应计算单元的输入端分别与自适应码本滤波器选择单元、4子帧ISP插值单元A以及4子帧ISP插值单元B的输出端相连，其输出端分别与最佳基音延迟和增益搜索单元以及固定码本搜索单元的输入端相连；所述的固定码本搜索单元的输入端分别与目标信号计算单

元、自适应码本滤波器选择单元以及冲激响应计算单元的输出端相连,其输出端输出固定码本增益索引并与增益量化单元的输入端相连;所述的增益量化单元的输入端分别与固定码本搜索单元以及自适应码本贡献计算单元的输出端相连,其输出端输出增益索引并与激励计算单元的输入端相连;所述的激励计算单元的输入端与增益量化单元的输入端相连,其输出端分别与滤波器状态更新单元和高频增益索引计算单元的输入端相连;所述的滤波器状态更新单元的输入端与激励计算单元的输出端相连;所述的高频增益索引计算单元的输入端输入采样率为16KHz的宽带语音并分别与4子帧ISP插值单元和激励计算单元的输出端相连,其输出端输出高频增益索引。

[0039] 如图8所示,AMR支持8种编码模式;如图9所示AMR-WB支持9种编码模式。在如下码流转换的具体步骤中,本发明以AMR 10.2kbps编码速率到AMR-WB23.85kbps编码速率下的码流转换为例进行介绍。

[0040] 一种将AMR窄带码流转换到AMR-WB宽带码流的转换装置及其方法,在进行码流在线转换之前,对于一种工作语种需要且仅需要一次“离线”建立转换时所需的各种映射关系;码流转换的具体步骤如下:

[0041] A、AMR解码

[0042] 将采样率为8KHz的语音信号经AMR 10.2kbps编码器进行编码,得到与之对应的窄带码流;将该窄带码流经AMR解码器进行解码。

[0043] A1、码流分离

[0044] 窄带码流分离单元将接收到的AMR窄带码流根据图10所示的比特分配表分离出VAD标志、LSP索引、基音索引、增益索引以及固定码本索引。

[0045] A2、LSP解码

[0046] 根据窄带码流分离单元输出的LSP量化索引,通过查表可以重建出量化后的LSP矢量。

[0047] A3、LSP四子帧插值

[0048] 由A2解码得到的LSP矢量作为第四子帧的LSP系数,并通过相邻帧之间LSP系数的内插得到第一、二和第三子帧的LSP系数,内插过程如式(1)、式(2)和式(3)所示。

$$[0049] \quad \hat{q}_1^{(n)} = 0.75\hat{q}_4^{(n-1)} + 0.25\hat{q}_4^{(n)} \quad (1)$$

$$[0050] \quad \hat{q}_2^{(n)} = 0.5\hat{q}_4^{(n-1)} + 0.5\hat{q}_4^{(n)} \quad (2)$$

$$[0051] \quad \hat{q}_3^{(n)} = 0.25\hat{q}_4^{(n-1)} + 0.75\hat{q}_4^{(n)} \quad (3)$$

[0052] 其中, $\hat{q}_4^{(n-1)}$ 是解码得到的上一帧第四子帧的LSP系数, $\hat{q}_4^{(n)}$ 是解码得到当前帧第四子帧的LSP系数, $\hat{q}_1^{(n)}$, $\hat{q}_2^{(n)}$ 以及 $\hat{q}_3^{(n)}$ 分别是内插得到的当前帧第一、二以及三子帧的LSP系数。

[0053] A4、LSP转换到A(z)

[0054] 在内插得到每一子帧的LSP系数之后,需要将其转换到线性预测系数 a_i ($i=1, 2, \dots, 10$)。循环变量 i 取值范围从1到5,每次增加1。每次变量 i 循环时

[0055] ① $f_1(i) = -2q_{2i-1}f_1(i-1) + 2f_1(i-2)$ 。

[0056] ②循环变量 j 取值范围从 $i-1$ 到1,每次循环变量 j 循环时,执行 $f_1^{[i]} = f_1^{[i-1]}(j) - 2q_{2i-1}f_1^{[i-1]}(j-1) + f_1^{[i-1]}(j-2)$ 操作。

[0057] 其中, $f_1(0)=1, f_1(-1)=0$ 。将 q_{2i-1} 替换成 q_{2i} 即可得到 $f_2(i)$ 。

$$[0058] \begin{cases} f'_1=f_1(i)+f_1(i-1) & i=1,\dots,5 \\ f'_2=f_2(i)-f_2(i-1) & i=1,\dots,5 \end{cases} \quad (4)$$

$$[0059] a_i = \begin{cases} 0.5f'_1(i)+0.5f'_2(i), & i=1,\dots,5 \\ 0.5f'_1(11-i)-0.5f'_2(11-i), & i=6,\dots,10 \end{cases} \quad (5)$$

[0060] A5、自适应码本解码

[0061] A51、基音周期解码

[0062] 根据A1分离出的基音索引P1来寻找基音周期T1的整数部分和分数部分。通过P1/P3获取第一/三子帧基音周期的整数部分 $\text{int}(T_1)/\text{int}(T_1)$ 和分数部分 $\text{frac1}/\text{frac3}$ 的步骤如下:

$$\begin{aligned} \text{if } P1 < 197 \\ \text{int}(T_1) &= (P1 + 2)/3 + 19 \\ \text{frac1} &= P1 - 3\text{int}(T_1) + 58 \end{aligned}$$

[0063] *else* (6)

$$\begin{aligned} \text{int}(T_1) &= P1 - 112 \\ \text{frac1} &= 0 \end{aligned}$$

end

$$\begin{aligned} \text{if } P3 < 197 \\ \text{int}(T_3) &= (P3 + 2)/3 + 19 \\ \text{frac3} &= P3 - 3\text{int}(T_3) + 58 \end{aligned}$$

[0064] *else* (7)

$$\begin{aligned} \text{int}(T_3) &= P3 - 112 \\ \text{frac3} &= 0 \end{aligned}$$

end

[0065] 第二/四子帧基音周期的整数部分和分数部分通过 $t_{\min 2}/t_{\min 4}$ 获得,其中, $t_{\min 2}/t_{\min 4}$ 可以通过如下的递推关系得到:

$$\begin{aligned} t_{\min 2} &= \text{int}(T_1) - 5 \\ \text{if } t_{\min 2} < 20 & \text{ then } t_{\min 2} = 20 \end{aligned}$$

$$t_{\max 2} = t_{\min 2} + 9$$

[0066] *if* $t_{\max 2} > 143$ *then* (8)

$$t_{\max 2} = 143$$

$$t_{\min 2} = t_{\max 2} - 9$$

end

$t_{\min 4} = \text{int}(T_3) - 5$
 if $t_{\min 4} < 20$ then $t_{\min 4} = 20$
 $t_{\max 4} = t_{\min 4} + 9$
 [0067] if $t_{\max 4} > 143$ then
 $t_{\max 4} = 143$
 $t_{\min 4} = t_{\max 4} - 9$
 end
 (9)

[0068] 则,第二/四子帧的基音周期 T_2/T_4 为:

$$[0069] \quad \text{int}(T_2) = (P_2 + 2) / 3 - 1 + t_{\min} \quad (10)$$

$$[0070] \quad \text{frac}_2 = P_2 - 2 - 3((P_2 + 2) / 3 - 1) \quad (11)$$

$$[0071] \quad \text{int}(T_4) = (P_4 + 2) / 3 - 1 + t_{\min} \quad (12)$$

$$[0072] \quad \text{frac}_4 = P_4 - 2 - 3((P_4 + 2) / 3 - 1) \quad (13)$$

[0073] A52、自适应码本解码

[0074] 解码得到基音周期后,可以通过内插过去的激励 $u(n)$ 得到自适应码本矢量 $v(n)$:

[0075]

$$v(n) = \sum_{i=0}^9 u(n-k-i) b_{60}(t+i \cdot 6) + \sum_{i=0}^9 u(n-k+1+i) b_{60}(6-t+i \cdot 6) \quad (14)$$

[0076] 其中,内插滤波器(截止频率为3.6KHz) b_{60} 是在 ± 59 处截断的汉明窗截取样函数 $\sin(x)/x$ 得到, $b_{60} = 0$ 。

[0077] A6、固定码本解码

[0078] 根据A1分离出的固定码本索引可以得到固定码本的脉冲位置、符号以及固定码本矢量。若该子帧基音周期整数部分小于子帧长度40,则需要修正固定码本矢量

$$[0079] \quad \mathbf{c}(n) = \mathbf{c}(n) + \hat{g}_p \mathbf{c}(n) \quad (15)$$

[0080] 其中, \hat{g}_p 是A71解码得到自适应码本增益。

[0081] A7、增益解码

[0082] A71、自适应码本增益解码

[0083] 根据A1分离出的增益索引从相应的量化表中查找相应的自适应码本增益 \hat{g}_p 和固定码本增益修正因子 \hat{g}_{gc} 。

[0084] A72、固定码本增益解码

[0085] 首先,计算预测能量

$$[0086] \quad \tilde{E}(n) = \sum_{i=1}^4 b_i E(n-i) \quad (16)$$

[0087] 然后,计算平均固定码本能量

$$[0088] \quad E_f = 10 \lg \left(\frac{1}{N} \sum_{j=0}^{N-1} c^2(j) \right) \quad (17)$$

[0089] 则,预测增益为:

$$[0090] \quad g_c' = 10^{0.05(\tilde{E}(n) + \bar{E} - E_f)} \quad (18)$$

[0091] 其中, \bar{E} 是固定码本的平均能量,在10.20kbps编码速率下为33。最后,量化后的固

定码本增益 \hat{g}_c 为:

$$[0092] \quad \hat{g}_c = \gamma_{gc} g'_c \quad (19)$$

[0093] A8、激励信号重建

[0094] 激励信号 $u(n)$ 可以由自适应码本激励和固定码本激励通过式(19)计算得到:

$$[0095] \quad u(n) = \hat{g}_p v(n) + \hat{g}_c c(n) \quad (20)$$

[0096] 根据自适应码本的贡献对激励信号进行修正:

$$[0097] \quad \hat{u}(n) = \begin{cases} u(n) + 0.25 \beta \hat{g}_p v(n), & \hat{g}_p > 0.5, 12.2 \text{ kbit/s 模式} \\ u(n) + 0.5 \beta \hat{g}_p v(n), & \hat{g}_p > 0.5, \text{其它模式} \\ u(n), & \hat{g}_p \leq 0.5 \end{cases} \quad (21)$$

[0098] 自适应增益控制(AGC)用来补偿未加重激励 $u(n)$ 和预加重激励 $\hat{u}(n)$ 之间的增益差, 预加重激励的增益缩放因子 η 为:

$$[0099] \quad \eta = \begin{cases} \sqrt{\frac{\sum_{n=0}^{39} u^2(n)}{\sum_{n=0}^{39} \hat{u}^2(n)}}, & \hat{g}_p > 0.5 \\ 1.0, & \text{其它} \end{cases} \quad (22)$$

[0100] 则增益缩放后的预加重激励信号为

$$[0101] \quad \hat{u}' = \eta \hat{u}(n) \quad (23)$$

[0102] A8、合成滤波

[0103] 一子帧(40个采样点)的重构语音为

$$[0104] \quad \hat{s}(n) = \hat{u}'(n) - \sum_{i=1}^{10} \hat{a}_i \hat{s}(n-i), n=0, 1, \dots, 39 \quad (24)$$

[0105] A9、后置滤波

[0106] A8得到的重构语音需要通过一个后置滤波器, 该滤波器是共振峰后置滤波器和谱倾斜补偿滤波器的级联。后置滤波器需要每5ms修改一次。其中, 共振峰滤波器 $H_f(z)$ 为

$$[0107] \quad H_f(z) = \frac{\hat{A}(z/\gamma_n)}{\hat{A}(z/\gamma_d)} \quad (25)$$

[0108] 其中, $\hat{A}(z)$ 为线性预测的逆滤波器, γ_n 和 γ_d 用来控制共振峰后置滤波器的阶数。

谱倾斜补偿滤波器 $H_t(z)$ 为

$$[0109] \quad H_t(z) = 1 - \mu z^{-1} \quad (26)$$

[0110] 其中

$$[0111] \quad \mu = \gamma_t \frac{r_h(1)}{r_h(0)} \quad (27)$$

$$[0112] \quad r_h(i) = \sum_{j=0}^{L_h-i-1} h_f(j) h_f(j+i) \quad (28)$$

[0113] 在10.20kbit/s编码速率下, $\gamma_n = 0.7$, $\gamma_d = 0.75$,

$$[0114] \quad \gamma_i = \begin{cases} 0.8, & r_k(1) > 0 \\ 0, & \text{其它} \end{cases}$$

[0115] B、参数提取

[0116] B1、VAD标志提取

[0117] A1码流分离出的前8个比特即为所需的VAD标志

[0118] B2、LSP提取

[0119] 所需的LSP为A3LSP四子帧插值的结果。

[0120] B3、基音提取

[0121] 所需的开环基音周期为A51解码得到的第一、三子帧基音周期的整数部分。

[0122] B4、固定码本提取

[0123] 所需的固定码本是A6解码得到的固定码本脉冲位置。

[0124] B5、窄带语音能量计算

[0125] 计算每一帧合成语音 $\hat{s}(n)$ 的对数域能量nb_ener_log,计算过程如下:

$$[0126] \quad nb_ener = \sum_{i=0}^{L_FRAME-1} synth^2(i) \quad (30)$$

$$[0127] \quad nb_ener_log = \log_2(nb_ener) \quad (31)$$

[0128] 其中,L_FRAME为语音帧的帧长,在AMR里L_FRAME=160。

[0129] 宽带参数扩展

[0130] C1、VAD参数扩展

[0131] 由于VAD参数主要是用来表征语音有无的,与带宽无关,所以将AMR解码得到的VAD参数直接映射到AMR-WB的编码端,从而省去了编码端VAD参数的计算。

[0132] C、C2、ISP参数扩展

[0133] 将窄带语音解码得到的10维LSP参数通过F1训练得到SVR模型进行预测,预测器的输出即为16维的ISP参数

[0134] C3、开环基音周期扩展

[0135] 由于,10.20kbps编码速率下的AMR与23.85kbps编码速率下的AMR-WB基音周期的分辨率不同;所以,若采用基音周期的直接扩展将会引起合成语音质量的严重下降。因此,对于该参数的扩展需要借助AMR解码器输出的合成语音,以及AMR-WB的基音周期搜索过程。首先,将AMR解码器端得到的第一/三子帧的开环基音周期作为F22训练得到的映射函数输入:

$$[0136] \quad T_{op1_wb} = T_{01} * 0.819 + 31.452, \quad (32)$$

$$[0137] \quad T_{op3_wb} = T_{03} * 0.728 + 30.339, \quad (33)$$

[0138] 这里, T_{op1_wb}/T_{op3_wb} 即为与之对应的宽带语音第一/三子帧的开环基音周期;为了保证合成语音质量,没有将该参数直接作为宽带语音开环基音搜索的结果,而是通过该参数来限制开环基音周期搜索的频率范围,这样在保证语音质量的同时,降低了开环基音搜索的计算量。

[0139] 具体实现过程是:将映射得到的开环基音周期减去一个常数作为开环基音周期搜索的下界;而将开环基音周期加上一个常数作为开环基音周期搜索的上界。该常数的选择

需要在计算量和语音质量之间做一个折中:大的搜索范围意味着较高的合成语音质量和较大的计算量,小的搜索范围意味着较低的合成语音质量和较小的计算量。在本发明中,该常数设为2.

[0140] C4、高频增益索引扩展

[0141] 高频增益索引的扩展时通过函数映射实现的。将AMR解码端得到的窄带语音能量作为F4训练得到的映射函数的输入,所得到的函数值即为宽带语音的高频增益索引值。

[0142] C5、宽带固定码本扩展

[0143] AMR10.20kbps与AMR-WB23.85kbps的固定码本结构相差较多,且这种CELP的编码模式对于固定码本的误差非常敏感,所以,为了保证合成语音的质量,采用的是同开环基音周期扩展相同的方法。

[0144] 首先,将AMR解码得到的窄带固定码本进行码本搜索,得到窄带码本索引;然后,将该索引映射到与之对应的宽带固定码本(此处的映射码本由F3训练得到),将索引所在的行矢量输出,即为与窄带对应的宽带固定码本。

[0145] 为了使合成语音质量降低不严重,根据映射得到的宽带码本,求取每个轨道脉冲位置的最大,最小值,该步的算法流程图如图11所示。轨道脉冲位置确定以后,当AMR-WB编码器搜索每个轨道脉冲的时候,不再进行16个位置的全搜索,而只需要搜索该轨道脉冲位置最大和最小值之间的位置。该方法在保证语音质量下降不明显的前提下,有效地缩减了脉冲搜索的范围,从而降低了固定码本搜索的计算量。

[0146] D、宽带参数部分编码

[0147] D1、ISP编码

[0148] D11、ISP到ISF转化

[0149] 将C2得到的ISP参数利用式(31)转换成ISF系数 f_i ($i=0,1,\dots,15$)

$$[0150] \quad f_i = \begin{cases} \frac{f_s}{2\pi} \arccos(q_i), & i = 0, \dots, 14 \\ \frac{f_s}{4\pi} \arccos(q_i), & i = 15 \end{cases} \quad (34)$$

[0151] 其中, $f_s=12800\text{kHz}$ 为采样率。

[0152] D14、ISF量化

[0153] 假设 $z(n)$ 是第 n 帧去均值后的ISF矢量,则预测残差矢量 $r(n)$ 可表示为

$$[0154] \quad r(n) = z(n) - p(n) \quad (35)$$

[0155] 其中, $p(n)$ 是第 n 帧根据式(5.10)预测得到的LSF矢量

$$[0156] \quad \mathbf{p}(n) = \frac{1}{3} \hat{\mathbf{r}}(n-1) \quad (36)$$

[0157] 其中, $\hat{\mathbf{r}}(n-1)$ 是上一帧的量化残差矢量。

[0158] 使用分裂多阶矢量量化器对 $r(n)$ 进行量化。首先,将矢量 $r(n)$ 分成9维的矢量 $r_1(n)$ 和7维的矢量 $r_2(n)$ 。然后,通过两级操作对两个子矢量进行量化。在第一级操作过程中,对 $r_1(n)$ 和 $r_2(n)$ 进行8bit量化;在第二级操作过程中,对两个子矢量进行二次分裂后根据编码模式进行量化。

[0159] D2、基音周期编码

[0160] D21、ISP四子帧插值

[0161] 将C2扩展得到的ISP作为第四子帧ISP,根据当前帧第四子帧的ISP系数 q_4 和前一帧第四子帧的ISP系数 $q_4^{(n-1)}$,插值得到当前帧第1、2、3子帧的ISP系数。插值过程同A3。

[0162] 在内插得到每一子帧的ISP系数之后,需要将其按照A4所述的步骤转换到线性预测系数 $a_i(i=1,2,\dots,16)$ 。

[0163] D22、计算加权语音

[0164] 将升采样后的合成语音通过式(37)所示的感知加权滤波器:

$$[0165] \quad W(z) = A(z/\gamma_1)H_{de-emph} \quad (37)$$

[0166] 其中,

$$[0167] \quad A(z/\gamma_1) = 1 + \sum_{i=1}^{16} \gamma_1^i a_i z^{-i}, \quad (38)$$

$$[0168] \quad H_{de-emph} = \frac{1}{1 - \beta_1 z^{-1}}, \quad (39)$$

[0169] 其中, $\beta_1 = 0.68$ 。

[0170] 对于长度为L的子帧,加权语音 $s_w(n)$ 为:

$$[0171] \quad s_w(n) = s(n) + \sum_{i=1}^{16} a_i \gamma_1^i s(n-i) + \beta_1 s_w(n-1), n=0,1,\dots,L-1, \quad (40)$$

[0172] D23、开环基音周期搜索

[0173] 第一子帧加权语音的相关函数为:

$$[0174] \quad C_1(d) = \sum_{n=0}^{63} s_{wd}(n) s_{wd}(n-d) w(d), d = T_{op1_wb} - 2, \dots, T_{op1_wb} + 2, \quad (41)$$

[0175] 第三子帧加权语音的相关函数为:

$$[0176] \quad C_3(d) = \sum_{n=0}^{63} s_{wd}(n) s_{wd}(n-d) w(d), d = T_{op3_wb} - 2, \dots, T_{op3_wb} + 2, \quad (42)$$

[0177] 其中, $w(d)$ 是加权函数。开环基音周期就是使得 $C_1(d)/C_3(d)$ 最大的 d 值。

$$[0178] \quad w(d) = w_1(d) w_n(d), \quad (43)$$

$$[0179] \quad w_1(d) = cw(d), \quad (44)$$

[0180] 其中, $cw(d)$ 的取值见定点计算描述表。

$$[0181] \quad w_n(d) = \begin{cases} cw(|T_{old} - d| + 98), & v > 0.8 \\ 1.0, & \text{其它} \end{cases}, \quad (45)$$

[0182] 开环基音增益 g 的计算公式为:

$$[0183] \quad g = \frac{\sum_{n=0}^{63} s_{wd}(n) s_{wd}(n-d_{\max})}{\sqrt{\sum_{n=0}^{63} s_{wd}^2(n) \sum_{n=0}^{63} (n-d_{\max})}}, \quad (46)$$

[0184] 其中, d_{\max} 是使得 $C(d)$ 取最大值的基音延迟; T_{old} 是前5个半帧的基音延迟的中值滤波值。 v 是自适应因子。如果当前帧的开环基音增益 $g > 0.6$,则认为该帧是浊音帧,下一帧的 v 设为1.0;否则, $v = 0.9v$ 。

[0185] D24、量化后的ISP系数4子帧差值

[0186] 将ISF量化单元输出的量化后的ISF系数通过式(46)转换到ISP系数,量化后的LSP系数4子帧插值过程同D21。

[0187] D25、ISP系数到线性预测系数转换

[0188] 在内插得到每一子帧的ISP系数之后,需要将其转换到线性预测系数 $\hat{a}_i(i=1,2,\dots,16)$ 。ISP系数 $q_i(i=1,2,\dots,16)$ 到线性预测系数 $a_i(i=1,2,\dots,16)$ 的转换过程如下:

[0189] 在已知内插的ISP系数的前提下,可以通过式(84)和(85)得到 $F_1(z)$ 和 $F_2(z)$,用 $q_i(i=1,2,\dots,16)$ 可以迭代计算 $f_1(z)$

```

for i=2 to m/2
    f1(i)=-2q2i-2f1(i-1)+2f1
    for j=i-1 downto 2
[0190]     f1(j)=f1(j)-2q2i-2f1(j-1)+f1(j-2)
    end
    f1(1)=f1(1)-2q2i-2
end

```

[0191] 初始值为 $f_1(0)=0, f_1(1)=-2q_0$ 。同理,用 q_{2i-1} 代替 $q_{2i-2}, m/2-1$ 代替 $m/2$ 且 $f_2(0)=1, f_2(1)=-2q_1$,可以计算得到 $f_2(z)$ 。

[0192] 在获得 $f_1(z)$ 和 $f_2(z)$ 之后,将 $F_2(z)$ 乘以 $1-z^{-2}$ 可以得到 $F_2'(z)$

[0193] $f_2'(i)=f_2(i)-f_2(i-2), i=2,\dots,m/2-1$ (47)

[0194] $f_1'(i)=f_1(i), i=0,\dots,m/2$ (48)

[0195] 则,线性预测系数 $a_i(i=1,2,\dots,16)$ 为

[0196]
$$a_i = \begin{cases} 0.5f_1'(i)+0.5f_2'(i), & i=1,\dots,m/2-1 \\ 0.5f_1'(i)-0.5f_2'(i), & i=m/2+1,\dots,m-1 \\ 0.5f_1'(m/2), & i=m/2 \\ q_{m-1} & i=m \end{cases} \quad (49)$$

[0197] D26、自适应码本目标信号计算

[0198] 线性预测残差信号 $r(n)$ 为:

[0199]
$$r(n)=s(n)+\sum_{i=1}^{16}\hat{a}_i s(n-i), n=0,1,\dots,63 \quad (50)$$

[0200] 则,自适应码本搜索的目标信号 $x(n)$ 是通过合成滤波器 $1/\hat{A}(z)$ 和加权滤波器 $A(z/\gamma_1)H_{de_exp}(z)$ 的输出。

[0201] D27、脉冲响应计算

[0202] 在AMR-WB编码中要计算的脉冲响应 $h(n)$ 是指感知加权合成滤波器

[0203]
$$H_w(z)=\frac{A(z/\gamma_1)H_{de_exp}(z)}{\hat{A}(z)} \quad (51)$$

[0204] 的单位脉冲响应。

[0205] D28、闭环基音搜索

[0206] 闭环基音搜索准则是使原始语音和重构语音之间均方加权误差最小,即使 T_k 最

大, T_k 为:

$$[0207] \quad T_k = \frac{\sum_{n=0}^{63} x(n)y_k(n)}{\sqrt{\sum_{n=0}^{63} y_k(n)y_k(n)}} \quad (52)$$

[0208] 其中, $x(n)$ 是D25得到目标信号, $y_k(n)$ 是滤波激励,其表达式为:

$$[0209] \quad y_k(n) = y_{k-1}(n-1) + u(-k)h(n) \quad (53)$$

[0210] 其中, $u(n)$, $n = -(231+17), \dots, 63$ 是激励缓冲器的值; $h(n)$ 是感知加权合成滤波器的脉冲应。在搜索阶段, $u(n)$, $n = 0, \dots, 63$ 是未知的,只有当基音延迟小于64的时候才需要。为了简化搜索,将线性预测残差存在 $u(n)$ 里使得(52)所示的关系对所有延时有效。确定最佳整数基音周期后,在该基音周期附近的分数从 $-3/4$ 到 $3/4$ 以步长 $1/4$ 作测试。内插 T_k ,并搜索其最大值得到分数基音周期。

[0211] D3、基音周期增益

[0212] 分数延迟确定后,在给定段内插过去的激励信号 $u(n)$ 得到 $v'(n)$ 。该内插操作是通过两个FIR滤波器实现的,其中一个是在 ± 17 处截断的汉明窗截取样函数,另一个是在 ± 63 处截断的汉明窗截取样函数。

[0213] 自适应码本 $v(n)$ 为:

$$[0214] \quad v(n) = \sum_{i=-1}^1 b_{LP}(i+1)v'(n+i) \quad (54)$$

[0215] 其中, $b_{LP} = [0.18, 0.64, 0.18]$ 。则自适应码本增益 g_p 为:

$$[0216] \quad g_p = \frac{\sum_{n=0}^{63} x(n)y(n)}{\sum_{n=0}^{63} y(n)y(n)}, \quad 0 \leq g_p \leq 1.2 \quad (55)$$

[0217] 其中, $x(n)$ 是目标信号, $y(n) = v(n)*h(n)$ 是自适应码本矢量滤波后的结果。

[0218] D4、固定码本搜索

[0219] D41、自适应码本贡献计算

[0220] 自适应码本贡献 $\varphi(n)$ 为

$$[0221] \quad \varphi(n) = g_p y(n) \quad (56)$$

$$[0222] \quad y(n) = y(n)*h(n) \quad (57)$$

[0223] D42、固定码本搜索目标信号

[0224] 固定码本搜索目标信号 $x_2(n)$ 为

$$[0225] \quad x_2(n) = x(n) - \varphi(n) \quad (58)$$

[0226] 如果 c_k 是第 k 个固定码本矢量,使 Q_k 最大的矢量即为所求,

$$[0227] \quad Q_k = \frac{(\mathbf{x}_2^t \mathbf{H} \mathbf{c}_k)^2}{\mathbf{c}_k^t \mathbf{H}^t \mathbf{H} \mathbf{c}_k} = \frac{(\mathbf{d}^t \mathbf{c}_k)^2}{\mathbf{c}_k^t \boldsymbol{\varphi} \mathbf{c}_k} = \frac{(C_k)^2}{E_k} \quad (59)$$

[0228] 其中, H 为对角线元素为 $h(0)$ 的下三角Toeplitz卷积矩阵,且对角线一次往下的元素是 $h(1), \dots, h(63)$;

$$[0229] \quad C = \sum_{i=0}^{N_p-1} a_i d(m_i) \quad (60)$$

[0230] 其中, m_i 是第 i 个脉冲的位置, a_i 为其幅度, $N_p = 24$ 为 23.85kbps 编码速率下的脉冲个数。

$$[0231] \quad E = \sum_{i=0}^{N_p-1} \varphi(m_i, m_i) + 2 \sum_{i=0}^{N_p-2} \sum_{j=i+1}^{N_p-1} a_i a_j \varphi(m_i, m_j) \quad (61)$$

[0232] 为了简化搜索过程, 用适当的量化信号 $b(n)$ 先做幅度脉冲符号的预判决

$$[0233] \quad b(n) = \sqrt{\frac{E_d}{E_r}} r_{LTP}(n) + \alpha d(n) \quad (62)$$

$$[0234] \quad d(n) = \sum_{i=n}^{63} x_2(n) h(i-n), \quad n = 0, \dots, 63 \quad (63)$$

[0235] 其中, r_{LTP} 为长时预测的残差信号, E_r 为其能量, E_d 为 d 能量, α 为扩展因子, 编码速率越大, α 越小, 在 23.85kbps 编码速率下, $\alpha = 0.5$ 。

[0236] AMR-WB 在 23.85 编码速率下的固定码本搜索流程图如图 12 所示。在进行脉冲搜索的时候只在 C5 确定的所在轨道脉冲位置最大值最小值之间的搜索。

[0237] D5、固定码本增益

[0238] 固定码本增益 g_c 可以由式 (63) 给出

$$[0239] \quad g_c = \frac{\mathbf{x}_2^T \mathbf{z}}{\mathbf{z}^T \mathbf{z}} \quad (64)$$

[0240] 其中, x_2 为固定码本搜索的目标矢量, z 是固定码本矢量与感知加权合成滤波器的脉冲响应 $h(n)$ 的卷积, 即

$$[0241] \quad z(n) = \sum_{i=0}^n c(i) h(n-i), \quad n = 0, 1, \dots, 63 \quad (65)$$

[0242] 其中

$$[0243] \quad h(n) = h(n) - \beta h(n-T), \quad n = T, T-1, \dots, 63 \quad (66)$$

[0244] 其中, T 是这一子帧的基音分数延时的最大整数部分, β 为量化后的基音增益。D5、基音增益和固定码本增益量化

[0245] 在 23.85kbps 编码速率下, 基音增益和固定码本增益的量化是通过 7-bit 码本实现的。

[0246] 固定码本增益的量化是由系数固定的 MA 预测器。4 阶 MA 预测器是在固定能量 $E(n)$ 上实现的,

$$[0247] \quad E(n) = 10 \log \left(\frac{1}{N} g_c^2 \sum_{i=0}^{63} c^2(i) \right) - \bar{E} \quad (67)$$

[0248] 其中, $c(i)$ 是固定码本激励, $\bar{E} = 30 \text{dB}$ 为固定码本能量。预测能量 $\tilde{E}(n)$ 为:

$$[0249] \quad \tilde{E}(n) = \sum_{i=1}^4 b_i E(n-i) \quad (68)$$

[0250] 其中, $[b_1 b_2 b_3 b_4] = [0.5, 0.4, 0.3, 0.2]$ 为 MA 预测器系数, $E(1)$ 、 $E(2)$ 、 $E(3)$ 、 $E(4)$ 分别为当前帧第 1、2、3、4 子帧的固定能量, $E(-1)$ 、 $E(-2)$ 、 $E(-3)$ 、 $E(-4)$ 分别为前一帧第 1、2、3、4 子帧的固定能量。

[0251] 预测固定码本增益 g_c 可以通过预测能量 $\tilde{E}(n)$ 计算得到, 具体实现如下:

[0252] 首先,计算平均固定能量 E_i

$$[0253] \quad E_i = 10 \log \left(\frac{1}{N} \sum_{i=0}^{N-1} c^2(i) \right) \quad (69)$$

[0254] 则,预测的固定码本增益 g'_c 为

$$[0255] \quad g'_c = 10^{0.05(E + \bar{E} - E_i)} \quad (70)$$

[0256] 定义 γ 为 g_c 和 g'_c 之间的修正因子

$$[0257] \quad \gamma = \frac{g_c}{g'_c} \quad (71)$$

[0258] 定义预测误差为 $R(n)$,则有

$$[0259] \quad R(n) = E(n) - \tilde{E}(n) = 20 \log \gamma \quad (72)$$

[0260] 在23.85kbps编码速率下,基音周期增益 g_p 和修正因子 γ 用7比特码本进行联合矢量量化,既由 g_p 和 γ 构成一个二维矢量 $[g_p, \gamma]^T$,然后进行码本增益搜索。增益码本的搜索就是通过是原始语音和重建语音的均方误差最小

$$[0261] \quad E = \mathbf{x}^T \mathbf{x} + g_p^2 \mathbf{y}^T \mathbf{y} + g_c^2 \mathbf{z}^T \mathbf{z} - 2g_p \mathbf{x}^T \mathbf{y} - 2g_c \mathbf{x}^T \mathbf{z} + 2g_p g_c \mathbf{y}^T \mathbf{z} \quad (73)$$

[0262] 其中, x 是目标矢量, y 为滤波后的自适应码本矢量, z 为滤波后的固定码本矢量。

[0263] E、宽带码流生成

[0264] 将C和D扩展得到的各参数索引按照图13的顺序,写到码流中,即可得到与AMR-WB23.85kbps解码器兼容的宽带码流。

[0265] F、映射关系训练

[0266] 将与A窄带语音相对应的,采样率为16KHz的宽带语音信号作为输入,由AMR-WB编码器在-dtx模式、23.85kbps的编码速率下进行编码并提取相关参数。

[0267] F1、ISP系数映射关系训练

[0268] F11、ISP系数提取

[0269] F111、预处理

[0270] 将输入的采样率为16KHz的16比特线性PCM语音信号经过一个如式(72)所示的高通滤波器和式(73)所示的预加重处理

$$[0271] \quad H_{hl}(z) = \frac{0.989502 - 1.979004z^{-1} + 0.989502z^{-2}}{1 - 1.978882z^{-1} + 0.979126z^{-2}} \quad (74)$$

$$[0272] \quad H_{pre_emph} = 1 - 0.68z^{-1} \quad (75)$$

[0273] F112、加窗和自相关计算

[0274] 加窗后的语音信号 $s_w(n)$ 为

$$[0275] \quad s_w(n) = w(n)s(n), n=0, 1, \dots, 383 \quad (76)$$

[0276] 其中, $s(n)$ 为F111预加重处理后的语音信号, $w(n)$ 为

$$[0277] \quad w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{2L_1 - 1}\right), & n=0, \dots, L_1 - 1 \\ \cos\left(\frac{2\pi(n - L_1)}{4L_2 - 1}\right), & n=L_1, \dots, L_1 + L_2 - 1 \end{cases} \quad (77)$$

[0278] 其中, $L_1 = 256, L_2 = 128$ 。 $s_w(n)$ 的自相关函数为

$$[0279] \quad r(k) = \sum_{n=k}^{383} s_w(n)s_w(n-k), \quad k=0,1,\dots,16 \quad (78)$$

[0280] 将 $r(k)$ 经滞后窗 w_{lag} 处理使其具有60Hz的带宽扩展,

$$[0281] \quad w_{lag}(i) = \exp\left[-\frac{1}{2}\left(\frac{2\pi f_0 i}{f_s}\right)^2\right], \quad i=1,2,\dots,16 \quad (79)$$

[0282] 其中, $f_0=60$, $f_s=12800$ 。此外对 $r(0)$ 乘以白噪声校正因子1.0001。

[0283] F113、用莱文逊-杜宾算法求解线性预测系数

[0284] 修正后的自相关函数为

$$[0285] \quad \begin{cases} r'(0)=1.0001r(0) \\ r'(k)=r(k)w_{lag}(k), \quad k=1,2,\dots,16 \end{cases} \quad (80)$$

[0286] 根据式(78)得到的 $r'(k)$ 借用莱文逊-杜宾算法可以求得线性预测系数 a_i ($i=1, 2, \dots, 16$),如式(81)和(82)所示

$$[0287] \quad k_i = \frac{r'(i) - \sum_{j=1}^{i-1} a_j^{(i-1)} r'(i-j)}{E_{i-1}}, \quad 1 \leq i \leq p \quad (81)$$

$$[0288] \quad \begin{cases} a_i^{(i)} = k_i \\ a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}, \quad 1 \leq j \leq i-1 \end{cases} \quad (82)$$

$$[0289] \quad E_i = (1 - k_i^2) E_{i-1} \quad (83)$$

[0290] 其中, $E_0=r'(0)$;求解结果为: $a_j=a_j^{16}, j=1,2,\dots,16$ 。

[0291] F114、线性预测系数到导谱对系数转换

[0292] 为了方便进行内插和量化,需要将线性预测系数 a_i ($i=1, 2, \dots, 16$)转换到导谱对系数 q_i ($i=1, 2, \dots, 16$)。ISP系数定义成差分多项式(80)和(81)的根。

$$[0293] \quad F'_1(z) = A(z) + z^{-16}A(z^{-1}) \quad (84)$$

$$[0294] \quad F'_2(z) = A(z) - z^{-16}A(z^{-1}) \quad (85)$$

[0295] 可证明这些多项式的所有解都在单位圆上交替出现, F'_2 有一个根 $z=-1$ ($\omega=\pi$),有一个根 $z=1$ ($\omega=0$)。通过定义新的多项式(84)和(85)可以消除这两个根:

$$[0296] \quad F_1(z) = F'_1(z) \quad (86)$$

$$[0297] \quad F_2(z) = F'_2(z)/(1-z^{-2}) \quad (87)$$

[0298] 其中, $F_1(z)$ 在单位圆上有8个共轭根而 $F_2(z)$ 在单位圆上有7个共轭根,因此,

$$[0299] \quad F_1(z) = (1+a[16]) \prod_{i=0,2,\dots,14} (1-2q_i z^{-1} + z^{-2}) \quad (88)$$

$$[0300] \quad F_2(z) = (1-a[16]) \prod_{i=0,2,\dots,14} (1-2q_i z^{-1} + z^{-2}) \quad (89)$$

[0301] 其中, $a[16]$ 是最后一个线性预测系数, $q_i = \cos(\omega_i)$, ω_i 是导谱频率(immittance spectral frequencies, ISF)并且满足

$$[0302] \quad 0 < \omega_1 < \omega_2 < \dots < \omega_{10} < \pi \quad (90)$$

[0303] 因为 $F_1(z)$ 和 $F_2(z)$ 都是对称多项式,所以只需要计算出每个多项式的前8个和前7

个系数以及最后一个线性预测系数。这些多项式的系数可以由一下递推关系得到

[0304] for i=0 to 7

[0305] $f_1(i) = a_i + a_{m-i}$

[0306] $f_2(i) = a_i - a_{m-i} + f_2(i-2)$

[0307] $f_1(8) = 2a_8$

[0308] 其中, $m=16$ 为预测器阶数, $f_2(-2) = f_2(-1) = 0$ 。当 $z = e^{j\omega}$ 时, 则有:

[0309] $F_1(\omega) = 2e^{-j8\omega} C_1(x)$ (91)

[0310] $F_2(\omega) = 2e^{-j7\omega} C_2(x)$ (92)

[0311] 其中,

[0312] $C_1(x) = \sum_{i=0}^7 f_1(i) T_{8-i}(x) + f_1(8)/2$ (93)

[0313] $C_2(x) = \sum_{i=0}^6 f_2(i) T_{8-i}(x) + f_2(7)/2$ (94)

[0314] 其中, $T_m = \cos(m\omega)$ 是 m 阶 Chebyshev 多项式。 $f(i)$ ($i=1, 2, \dots, 5$) 是 $f_1(z)$ 或者 $f_2(z)$ 的系数。当 $x = \cos(\omega)$ 时, $C(x)$ 的递推关系为

[0315] *for* $k = n_f - 1$ *down to* 1
 $b_k = 2xb_{k+1} - b_{k+2} + f(n_f - k)$
end
 $C(x) = xb_1 - b_2 + f(n_f)/2$

[0316] 其中, 当 $n_f = 8$ 时 $C(x) = C_1(x)$; 当 $n_f = 7$ 时 $C(x) = C_2(x)$ 。 $b_{n_f} = f(0)$, $b_{n_f+1} = 0$ 。

[0317] 至此, 已经求解得到了宽带语音的导谱对系数 q_i ($i=1, 2, \dots, 16$)

[0318] F12、10 维 LSP 参数到 16 维 ISP 参数映射关系的训练

[0319] 本专利通过引入支持向量回归 (SVR) 模型, 完成从窄带语音 LSP 系数 (A2 解码得到) 到宽带语音 ISP 系数 (F11 得到) 的预测。预测的准确性与预测数据本身的特性和模型训练过程的参数设置有关, 尤其是后者。由于 ISP 各维之间的相关性比较弱, 所以可以分别进行 10 维 LSP 到一维 ISP 的模型训练 (共需 16 次)。本发明以 A2 解码得到的 10 维 LSP 到 F11 得到的第一维 ISP 为例, 介绍 SVR 模型的训练过程。

[0320] 首先, 要对 A2 解码得到的 10 维 LSP 进行归一化。此处归一化的方法有多种, 本专利选用的是按维 (列) 归一化。具体实现过程如下:

[0321] (1) 分别计算每一维的最大值 \max_i

[0322] $\max_i = \max_{0 < j \leq \text{frame_num}} LSP_i^j, i=1, 2, \dots, 10$ (95)

[0323] 其中, frame_num 为帧数, LSP_i^j 表示第 j 帧第 i 维的 LSP 系数。

[0324] (2) 按维归一化

[0325] $LSP_i^j_norm = \frac{LSP_i^j}{\max_i}, i=1, 2, \dots, 10; j=1, 2, \dots, \text{frame_num}$ (96)

[0326] 然后, 将归一化后的 frame_num 帧 10 维的 LSP 系数作为训练模型的输入; frame_num 帧第一维的 ISP 系数 (F11 得到) 作为训练模型的目标输出, 由 SVR 进行训练得到模型一个由 10 维矢量到一维标量的预测模型 model 。本章训练过程 SVR 参数设置如图 14 所示;

[0327] F2、开环基音周期映射关系训练

[0328] F21、宽带开环基音周期提取。

[0329] F211、ISP系数到ISF系数转换同D11。

[0330] F212、ISF系数量化同D12。

[0331] F213、4子帧ISP系数插值同D21。

[0332] F214、4子帧量化的ISP系数插值同D21。

[0333] F215、ISP系数到线性预测系数转换同D25。

[0334] F216、感知加权同D22。

[0335] F217、开环基音搜索同D23。

[0336] 至此,已经求解得到了宽带语音的开环基音F22、开环基音周期映射关系训练分别以A51解码得到窄带语音的第一/三子帧的开环基音周期 T_{01}/T_{03} 作为函数输入,D215搜索得到的宽带语音的第一/三子帧的开环基音周期 T_{op1_wb}/T_{op3_wb} 作为函数输出,应用最小二乘法拟合LEN帧二者之间的函数关系:

$$[0337] \quad T_{wb} = cT + d, \quad (97)$$

[0338] 利用最小二乘法拟合该函数关系的系数化简结果为

$$[0339] \quad c = \frac{\varphi_1 - \varphi_2 \varphi_3}{\varphi_4 - \varphi_2^2}, \quad (98)$$

$$[0340] \quad d = \frac{\varphi_3 \varphi_4 - \varphi_2 \varphi_1}{\varphi_4 - \varphi_2^2}, \quad (99)$$

[0341] 其中 $\varphi_1 = \frac{1}{LEN} \sum_{i=0}^{LEN-1} T_{wb} T$ $\varphi_2 = \frac{1}{LEN} \sum_{i=0}^{LEN-1} T$ $\varphi_3 = \frac{1}{LEN} \sum_{i=0}^{LEN-1} T_{wb}$ $\varphi_4 = \frac{1}{LEN} \sum_{i=0}^{LEN-1} (T_{wb})^2$ 拟合得到的第一子帧之间的映射关系为

$$[0342] \quad T_{op1_wb} = T_{01} * 0.819 + 31.452 \quad (100)$$

[0343] 第三子帧之间的映射关系为:

$$[0344] \quad T_{op3_wb} = T_{03} * 0.728 + 30.339 \quad (101)$$

[0345] F3、固定码本映射关系训练。

[0346] F31、固定码本参数提取。

[0347] F311、自适应码本目标信号计算同D26。

[0348] F312、冲激响应计算同D27。

[0349] F312、闭环基音搜索同D28。

[0350] F31、2闭环基音搜索同D3。

[0351] F313、自适应码本贡献计算同D41。

[0352] F314、固定码本搜索目标计算同D42,只是此处需要搜索脉冲所在轨道的所有位置。

[0353] F32、固定码本参数映射关系训练

[0354] 在本发明中宽带固定码本扩展是通过码本映射完成的,所以需要离线建立一对一对应的映射码本。其中窄带码本包括A6解码得到的8维窄带语音脉冲位置矢量,宽带码本包括F314搜索得到的24维宽带语音脉冲位置矢量。按照先是8维窄带语音脉冲位置,然后是24维宽带语音脉冲位置的顺序,组合成一个32维矢量。

[0355] 其中窄码本生成采用动态聚类中的C-均值算法,宽带码本生成采用一种加权求平均的方法。

[0356] F321、窄带码本生成

[0357] 通过C-均值聚类方法可以聚类得到低频包络码本。设码本容量(即聚类数)为N,以24维矢量的前8维为聚类对象进行聚类处理,得到每一类的质心矢量,所有这些质心矢量的集合构成低频码本。若码本容量N过大,则计算量太大;若N过小,则码本增益过小,恢复出来的宽带语音信号效果较差。所以需要在计算复杂度和扩展语音质量之间寻求一种折中。在本章中,N取为2048。

[0358] F322、宽带码本生成

[0359] 对于前8维聚类处理后的每一类,采用加权求平均的方法来计算后24维的中心矢量。该方法具体实现步骤如下:

[0360] (1)计算第i类初始质心 $aver0[i][k]$

$$[0361] \quad aver0[i][k] = \frac{1}{n} \sum_{j=0}^n x[j][k], \quad i = ind[j], \quad k = 10, 11, \dots, 37 \quad (102)$$

[0362] 其中, $x[j][k]$ 表示一个28维高频时域、频域包络矢量,n为某一类中的高频时域、频域包络矢量数, $ind[j]$ 表示矢量 $x[j][k]$ 所在类的类别号。

[0363] (2)计算第j个矢量 $x[j][k]$ 与所在类质心的距离 $dist[j]$

$$[0364] \quad dist[j] = \sum_{k=0}^M (x[j][k] - aver0[ind[j]][k])^2, \quad k = 1, 2, \dots, 28 \quad (103)$$

[0365] (3)计算第i类中所有矢量与质心之间距离倒数之和 $w[i]$

$$[0366] \quad w[i] = \sum_{ind[j]=i} \frac{1}{dist[j]} \quad (104)$$

[0367] (4)计算第i类新质心 $aver[i][k]$

$$[0368] \quad aver[i][k] = \sum_{ind[j]=i} \frac{1}{dist[j]} \sum_{k=0}^{M_i} \frac{x[j][k]}{w[i]} \quad (105)$$

[0369] 其中, M_i 是第i类的矢量数。

[0370] (5)分别计算初始质心 L_1 范数 $sum0$ 和新质心 L_1 范数 sum

$$[0371] \quad sum0 = \sum_{k=0}^M |aver0[i][k]| \quad (106)$$

$$[0372] \quad sum = \sum_{k=0}^M |aver[i][k]| \quad (107)$$

[0373] (6)判断每一类新质心和初始质心之间的距离是否小于预定门限T,即是否满足式(108)

$$[0374] \quad \frac{|sum0 - sum|}{sum} \leq T \quad (108)$$

[0375] 若式(108)不满足,则令 $aver0[i][k] = aver[i][k]$,并返回到步骤(2),直到所有分类质心都满足式(108)。

[0376] 迭代结束之后,所得到的质心即为高频时域包络以及频域包络聚类质心,所有这

些质心组成高频包络码本。

[0377] 在高频码本生成过程中,门限T的选择相当重要,若T太大,则不能有效减少一些特殊点对质心的影响;若T太小,则计算量会明显增大。因为,在本发明中,码本生成过程是离线进行的,所以T可以选的尽可能小

[0378] F4、高频增益索引映射关系训练

[0379] F41、高频增益索引提取

[0380] F411、固定码本增益计算

[0381] 同D5

[0382] F412、基音增益计算

[0383] 同D3

[0384] F413、增益量化

[0385] 同D6

[0386] F414、激励计算

[0387] 当前帧的激励信号 $u(n)$ 为

$$[0388] \quad u(n) = \hat{g}_p v(n) + \hat{g}_c c(n) \quad (109)$$

[0389] 其中, \hat{g}_p 和 \hat{g}_c 分别为F413中量化后的基音增益和固定码本增益。

[0390] F415、高频增益计算

[0391] 在23.85kbps编码速率下,高频增益 g_{HB} 为

$$[0392] \quad g_{HB} = \frac{\sum_{i=0}^{63} (s_{HB}(i))^2}{\sum_{i=0}^{63} (s_{HB2}(i))^2} \quad (110)$$

[0393] 其中, $s_{HB}(i)$ 为输入宽带语音经带通滤波器(通带为6.4到7KHz)滤波结果, $s_{HB2}(i)$ 为高频带激励信号 $u_{HB2}(i)$ 经高频带合成滤波器 $A_{HB}(i)$ 滤波的结果

$$[0394] \quad A_{HB}(i) = \hat{A}(z/0.8) \quad (111)$$

[0395] $\hat{A}(z)$ 由采样率为12.8KHz的信号分析得到,而解码的是16KHz的信号,所以

$$[0396] \quad FR_{16}(f) = FR_{12.8}\left(\frac{12.8}{16}f\right). \quad (112)$$

[0397] 其中, $FR_{12.8}(f)$ 为 $\hat{A}(z)$ 的频响。这说明12.8KHz采样率下的5.1KHz-5.6KHz将会映射成16KHz采样率下的6.4-7.0KHz。

[0398] F42、高频增益索引映射关系训练

[0399] 窄带语音能量与高频增益索引的映射关系,可以由F22所介绍的最小二乘法进行线性拟合得出。将B5得到的窄带语音能量 nb_ener_log 作为输入,F415得到的高频增益索引 g_{HB} 作为输出,利用最小二乘法进行线性拟合,可以得到二者之间的映射关系为:

$$[0400] \quad g_{HB} = 0.535nb_ener_log + 1310.7 \quad (113)$$

[0401] 以上所述,仅为本发明较佳的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,根据本发明的技术方案及其发明构思加以等同替换或改变,都应涵盖在本发明的保护范围之内。

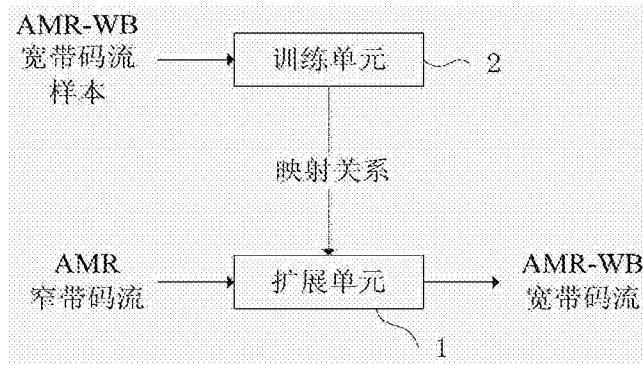


图1

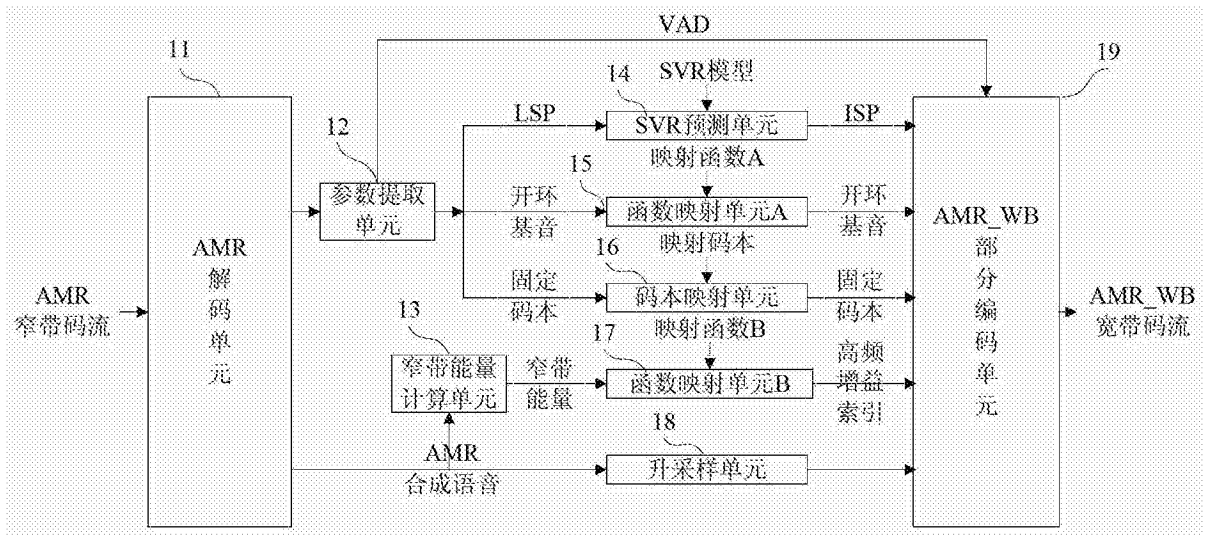


图2

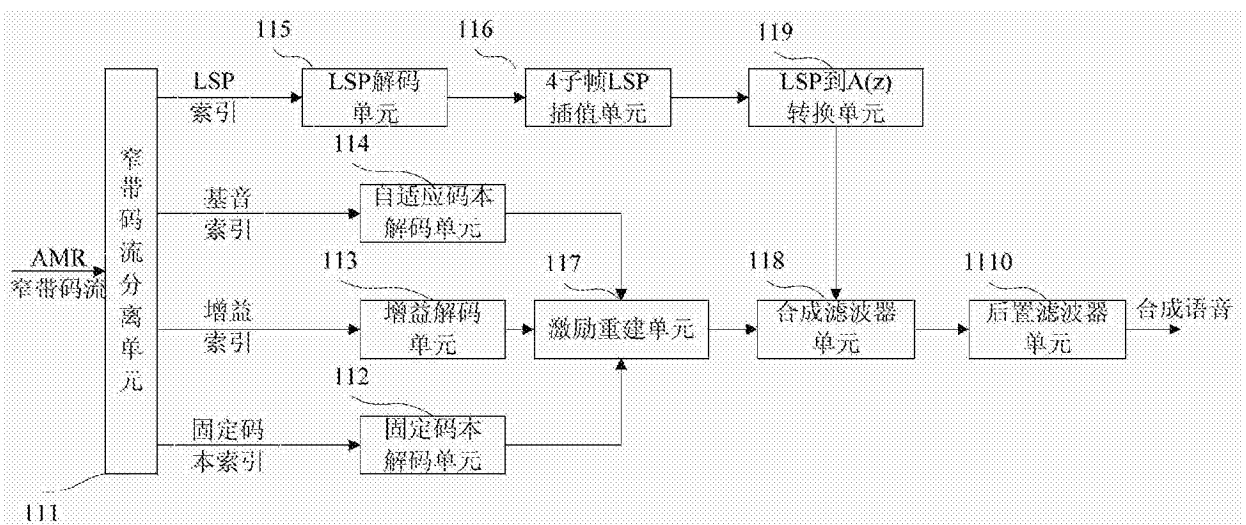


图3

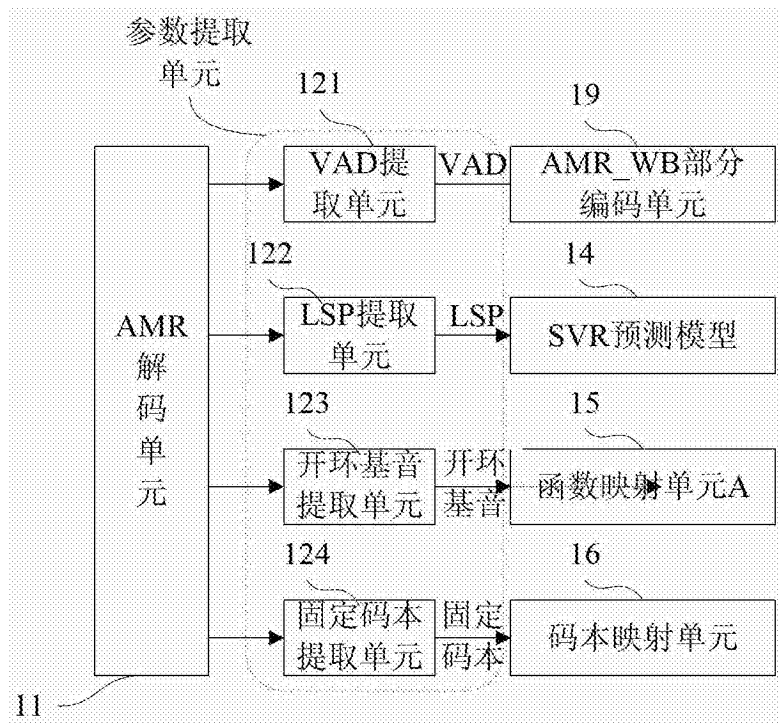


图4

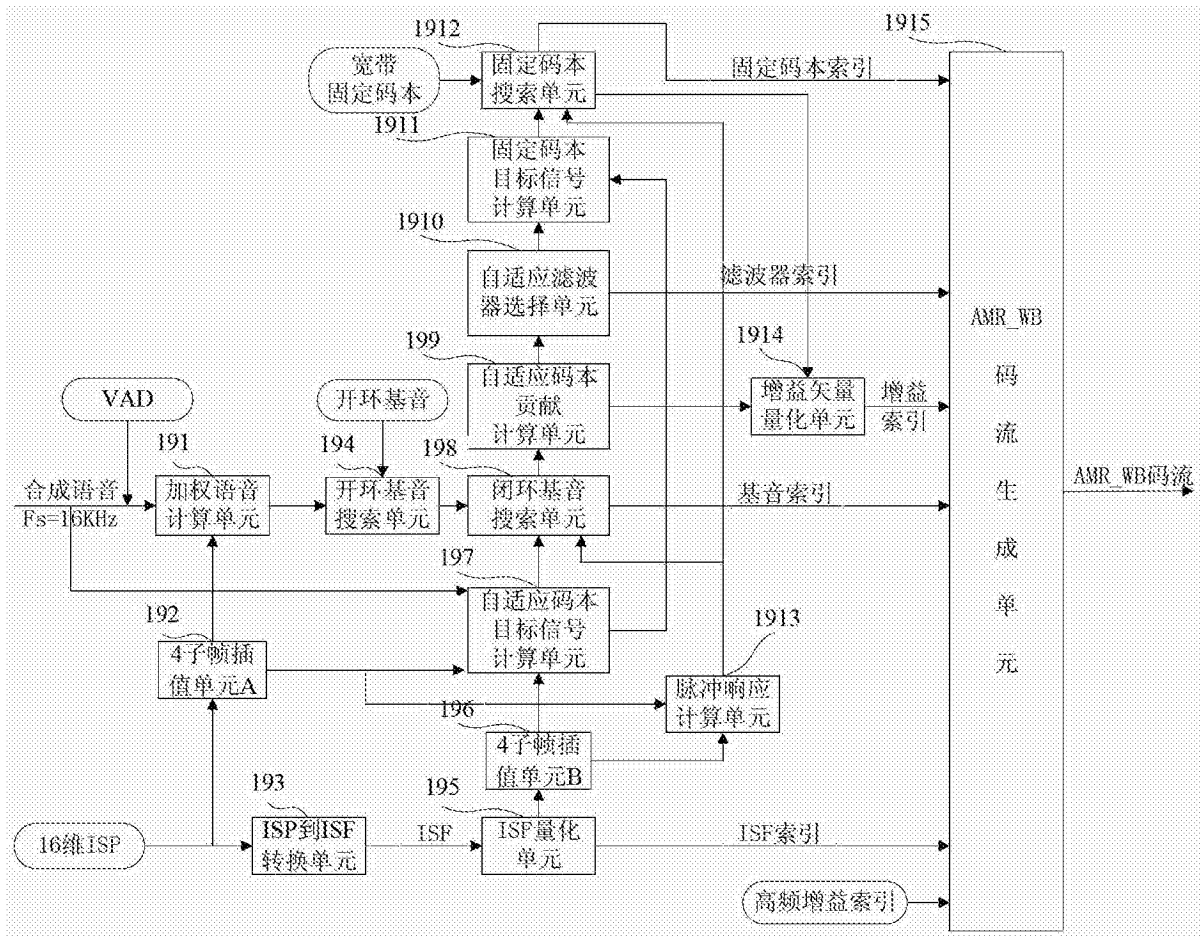


图5

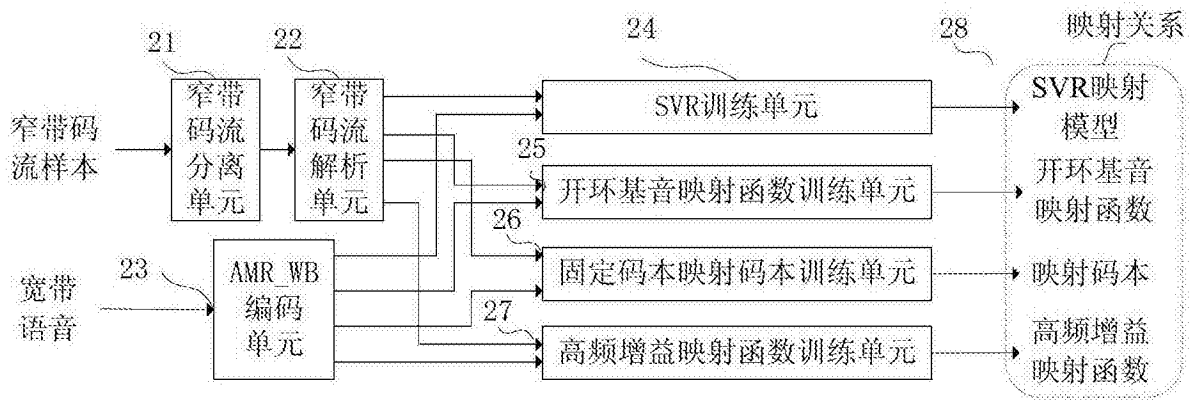


图6

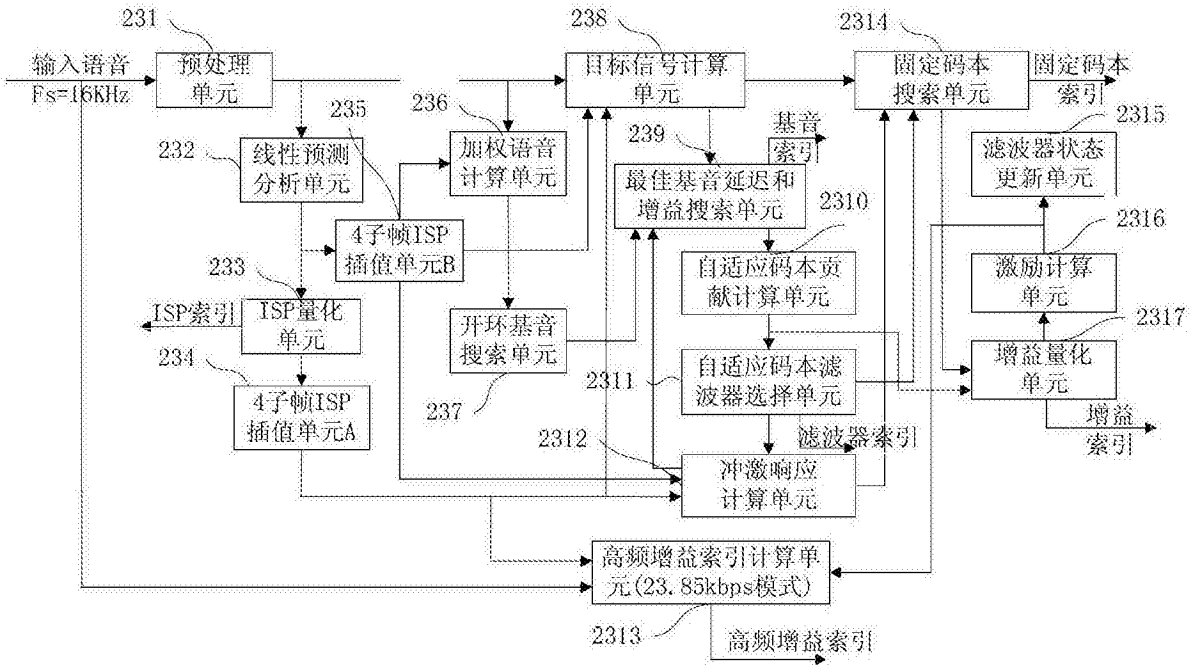


图7

编码模式	编码速率
AMR_12.20	12.20kbps
AMR_10.20	10.20kbps
AMR_7.95	7.95kbps
AMR_7.40	7.40kbps
AMR_6.70	6.70kbps
AMR_5.90	5.90kbps
AMR_5.15	5.15kbps
AMR_4.75	4.75kbps
AMR_SID	1.80kbps

图8

编码模式	编码速率
0	6.60kbps
1	8.85kbps
2	12.65kbps
3	14.25kbps
4	15.85kbps
5	18.25kbps
6	19.85kbps
7	23.05kbps
8	23.85kbps

图9

比特(MSB-LSB)	描述
s1-s8	VAD 标志
s9-s17	第 1 个 ISP 子矢量索引
s18-s26	第 2 个 ISP 子矢量索引
s18-s23	第 3 个 ISP 子矢量索引
子帧 1	
s27-s34	自适应码本索引
s35	第 1 个和第 5 个脉冲的符号信息
s36	第 2 个和第 6 个脉冲的符号信息
s37	第 3 个和第 7 个脉冲的符号信息
s38	第 4 个和第 8 个脉冲的符号信息
s39-s48	第 1, 2, 5 个脉冲的位置
s49-s58	第 3, 6, 7 个脉冲的位置
s59-s65	第 4, 8 个脉冲的位置
s66-s72	码本增益
子帧 2	
s73-s77	自适应码本索引
s78-s115	同 s35-s72
子帧 3	
s116-s161	同 s27-s72
子帧 4	
S162-s204	同 s73-s115

图10

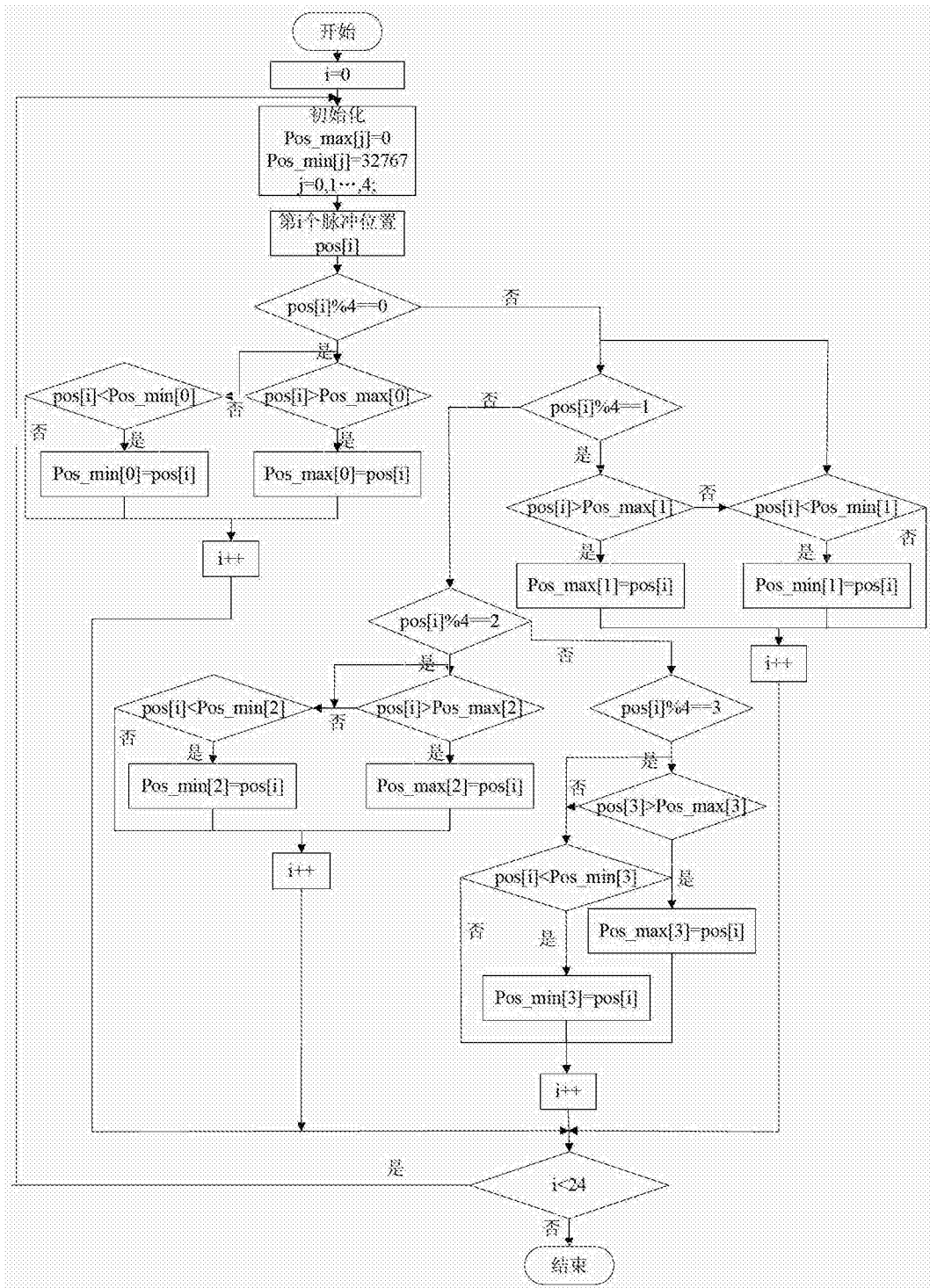


图11

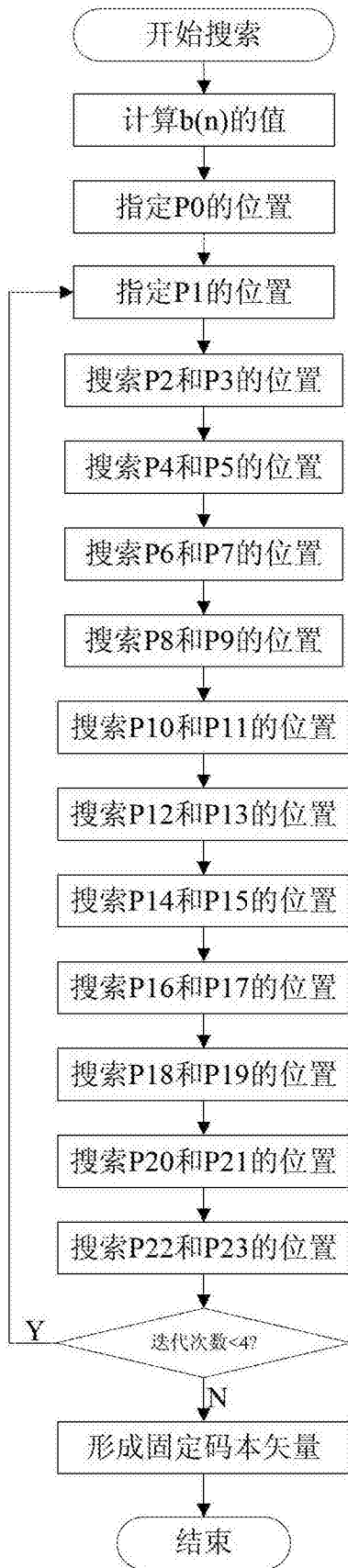


图12

比特(MSB-LSB)	描述
s1	VAD 标志
s2-s9	第 1 个 ISP 子矢量索引
s10-s17	第 2 个 ISP 子矢量索引
s18-s23	第 3 个 ISP 子矢量索引
s24-s30	第 4 个 ISP 子矢量索引
s31-s37	第 5 个 ISP 子矢量索引
s38-s42	第 6 个 ISP 子矢量索引
s43-s47	第 7 个 ISP 子矢量索引
子帧 1	
s48-s56	自适应码本索引
s57	长时预测滤波器标志
s58-s68	轨道 1 的码本索引 1
s69-s79	轨道 2 的码本索引 1
s80-s90	轨道 3 的码本索引 3
s91-s101	轨道 4 的码本索引 1
s102-s112	轨道 1 的码本索引 2
s113-s123	轨道 2 的码本索引 2
s124-s134	轨道 3 的码本索引 2
s135-s145	轨道 4 的码本索引 2
s146-s152	码本增益
s153-s156	高频带能量
子帧 2	
s157-s162	自适应码本索引
s163-s262	同 s57-s15
子帧 3	
s263-s371	同 s48-s156
子帧 4	
s372-s477	同 s157-s262

图13

参数	标志	本发明的设置	备注
-s	SVM 模型	3	epslion-SVR
-t	核函数类型	2	RBF 核函数
-c	损失函数	2.2	
-g	属性数目的倒数	2.8	
-p	epslion-SVR 中 epslion 的值	0.01	

图14