

摘 要

DRM 即数字调幅广播，是数字化广播的候选方案之一。它提供新业务，同时仅改变发射系统的信源、编码和调制模块。通常数字音频质量由数据率决定，然而，高数据率需要高传输带宽。因此，需要对数字音频信号进行压缩，即信源编码。信源编码是 DRM 的关键技术之一，编码质量直接决定了声音信号的质量和系统传输所需要的频带宽度。

本文阐述了 MPEG-4 音频编码中 CELP 编解码的原理、算法和验证。作者阅读了大量 MPEG-4 中 CELP 的语音编码标准及有关文献。并在 VC 的编程环境下，用 C 语言编程实现 MPEG-4 中码激励线性预测的编解码。

在论文安排上，首先，作者简单介绍了论文背景和所做工作；其次，较深入地阐述了 CELP 编码原理；然后，详细讲述了 MPEG-4 音频编码中 CELP 的实现原理；接着，讲述了 CELP 编解码在 PC 机上的详细实现过程；最后，对实验结果进行了分析和主观评估。

关键词：语音编码；CELP；MPEG-4

Abstract

DRM, which means Digital AM Broadcasting, is one of the digital broadcasting schemes. DRM system supports new business through changing only message resource, coding and modulation modules. Although digital audio quality is determined by data rate, high data rate requires high transfer bandwidth. So, we need compress digital audio signal, that is source coding. Source coding is one of the key technologies in DRM, the quality of source coding determines the quality of the speech and the required system bandwidth.

This paper puts emphasis on theory, algorithm and verifications of CELP codec in speech coding of MPEG-4. The author read a great amount of literature and standard papers and then programmed and made the realization of CELP codec.

In this paper, the author firstly introduces the background and his work. Then, he details the theory of CELP coding. Thirdly, the author discusses realization theory of CELP in MPEG-4 speech coding. Next, he presents the realization process of CELP coding at length. Finally, the author analyses and subjectively evaluates the result.

Key Words: speech coding; CELP; MPEG-4

目 录

第 1 章 引言	1
1.1 应用背景	1
1.2 中低速率语音编码技术	3
1.2.1 多脉冲线性预测编码	3
1.2.2 码激励线性预测编码	3
致谢	5
参考文献	6
附录	7

第 1 章 引言

1.1 应用背景

模拟的调幅广播发明于二十世纪二十年代，其后虽然也有一些技术上的进步，但系统体系基本没变，存在着传输质量差、业务单一、易被干扰等比较明显的缺点，随着 FM、电视和互联网等新型媒体的出现，模拟调幅广播的地位受到了重大挑战，用户规模迅速下降。因此如何利用现有频段和技术对调幅广播系统进行改造，使之能够满足时代的需要成为了各国广播运行部门和技术部门的一个重要课题。

普遍的共识是利用数字通信技术对发射系统进行数字化改造，提供新业务的同时保持模拟调幅业务的平滑过渡，仅改变发射系统的信源、编码和调制模块，尽可能利用现有的各种发射机，通常这种方案被称为数字调幅广播，或者数字 AM 广播。

目前虽然对于数字化广播有多种候选方案，但最成熟的并且已经成为国际标准的是所谓的 DRM 数字广播技术，DRM 是研发该技术的 Digital Radio Mondiale（法语单词）的简称，DRM 数字广播于 2001 年成为 ITU 标准，2002 年通过 IEC 审核进入实施阶段。DRM 系统是经过严格试验的、技术成熟的系统，并且 DRM 是目前唯一的非专利的数字广播系统，各国在实施中不会遇到所谓的专利壁垒。和其他数字广播系统相比 DRM 数字调幅广播具有以下优点：

- 大幅度降低广播功率；
- 显著提高信号传送的音质；
- 大大提高调幅频段信号传送的可靠性；
- 与现有模拟信号兼容，可以实现模拟数字信号同播，实现向全数字化的平滑过渡；
- 不改变现有频段分配，充分利用现有中短波频谱资源；
- 对现有发射机进行成本有限的改造，不需要增加新的发射机；
- 提供新的附加业务和数字传输

在系统设计上 DRM 采用了以下设计思想：

- 采用高效的信源变码技术以满足声音质量要求的前提下有效降低码率；

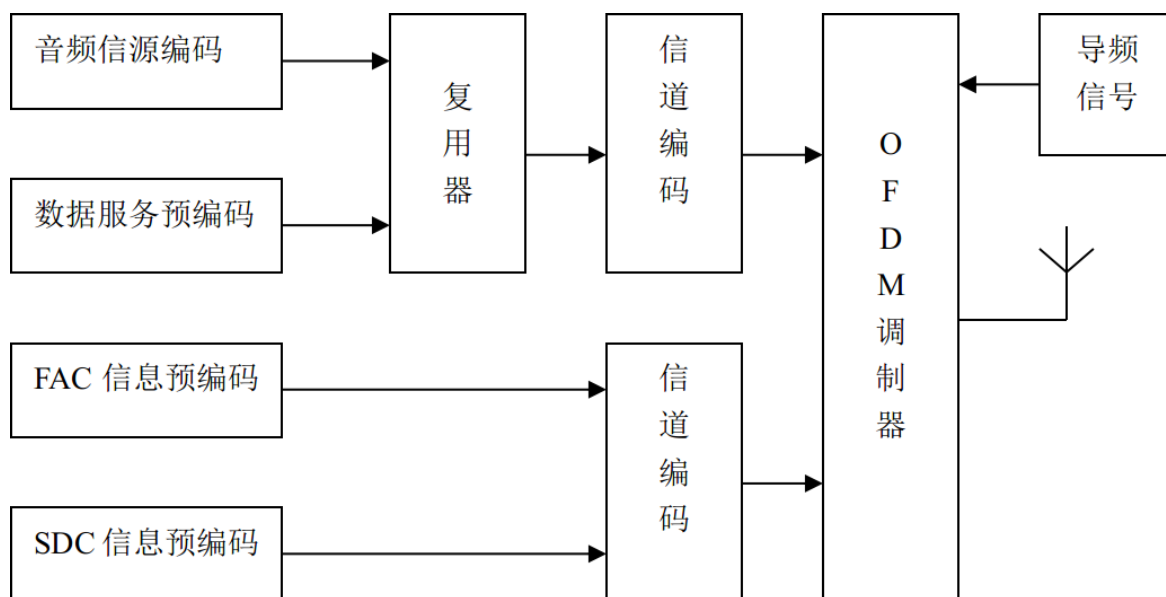


图 1.1 DRM 系统的原理框图

- 采用高效的信道编码和调制技术对抗多径传输和频率选择性衰减等信道影响；
- 提供灵活高效的信源复用方式满足不同的业务需求。

系统的原理框图如1.1所示。

通常数字音频的质量由数据率决定，但高的数据率需要高的传输带宽，数据压缩技术的出现解决了数字广播中音频质量和带宽的矛盾。DRM 系统中的数据压缩体现于语音编码（信源编码），其质量直接决定了声音信号的质量和系统传输所需要的频带宽度。

在 DRM 标准中根据不同的业务需要定义了三种不同的语音编码技术：先进音频编码 AAC、码激励线性预测 CELP 和谐波矢量激励编码 HVXC，除此之外 DRM 系统还采用了 SBR 频带复制技术，它可以在进行音频编码之前将带宽压缩为原来的 1/2，解码后通过低频段波形和附加信息恢复出高频信息，在保证音质的同时可以降低 30% 的数据率。DRM 系统要求在比较低的数据率下达到或超过 FM 的音质，系统中使用的三种语音编码都源于 MPEG-4 的子集。MPEG-4 标准的音频编码部分给出了这三种编码的算法。本文定位于研究 MPEG-4 音频编码中 CELP 编码的原理、算法验证。

1.2 中低速率语音编码技术

1.2.1 多脉冲线性预测编码

对余量信号进行深入研究后,与原始语音信号相反,余量信号中的小信号对合成语音的质量影响不大,如果对余量信号进行削波处理,即将幅度低于某一阈值的所有信号全置成零,这样,只要适当调整阈值就可以使余量信号的 90% 的样点值为零,用余下的幅度较大的信号作为激励信号源,其合成语音并未产生明显的畸变。这就提供了一种新的编码途径。1982 年, Bishnu S. Atal 和 Joel R. Remde 首先提出多脉冲线性预测编码方案。在此方案中首先规定激励脉冲序列在一定时间间隔中只能出现数目有限的非零脉冲,然后每个脉冲的位置和幅度用合成分析法和感觉加权均方误差最小的判决准则进行优化,最后用优化的脉冲序列表示余量信号作为激励信号源。这样既压缩了编码速率,有改善了合成语音的质量。这样的编码系统称为多脉冲线性预测编码 (Multi-Pulse Linear Predictive Coding), 简称 MPLPC。在这种编码方案中,为了求出最佳的脉冲序列,其运算量相当大,而且传送脉冲幅度和相位的信息量也比较大。由于一般多脉冲激励每 10ms 时间内至少需要 8 个非零脉冲,每个脉冲需要用幅度和相位两个参数来描述,这样非零激励脉冲传输信息为 2×800 个参数/秒。MPLPC 可以在 9.6 16kbps 范围内获得较好的合成语音质量。如果再降低编码速率,则合成语音质量将会变的很差,与多脉冲激励编码相似,但是更实用的编码方法是规则脉冲激励语音编码 (Regular-Pulse Excitation Coding), 简称 RPE-LPC, RPE-LPC 是 Ed.F. Deprettere 和 Peter Kroon 在 1985 年提出的,它用一组间距一定的非零的规则脉冲代替余量信号,该脉冲的相位(即第一个非零脉冲出现的位置)和每个非零脉冲的幅度可以按照 MPLPC 同样的方法进行优化。在 RPE-LPC 的激励脉冲序列中,因为各个非零脉冲的相互位置是固定的,所以它的计算量和编码速率与 MPLPC 相比都要小很多。

1.2.2 码激励线性预测编码

众所周知,在速率低于 1bit/采样的情况下,采用矢量量化(VQ)技术表示波形是非常有效的。1985 年, Manfred R. Schroeder 和 Bishnu S. Atal 提出了用 VQ 码本中的码字作为激励源的线性预测编码技术。码本中每一个存储的码字矢量

表 1.1 水系聚氨酯分类

类别	水溶型	胶体分散型	乳液型
状态	溶解 ~ 胶束	分散	白浊
外观	水溶型	胶体分散型	乳液型
粒径 / μm	< 0.001	$0.001 - 0.1$	> 0.1
重均分子量	$1000 \sim 10000$	数千 ~ 20万	> 5000

（简称码矢量）都可以代替余量信号作为可能的激励信号源。在编码时，对码本中的码矢量逐个进行搜索，找到与输入语音误差最小的合成语音的码矢量，只要将该码矢量的标号传送给接收端，在接收端用存储的同样的码本根据收到的标号恢复出相应的码矢量作为激励。这样的编码系统，称为码激励线性预测编码 (Code Excited Linear Predictive Coding)，简称 CELP，CELP 在 4.8-16kbps 范围内可以获得质量相当高的合成语音，并且抗噪声和多次转接性能也很好。

水系聚氨酯的形态对其流动性，成膜性及加工织物的性能有重要影响，一般分为三种类型^[1]，如表 1.1所示。

致谢

本论文的工作是在导师……。

参考文献

- [1] 姜敏, 彭少贤, 酆华兴. 形状记忆聚合物研究现状与发展[J]. 现代塑料加工应用, 2005, 17 (2): 53-56.

附录

附录相关内容...