

代号 10701
分类号 TN912

学号 0708420202
密级

西安电子科技大学

硕士学位论文



题(中、英文)目 音频重采样算法设计与实现

Design and Implementation of
Audio Resampling Algorithms

作者姓名 韩涛 指导教师姓名、职务 马鸿飞 副教授

学科门类 工学 学科、专业 通信与信息系统

提交论文日期 二零一零年一月

西安电子科技大学 学位论文创新性声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中做了明确的说明并表示了谢意。

申请学位论文与资料若有不实之处，本人承担一切的法律责任。

本人签名： 韩涛 日期 2010.1.12

西安电子科技大学 关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属西安电子科技大学。学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容，可以允许采用影印、缩印或其它复制手段保存论文。同时本人保证，毕业后结合学位论文研究课题再撰写的文章一律署名单位为西安电子科技大学。

（保密的论文在解密后遵守此规定）

本人签名： 韩涛 日期 2010.1.12

导师签名： 马海生 日期 2010.1.13

摘 要

随着数字信号处理理论和算法的发展，多抽样率信号处理在多媒体信号处理领域显得越来越重要。为了减少计算复杂度和存储复杂度，采样速率转换技术是十分必要的，音频重采样算法可以用来实现音频信号任意采样速率之间的转换。

本文首先简要介绍了多抽样率信号处理的基本概念和原理，设计了三种音频重采样算法，分别是基于线性插值的音频重采样算法、基于拉格朗日插值的音频重采样算法以及基于正弦插值的音频重采样算法，并用这三种算法实现了音频信号的重采样。线性插值基于插值点相邻两点的幅值来线性计算插值点的具体幅值，拉格朗日插值是利用拉格朗日插值多项通过设置好长度的窗口来计算插值点的具体幅值，正弦插值是用三角多项式通过预先设置好长度的窗口来求得插值点的幅值。本文实现了所设计的三种算法，这些算法可实现任意采样速率之间的转换，并支持上采样和下采样。

在音频信号转换时间开销、信噪比和段信噪比等方面的测评结果表明所设计的三种算法既可以保证较高的重构信号质量，同时还具有较高的转换效率。

关键词： 数字信号处理 多抽样率信号处理 音频重采样 插值 抽取

Abstract

With the development of theory and algorithms of digital signal processing, multi-rate signal processing becomes more and more important in multimedia signal processing. In order to reduce both computational and storage complexities, sampling rates conversion techniques are definitely required, audio resampling algorithm can be used to convert audio signal between arbitrary sampling rates.

Firstly, the basic concepts and the principle of multi-rate signal processing is introduced briefly in the paper; then three audio resampling algorithms, i.e., direct interpolation, Lagrange interpolation and sine interpolation algorithms, are designed and implemented. Direct interpolation method calculates the magnitude of the interpolation points linearly based on the two adjacent points, while Lagrange interpolation algorithm calculates the magnitude of the interpolation points using the Lagrange polynomial with a predefined window length, the sine interpolation method calculates the magnitude of the interpolation points based on the trigonometric polynomial with the dynamically chosen window length. Arbitrary sampling rates conversions have been implemented with the designed algorithms supporting both up-sampling and down-sampling.

Results of the evaluations, including time consumption, SNR and SegSNR, have shown that the three proposed algorithms can assure both high quality of the reconstruction signals and high conversion efficiency.

Keyword: Digital Signal Processing Multi-Rate Signal Processing

Audio Resampling Interpolation Decimation

目 录

第一章 绪论.....	1
1.1 多抽样率信号处理的引入及发展	1
1.2 多抽样率信号处理在多媒体技术中的应用	2
1.3 论文研究内容	3
1.4 论文结构安排.....	3
第二章 重采样技术简介.....	5
2.1 重采样.....	5
2.2 插值及抽取.....	6
2.2.1 整数倍抽取.....	7
2.2.2 整数倍插值.....	10
2.2.3 整数倍抽取和内插的应用	13
2.3 抽样系统中的恒等变换	14
2.4 多抽样系统中的易位	15
2.5 滤波器的实现	17
2.5.1 重采样的滤波器实现.....	17
2.5.2 多级实现.....	20
第三章 音频重采样算法设计与实现.....	25
3.1 音频重采样	25
3.2 音频重采样算法设计	26
3.2.1 线性重采样.....	26
3.2.2 拉格朗日重采样.....	27
3.2.3 正弦重采样.....	29
3.3 重采样算法的实现	30
3.3.1 线性重采样算法.....	30
3.3.2 拉格朗日重采样算法.....	36
3.3.3 正弦重采样算法.....	42
3.4 本章小结	47
第四章 性能评估.....	49
4.1 音频测试.....	49
4.2 性能评估与结论.....	50
4.2.1 时间开销.....	50
4.2.2 信噪比.....	52
4.2.3 段信噪比.....	55
4.3 小结	58

第五章 结论	59
致谢	61
参考文献	63
研究成果	67

第一章 绪论

1.1 多抽样率信号处理的引入及发展

随着数字信号处理的发展，信号的处理、编码、传输和存储等工作量越来越大。为了节省计算工作量及存储空间，在一个信号处理系统中常常需要不同的采样率及其相互转换，在这种需求下，多抽样率信号处理产生并发展起来。它的应用带来许多好处，例如：可降低计算复杂度、降低传输速率、减少存储量等^[1]。

在信号处理领域，多抽样率信号处理^[2-5]最早于20世纪70年代提出，在多抽样率信号处理发展中，一个突破点是70年代两通道正交镜像滤波器组应用于语音信号的压缩。在该方法中，信号通过分析滤波器组被分成低通和高通两个子带，每个子带经过2倍抽取和量化后再进行压缩，之后可以通过综合滤波器组近似地重建出原始信号，重建的近似误差一部分源于子带信号的压缩编码，一部分是由分析和综合滤波器组产生的误差，其中最主要的误差是混叠误差，它是由分析滤波器组不是理想带限而引起的。在很多应用系统中，混叠误差存在一定程度的影响，因此就需要对其进行改进，从此多抽样率数字信号处理得到了众多学者的重视。从20世纪80年代开始，多抽样率信号处理理论在各个领域得到了蓬勃发展，各种理论研究成果和应用层出不穷，促使了整个信号处理领域的发展。20世纪末，关于消除混叠和准确重建的理论已经得到了充分的发展。

1981年Crochier R.E.和Rabiner L.R.发表了一篇著名的关于多抽样率信号处理系统的基本模块——内插和抽取的综述性文章^[6]。随后，Vaidyanathan P.P.发表了许多与多抽样率信号处理系统相关内容的著作。从此，这一领域得到了快速的发展，特别是在多抽样率滤波器组的设计方面，涌现了多种准确重建滤波器的形式。在文献[1]中提到了多抽样率系统应用于通信、语音信号处理、谱分析、雷达系统和天线系统，以及在数字音频系统、子带编码技术(用于声音和图像的压缩)和模拟语音个人系统(如标准电话通信)等方面的应用。另外文献[1]中还提出了多相理论和多抽样率系统在一些非传统领域的应用，包括：高效率信号压缩的多抽样率理论；高效窄带滤波器的脉冲响应序列的编码新技术的推导；可调整的多级FIR滤波器的设计等。基于上述研究的发展，从20世纪80年代初开始，多抽样率处理技术在工程实践中得到广泛的应用，主要用于通信系统、语音、图像压缩、数字音频系统、统计和自适应信号处理、差分方程的数值解等。

多抽样率是指在一个系统中存在着两个或两个以上的抽样率。使抽样率降低的抽样率转换称为抽取，亦称抽样率压缩；使抽样率升高的抽样率转换称为内插，亦称抽样率扩张。传统的完成信号抽样率的转换有两种方法：模拟方法和数字方

法。直观来讲任何抽样率的变换都可以通过将抽样信号 $x(m)$ 经过D/A转换还原到带限的模拟信号 $x_a(t)$ ，再对它进行不同的速率采样(经过模数转换变成数字信号)得到新的离散信号 $x'(n)$ ，从而完成抽样率从 F_1 到 F_2 的转换。这种传统方法的过程比较复杂，而且由于量化噪声等的引入，容易造成信号失真。这种传统方法的具体步骤如图1.1所示。

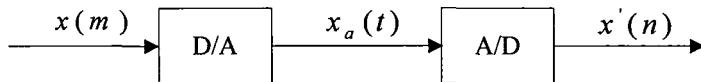


图1.1 传统的抽样率转换的方法

因此人们采用数字方法来变换抽样率，所谓数字方法就是完全用数字处理的方法完成抽样率的转换，而不必将信号在数字域和模拟域之间不断的转换。这种采用数字方法实现抽样率转换^[7-9]的方法就是多抽样数字信号处理。

1.2 多抽样率信号处理在多媒体技术中的应用

多抽样率信号处理中的滤波器组^[10,11]理论可广泛地应用于音/视频的压缩、编码和识别。滤波器组最初用于语音压缩，后来逐渐应用于图像、视频的压缩。将信号分成不同的子带，然后根据不同子带包含的信息量的不同分别处理，通过舍弃包含不重要信息的子带、保留含有重要信息的子带来实现压缩。多年来，滤波器组理论在音/视频信号领域一直备受关注。

在音/视频编码方面，滤波器组最重要的使用形式就是子带编码。在音频子带编码中，使用一维滤波器将语音信号分解成一系列子带；通过内插、滤波和叠加单个子带可以无失真地重建原始语音信号。在图像和视频的子带编码中使用的是二维可分离滤波器，即通过将滤波器先应用于某一维(如垂直向)，再应用于另一维(如水平向)来实现。通常图像编码的方法是：把一个图像的傅里叶频谱分成若干个互不重叠的子带，然后对每一子带进行反变换，得到一组带通图像；再对各个带通图像进行二次采样，并用不同的比特率进行编码，这个比特率的选择要和各个子带所包含的信息量以及主观视觉能力的要求相匹配。

音/视频的识别技术从20世纪90年代开始成为研究热点，主要包括语音识别和指纹识别。各类识别技术普遍采用的方法就是将滤波器组对原始信号进行频带分割，然后提取相关特征，与预设的模板进行匹配。Sang-II Park等提出了一种方向滤波器组的改进结构，用高效树形结构来实现。这种结构保持了子带域的可见信息，可实现高效计算，并且可以达到视觉上的完全重建。Chul-Hyun Park等提出了一个基于方向滤波器组的指纹匹配方法：将图像的方向能量作为一个显著的特征来进行指纹匹配，用方向滤波器组将一个指纹图像分解为几个方向子带输出信号，从子带输出中获得每一块的方向能量分布，此方法能极大地减少所需存储器容量。

和匹配时间。

多抽样率信号处理还可以应用在多媒体版权保护方面，如音频数字水印技术，通过在原始音频数据中嵌入特殊的辨识信息(水印)，作者的著作权可以得到保护。使用正交镜像滤波器(QMF: Quadrature Mirror Filter)组对音频原数据进行频带分割，选择包含大量能量的低频子带嵌入水印，可以减少噪声影响并提高该音频数字水印的安全性和鲁棒性。

1.3 论文研究内容

本文研究了多抽样率信号处理及其关键技术，在此基础上详细研究了多抽样率信号处理中的插值和抽取技术，并将这两种技术应用在音频重采样中。设计了三种音频重采样算法：线性重采样、拉格朗日重采样、正弦重采样，详细介绍了三种重采样算法的设计思想以及三种重采样算法的异同。每种重采样算法都针对下采样和上采样的不同情况，提出了具体的解决方案。并分别用这三种不同的重采样算法，在实现音频信号在常用速率之间转换的同时，保证了经过重采样的音频信号的质量。

本文进一步对采用三种重采样算法生成音频信号的质量，具体从时域和频域进行了对比，总结了每种算法适合应用于哪种音频信号的处理；记录每种算法在上下采样时的时间开销，阐明了三种重采样算法在时间开销方面的优劣，结合三种重采样算法的算法复杂度，总结了三种重采样算法分别在上采样和下采样时的优缺点。提出了三种重采样算法在处理哪种音频信号的时候，既能得到好的重采样效果，又能在时间的开销方面得到优化。

最后针对三种重采样算法生成的音频信号进行了客观测试，将每种算法具体在时间开销，信噪比，段信噪比方面进行了计算。分析比较了每种算法在各常用速率之间转换的时间开销，信噪比，段信噪比，依据具体的数据，针对每种重采样算法应用于不同速率之间转换时生成音频信号的效果进行评估，得出结论。

1.4 论文结构安排

本文研究了如何将重采样技术应用于在音频信号处理中，在此基础上设计了三种音频重采样算法：线性重采样、拉格朗日重采样、正弦重采样。对每种算法进行重采样生成的音频信号进行研究、分析和总结。根据三种算法的特点，将其应用于不同的音频重采样信号。具体论文结构如下：

第一章：绪论，简要介绍多抽样率的发展以及多抽样率在多媒体音\视频方面的应用。

第二章：重采样技术简介，介绍了重采样的原理，要实现重采样具体涉及到

的一些关键技术，如何解决重采样时遇到的具体问题，以及实现重采样的意义。

第三章：音频重采样算法设计与实现，设计了音频重采样的三种具体算法，介绍了三种重采样算法的思想与原理，分别采用三种重采样算法，实现了音频信号在各常用速率之间的转换。针对每种重采样算法，详细介绍了上采样和下采样的基本原理和算法，对每种算法在各速率之间转换的时间开销进行了统计。从时域和频域分析比较了三种算法在各常用速率之间转换的优劣。

第四章：性能评估，介绍了音频性能测试分析的基本方法：主观测试，客观测试。对三种重采样算法生成的音频信号进行了客观测试，对用每种算法在实现音频信号采样率转换的时间开销进行了分析和比较，从算法设计方面分析了具体时间开销大小的原因；还从信噪比和段信噪比方面对每种算法进行测试，以及将三种算法的信噪比和段信噪比进行对比分析。

第五章：总结，总结了本文的主要研究和工作。

第二章 重采样技术简介

重采样^[12,13]即将原始的采样频率变换为新的采样频率以适应不同采样率的要求。重采样系统的基本模块是抽取器和内插器，它们很早就应用于数字系统的设计。由于抽取可能产生混叠，内插产生镜像，所以需要在抽取前进行抗混叠滤波，在内插后进行抗镜像滤波。为简化系统的分析和设计，可以根据抽取和插值的等效变换定理，将抽取器(插值器)和抗混叠(去除镜像)滤波器位置等效互换，引入多相分析，获得一个等效的滤波器。信号的整数倍抽样率转换一般由抽取器(内插器)和抗混叠(抗镜像)滤波器构成，因此均匀滤波器组可以用多项矩阵-FFT 结构滤波器组来实现。但是，在工程实践中，常常需要对信号进行分数倍抽样转换，这一过程通常是由整数倍抽取器、内插器和滤波器结合实现的。对于分数倍采样因子的系统，还需要设计最小延迟的有效结构进行延迟分析。另外，在工程实践中，还可以通过使用频率响应屏蔽技术(FRM: Frequency-Response-Masking Technique)的变体来降低因抽取或者插值滤波器的传输带宽的严格要求而引入的复杂性。

2.1 重采样

在对信号进行采样时，只要满足Nyquist采样定理^[14,15]（采样频率大于等于信号最高频率的两倍），就可以通过插值函数完整地将信号恢复出来。设有一带限信号 $x(t)$ ，其最高频率为 F_0 ，对其进行采样，采样周期为 T ，并且满足Nyquist采样定理，产生的采样序列为 $x(n)$ 。由于满足Nyquist采样定理，我们可以从 $x(n)$ 中恢复出 $x(t)$ ，一个理想的重构系统如2.1图所示。

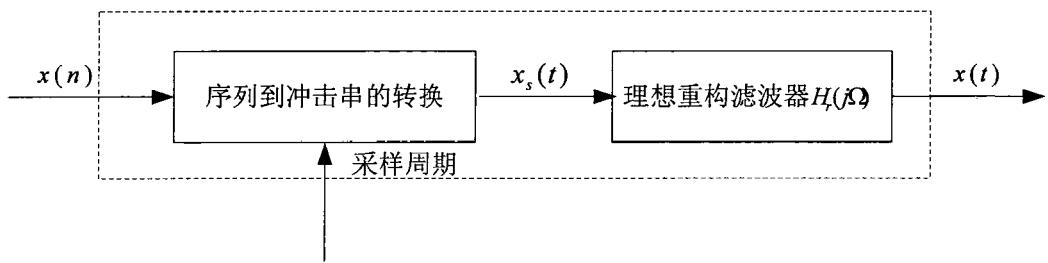


图2.1 理想重构系统

其中， $x(t)$ 是 $x(n)$ 由形成的冲激串：

$$x(t) = \sum_{n=-\infty}^{+\infty} x(n)h_r(t-nT) \quad (2-1)$$

滤波器的频率响应 $H_r(j\Omega)$ 如图 2.2 所示。

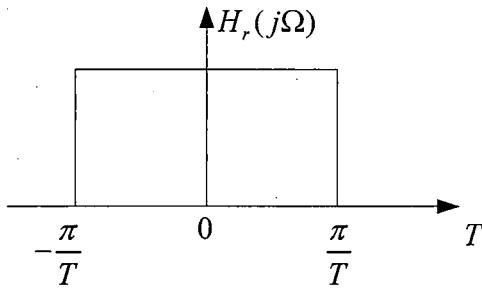


图 2.2 重构滤波器的频率响应

根据重构滤波器的频率响应，可以得到 $H_r(j\Omega)$ 的傅立叶反变换。

$$h_r(t) = \frac{\sin(\pi t/T)}{\pi t/T} \quad (2-2)$$

由公式(2-1)和(2-2)，可以得到重构公式

$$x(t) = \sum_{n=-\infty}^{+\infty} x(n) \frac{\sin[(\pi(t-nT)/T)]}{\pi(t-nT)/T} \quad (2-3)$$

现在对信号以采样周期 T' 进行重采样，产生的新序列为 $x'(n) = x(nT')$ ，根据公式(2-3)，可以得到 $x'(n)$ 的重构公式。

$$x'(n) = x(nT') = \sum_{n=-\infty}^{+\infty} x(n) \frac{\sin[\pi(nT' - nT)/T]}{\pi(nT' - nT)/T} \quad (2-4)$$

实际中，在进行插值重构时，不可能用所有的采样点来进行插值重构，而是对于时域上任意一点的插值，通过对其周围一定数量的采样点来实现。

$$x'(n) = x(nT') = \sum_{n=M-N}^{M+L+N} x(n) \frac{\sin[\pi(nT' - nT)/T]}{\pi(nT' - nT)/T} \quad (2-5)$$

其中， M 是距离时刻 nT' 最近的采样点的标号，并且 M 不超过 nT' 。对于 $x(nT')$ ，用 nT' 时刻周围的 $2N+2$ 个采样点的值来对其进行重构。

2.2 插值及抽取

为了更加有效地传输和存储音频信息，音频信号在处理中，常常需要改变信号的采样速率。如对窄带带通信号，若原来采样速率过高，可通过降低采样速率(即：抽取^[16-19])来减少数据冗余，降低计算量；而对采样 A/D 变换等应用，通过对信号的升高采样速率(即：插值^[20-24])可以获得更好的性能。数字系统存在不同采样速率的数字信号，因此需要将一种采样速率的信号转换为另一种采样速率的信号，例如：符合 AC97 标准的声卡最后输出采样频率都定在 48kHz，而实际中 CD 的音频是 44.1kHz 的。从 CD 转换过来的 MP3 等大多部格式也是 44.1kHz 的，也就是说

在这种类型的声卡上，在数字信号转成模拟信号之前，还要对采样频率做一次转换。

在数字信号处理领域，常常遇到采样率变换的问题，如采样率升高 M 倍或者降低 L 倍，这样的问题就是重采样问题。

某带限时间波形 $x(t)$ ，以采样频率 f_s (在满足采样定律的条件下)进行采样，得到时间序列 $x(k)$ 。对 $x(k)$ 进行采样率变换，如升高采样率，提高采样率 M (正整数)倍，升高采样率后的序列用 $y(m)$ 表示，那么，升高采样率前后序列的关系为。

$$y(m) = \sum_{k=-\infty}^{\infty} x(k)h(m - kM) \quad (2-6)$$

$$h(m) = \sin c(m / M) \quad (2-7)$$

式(2-6)表明了首先对输入序列的每一样点之间添加 $(M - 1)$ 个零，然后通过截止频率为 $1/M$ (归一化频率)的低通滤波器，即可实现升高采样率变换。

如果进行降低采样率，如采样率降低 L (正整数)倍(当然降低后的频率还是满足采样定理)，降低采样率后的序列用 $y(m)$ 表示，那么，降低采样率前后序列的关系为：

$$y(m) = w(mL) = \sum_{m=-\infty}^{\infty} x(k).h(mL - k) \quad (2-8)$$

$$h(m) = \sin c(m / L) \quad (2-9)$$

式(2-8)表明了首先让输入序列通过截止频率为 $1/L$ (归一化频率)的低通滤波器，然后每 L 样点抽取一个样点，即可以实现降低采样率变换。

这里所讨论的升高采样率和降低采样率的倍数均是正整数的情况，对于有理数的情况，可以用 M/L 去逼近。

综合式(2-6)和式(2-8)，我们可以得到。

$$y(m) = w(mL) = \sum_{K=-\infty}^{\infty} x(k)h(mL - kM) \quad (2-10)$$

$$h(m) = \sin c(m/M_l) \quad (2-11)$$

$$M_l = \max(L, M) \quad (2-12)$$

式(2-10)就是重采样的计算式，如均匀到均匀，非均匀到非均匀的插值，这就是重采样插值的原理。但是用上式直接实现重采样插值计算时，由于 $h(m)$ 为无限长序列，上式的计算量很大。

2.2.1 整数倍抽取

当信号的抽样数据量太大时，为了减少数据量以便于处理和计算，把抽样数据每隔 $D - 1$ 个取一个，这里 D 是一个整数。这样的抽取称为整数倍抽取， D 为抽

取因子。图 2.3 为一般的整数倍抽取示意图。

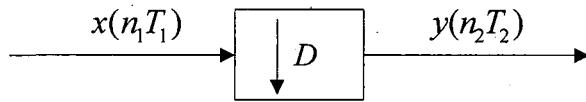


图 2.3 整数倍抽取示意图

$x(n_1 T_1)$ 的抽样周期为 T_1 , 相应的抽样率为 F_1 , 进行整数倍抽取后所得的新序列为 $y(n_2 T_2)$, $y(n_2 T_2)$ 的抽样周期为 T_2 , 相应的抽样率为 F_2 。由于每隔 D 个 T_1 抽取一个数据, 所以 $T_2 = D T_1$

以上是从时域中讨论整数倍抽取是怎样实现的, 看起来极为简单, 只要每隔 $D-1$ 个样值抽取一个就行了, 其实并非如此, 因为如果 $x(n_1 T_1)$ 是模拟信号 $x(t)$ 的抽样信号, 则 $x(t)$ 与 $x(n_1 T_1)$ 的傅里叶变换 $X(j\Omega)$ 与 $X(e^{j\Omega T_1})$ 将分别是

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t) e^{-j\Omega t} dt \quad (2-13)$$

$$X(e^{j\Omega T_1}) = \sum_{n_1=-\infty}^{\infty} x(n_1 T_1) e^{-j\Omega T_1 n_1} \quad (2-14)$$

而 $X(e^{j\Omega T_1})$ 与 $X(j\Omega)$ 的关系式。

$$X(e^{j\Omega T_1}) = \frac{1}{T_1} \sum_{k=-\infty}^{\infty} X(j\Omega - jk \frac{2\pi}{T_1}) \quad (2-15)$$

式中:

$$\Omega = 2\pi f \quad (2-16)$$

如果定义:

$$\omega_1 = \Omega T_1 = 2\pi(f / F_1) \quad (2-17)$$

$$\text{则(2-15)式可以写成: } X(e^{j\omega_1}) = \frac{1}{T_1} \sum_{k=-\infty}^{\infty} X(j\Omega - jk\Omega_{sa1}) \quad (2-18)$$

$$\text{式中 } \omega_1 \text{ 称为归一化角频率: } \Omega_{sa1} = \frac{2\pi}{T_1} \quad (2-19)$$

图 2.4 为原始模拟信号, 在满足抽样定理情况下的原抽样信号以及将抽样率降低 D 倍的信号以及它们在频率域的频谱。图 2.4 (a) 中 $x(t)$ 为原始模拟信号, $|X(j\Omega)|$ 为其在频率域的频谱, 图 2.4 (b) 中 $x(n_1 T_1)$ 为满足抽样定理的时域信号, $X(e^{j\Omega T_1})$ 为其在频率域的频谱, 图 2.4 (c) 中 $y(n_2 T_2)$ 为抽样率降低 D 倍后, 不满足抽样定理的时域信号, $Y(e^{j\Omega T_2})$ 为其在频率域的频谱。

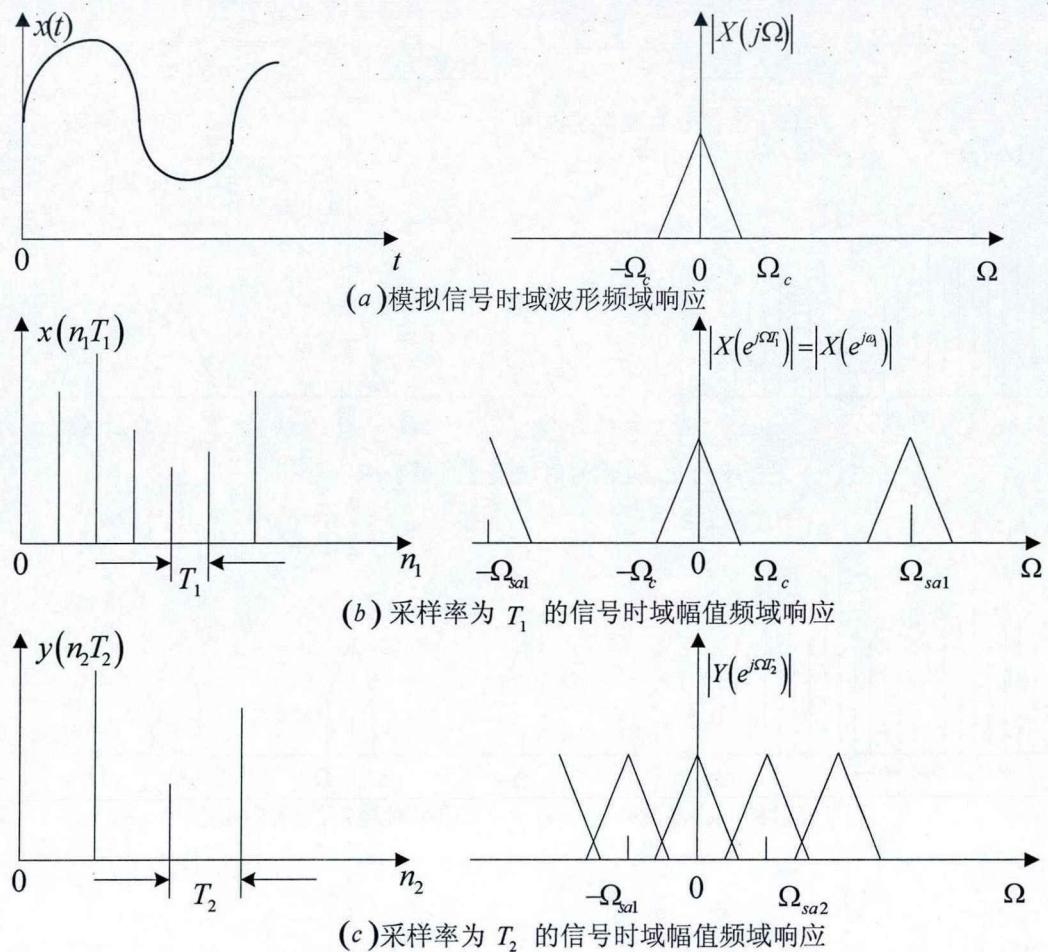


图 2.4 模拟信号及抽样信号时域波形，频域响应

从图 2.4 可以看出，只有在抽取之后的抽样率仍然满足抽样定理的要求时才能会恢复出原始信号，否则就会产生混叠，难以恢复出原始信号。

这时需要抗混叠滤波，抗混叠滤波抽取系统如图 2.5 所示。

$x(n_1 T_1)$ 为满足抽样定理的时域信号， $X(e^{j\Omega T_1})$ 为其在频率域的频谱， $v(n_1 T_1)$ 为在满足抽样定理时，经过抗混叠滤波的时域信号， $V(e^{j\Omega T_1})$ 为其在频率域的频谱， $y(n_2 T_2)$ 为 $v(n_1 T_1)$ 经过 D 倍抽取，不满足抽样定理的时域信号， $Y(e^{j\Omega T_2})$ 为其在频率域的频谱。

要想降低抽样率又能很好的恢复出原始信号就得寻找其它方法，通常采取的是抗混叠滤波。所谓的抗混叠滤波就是在抽取之前，对信号进行低通滤波，把信号的频带限制在 $\Omega_{sa2}/2$ 以下。

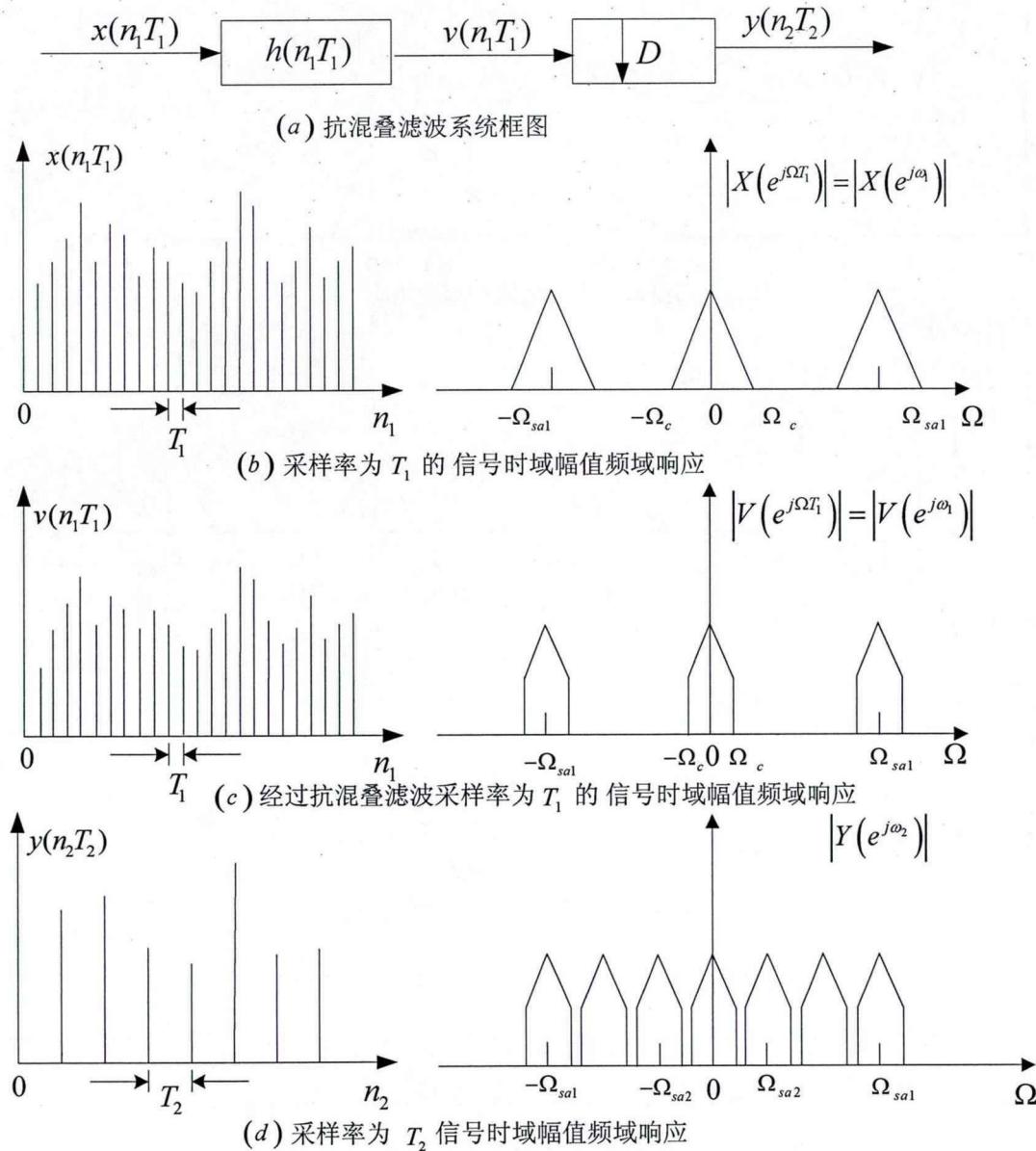


图 2.5 抗混叠系统框图

2.2.2 整数倍插值

整数倍内插是先在已知抽样序列 $x(n_1T_1)$ 的相邻抽样点之间等间距地插入 $(I-1)$ 个零值，然后进行低通滤波，即可求得 I 倍内插的结果。

$x(n_1T_1)$ 零值内插后，得到 $v(n_2T_2)$ 。 $v(n_2T_2)$ 经过 $h(n_2T_2)$ 低通滤波变成 $y(n_2T_2)$ 。具体过程如图 2.6。

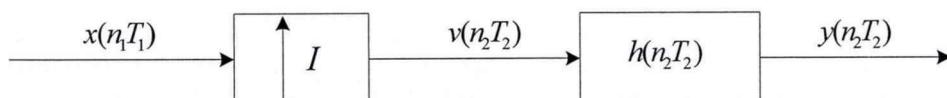


图2.6 内插系统框图

$x(n_1 T_1)$ 采样率为 T_1 的信号, $v(n_2 T_2)$ 为对 $x(n_1 T_1)$ 插零值后的信号, $y(n_2 T_2)$ 为 $v(n_2 T_2)$ 经过滤波后的信号, 具体在时域幅值如图 2.7 所示。

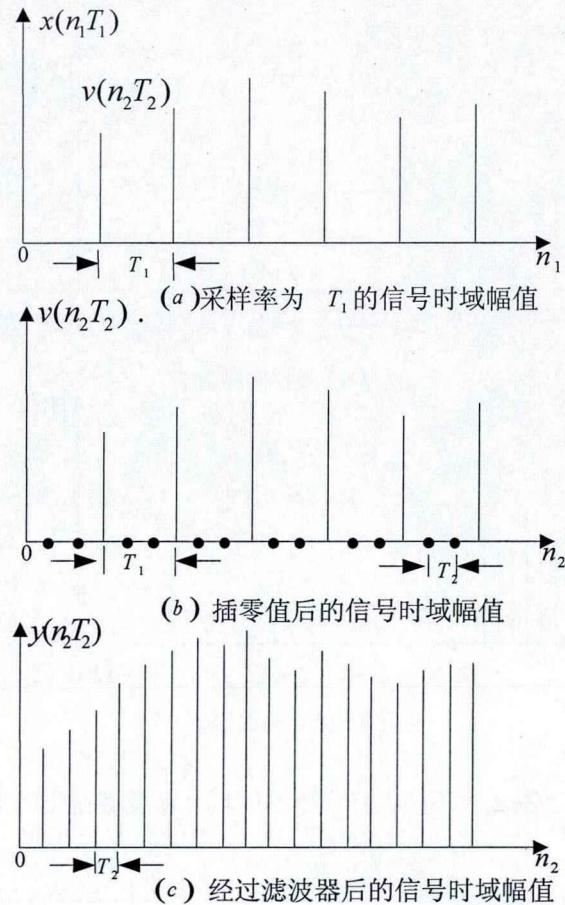


图 2.7 $x(n_1 T_1)$, $v(n_2 T_2)$ 及 $y(n_2 T_2)$ 时域幅值

先将 $x(n_1 T_1)$ 插零值变成 $v(n_2 T_2)$, 再把 $v(n_2 T_2)$ 通过低通滤波变成 $y(n_2 T_2)$ 。为何 $v(n_2 T_2)$ 经过低通滤波就能得到 $y(n_2 T_2)$?

回到 $x(n_1 T_1)$ 所代表的原始模拟信号 $x(t)$ 。设 $x(t)$ 及其傅里叶变换 $X(j\Omega)$ 如图 2.8 所示。

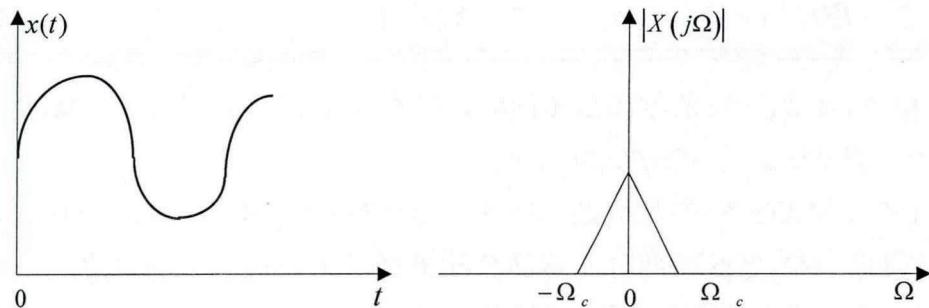


图 2.8 模拟信号时域波形及其傅里叶变换

设 $x(n_1 T_1)$ 和 $y(n_2 T_2)$ 分别为以抽样时间间隔 T_1 和 T_2 对 $x(t)$ 进行抽样, 则傅里叶变换为 $X(e^{j\omega_1})$ 和 $Y(e^{j\omega_2})$ 。 $X(e^{j\omega_1})$ 和 $Y(e^{j\omega_2})$ 都是周期函数, 如果用真实角频率 Ω 表

示，则 $X(e^{j\omega_1}) = X(e^{j\Omega T_1})$ ，其周期为 $\Omega_{sa1} = 2\pi/T_1$ ；同理， $Y(e^{j\omega_2}) = Y(e^{j\Omega T_2})$ ，其周期为 $\Omega_{sa2} = 2\pi/T_2 = I(2\pi/T_1) = I\Omega_{sa1}$ 。 $X(e^{j\omega_1})$ 和 $Y(e^{j\omega_2})$ 如图 2.9 所示。

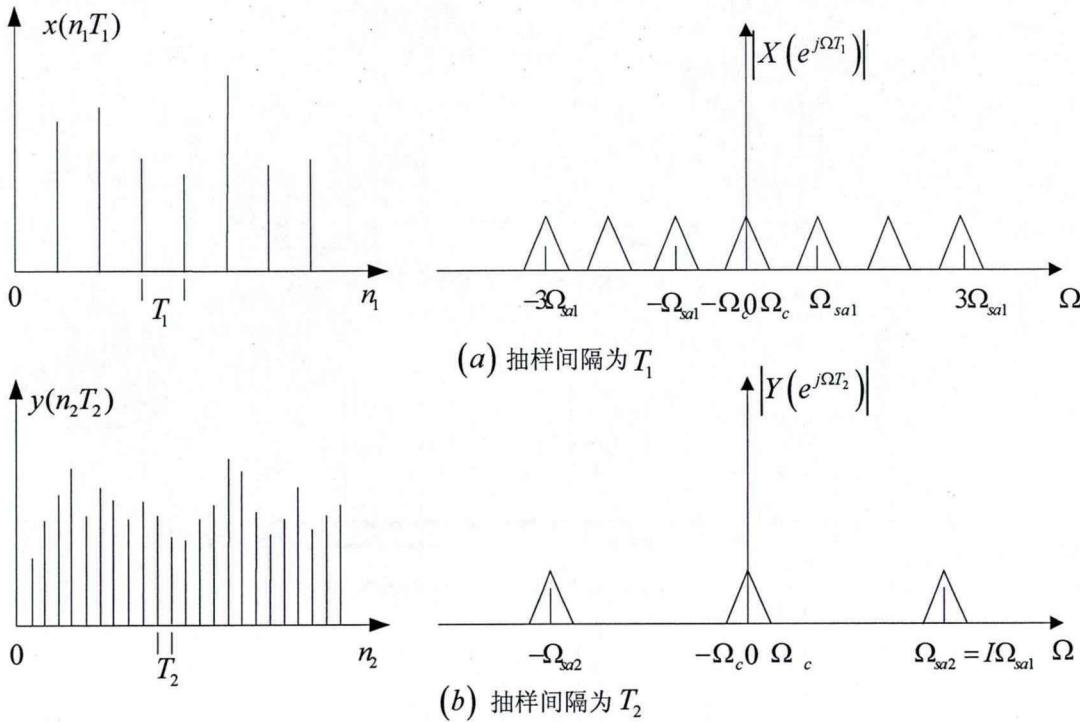


图 2.9 不同抽样率下的时域幅值及其傅里叶变换

$$v(n_2 T_2) = \begin{cases} x(n_2 \frac{T_1}{I}) & \text{当 } n_2 = 0, \pm I, \pm 2I, \dots \\ 0 & \text{其它} \end{cases} \quad (2-20)$$

于是 $V(e^{j\omega_2}) = \sum_{n_2=-\infty}^{\infty} v(n_2 T_2) e^{-j\omega_2 n_2} = \sum_{n_2=-\infty}^{\infty} x(\frac{n_2}{I} T_1) e^{-j\Omega T_1 n_2 / I} \quad (2-21)$

$$\frac{n_2}{I} = n_1 \quad (2-22)$$

所以： $V(e^{j\omega_2}) = \sum_{n_1=-\infty}^{\infty} x(n_1 T_1) e^{-j\Omega T_1 n_1} = X(e^{j\omega_1}) \quad (2-23)$

可见 $V(e^{j\omega_2})$ 和 $X(e^{j\omega_1})$ 的频谱是相同的，只不过 $X(e^{j\omega_1})$ 是以 $\Omega_{sa1} = 2\pi/T_1$ 为周期，而 $V(e^{j\omega_2})$ 则是以 $\Omega_{sa2} = 2\pi/T_2$ 为周期。

既然 $V(e^{j\omega_2})$ 和 $X(e^{j\omega_1})$ 的频谱是相同的，那么 $V(e^{j\omega_2})$ 相较于 $Y(e^{j\omega_2})$ 而言，就多出了 Ω_c 到 $\Omega_{sa2} - \Omega_c$ 之间的部分。这部分频谱好像是在 $\Omega = \Omega_c$ 处树立了一面镜子，在镜子里现出来的频谱，所以称这部分频谱为镜像。现在可以看出要想从 $V(e^{j\omega_2})$ 得到 $Y(e^{j\omega_2})$ 并不困难，只需将 $V(e^{j\omega_2})$ 通过以 Ω_c 为通带边缘的滤波器即可。图 2.10 为滤波器理想频率幅值响应。

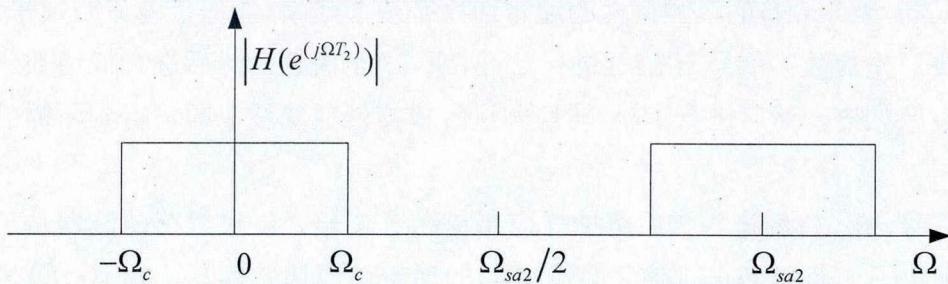


图 2.10 低通滤波器的理想幅频特性

2.2.3 整数倍抽取和内插的应用

在对模拟信号 $x(t)$ 进行 A/D 转换的困难在于对抗混叠滤波器的要求太高，要求滤波器的过渡带太窄，这个滤波器的阶次太高，难以设计和实现这个滤波器。但是如果将过渡带放宽之后，则滤波后的信号 $v(t)$ 中又必然有 $f > f_h$ 的成分，只是用抽样率 $F = 2f_h$ 进行 A/D 转换后， $v(nT)$ 的频谱发生混叠。如果使用高抽样率进行抽样则又会使得信号的数据量太大，引起处理信号时计算量和存储量的增加。

模拟滤波器技术要求太高的原因是因为 A/D 转换所用的抽样速率太低，要求 $h(t)$ 的过渡频带为 0，否则将产生严重混叠。于是针对这个问题加大抽样率以使模拟滤波器 $g(t)$ 易于设计和实现。至于抽样率太高的问题则用抽取的方法解决。

一个数字音频系统的改进方案：电话系统中 4kHz 以下的声音频率是我们要保证的，但是送话器发出声音的频谱比 4kHz 宽得多。因此在将模拟声音信号进行 A/D 转换变为抽样信号之前需要进行模拟低通滤波。如果 A/D 转换的抽样率只用 8kHz，则要求模拟滤波器的过度带宽为 0，但是如果将抽样率提高到 16kHz，经过 A/D 后再把抽样信号经过 2 倍抽取把抽样率降到 8kHz，则模拟滤波器的过渡带是从 4kHz 到 12kHz。这样做，在 4kHz 到 12kHz 的频带中是有混叠的，但这部分的混叠经过抽取器前的数字滤波后便被滤掉了。A/D 后的抽样信号则这样把设计模拟滤波器的困难改由设计数字滤波器来解决，这是解决问题的关键所在，所用的方法则是先提高 A/D 的抽样率然后再进行抽取，所以数据量并没有增加。

当数字信号 $y(n_2 T_2)$ 经过传送到达接收端以后，变成 $\hat{y}(n_2 T_2)$ ，又会遇到从数字变为模拟信号的问题，即 D/A 转换问题。

将 $\hat{y}(n_2 T_2)$ 经过 D/A 转换器转换为 $\hat{y}(t)$ 有两种不同的想法。一种是对 $\hat{y}(n_2 T_2)$ 的每一个抽样值进行保持得出台阶形的连续信号 $\hat{y}(t)$ ，然后对 $\hat{y}(t)$ 进行低通滤波滤除高频噪声，得出 $\hat{x}(t)$ 。还有一种想法就是先对 $\hat{y}(n_2 T_2)$ 进行零值内插，并使零值内插因子 $I \rightarrow \infty$ ，这样便得出另一形式的连续信号 $\hat{y}(t)$ ，将 $\hat{y}(t)$ 通过低通滤波器 $\hat{h}(t)$ 进行滤波便可得到 $\hat{x}(t)$ 。

使用模拟滤波器 $\hat{h}(t)$ 滤出 $|\Omega| < \Omega_{sa2}/2$ 区间的频谱是有困难的。困难在于对 $\hat{h}(t)$

技术要求太高, 要求 $\hat{h}(t)$ 的幅频特性的通带边缘必须很陡峭, 而通带内又要很平坦。如何来解决? 先对数字信号 $\hat{y}(n_2 T_2)$ 进行 2 倍(或 3、4)内插, 然后进行低通数字滤波, 滤去镜像频谱, 然后进行 D/A 转换并进行模拟低通滤波, 即可得到所需的模拟信号。

数字信号 $\hat{y}(n_2 T_2)$ 经过内插后将抽样速率提高了 2 倍, 经过数字滤波器 $\hat{h}(n_1 T_1)$ 得到信号 $\hat{v}(n_1 T_1)$, 设计 $\hat{h}(n_1 T_1)$ 使之具有陡峭的通带边缘比较容易。对 $\hat{v}(n_1 T_1)$ 进行 D/A 转换, 使:

$$\hat{v}(t) = \begin{cases} \hat{v}(n_1 T_1) & t = n_1 T_1 \\ 0 & t \neq n_1 T_1 \end{cases} \quad (2-24)$$

显然, $\hat{v}(t)$ 的频域表示 $\hat{V}(j\Omega)$ 的形式与 $\hat{v}(n_1 T_1)$ 的频域形式 $\hat{V}(e^{j\omega_1})$ 相同, 不过其周期是无限长的。所以 $\hat{V}(j\Omega)$ 是个非周期的连续频谱。

对 $\hat{v}(t)$ 进行模拟低通滤波, 这时模拟滤波器 $\hat{h}(t)$ 的通带边缘为 $|\Omega_p| = \pi/2T_1$, 过渡带为 $\pi/2T_1 < \Omega < 3\pi/2T_1$, 阻带为 $\Omega > 3\pi/2T_1$ 。滤波器的输出为模拟信号 $\hat{x}(t)$ 。这个模拟低通滤波器的过渡带较宽, 在设计上就不困难了。滤波器的输出 $\hat{x}(t)$ 就是所要的模拟信号了。

2.3 抽样系统中的恒等变换

多抽样系统中因为希望抽样率转换系统中的计算工作量最小, 每秒钟内的乘法次数最少, 所以在多抽样率网络的具体实现中常常把乘法安排在低抽样率的一侧, 因为这样的安排使计算的效率最高。

典型的两种网络恒等变换, 如图 2.11 和图 2.12 所示。

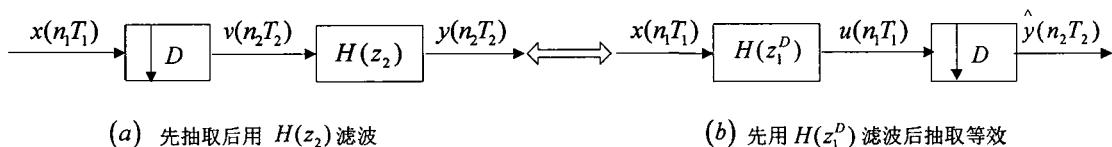


图 2.11 先抽取后用 $H(z_2)$ 滤波等效于先用 $H(z_1^D)$ 滤波后抽取

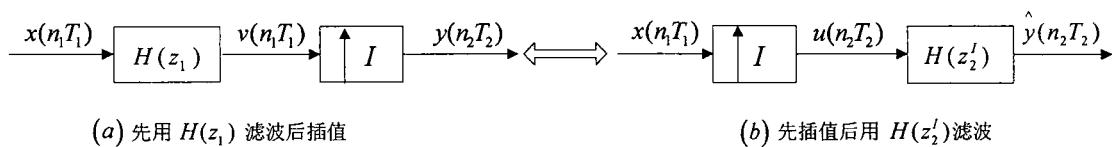


图 2.12 先用 $H(z_1)$ 滤波后插值等效于先插值后用 $H(z_2^I)$ 滤波

以图 2.11 中图(a)抽取为例, 从图 Z 变换可以看出其等效。

$$Y(z_2) = H(z_2)V(z_2) \quad (2-25)$$

$$V(z_2) = \frac{1}{D} \sum_{l=0}^{D-1} X(z_1 W^l) \quad (2-26)$$

所以

$$Y(z_2) = H(z_2) \frac{1}{D} \sum_{l=0}^{D-1} X(z_1 W^l) \quad (2-27)$$

在图 2.11 的图(b)中输入输出的关系为

$$\hat{Y}(z_2) = \frac{1}{D} \sum_{l=0}^{D-1} U(z_1 W^l) \quad (2-28)$$

$$U(z_1) = X(z_1) H(z_1^D) \quad (2-29)$$

$$U(z_1 W^l) = X(z_1 W^l) H\left[\left(z_1 W^l\right)^D\right] = X(z_1 W^l) H(z_1^D) \quad (2-30)$$

所以

$$\hat{Y}(z_2) = H(z_1^D) \frac{1}{D} \sum_{l=0}^{D-1} X(z_1 W^l) \quad (2-31)$$

比较 $\hat{Y}(z_2)$ 和 $Y(z_2)$ 可知，如果 $H(z_1^D) = H(z_2)$ 则 $\hat{Y}(z_2)$ 和 $Y(z_2)$ 相等。而又由于 $z_1^D = e^{j\omega_1 D} = e^{j\Omega(T_1 D)} = e^{j\omega_2} = z_2$ 。图 2.11 中的图(a),(b)是等效的。

用同样的方法可以证明图 2.12 中插值的两图是等效的。

既含抽取又含插值的等效变换：

先零值内插后抽取，且内插因子 I 等于抽取因子 D 时，零值内插与抽取的总效应等于直通，如图 2.13(a)，即总效应等于乘以常数 1. 但不等于先抽取后内插，如图 2.13(b)所示。

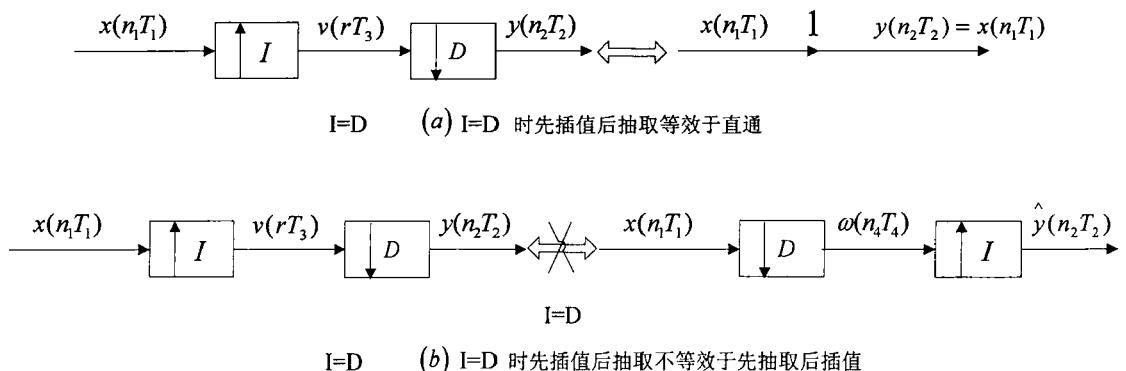


图 2.13 先插值与后抽和取先抽取后插值不同

2.4 多抽样系统中的易位

对线性时不变系统，易位是指把信号流的方向翻转过来，保留各支路的传递因子不变，同时将输入、输出信号调换位置。

对带有抽样率转换的系统，易位是指：(1)把各支路中的信号流方向翻转过来，保留各支路中的传递算子不变，(2)并保留网络中各节点处的抽样率不变，(3)还须把抽样率转换元件(抽取器以及内插器)用它的易位元件所代替，(4)还要把输入输出调换位置。这里所谓的易位元件是指：具有抽取因子 D 的抽取器的易位元件是具有相同因子的内插器。即，抽取的易位元件是插值，且 $I = D$ ；或具有内插因子 I 的内插器的易位元件是具有相同因子的抽取器，即插值的易位元件是抽取，且 $D = I$ 。

未经过易位的系统称为原系统，经过易位后的系统称为易位系统。对于单一抽样率的线性时不变系统，原系统和易位系统具有相同的转移函数。对于多抽样率的系统，原系统与易位系统具有对偶运算性质，一般不能保持其具有两种频率的系统函数(可称之为双频率系统函数)不变。

多抽样网络 S 如图 2.14 的图(a)，如何求出网络 S 的易位网络 \hat{S} ，以及 S 与 \hat{S} 的输入输出关系。

按照易位网络的步骤，画出 S 的易位网络 \hat{S} ，将 S 中的各支路传输方向翻转过来，将抽取器与内插器互换而不改变其抽样率转换因子的数值，保持各点的抽样率不变，并将输入输出易位，便得到了网络 \hat{S} ，如图 2.14 的图(b)。

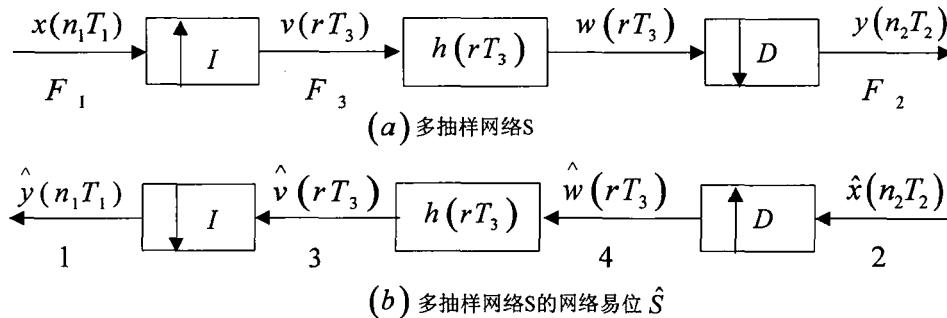


图 2.14 多抽样网络及其易位网络

易位网络 \hat{S} 中，输入为 $\hat{x}(n_2T_2)$ 经内插变成 $\hat{w}(rT_3)$ ，内插因子为 D ，于是 $T_3 = 1/DT_2$ ， $\hat{w}(rT_3)$ 经过低通滤波后变为 $\hat{v}(rT_3)$ ， $\hat{v}(rT_3)$ 经抽取后得到输出 $\hat{y}(n_1T_1)$ ，抽取因子为 I 。原网络 S 及易位网络 \hat{S} 中的运算过程的时域幅值分别如图 2.15 和图 2.16 所示。从图中可以看出原网络 S 的抽样率转换比值时 $f_{sa2}/f_{sa1} = T_1/T_2 = 2/4$ ，而易位网络 \hat{S} 的比值是 $f_{sa1}/f_{sa2} = T_2/T_1 = 4/2$ ，这里， f_{sa1} 和 f_{sa2} 分别是网络节点 1 和 2 处的抽样率，所以原网络是把抽样率从小变大，而易位网络是把抽样率从大变小，这是一种对偶的或互补的运算过程，而不是等效的过程。

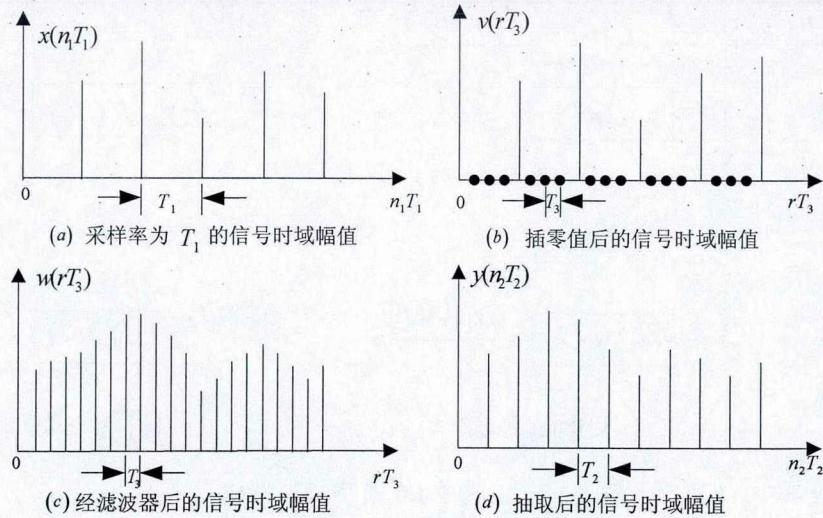
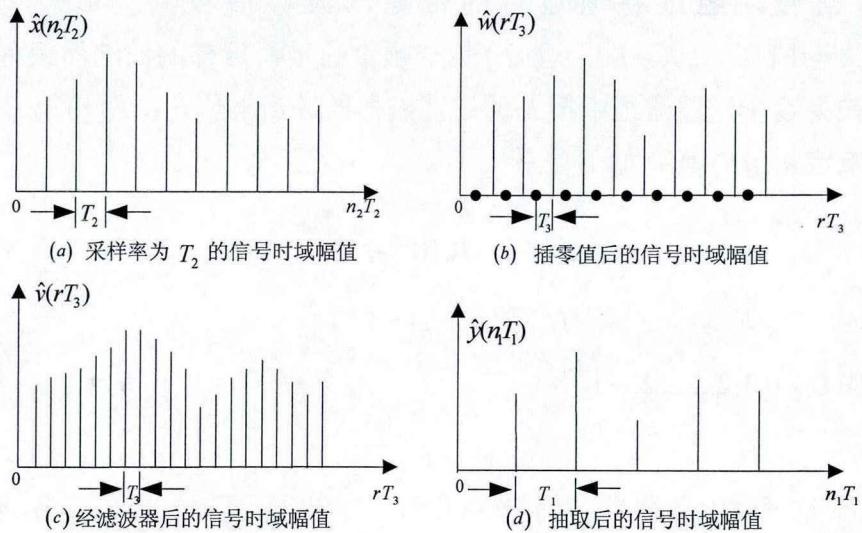


图 2.15 多抽样网络 S 具体的各抽样点的值

图 2.16 易位网络 \hat{S} 具体的各抽样点的值

2.5 滤波器的实现

2.5.1 重采样的滤波器实现

具有一个共同输入信号或一个共同输出信号的一组滤波器称为滤波器组。通常见到的最典型的两种滤波器组如图 2.17 所示。

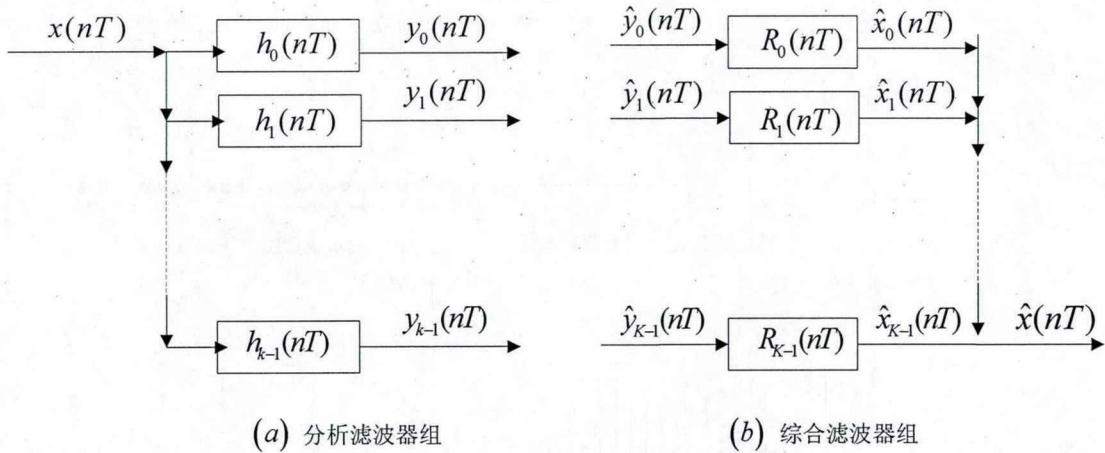


图 2.17 典型滤波器组

图 2.17 中(a)所示是一个具有相同输入信号的滤波器组。输入信号 $x(nT)$ 进入 K 个通道，每个通道中有一个滤波器 $h_k(nT), k = 0, 1, 2, \dots, K-1$ 。设 $x(nT)$ 为一个宽频带信号，经过各通道中的带通滤波器后被分为 K 个子频带信号 $y_k(nT), k = 0, 1, 2, \dots, K-1$ 。 $y_k(nT)$ 是窄频带信号，这样的滤波器组称为分析滤波器组。如果这个滤波器组中第 K 通道滤波器的冲击响应 $h_k(nT)$ 与第 0 通道滤波器的冲激响应 $h_0(nT)$ 具有如下关系

$$h_k(nT) = h_0(nT) e^{\frac{j2\pi}{K} kn} \quad (2-32)$$

或者

$$H_k(z) = H_0(zW^k) \quad (2-33)$$

式中 $k = 0, 1, 2, \dots, K-1$ 及

$$W = e^{-\frac{j2\pi}{K}} \quad (2-34)$$

则称这个滤波器组为均匀 DFT 滤波器组。滤波后各通道的信号 $y_k(nT), k = 0, 1, 2, \dots, K-1$ ，是窄带信号，因此他们的抽样率可降低。如果 $x(nT)$ 是一个满带信号，即 $X(e^{j\omega})$ 的频谱占满 $-\pi$ 到 π 的区域，而各通道的信号 $y_k(nT)$ 都具有相同的带宽 B ，则：

于是抽样率最多可降低到 $1/KT$ 。如果抽样率低于 $1/KT$ 则必然会出现混叠现象。这表明各通道滤波后的信号可进行抽取因子 D 等于或小于 K 的抽取。因此， $D = K$ 的抽取称为最大抽取，如图 2.18 所示。

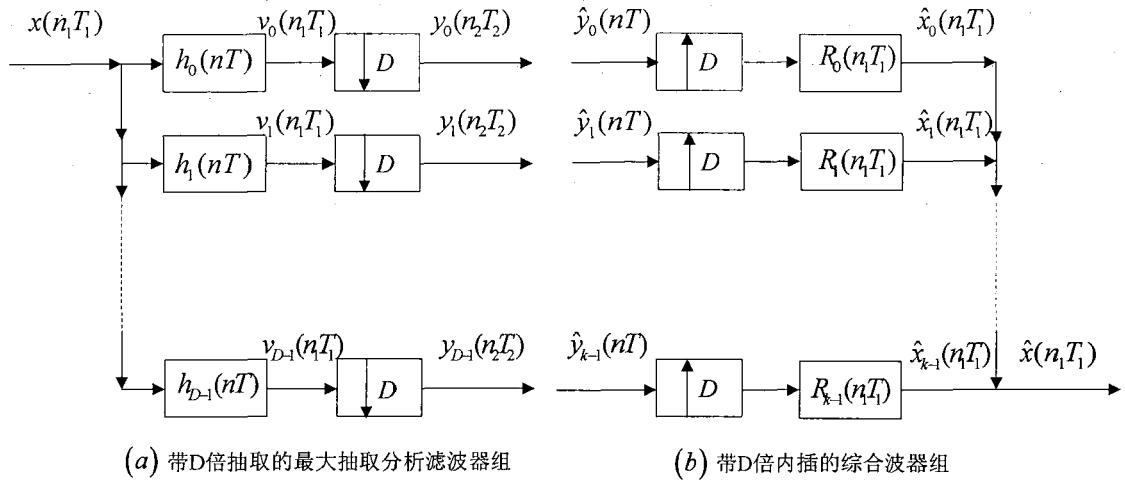


图 2.18 最大抽取滤波器组和综合滤波器组

在综合滤波器组中输入信号 $\hat{y}_k(n_2 T_2)$, $k = 0, 1, 2, \dots, D-1$. 先进行零值内插, 经综合滤波器 $R_k(n_1 T_1)$ 后保留了所需要的子频带, 得到相应子频带信号 $\hat{x}_k(n_1 T_1)$, $k = 0, 1, 2, \dots, D-1$ 。把所有的 $\hat{x}_k(n_1 T_1)$ 相加起来, 就得到所求的综合信号 $\hat{x}(n_1 T_1)$ 即

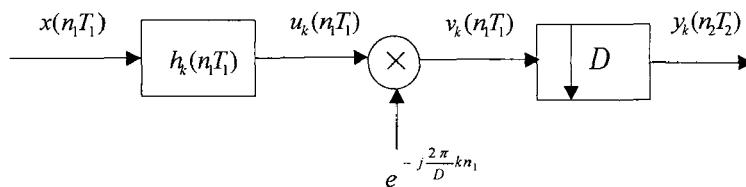
$$\hat{x}(n_1 T_1) = \sum_{k=0}^{D-1} \hat{x}_k(n_1 T_1) \quad (2-35)$$

及 $\hat{X}(e^{j\omega_1}) = \sum_{k=0}^{D-1} \hat{X}_k(e^{j\omega_1}) \quad (2-36)$

分析滤波器组中的每一个通道可以理解为输入信号 $x(n_1 T_1)$ 先经过一个带通滤波器 $h_k(n_1 T_1)$ 滤出一个信号 $u_k(n_1 T_1)$, 然后对 $u_k(n_1 T_1)$ 进行抽取, 由于 $u_k(n_1 T_1)$ 它是一个窄带信号, 为了降低它的抽样率, 先将 $u_k(n_1 T_1)$ 它的频谱中心平移到坐标原点, 得到 $v_k(n_1 T_1)$, 然后再进行 D 倍抽取得到第 K 通道的输出 $y_k(n_2 T_2)$, 其过程示于图 2.19 所示。

$$h_k(n_1 T_1) = e^{j\frac{2\pi}{D}kn_1} h_0(n_1 T_1) \quad (2-37)$$

$$H_k(e^{j\omega_1}) = H_0[e^{j\omega_1} W^k] \quad (2-38)$$

图 2.19 分析滤波器组中的第 K 个通道

2.5.2 多级实现

抽样率转换器(SRC:sampling rate converter), 当抽样率转换器转换因子(抽取器的周期因子 D 或内插器的内插因子 I)较大时, 直接把抽样率转换工作一次完成, 从计算工作量或者存储量来说, 往往不如经过两次或两次以上转换来的经济。例如, 给定信号 $x(n_0T_0)$ 其抽样率为 $F_0 = 1/T_0$, 要求进行抽取得到的结果为 $y(n_2T_2)$, $y(n_2T_2)$ 的抽样率为 $F_2 = 1/T_2$ 。设 $F_0 = 100F_2$ 即抽取因子 D 为 100, 这样的情况下, 用一次抽取来完成所需的计算工作量比两次抽取(第一次以抽取因子 $D_1 = 50$ 进行抽取, 然后以抽取因子 $D_2 = 2$ 对第一次的结果进行抽取)所需的计算量要大得多。我们把一次抽取(或者内插)完成所需的抽样率转换称为抽样率转换的单级实现, 把两次或者两次以上的抽取(或内插)称为多级抽取(或内插)。

抽取器的单级实现框图和两级实现框图分别如图 2.20 和 2.21 所示。

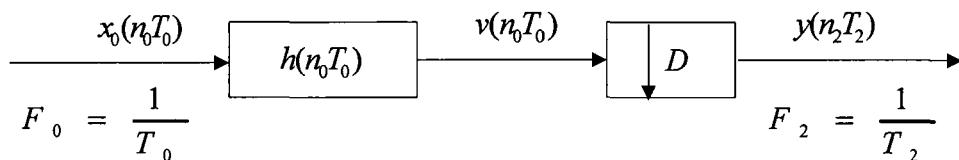


图2.20 抽取器的单级实现框图

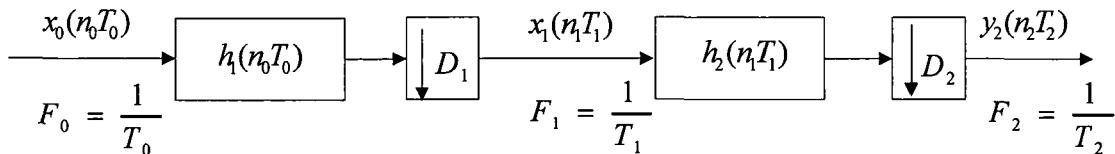


图 2.21 抽取器的两级实现框图

$h_1(n_0T_0)$ 和 $h_2(n_1T_1)$ 分别为第一级和第二级的抗混叠滤波器, 其所处的抽样率分别为 F_0 和 F_1 。

如果单级抽取和两级抽取同样完成抽样因子 $D = 50$ 的抽取任务, 则有一下可能情况: 即 $D_1 = 25, D_2 = 2$; $D_1 = 10, D_2 = 5$; $D_1 = 5, D_2 = 10$ 等情况。

依照上述的想法推论下去, 抽取器的 K 级实现框图如图 2.22 所示。

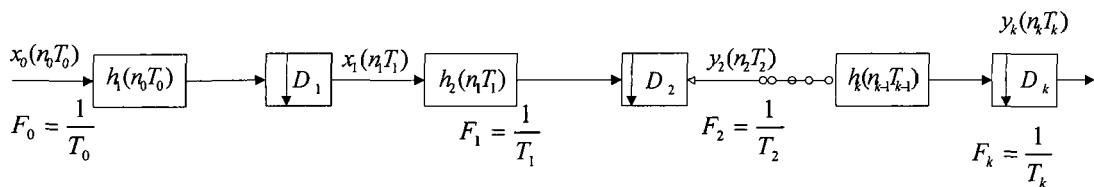


图2.22 抽取器的K级实现框图

多级实现的意义在于多级实现可以节省大量工作量。通过实例来更直观的解释, 假设 $x_0(n_0T_0)$ 是给定的信号, 它的抽样率为 $F_0 = 5000Hz$, 经过 50 倍的抽取

($D = 50$)后变成 $y(n_2 T_2)$, $y(n_2 T_2)$ 的抽样率为 $F_2 = 100 \text{ Hz}$, 设信号所含的最高频率为 50 Hz , 信号能量的绝大部分集中在 40 Hz 以下, 当使用单级实现时, 应当考虑抗混叠滤波器 $h(n_0 T_0)$ 的通带边缘频率 f_p 为 40 Hz , 阻带频率为 50 Hz 。如果我们要求此滤波器的通带误差容限为 $\delta_p = 0.01$, 阻带误差容限为 $\delta_s = 0.001$ (设此滤波器的幅度在零频率处为 1.0)并假定所涉及的滤波器是 FIR 最佳等波纹滤波器, 则滤波器的阶数 N (即有限脉冲响应(FIR: Finite Impulse Response)滤波器的长度)可估计为

$$N \cong \frac{D_\infty(\delta_p, \delta_s)}{\frac{f_s - f_p}{F_0}} \quad (2-39)$$

式中

$$D_\infty(\delta_p, \delta_s) = [a_1 \lg(\delta_p)^2 + a_2 \lg \delta_p + a_3] \lg \delta_s + [a_4 \lg(\delta_p)^2 + a_5 \lg \delta_p + a_6] \quad (2-40)$$

其中 $a_1 = 0.005309$, $a_2 = 0.07114$, $a_3 = -0.4761$, $a_4 = -0.00266$, $a_5 = -0.5946$, $a_6 = -0.4278$

在确定了滤波器的阶数 N 后, 滤波器的输出 $v(n_0 T_0)$ 为

$$v(n_0 T_0) = \sum_{m=0}^{N-1} h(m T_0) x[(n_0 - m) T_0] \quad (2-41)$$

即, 为求得每一个 $v(n_0 T_0)$ 应进行 N 次乘法和 $N-1$ 此加法, 而每秒钟内 F_0 次抽样, 所以每秒钟的乘法次数(MPS: Meters Per Second)为 $N F_0$ 。考虑实现时使用高效结构, 乘法次数可降低 D 倍(D 为抽取因子), 再考虑 FIR 滤波器的线性相位约束, 乘法次数又减少一倍, 如果每秒乘法次数用 R_T^* 表示, 则

$$R_T^* = \frac{N F_0}{2 D} \quad (2-42)$$

式(2-42)就是式(2-41)单级实现的主要工作量。

按照式(2-39), 单级实现时, 滤波器的阶数 N 为

$$N = \frac{D_\infty(0.001, 0.001)}{\frac{50 - 40}{5000}} = 1271 \quad (2-43)$$

$$\text{每秒钟的乘法次数: } R_T^* = \frac{N F_0}{2 D} = \frac{1271(5000)}{2(50)} = 63550 (\text{MPS}) \quad (2-44)$$

从以上的分析可以看出, 计算工作量既受通带和阻带的误差容限(δ_p, δ_s)的影响又受到过渡带宽 $f_s - f_p$ 和抽样率 F_0 的影响, 还有抽取因子 D 的影响。其中过渡带宽的影响是很大的。

值得注意的是, 虽然滤波器的阻带是从 $f_s = 50 \text{ Hz}$ 到 $F_0/2 = 2500 \text{ Hz}$, 由式(2-39)可以看出, $f_s - f_p / F_0$ 比值小就意味着滤波器的阶次 N 高, 这样就会增加计算工作

量。但是单级实现的情况下却不允许增加过渡带宽，其原因是抽取后的抽样率 $F_2 = 100Hz$ ，而信号中的最高频率只能小于等于 $F_2/2$ ，所以 $f_s = 50Hz$ 是阻带开始的最高限，如果 f_s 高于 $50Hz$ 则出现混叠现象， f_s 低于 $50Hz$ 则滤波器的阶次将更高，计算工作量将更大。过渡带与抽样率相比较太小，这是工作量大的根本原因。

看看如何两级实现，实现框图如图 2.23 所示。

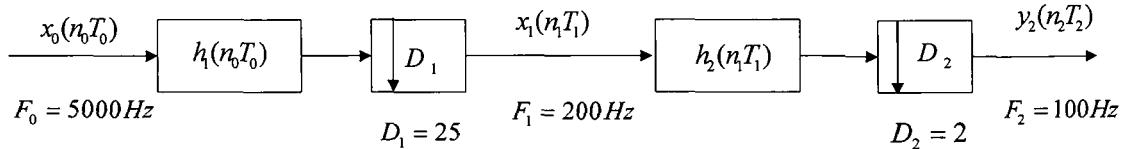


图 2.23 两级滤波器实现采样率转换框图

按照第一级滤波器 $h_1(n_0 T_0)$ 的技术要求，如果将第一级滤波器的频响响应设计成图 2.24 所示。

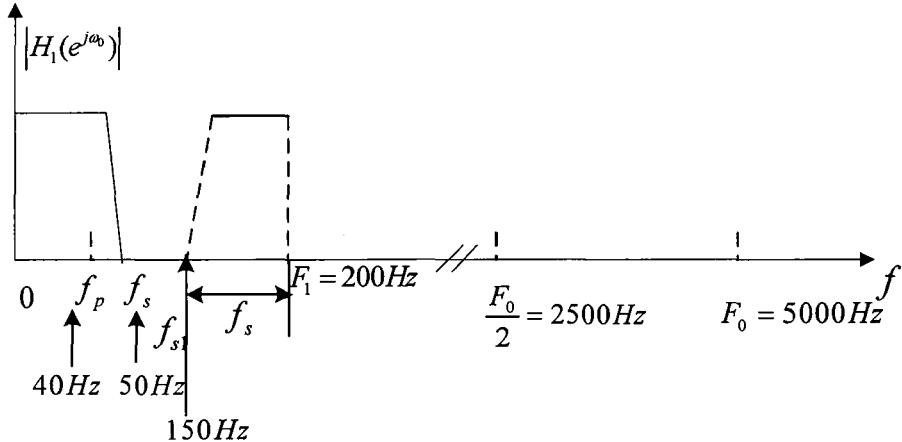


图 2.24 两级实现经抽取后的第一级滤波器的技术要求

则第一级滤波器 $h_1(n_0 T_0)$ 的阻带边缘频率 $f_{s1} = F_1 - f_s = 200 - 50 = 150Hz$ ，这里 F_1 为 $h_1(n_0 T_0)$ 输出信号经抽样后的抽样率， $F_1 = f_0/D_1 = 200Hz$ 。为了使信号在 $0 \leq f \leq f_s$ 频带内受到保护，过渡带可以设计为 $f_p \leq f \leq f_{s1}$ ，这样第一级滤波器的频率响应为图 2.25 所示。

图中画出了经过抽取后 $|H_1(e^{j\omega_0})|$ 从 0 到 F_1 为一个周期的频率响应。这时 $H_1(e^{j\omega_0})$ 的过渡带宽为 $f_{s1} - f_p = 150 - 40 = 110Hz$ ，它与单级实现时相比增宽了 10 倍，从而是滤波器 $h_1(n_0 T_0)$ 的阶数大大降低。

至于第二级中的滤波器 $h_2(n_1 T_1)$ 虽然其中过渡带宽只有 $f_s - f_p = 10Hz$ 但这时的抽样率为 $F_1 = 200Hz$ ，所以其归一化的过渡带宽度 $(f_s - f_p)/F_1 = 1/20$ ，这个值远大于单级实现时的归一化过渡带宽度 $(f_s - f_p)/F_0 = 1/500$ 。因而 $h_2(n_1 T_1)$ 的阶数 N_2 也将远低于单级实现时的滤波器阶数。

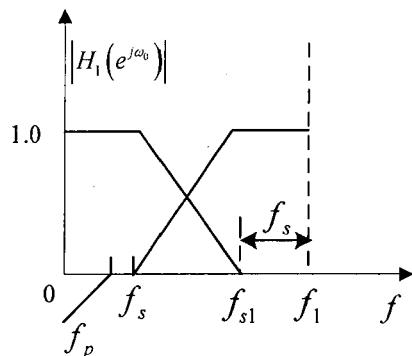


图 2.25 满足第一级抽取后 $0 \leq f \leq f_s$ 频带受到保护的滤波器频率响应

为了有个节省计算量多少的概念，现将两级实现的每秒乘法次数计算如下：

第一级滤波器阶数

$$N_1 = \frac{D_\infty(\delta_{p1}, \delta_s)}{\frac{f_{s1} - f_p}{F_0}} = \frac{2.7614}{\frac{150 - 40}{5000}} \cong 126 \quad (2-45)$$

第一级中的每秒乘法次数

$$R_{T1}^* = \frac{N_1 F_0}{2 D_1} = \frac{126(5000)}{2(25)} = 12600 \text{ (MPS)} \quad (2-46)$$

第二级滤波器阶数

$$N_2 = \frac{D_\infty(\delta_{p2}, \delta_s)}{\frac{f_{s2} - f_p}{F_1}} = \frac{2.7614}{\frac{50 - 40}{200}} \cong 55 \quad (2-47)$$

第二级中的每秒乘法次数

$$R_{T2}^* = \frac{N_2 F_1}{2 D_2} = \frac{55(200)}{4} = 2750 \text{ (MPS)} \quad (2-48)$$

所以两级实现系统中的每秒乘法次数 R_T^* 为

$$R_T^* = R_{T1}^* + R_{T2}^* = 15350 \text{ (MPS)} \quad (2-49)$$

与单级实现的每秒乘法次数相比约 4.1:1 这当然是很大的节省。

第三章 音频重采样算法设计与实现

本章介绍了音频重采样的概念，阐述了音频重采样的意义，设计了三种音频重采样的算法，介绍了三种算法的设计思想，针对上采样和下采样分别设计了不同的方法。分别用三种算法实现了各常用音频采样率的相互转换，统计了每种算法在不同速率之间转换的时间开销，分析了三种算法在上采样和下采样时的效果。

3.1 音频重采样

采样频率是数字音频技术中的一个重要概念，它表示模数转换时每秒钟抽取样值的个数。采样频率越高，其模拟原值的精度越高，保真度也就越高。当然，采样频率越高，A/D 转换时所需空间越大，硬件的投入也就越高。因此，只有在条件允许的情况下，采样频率越高越好。我们日常使用的音频工作站的声卡的采样频率大多定在 48kHz，量化精度一般为 16bit，这就是我们通常所说的广播级声卡的技术规格，它优于普通 CD 机，具有较高的技术指标和性价比。现在的数字音频文件具有许多格式，例如，WAV 文件、MPEG 文件、MP3 文件、WMA 文件、RM 文件、.S48 文件等等。各种数字音频文件又根据各自需要，具有各自不同的采样频率：广播用的.S48 文件使用的是 48kHz，而电信的语音、手机用的彩铃所使用的 WAV 文件使用的则是 8kHz。现在比较高档的 A/D、D/A 转换器、数字调音台及各种数字信号处理(DSP: Digital Signal Process)的采样频率都能达到 48kHz，并且有多种输出采样频率，具有自己的内部计时钟，既可以设置内部计时钟为主时钟，让其它设备与自己保持同步；也可以让其自动跟踪，与外接信号源同步。同步在数字电路中至关重要，统一的采样频率是数字电路正常工作的前提，不同的采样频率存在于一个系统中则可能产生许多不可预料的问题。

音频重采样就是对数字音频流进行重新的采样，产生不同与原速率的音频数字流。现代多媒体系统中，音频信号的采样频率通常为 48 kHz、44.1 kHz、32 kHz 或 16 kHz。如 CD 的采样速率为 44.1 kHz 而 SBLive 系列/Audigy 系列/AC97 兼容的声卡都已经将输出的采样频率固定为 48kHz，要实现 CD 中的音频信号也可以用上述声卡来播放，在最终输出到声卡之前，都需要由驱动程序或者 Windows 混音器来对 CD 中的音频数据流进行重采样，转换成声卡可以接受的(48 kHz)速率，即要实现音频速率转换^[25-28]。在设置重采样速率的时候，需要根据不同声卡的具体参数来采用不同的速率。

音频重采样分为上采样和下采样两种：上采样就是插值，是多速率信号处理中的基本内容之一，上采样可以应用于不同的场合；下采样就是抽取，也是多速率信号处理中的基本内容之一，下采样可以带来许多相应的好处。

3.2 音频重采样算法设计

从数学的角度来看，音频重采样^[29-32]是插值问题；从信号处理的角度来看，对于音频信号来说重采样是滤波处理，插值多项式是最典型的和自然的方法。

假设已对音频信号 $x(n)$ ($1 \leq n \leq N$) 以频率 L 进行采样，现在要对它以新频率 M 进行采样， L 和 M 都是实数， M 可以比 L 大或者小。如果 $M > L$ ，那么就是上采样，反之，就是下采样。让 $y(k)$ ($1 \leq k \leq K$) 表示重采样后的信号，明显必须满足下面的关系

$$K = (M / L) * N \quad (3-1)$$

要导出 $y(k)$ 的一些基本问题必须考虑：(1)需要多少个 $x(n)$ 采样点去导出一个 $y(k)$ 的采样点幅值，这便涉及到了窗口大小的问题；(2)什么插值函数比较合适与如何估计逼近精度，这涉及到逼近函数的问题。

对于第一个问题，每个音频信号的采样仅仅决定于一些与它邻近的采样值。实际上窗是非常小的，通常仅仅包含几个采样值，以节省计算量和满足时间要求。

第二个问题，内插函数和逼近函数是数学问题。众所周知任何连续函数可以被任意近似均匀的多项式或三角多项式表示。逼近的精度可以用泰勒级数和原函数的高阶函数来估计。所以理论上可以选择合适的多项式或者三角多项式作为音频重采样的插入函数。

只要窗的大小和插值函数被确定，重采样的性能便相应的被确定了下来。

3.2.1 线性重采样

线性重采样基本设计思想就是用一条通过两个固定点的直线去逼近所要采样的这个点，就意味着所要求的点和已经存在的两个点，这三个点位于同一条直线上。线性重采样的基本思想是：

(1) 将音频信号的原始采样率 L ，要转换到的新采样率 M ，原始信号的长度 N ，利用式(3-1)求出新采样率下的音频信号长度 K 。

(2) 对于每个离散时间值： k ($1 \leq k \leq K$)，让实值 n_k 的值为：

$$n_k = (L/M) * k \quad (3-2)$$

n_k 为在原始采样间隔的情况下，要进行插值的具体位置。 L 为原音频信号的采样频率， M 为需要转换的新采样频率， L/M 是频率缩放因子。为了使 n_k 有意义，对于 n_k 这里必须存在一个时间指数 n ($1 \leq n \leq N$) 而且 n_k 和 n 的关系必须满足：

$$n \leq n_k \leq n + 1 \quad (3-3)$$

(3) 确定两个加权系数的值，利用第二步计算得到的 n_k 值，求得两个权值。只有选择了恰当的权值，才能让线性插值所要插取的值更加接近插值点的理想幅值，让 w_1 和 w_2 分别代表两个重采样权值：

$$w_1 = n_k - n \quad (3-4)$$

$$w_2 = 1 - w_1 \quad (3-5)$$

(4) 将 w_1 , w_2 两个重采样权值与要插入的点相邻的两点幅值线性结合, 求得将要插入点的具体的幅值, 所以 $y(k)$ 可以写为:

$$y(k) = w_1 x(n+1) + w_2 x(n) \quad (3-6)$$

从式(3-6)可以看出, 为了计算得到每一个所要插入的 $y(k)$ 的值, 仅需要从原始的采样值中提取两个采样值, 并且 $y(k)$ 的幅值等于所提取的这两个点幅值的加权均值, 即 $y(k)$ 和原始的两个采样值位于同一条直线上, 满足线性关系。

以上是线性重采样的设计思想, 但是采用线性重采样进行采样频率转换时, 会遇到上采样和下采样两种不同情况, 上采样会产生镜像, 下采样会产生混叠, 为了消除镜像和混叠, 就需要将信号通过低通滤波器。为了解决不同采样率时遇到的不同情况, 低通滤波器的位置将不同。

上采样时, 在频域会产生镜像, 将新采样频率下生成的音频信号 $y(k)$ 通过低通滤波器, 这样就会消除上采样所带来的镜像, 具体过程如图 3.1 所示。

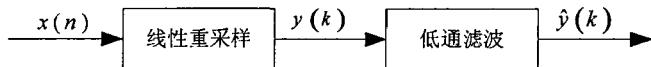


图 3.1 线性重采样上采样

下采样时, 在频域会产生混叠, 如果像上采样一样, 将滤波器的位置放在新采样频率下生成的音频信号 $y(k)$ 后面, 将不能消除混叠现象, 所以为了解决混叠, 要在原始信号进行采样率转换之前进行低通滤波, 避免产生混叠。具体过程如图 3.2 所示。

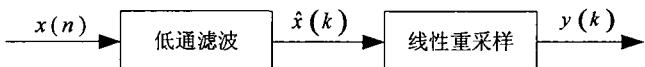


图 3.2 线性重采样下采样

3.2.2 拉格朗日重采样

线性重采样的设计思想是利用插值点前后两点的值来计算插值点的具体幅值, 由于线性重采样进行插值点幅值计算时的窗长是固定的, 所以在需要更精确的进行插值时, 线性重采样往往不能满足要求, 所以需要设计一种可以灵活选取窗口长度的重采样算法。

拉格朗日重采样的设计思想就是用多项式去逼近原始音频信号并且在拉格朗日内插多项式的基础上计算重采样样值。具体步骤为:

(1) 将音频信号的原始采样率 L , 要转换到的新采样率 M , 原始音频信号的长

度 N ，依据式(3-1)求出新采样率下的音频信号长度 K 。

(2)对于每个离散时间值： $k (1 \leq k \leq K)$ ，让实值 n_k 的值为式(3-2)。 n_k 为在原始采样间隔的情况下，要进行插值的具体位置。 L 为原始音频信号的原始采样频率， M 为需要转换的新采样频率， L/M 是频率缩放因子。为了使 n_k 有意义，对于 n_k 这里必须存在一个时间指数 $n (1 \leq n \leq N)$ 而且 n_k 和 n 的关系必须满足式(3-3)。

(3)设置一个窗口长度，窗口长度决定用多少点的幅值进行计算，以确定将要插值点的幅值。假设拉格朗日内插多项式的最高阶数是 $2 * w$ ，那么窗函数的大小就是 $2 * w + 1$ ，这 $2 * w + 1$ 个原始采样值的时间因子如式(3-7)所示。

$$n-w, \dots, n-1, n, n+1, \dots, n+w \quad (3-7)$$

窗口的长度选择的越长，计算的复杂度就越高，所要重采样点的幅值就会越精确，但同时在时间开销方面，选择的窗口长度越长，时间的开销量就会随之增大。

(4)对在该窗口长度内的点利用拉格朗日重采样权值公式(3-8)进行计算，确定每个进行估值点的重采样权值 q_i 的值。

$$q_i = \frac{\prod_{j=-w}^w (n_k - (n-j))}{\prod_{j=-w}^w ((n-i) - (n-j))} \quad (3-8)$$

这里下标 j 的范围为 $-w$ 到 w 并且不等于 i 。(3-8)式的分母可以表示为：

$$\prod((n-i) - (n-j)) = \prod(j-i) = (-1)^{w+i} * (w-i)! * (w+i)! \quad (3-9)$$

(5)将每个点的重采样权值 q_i 和每点的幅值结合，求得插值点的具体幅值， $y(k)$ 的值便为带权值的 $2 * w + 1$ 个原始样值的多项式按式(3-10)计算得到：

$$y(k) = \sum_{i=-w}^w \frac{\prod_{j=-w}^w (n_k - (n-j))}{\prod_{j=-w}^w ((n-i) - (n-j))} x(n-i) = \sum_{i=-w}^w q_i x(n-i) \quad (3-10)$$

以上是拉格朗日重采样的设计思想，与线性重采样类似，在用拉格朗日重采样进行采样频率转换时，会遇到上采样和下采样两种不同情况，上采样会产生镜像，但是拉格朗日的窗口长度是可以灵活选取的。实验的结果证明：在用拉格朗日重采样进行速率转换时，选取较大的窗口长度不仅可以更精确的得到插值点的幅值，还可以有效的消除镜像。为了更好的消除重采样产生的混叠和镜像，就需要在适当的位置进行低通滤波。为了解决不同采样率时遇到的不同情况，低通滤波器的位置将不同。

上采样时，在频域会产生镜像，将新采样频率下生成的信号 $\hat{y}(k)$ 通过低通滤波器，这样就会消除上采样所带来的镜像现象，具体过程如图 3.3 所示。

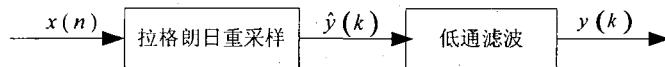


图 3.3 拉格朗日重采样上采样

下采样时，在频域会产生混叠，如果像上采样一样，将滤波器的位置放在新采样频率下生成的信号 $y(k)$ 之后将不能消除混叠现象；所以为了消除混叠，可以在原始信号 $x(n)$ 进行采样率转换之前进行低通滤波，避免产生混叠。具体过程如图 3.4 所示。

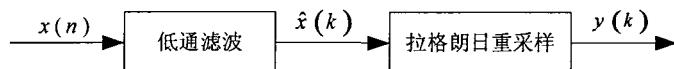


图 3.4 拉格朗日重采样下采样

3.2.3 正弦重采样

通过香农理论可以知道一个连续的音频信号 $x(t)$ 可以经由离散值 $x(n)$ 通过式(3-11)的公式重构：

$$x(t) = \sum_{n=-\infty}^{+\infty} \frac{\sin(\pi(t-nT)/T)}{\pi(t-nT)/T} x(n) \quad (3-11)$$

依据香农理论，实施类似拉格朗日方法，利用三角函数多项式—正弦多项式去逼近原始连续音频信号并且在正弦内插多项式的基础上计算重采样样值。具体步骤为：

(1) 将音频信号的原始采样率 L ，要转换到的新采样率 M ，原始音频信号的长度 N ，依据式(3-1)求出新采样率下的音频信号长度 K 。

(2) 对于每个离散时间值： $k (1 \leq k \leq K)$ ，让实值 n_k 的值为式(3-2)。 n_k 为在原始采样间隔的情况下，要进行插值的具体位置。 L 为原始音频信号的采样频率， M 为需要的新的采样频率， L/M 是频率缩放因子。为了使 n_k 有意义，对于 n_k 这里必须存在一个时间指数 $n (1 \leq n \leq N)$ 而且 n_k 和 n 的关系必须满足式(3-3)。

(3) 设置一个窗口长度，窗口长度决定用多少点的幅值进行计算来确定将要插值点的幅值。假设正弦内插多项式的最高阶数是 $2*w$ ，那么窗函数的大小就是 $2*w+1$ ，这 $2*w+1$ 原始样值的时间因子如式(3-7)。同拉格朗日重采样一样，窗口的长度选择的越长，计算的复杂度就越高，所要重采样点的幅值就会越精确，但同时在时间开销方面，选择的窗口长度越长，时间的开销量就会越大。

(4) 对在该窗口长度内的点利用正弦插值权值公式(3-12)进行计算，确定每个进行估值点的重采样权值 p_i 的值。

$$p_i = \frac{\sin(\pi(n_k - n + i))}{\pi(n_k - n + i)} \quad (3-12)$$

(5) 将每点的重采样权值 p_i 和幅值进行结合, 求得插值点的具体幅值, $y(k)$ 的值便为带权值的 $2 * w + 1$ 个原始样值的多项式按式(3-13)计算得到:

$$y(k) = \sum_{i=-w}^w \frac{\sin(\pi(n_k - n + i))}{\pi(n_k - n + i)} x(n - i) = \sum_{i=-w}^w p_i x(n - i) \quad (3-13)$$

以上是正弦重采样的设计思想, 与拉格朗日重采样类似, 可以选取窗口长度来计算插值点的幅值, 在用正弦重采样进行采样速率转换时, 同样会遇到上采样和下采样两种不同情况, 上采样会产生镜像, 正弦重采样的窗口长度虽然可以灵活选取, 但是实验的结果证明: 在用正弦重采样进行速率转换时, 选取较大的窗口长度消除镜像的效果没有拉格朗日重采样算法明显。为了更好的消除镜像, 同样可以将新采样频率下生成的信号 $y(k)$ 通过低通滤波器。下采样时会产生混叠, 为了消除混叠, 就需要将原始音频信号 $x(n)$ 通过低通滤波器, 然后进行重采样。

上采样时, 在频域会产生镜像, 将新采样频率下生成的信号 $\hat{y}(k)$ 通过低通滤波器, 这样就会消除上采样所带来的镜像现象, 具体过程如图 3.5 所示。

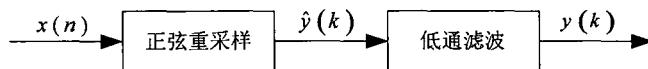


图 3.5 正弦重采样上采样

下采样时, 在频域会产生混叠, 将滤波器的位置放在新采样频率下生成的信号 $y(k)$ 之后将不能消除混叠现象。所以为了更好的重构音频信号, 就必须消除混叠, 所以在原始信号进行采样率转换之前先进行低通滤波, 避免产生混叠。具体过程如图 3.6 所示。

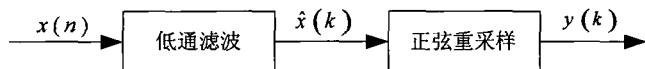


图 3.6 正弦重采样下采样

3.3 重采样算法的实现

3.3.1 线性重采样算法

线性重采样的基本思想是用在同一条直线上的两个点的线性加权值去逼近所要重采样的点。

用线性重采样进行下采样时的框图如图3.7所示。

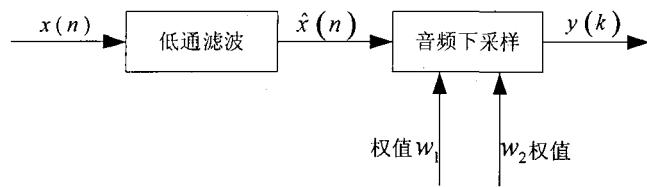


图 3.7 线性重采样下采样框图

为了避免下采样时产生混叠，原始音频信号 $x(n)$ 首先通过低通滤波器，得到音频信号 $\hat{x}(n)$ 。线性重采样的两个重采样权值 w_1 ， w_2 与经过低通滤波的音频信号 $\hat{x}(n)$ 同时进入音频下采样模块，按照线性重采样算法计算，得到音频重采样信号 $y(k)$ 。

基于线性重采样算法的下采样流程图如图 3.8 所示。

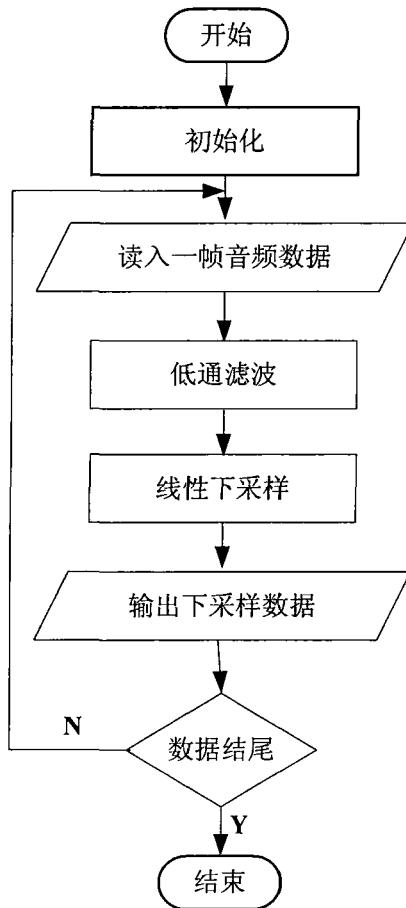


图 3.8 线性重采样下采样流程图

(1)首先初始化原始音频信号 $x(n)$ ，将原始音频信号的长度 N ，原采样频率 L ，新采样频率 M ，依式(3-1)求得将要生成的新音频信号的长度 K ，为即将生成的新音频信号分配长度为 K 的short型内存单元。

(2)按帧读入原始音频信号，求得线性重采样的具体位置 n_k 和线性重采样的权值 w_1 ， w_2 。

(3)为了避免出现混叠现象，在进行下采样之前，将原始音频信号通过低通滤波器。

(4)将经过低通滤波器的音频信号 $\hat{x}(n)$ 结合第(2)步得到的线性重采样权值 w_1 , w_2 ，按照式(3-6)进行计算，得到新采样频率下的音频信号 $y(k)$ ，将新的音频信号 $y(k)$ 写入已经分配的内存单元。

(5)判断原始音频信号是否处理完毕，如果没有，则进入第(2)步继续按流程进行处理，直到处理完原始音频信号。

下采样时，选取由48kHz转换到16kHz和44.1kHz到32kHz来说明用线性重采样的算法进行下采样时的情形。

图3.9(a)为原始采样率为48kHz的音频信号，图3.9(b)为由采样率48kHz转换到16kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

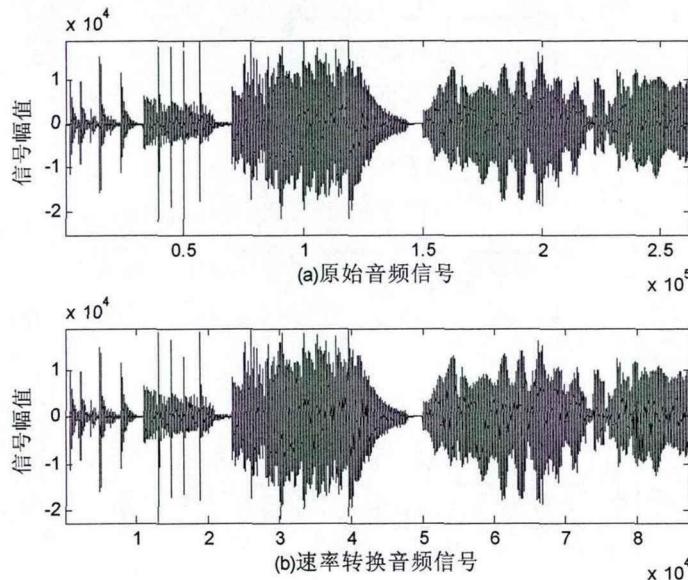


图3.9 48kHz到16kHz速率转换音频时域波形图

图3.10(a)为原始采样率为44.1kHz的音频信号，图3.10(b)为由采样率44.1kHz转换到32kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

从图3.9和图3.10可以看出：线性重采样在进行下采样时，重采样生成的音频信号在时域波形上没有明显的失真，这说明线性重采样能够很好的重构下采样音频信号。

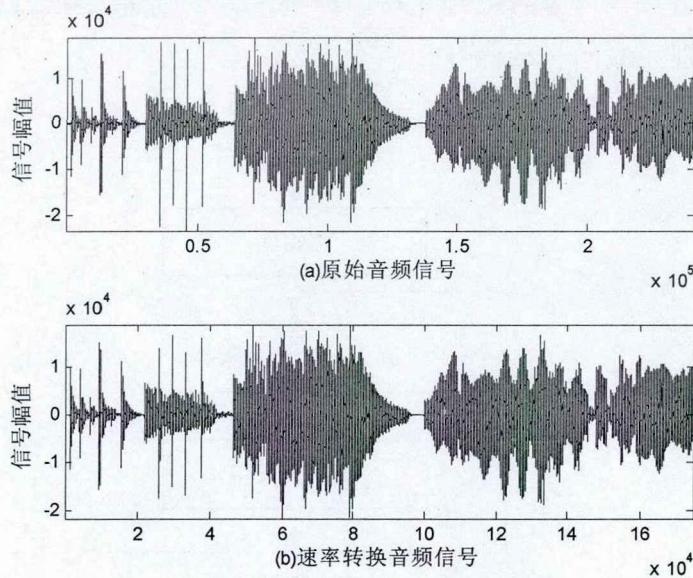


图3.10 44.1kHz到32kHz速率转换音频时域波形

用线性重采样进行上采样时的框图如图3.11所示。

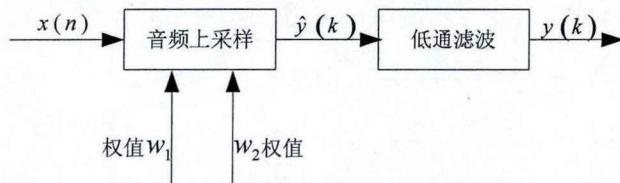


图 3.11 线性重采样上采样框图

原始音频信号 $x(n)$ 与线性重采样的两个权值 w_1 , w_2 , 同时进入音频上采样模块, 生成重采样音频信号 $\hat{y}(k)$, 将 $\hat{y}(k)$ 经过低通滤波去除上采样带来的镜像得到目标音频重采样信号 $y(k)$ 。

基于线性重采样算法的上采样流程图如图 3.12 所示。

(1)首先初始化原始音频信号 $x(n)$, 将原始音频信号的长度 N , 原采样频率 L , 新采样频率 M , 依式(3-1)求得将要生成的新音频信号的长度 K , 为即将生成的新音频信号分配长度为 K 的short型内存单元。

(2)按帧读入原始音频信号, 求得线性重采样的具体位置 n_k 和线性重采样的权值 w_1 , w_2 。

(3)将原始音频信号 $x(n)$ 结合第(2)步得到的线性重采样权值 w_1 , w_2 , 按照式(3-6)进行计算, 得到新采样频率下的音频信号 $\hat{y}(k)$, 将新的音频信号 $\hat{y}(k)$ 写入已经分配的内存单元。

(4)由于上采样会带来镜像现象, 将新的音频信号 $\hat{y}(k)$ 通过低通滤波器, 去除镜像, 得到音频信号 $y(k)$ 。

(5) 判断原始音频信号是否处理完毕, 如果没有, 则进入第(2)步继续按流程进行处理, 直到处理完原始音频信号。

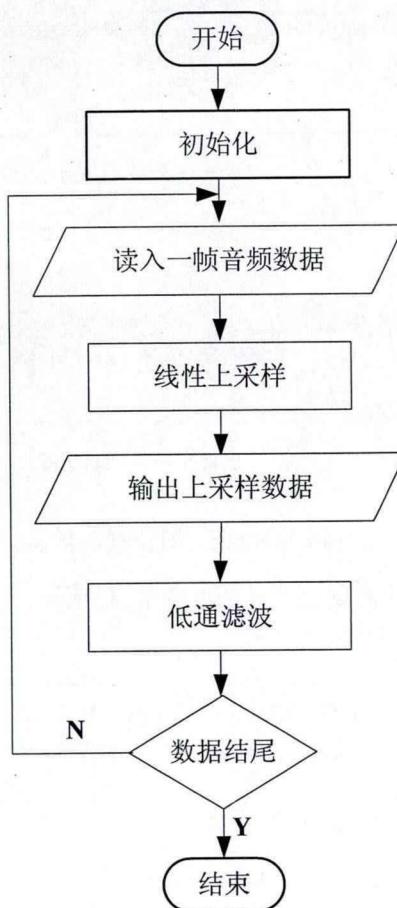


图 3.12 线性重采样上采样流程图

上采样时: 选取由32kHz转换到48kHz与16kHz到44.1kHz来说明用线性重采样的算法进行上采样时的情形。

图3.13(a)为原始采样率为32kHz的音频信号, 图3.13(b)为由采样率32kHz转换到48kHz的音频信号, 对比图中原始的音频信号和进行速率转换之后的音频信号, 发现在时域波形上没有明显的变化。

图3.14(a)为原始采样率为16kHz的音频信号, 图3.14(b)为由采样率16kHz转换到44.1kHz的音频信号, 将图中原始的音频信号和进行速率转换之后的音频信号进行对比, 发现在时域波形上没有明显的变化。

从图3.13和图3.14可以看出: 线性重采样在进行上采样时, 重采样生成的音频信号在时域波形上没有明显的失真, 有着不错的效果。

对音频信号使用线性重采样的算法进行重采样, 在48kHz, 44.1kHz, 32kHz, 16kHz之间的相互转换时, 无论是进行上采样还是进行下采样音频信号都得到了很好的重构。

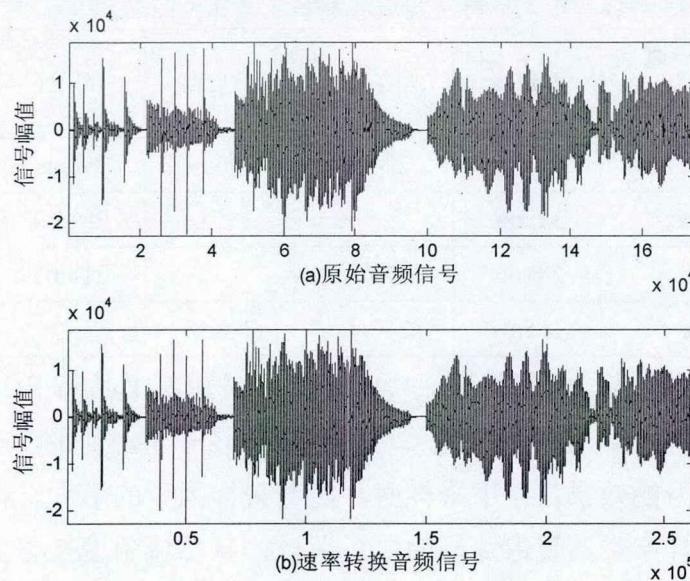


图3.13 32kHz到48kHz速率转换音频时域波形图

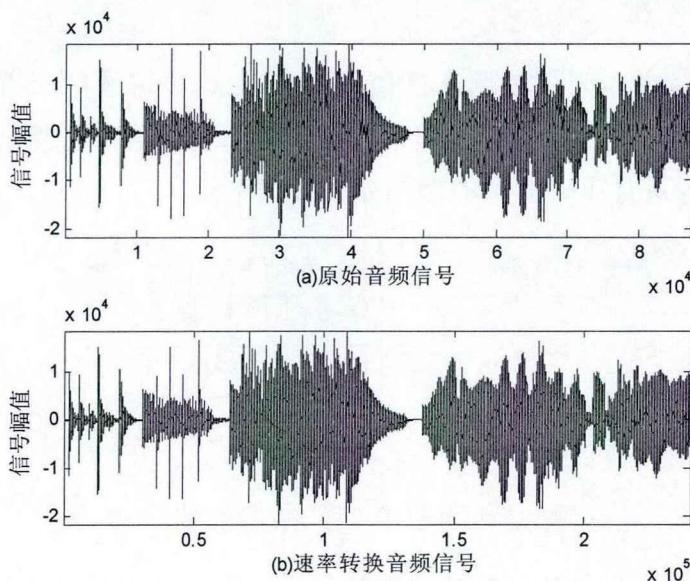


图3.14 16kHz到44.1kHz速率转换音频时域波形图

在CPU为AMD Sempron(tm) Processor 2500+, 主频为1.41GHz, 内存为512MB的PC机上, 软件开发平台为Microsoft Visual C++ 6.0上, 一段播放时长为5461ms的音频信号, 使用线性重采样算法, 测得各速率之间进行转换的时间开销如表3.1所示。

从表3.1中明显可以看出在各速率之间的转换时, 上采样比下采样在时间开销方面稍大, 这是因为上采样是增加了采样点个数, 下采样时减少了采样点个数。采用线性重采样算法增加采样点个数用去的时间比减少采样点个数要多; 同时可以看到在各个不同速率转换到同一个指定速率时, 在时间开销方面相差无几。

表3.1 线性重采样算法实现各速率之间转换的时间

原采样速率 转换后速率	48kHz	44.1kHz	32kHz	16kHz
48kHz		296ms	296ms	296ms
44.1kHz	312ms		265ms	281ms
32kHz	281ms	265ms		218ms
16kHz	265ms	250ms	218ms	

对比生成音频信号的频域，线性重采样在上采样时随着采样因子(原始采样率和新采样率的比值)的增大，在频域产生镜像的现象比较严重，如果没有滤波器，镜像会影响到生成信号的效果；在下采样时，随着采样因子的增大，在频域产生混叠的现象没有上采样产生的镜像那样明显。所以，从频域角度来看，线性重采样适合应用于下采样。

3.3.2 拉格朗日重采样算法

拉格朗日重采样的基本思想是用在插值点前后各 w (w 可选)个点的幅值通过结合利用拉格朗日算法求得的重采样权值 q_i ，来逼近所要插值点的幅值。

用拉格朗日重采样进行下采样时的框图如图3.15所示。

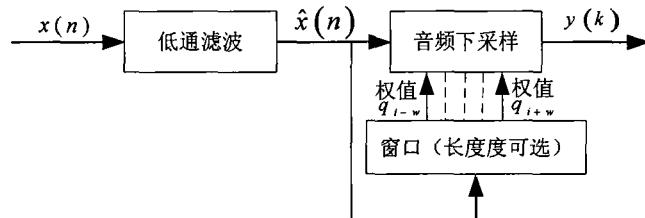


图 3.15 拉格朗日重采样下采样框图

为了避免下采样时产生混叠，原始音频信号 $x(n)$ 首先通过低通滤波器，滤波之后的信号为 $\hat{x}(n)$ ， $\hat{x}(n)$ 经过选择好 $2w+1$ 长度的窗口，在窗长范围内按照式(3-8)计算拉格朗日重采样权值 q_i ，得到插值权值后与经过低通滤波的音频信号 $\hat{x}(n)$ 同时进入音频下采样模块，按照拉格朗日重采样算法计算，得到音频重采样信号 $y(k)$ 。

基于拉格朗日重采样算法的下采样流程图如图 3.16 所示。

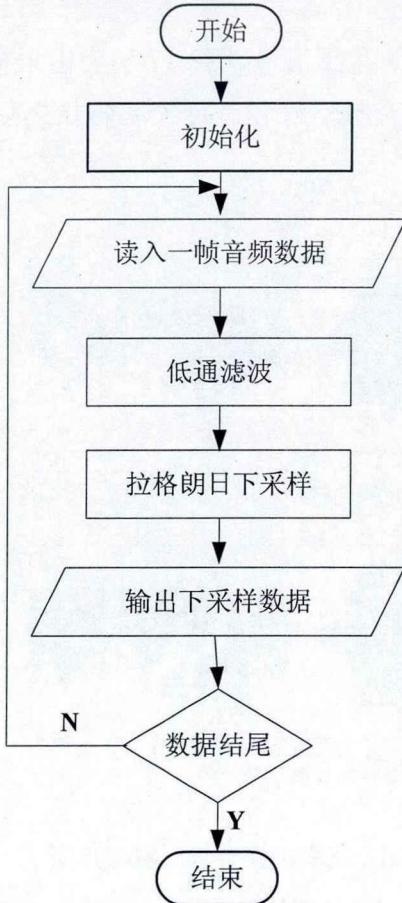


图 3.16 拉格朗日重采样下采样流程图

(1)首先初始化原始音频信号 $x(n)$ ，将原始音频信号的长度 N ，原采样频率 L ，新采样频率 M ，依式(3-1)求得将要生成的新音频信号的长度 K ，为即将生成的新音频信号分配长度为 K 的short型内存单元，选择窗口长度 $2w+1$ 。

(2)按帧读入原始音频信号，求得拉格朗日重采样的具体位置 n_k 。

(3)为了避免出现混叠现象，在进行下采样之前，将原始音频信号 $x(n)$ 通过低通滤波器。在选择好的窗口处长度范围内依式(3-8)计算拉格朗日重采样权值 q_i ，窗口长度为 $2w+1$ ，所以需要求出 $2w+1$ 个重采样权值。

(4)经过低通滤波器的音频信号 $\hat{x}(n)$ 结合第(3)步得到的拉格朗日重采样权值 q_i ，按照式(3-10)进行计算，得到新采样频率下的音频信号 $y(k)$ ，将新的音频信号 $y(k)$ 写入已经分配的内存单元。

(5)判断原始音频信号是否处理完毕，如果没有，则进入第(2)步继续按流程进行处理，直到处理完原始音频信号。

其中拉格朗日的窗口长度可以灵活选取，窗口长度越长插值点的幅值计算越精确，相应的计算复杂度越大，时间开销越大。

下采样时，选择插值点前后各 10 个点来计算，即 $w=10$ ，窗口长度为 $2 \times w + 1 = 21$ ，选取由 48kHz 转换到 16kHz 和 44.1kHz 到 320kHz 来说明用拉格朗日重

采样的算法进行下采样时的情形。

图3.17(a)为原始采样率为48kHz的音频信号, 图3.17(b)为由采样率48kHz转换到16kHz的音频信号, 对比图中原始的音频信号和进行速率转换之后的音频信号, 发现在时域波形上没有明显的变化。

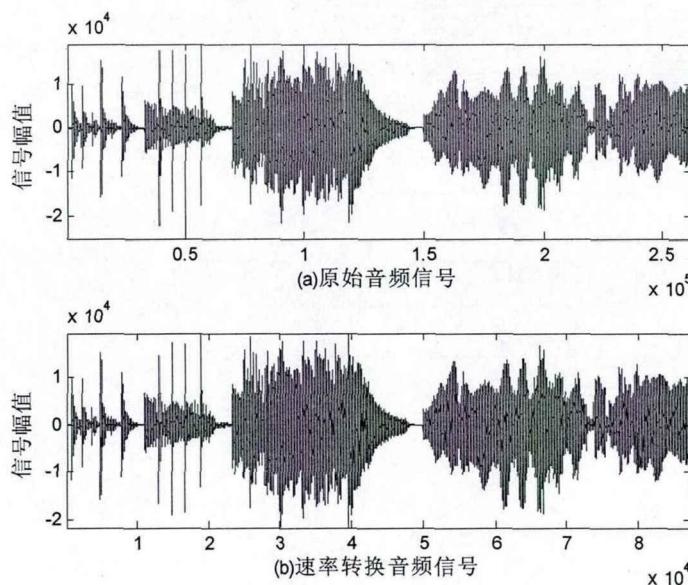


图3.17 48kHz到16kHz速率转换音频时域波形图

图3.18(a)为原始采样率为44.1kHz的音频信号, 图3.18(b)为由采样率44.1kHz转换到32kHz的音频信号, 对比图中原始的音频信号和进行速率转换之后的音频信号, 发现在时域波形上同样没有明显的变化。

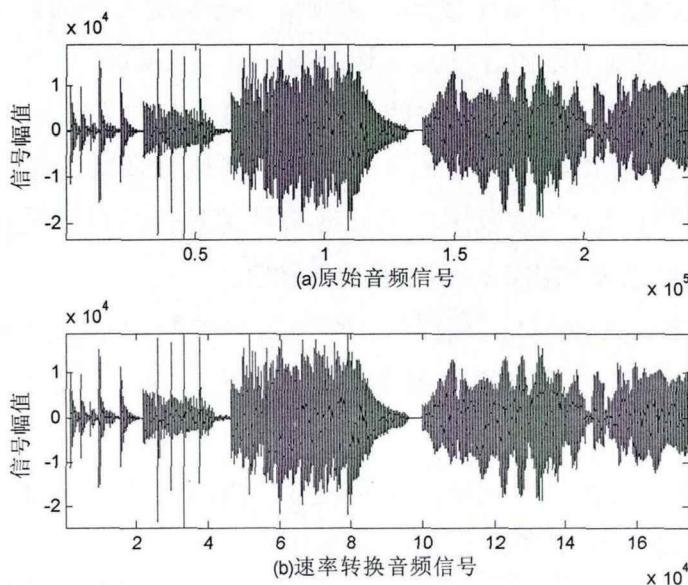


图3.18 44.1kHz到32kHz速率转换音频时域波形

从图3.17和图3.18可以看出: 拉格朗日重采样在进行下采样时, 重采样生成的音频信号在时域波形上没有明显的失真, 这说明拉格朗日重采样能够很好的重构

下采样音频信号。

用拉格朗日重采样进行上采样时的框图如图3.19所示。

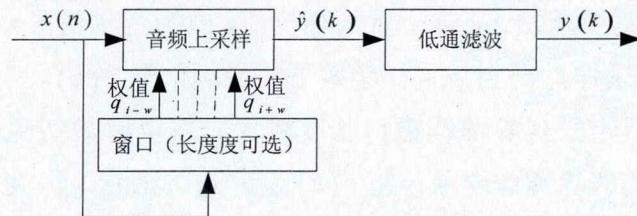


图 3.19 拉格朗日重采样上采样框图

原始音频信号 $x(n)$ 与拉格朗日重采样的权值 q_i 同时进入音频上采样模块，生成重采样音频信号 $\hat{y}(k)$ ，该信号经过低通滤波去除上采样带来的镜像得到目标音频重采样信号 $y(k)$ 。

基于拉格朗日重采样算法的上采样流程框图如图 3.20 所示。

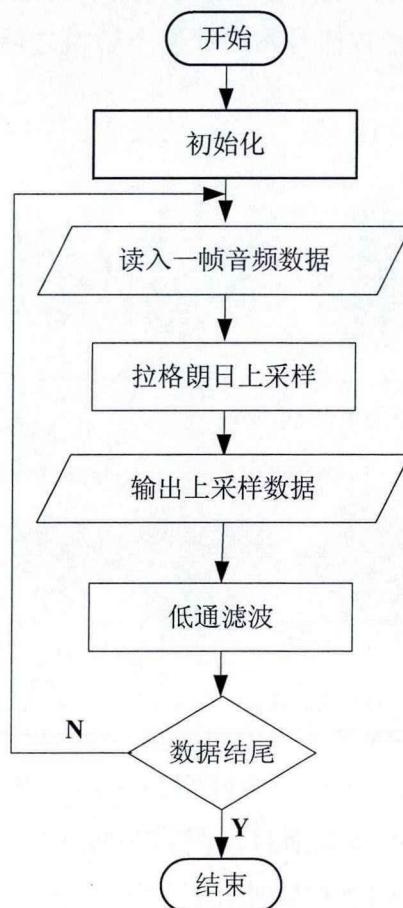


图 3.20 拉格朗日重采样上采样流程图

(1)首先初始化原始音频信号 $x(n)$ ，将原始音频信号的长度 N ，原采样频率 L ，新采样频率 M ，依式(3-1)求得将要生成的新音频信号的长度 K ，为即将生成的新音频信号分配长度为 K 的short型内存单元，选择窗口长度 $2w+1$ 。

(2)按帧读入原始音频信号，求得拉格朗日重采样的具体位置 n_k 。

(3)在选择好的窗口处长度范围内依式(3-8)计算拉格朗日重采样权值 q_i ，窗口长度为 $2w+1$ ，所以需要求出 $2w+1$ 个重采样权值，原始音频信号 $x(n)$ 结合权值 q_i ，按照式(3-10)进行上采样，得到新采样频率下的音频信号 $\hat{y}(k)$ 。

(4)虽然拉格朗日重采样算法在窗口长度较大时可以有效避免上采样带来的镜像，但是为了更好的消除镜像现象，将 $\hat{y}(k)$ 通过低通滤波，得到音频信号 $y(k)$ ，将 $y(k)$ 写入已经分配的内存单元。

(5)判断原始音频信号是否处理完毕，如果没有，则进入第(2)步继续按流程进行处理，直到处理完原始音频信号。

上采样时：选取由32kHz转换到48kHz和16kHz到44.1kHz，同样选择插值点前后各10个点来计算，即 $w=10$ ，窗口长度为 $2*w+1=21$ ，来说明用拉格朗日重采样的算法进行上采样时的情形。

图3.21(a)为原始采样率为32kHz的音频信号，图3.21(b)为由采样率32kHz转换到48kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

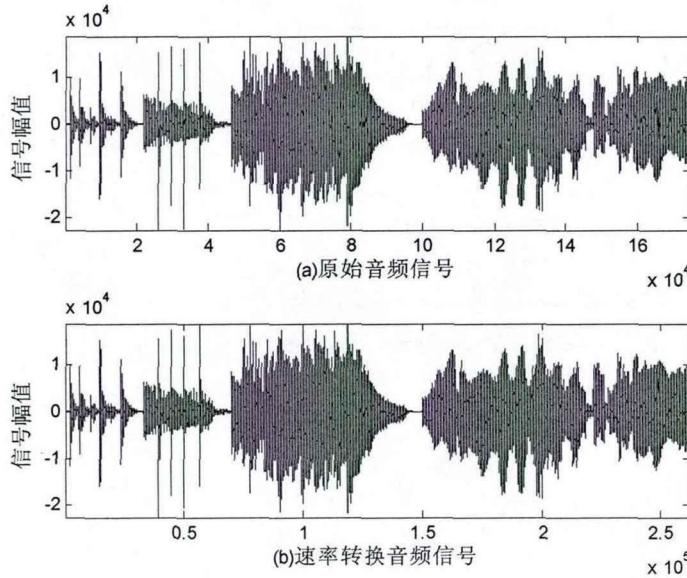


图3.21 32kHz到48kHz速率转换音频时域波形图

图3.22(a)为原始采样率为16kHz的音频信号，图3.22(b)为由采样率16kHz转换到44.1kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

从图3.21和图3.22可以看出：拉格朗日重采样在进行上采样时，重采样生成的音频信号在时域波形上没有明显的失真，有着不错的效果。

总之，对音频信号使用拉格朗日重采样的算法进行重采样，在48kHz, 44.1kHz, 32kHz, 16kHz之间的相互转换，无论是进行上采样还是进行下采样音频信号都得

到了很好的重构。

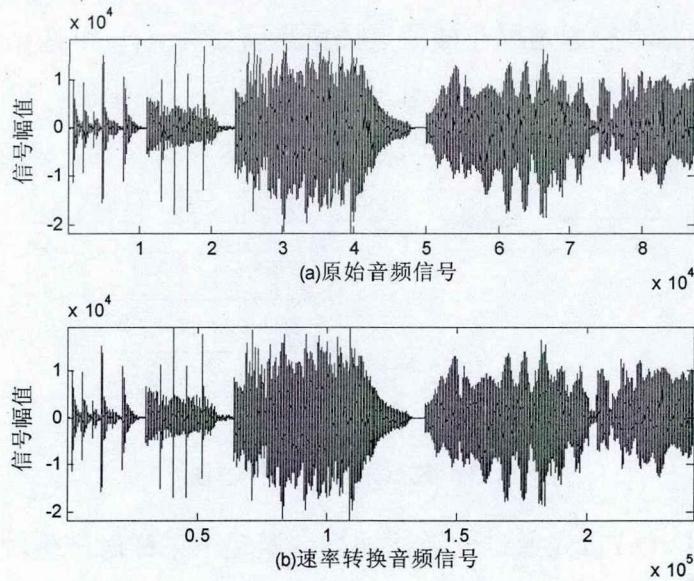


图3.22 16kHz到44.1kHz速率转换音频时域波形图

测试环境和音频信号与线性重采样一样，选择插值点前后各10个点来计算，即 $w=10$ ，窗口长度为 $2*w+1=21$ 的拉格朗日重采样算法处理音频信号，测得各速率之间进行转换的时间开销如表3.2所示。

表3.2 拉格朗日重采样算法实现各速率之间转换的时间

原采样速率 转换后速率	48kHz	44.1kHz	32kHz	16kHz
48kHz		4609ms	4593ms	4593ms
44.1kHz	4265ms		4234ms	4234ms
32kHz	3203ms	3218ms		3078ms
16kHz	1718ms	1750ms	1640ms	

从表3.2中明显可以看出在各速率之间转换时，上采样比下采样在时间开销方面要大，这是因为上采样是增加了采样点个数，下采样时减少了采样点个数。采用拉格朗日重采样算法增加采样点个数用去的时间比减少采样点个数要多；同时可以看到在各个不同速率转换到同一个指定速率时，在时间开销方面相差无几。

对比生成音频信号的频域，可以看到，在没有滤波的情况下，拉格朗日重采样在上采样时随着采样因子的增大，在频域产生镜像的现象能较好的得到控制；在下采样时，随着采样因子的增大，在频域产生混叠的现象也能较好的得到控制。但是在上采样时随着插值因子的变化比在下采样时的变化小。所以，从频域角度来看，拉格朗日重采样在上采样时会取得更好的结果。

3.3.3 正弦重采样算法

正弦重采样的基本思想是用在插值点插值点前后各 w (w 可选)个点的幅值，结合利用正弦算法求得的插值权值 p_i ，来逼近所要插值点的幅值。

用正弦重采样进行下采样时的框图如图3.23所示。

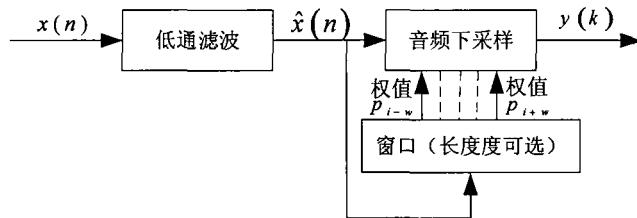


图 3.23 正弦重采样下采样框图

原始音频信号 $x(n)$ 首先通过低通滤波器，避免下采样时产生混叠，经过低通滤波的音频信号为 $\hat{x}(n)$ ， $\hat{x}(n)$ 经过选择好长度的窗口，在窗长范围内计算插值权值 p_i ， $\hat{x}(n)$ 与得到的重采样权值 p_i 同时进入音频下采样模块，进行音频信号重采样。

基于正弦重采样算法的下采样流程图如图 3.24 所示。

(1)首先初始化原始音频信号 $x(n)$ ，将原始音频信号的长度 N ，原采样频率 L ，新采样频率 M ，依式(3-1)求得将要生成的新音频信号的长度 K ，为即将生成的新音频信号分配长度为 K 的short型内存单元，选择窗口长度 $2w+1$ 。

(2)按帧读入原始音频信号，求得正弦重采样的具体位置 n_k 。

(3)为了避免出现混叠现象，在进行下采样之前，将原始音频信号通过低通滤波器。在选择好的窗口处长度范围内依式(3-12)计算正弦重采样权值 p_i ，窗口长度为 $2w+1$ ，所以需要求出 $2w+1$ 个插值权值。

(4)将经过低通滤波器的音频信号 $\hat{x}(n)$ 结合第(3)步得到的正弦重采样权值 p_i ，按照式(3-13)进行计算，得到新采样频率下的音频信号 $y(k)$ ，将新的音频信号 $y(k)$ 写入已经分配的内存单元。

(5)判断原始音频信号是否处理完毕，如果没有，则进入第(2)步继续按流程进行处理，直到处理完原始音频信号。

下采样时，选取由48kHz转换到16kHz和44.1kHz到320kHz，选择插值点前后各10个点来计算，即 $w=10$ ，窗口长度为 $2*w+1=21$ ，来说明用正弦重采样的算法进行下采样时的情形。

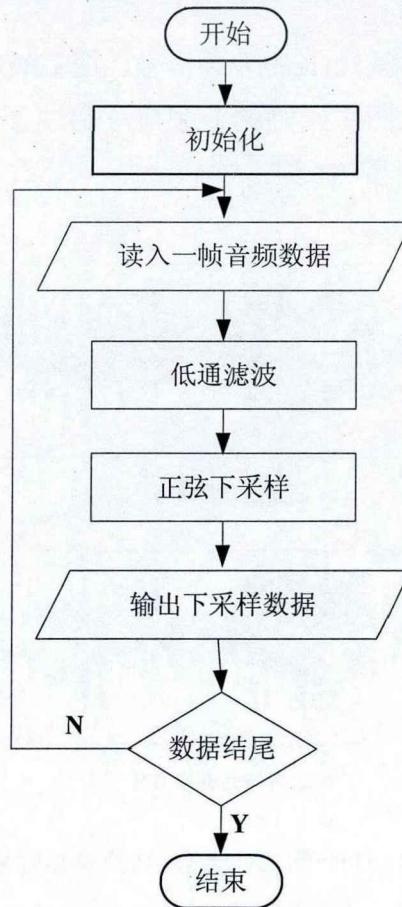


图 3.24 正弦重采样下采样流程图

图3.25(a)为原始采样率为48kHz的音频信号,图3.25(b)为由采样率48kHz转换到16kHz的音频信号,对比图中原始的音频信号和进行速率转换之后的音频信号,发

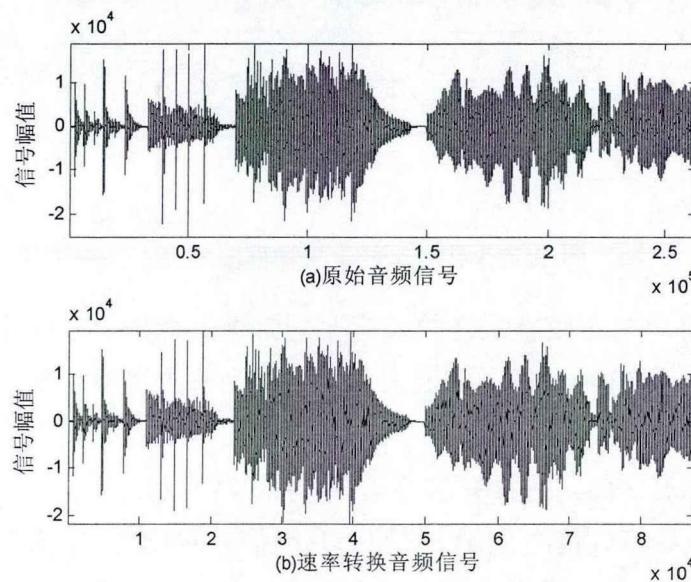


图3.25 48kHz到16kHz速率转换音频时域波形图

现在时域波形上没有明显的变化。

图3.26(a)为原始采样率为44.1kHz的音频信号，图3.26(b)为由采样率44.1kHz转换到32kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

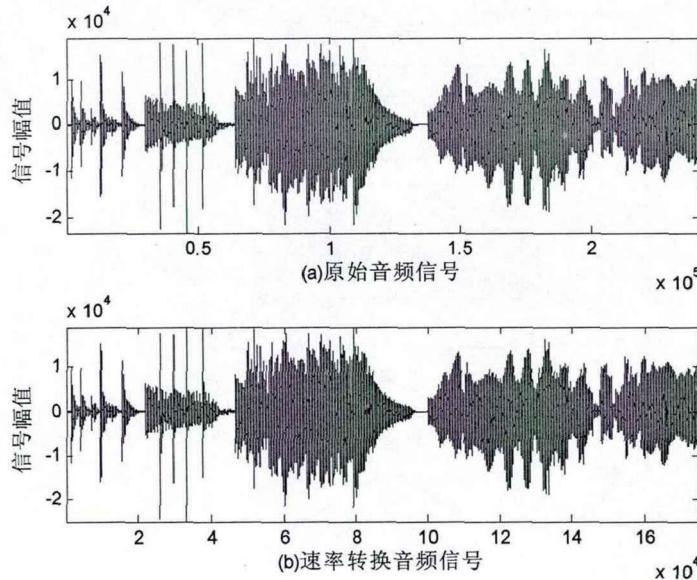


图3.26 44.1kHz到32kHz速率转换音频时域波形

从图3.25和图3.26可以看出：正弦重采样在进行下采样时，重采样生成的音频信号在时域波形上没有明显的失真，这说明正弦重采样能够很好的重构下采样音频信号。

用正弦重采样进行上采样时的框图如图3.27所示。

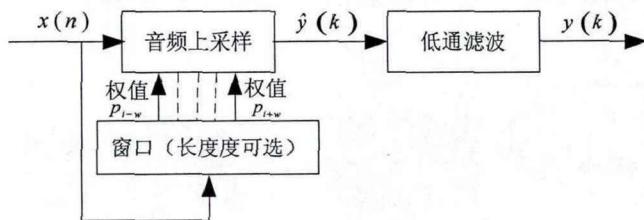


图 3.27 正弦重采样上采样框图

原始音频信号 $x(n)$ 与正弦重采样的权值 p_i 同时进入音频上采样模块，生成重采样音频信号 $\hat{y}(k)$ ， $\hat{y}(k)$ 经过低通滤波去除上采样带来的镜像得到目标音频重采样信号 $y(k)$ 。

基于正弦重采样算法的上采样算法流程图如图3.28所示。

(1)首先初始化原始音频信号 $x(n)$ ，将原始音频信号的长度 N ，输入原采样频率 L ，新采样频率 M ，依式(3-1)求得将要生成的新音频信号的长度 K ，为即将生成的新音频信号分配长度为 K 的short型内存单元，选择窗口长度 $2w+1$ 。

(2)按帧读入原始音频信号，求得正弦重采样的具体位置 n_k 。

(3)在选择好的窗口处长度范围内依式(3-12)计算正弦重采样权值 p_i ，窗口长度为 $2w+1$ ，所以需要求出 $2w+1$ 个插值权值，原始音频信号 $x(n)$ 结合权值 p_i 按照式(3-13)进行上采样，得到新采样频率下的音频信号 $\hat{y}(k)$ 。

(4)正弦重采样算法在窗口长度较大时可以有效避免上采样带来的镜像，但是其去除镜像的效果没有拉格朗日重采样明显。为了更好的消除镜像现象，将 $\hat{y}(k)$ 通过低通滤波，得到音频信号 $y(k)$ ，将 $y(k)$ 写入已经分配的内存单元。

(5)判断原始音频信号是否处理完毕，如果没有，则进入第(2)步继续按流程进行处理，直到处理完原始音频信号。

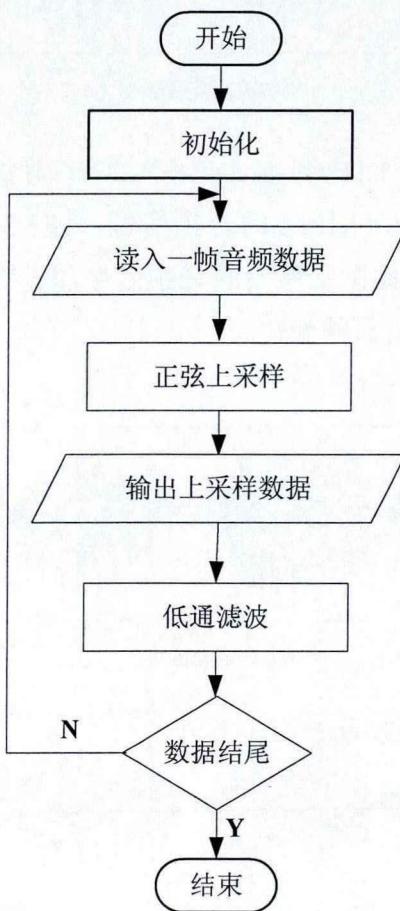


图 3.28 正弦重采样上采样流程框图

上采样时：选取由32kHz转换到48kHz和16kHz到44.1kHz，与下采样一样选择插值点前后各10个点来计算，即 $w=10$ ，窗口长度为 $2*w+1=21$ ，来说明用正弦重采样的算法进行上采样时的情形。

图 3.29(a)为原始采样率为 32kHz 的音频信号，图 3.29(b)为由采样率 32kHz 转换到 48kHz 的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

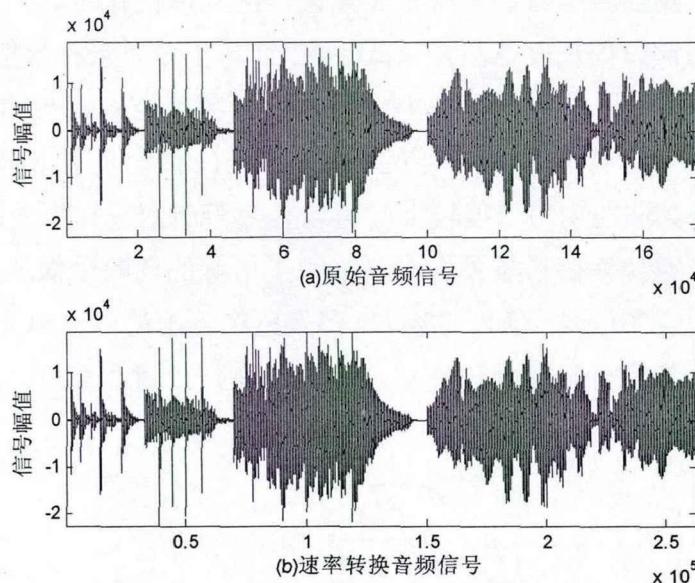


图3.29 32kHz到48kHz速率转换音频时域波形图

图3.30(a)为原始采样率为44.1kHz的音频信号，图3.30(b)为由采样率16kHz转换到44.1kHz的音频信号，对比图中原始的音频信号和进行速率转换之后的音频信号，发现在时域波形上没有明显的变化。

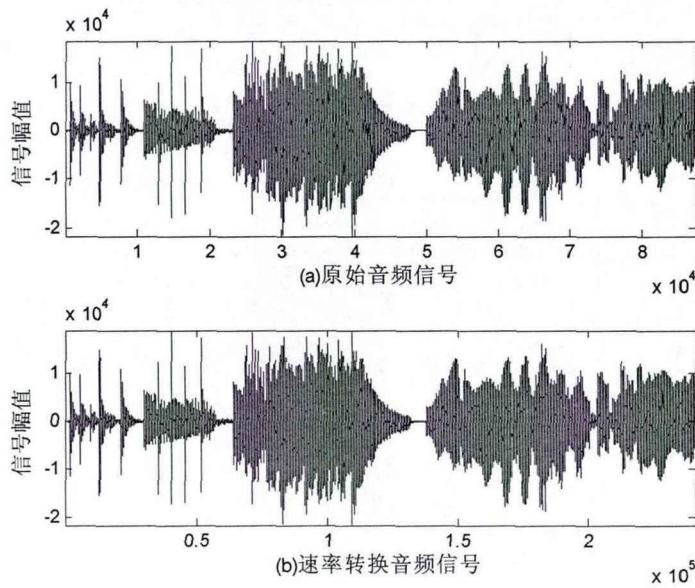


图3.30 16kHz到44.1kHz速率转换音频时域波形图

从图3.29和图3.30可以看出：正弦重采样在进行上采样时，重采样生成的音频信号在时域波形上没有明显的失真，这说明正弦重采样能够很好的重构上采样音频信号。

对音频信号使用正弦重采样的算法进行重采样，在48kHz, 44.1kHz, 32kHz, 16kHz之间的相互转换，无论是进行上采样还是进行下采样音频信号都得到了很好的重构。

测试环境和音频信号与线性重采样和拉格朗日重采样一样，同样选择插值点前后各10个点来计算，即 $w=10$ ，窗口长度为 $2*w+1=21$ 的正弦重采样算法处理音频信号，测得各速率之间进行转换的时间开销如表3.3所示。

表3.3 正弦重采样算法实现各速率之间转换的时间

原采样速率 转换后速率	48kHz	44.1kHz	32kHz	16kHz
48kHz		1203ms	1203ms	1187ms
44.1kHz	1125ms		1093ms	1109ms
32kHz	906ms	843ms		828ms
16kHz	562ms	531ms	500ms	

从表3.3中明显可以看出在各速率之间的转换时，上采样比下采样在时间开销方面要大，这是因为上采样是增加了采样点个数，下采样时减少了采样点个数。采用同样正弦重采样算法，增加采样点个数肯定用去的时间比减少采样点个数要多；同时可以看到在各个不同速率转换到同一个指定速率时，在时间开销方面相差无几。

在不使用滤波器的情况下，对比生成音频信号的频域，可以看到，正弦重采样在上采样时随着采样因子的增大，在频域产生镜像的现象能很好的得到控制；在下采样时，随着采样因子的增大，在频域产生混叠的现象要比线性重采样和拉格朗日重采样产生混叠严重。所以，从频域角度来看，正弦重采样更加适合应用于上采样。

3.4 本章小结

在重采样原理的基础上，设计了三种音频重采样的算法：线性重采样算法、拉格朗日重采样算法、正弦重采样算法。对比三种重采样算法的原理看出，线性重采样的原理在三种重采样算法中最为简单，只需要相邻两点的幅值就可确定插值点的幅值，然后通过简单的数学公式确定这两个幅值进行线性运算的权值；拉格朗日重采样算法在每求得一个采样点的幅值之前都要先通过复杂的数学公式求得重采样权值，求得重采样权值的数学公式十分复杂，很明显在算法复杂度上比线性重采样要复杂，在时间开销方面要比线性重采样花费更多的时间来求得重采样权值；正弦重采样在每求得一个插入幅值前，也必须确定重采样权值，权值也是通过复杂的数学公式来进行计算，但是从数学角度来分析，正弦重采样比拉格朗日重采样的权值因子在计算方会计算量小一些，所以在时间开销方面比拉格朗日重采样要小；但是正弦重采样比线性重采样要复杂得多，所以在时间开销方面正弦重采样要更费时。

从频域的角度来看，在上采样时，如果没有滤波器^[33-36]，正弦重采样和拉格朗日重采样算法自身就可以很好的抑制镜像的产生，重采样得到较好的音频信号，但是线性重采样在没有滤波器的时候，会发生严重的频率混叠现象，所以在上采样时尽量采用正弦重采样和拉格朗日重采样算法，虽然会更费时，但是会得到较好的重采样音频信号；在下采样时，如果没有滤波器，线性重采样在频域产生的混叠现象不是很明显，正弦重采样和拉格朗日重采样比较线性重采样而言，没有明显的改善情况，而且在时间开销上，线性重采样会有明显的优势，所以建议在下采样时，可以考虑用线性重采样算法。

从收敛性来看，为了提高拉格朗日重采样的精度，常常增加节点的个数，即提高拉格朗日重采样多项式的阶次。但是当阶次过高时，会产生龙格(Runge)现象(龙格现象是用高阶多项式进行多项式插值时所出现的问题。龙格在研究多项式插值的时候，发现有的情况下，并非取节点越多，多项式就越精确。著名的例子是 $f(x)=1/(1+25x^2)$ 。它的插值函数在两个端点处发生剧烈的波动，造成较大的误差。究其原因，是舍入误差造成的。)，即当节点加密时，拉格朗日重采样只能在较小的区间收敛，而在其他区间发散。正弦重采样实际上是三角多项式插值。我们知道许多复杂的波形可以分解为一系列谐波的叠加，所以正弦重采样当窗口长度过大时，也会出现龙格现象。这是所有基于多项式插值的缺点。

而线性重采样是线性插值，只是用相邻两点的值来估计插入点的幅值，所以线性重采样不会出现龙格现象。

第四章 性能评估

4.1 音频测试

国际电信联盟ITU(International Telecommunication Union)工作组提出音频质量的感知评估方法^[37](PEAQ:Perceptual Evaluation of Audio Quality)以及语音质量的感知评估方法^[38](PESQ:Perceptual Evaluation of Speech Quality)。音频质量的感知评估方法测试方法主要适用于高采样率的音频信号，而语音质量的感知评估方法测试方法主要适合于语音和窄带音频信号。对于音频质量的评价归纳起来大致分为：主观评价和客观评价。

主观评价方法是以人为主体来评价声音的质量，符合人耳对声音质量的感觉，其缺点是受人的主观影响较大，重复性和稳定性较差。最主要的主观评定方法由国际电信联盟无线电通信分部，在多家广播业者以及MPEG音频小组的协助下，制订了一套非常详细的听音测试规则：主观评定等级(Subjective Opinion Scale)，或称为平均意见分^[39](MOS: Mean Opinion Score)，这套规则的目的在于在最苛刻的环境下，由专家听者对编码器的表现打分。国际电信联盟无线电通信分部的测试过程要求有数量相对较大的测试主体听着，测试是双盲的，即在举行测试的房间里，包括听者在内的所有的人员都不知道所播放的压缩材料出自哪一种编码器。参与者比较不同的编码信号样本，并把它们同未经压缩的原始信号比较，之后对各编码器打分。最后评价结果以分为五级的平均意见分或比较平均意见分(CMOS: Comparison Mean Opinion Score)通常采用七级分制的形式给出。

主观评价，不仅费时费力，而且往往受种种测试条件的限制和测试人员主观因素的影响，使得测试结果的可靠性受到一定的影响。因此，人们就希望采用客观测评来评价音频系统的音质。它不但不受条件限制和人的主观因素影响，而且，应用起来也机动、灵活得多。不同时间和不同场合的测试结果也便于直接比较。

客观评定方法是用客观测量的手段来评价音频信号的质量，信噪比(SNR: Signal-to-Noise Ratio)就是常用的一种。信噪比是在整个音频范围上的测量尺度，是对音频信号的长期测度，不能充分反映音频信号的时变特征。客观评定方法还包括段信噪比(SegSNR: Segment Signal-to-Noise Ratio)、算法复杂度(时间开销)等，它们都建立在度量均方误差的基础上，对于增益变化和延迟等很敏感，计算简单，特点是客观直观。

所以本文采取了客观的评测方法，从时间开销、信噪比、段信噪比方面来评测三种重采样音频算法。

4.2 性能评估与结论

4.2.1 时间开销

测试环境：在CPU为AMD Sempron(tm) Processor 2500+，主频为1.41GHz，内存为512MB的PC机上，软件开发平台为Microsoft Visual C++ 6.0上，一段播放时长为5461ms的音频信号，拉格朗日重采样和正弦重采样窗口长度取插值点前后各10个点来计算插值点幅值，即 $w=10$ ，由式(3-7)知，窗口长度为 $2*w+1=21$ 与取插值点前后各5个点来计算插值点幅值，即 $w=5$ ，窗口长度为 $2*w+1=11$ 。

原始音频信号从48kHz采样率到44.1kHz、32kHz、16kHz采样率的下采样时，测得不同的重采样算法在时间开销方面的具体数据如表4.1。

表4.1 三种重采样算法下采样的时间开销

原采样速率为 48kHz						
重采样算法 转换后速率	线性重采样	拉格朗日重采样		正弦重采样		
		w = 10	w = 5	w = 10	w = 5	
44.1kHz	312ms	4265ms	1578ms	1125ms	734ms	
32kHz	281ms	3203ms	1093ms	906ms	640ms	
16kHz	265ms	1718ms	656ms	562ms	468ms	

从表4.1中的数据可以看出：在原始采样速率为48kHz时，对它进行下采样，分别转换到速率为44.1kHz、32kHz、16kHz，在时间开销方面，很明显线性重采样要比拉格朗日重采样和正弦采样花费的时间少，时间开销方面的优势很明显；对比拉格朗日重采样和正弦重采样我们可以看出，正弦重采样的时间开销要比拉格朗日重采样的时间开销小，但是优势没有线性插值明显。而且拉格朗日和正弦音频重采样在选择窗口长度为21时的时间开销比窗口长度为11时的时间开销大。之所以出现这种情况，这是因为增加了窗口长度便相应的增加计算量。

原始音频信号从16kHz采样率到48kHz、44.1kHz、32kHz采样率的上采样时，测得不同的重采样算法在时间开销方面的具体数据如表4.2。

表4.2 三种重采样算法上采样的时间开销

原采样速率为 16kHz						
重采样算法 转换后速率	线性重采样	拉格朗日重采样		正弦重采样		
		w = 10	w = 5	w = 10	w = 5	
48kHz	296ms	4593ms	1546ms	1187ms	769ms	
44.1kHz	281ms	4234ms	1468ms	1109ms	718ms	
32kHz	218ms	3078ms	1046ms	828ms	593ms	

从表4.2中的数据可以看出：在原始采样速率为16kHz时，对它进行上采样，分

别转换到速率为48kHz、44.1kHz、32kHz，在时间开销方面，如同下采样一样，很明显线性重采样要比拉格朗日重采样和正弦采样花费的时间少，时间开销方面的优势很明显；对比拉格朗日重采样和正弦重采样我们可以看出，正弦重采样的时间开销同样要比拉格朗日重采样的时间开销少，但是优势同样没有线性插值明显。同样，拉格朗日和正弦音频重采样算法窗口较大时，时间开销明显大，原因同下采样一样。

从三种算法在时间开销方面的具体数据对比可以看出，线性音频重采样算法的实时性更好，更适合在实时性要求比较高的情况下处理相关的音频信号；拉格朗日音频重采样与正弦音频重采样的时间开销明显比线性音频重采样都大，但是在相同窗口长度时，正弦音频重采样比拉格朗日音频重采样的时间开销小，这说明正弦重采样在保证重采样音频信号质量的同时，又可以较好的满足音频信号处理的实时性要求。

之所以在时间开销方面会出现这样的结果，是因为三种重采样算法在设计时不同的复杂度所决定的，无论上采样还是下采样，线性插值在时间开销上都具有明显的优势。从设计思想方面分析可以回答出现这种现象的具体原因：

(1)线性重采样算法的思想最为简单，它是利用在同一条直线上的两点的线性权值来逼近所需要重采样的点的幅值，其计算量小，任意点的插值仅仅需要相邻两点的幅度值，然后将其幅值进行线性加权，其中权值如何选择是这种重采样算法的关键所在。这种算法思想简单，用它来对瞬变值较小的音频信号进行重采样时，不仅可以得到较好的效果，而且由于算法的思想简单，所以在时间开销方面会有很大的优势。

(2)拉格朗日重采样算法首先要确定一个窗长，窗口的长度不仅决定了所要重采样点的幅值必须由在此窗长范围内的所有点的幅值的加权来确定，而且在加权之前，首先必须由这个窗长内的所有点的幅值通过复杂的数学运算，来确定每一个窗长内的点在进行重采样时所需要的权值，每一个权值的确定都需要这个窗长内的所有点的幅值进行复杂运算。等到确定了权值之后，然后将权值和窗长内的幅值组合，最终得到所要重采样点的幅度值。由于这种算法的窗口长度是可以进行调节，所以可以灵活的选取窗长以适应不同重采样质量要求。窗口的长度选择的越长，计算的复杂度就越高，所需要重采样点的幅值就会越精确，但是在窗口长度到达一定长度的时候，改善的效果就不会太明显；同时在时间开销方面，选择的窗口长度越长，时间的开销量也就越大。所以，在重采样的质量和重采样的时间开销方面必须做到一个平衡。这种算法相较于线性重采样，其思想复杂度很高，而且窗口的长度可以根据需要进行不同的选取，所以更加具有灵活性。其中，线性重采样可以看作是窗口长度选取为2的特殊的拉格朗日的重采样算法，但是拉格朗日重采样由于算法思想的复杂性决定了，对于任何音频信号的重采样，其

在时间开销方面要比线性重采样大得多。这也就是为了得到高质量的重采样的效果所必须付出的代价。拉格朗日重采样算法不仅可以将瞬变值很小的音频信号得到很好效果，而且在瞬变值大的音频重采样信号中也可以得到很好的效果。之所以很这样，是因为拉格朗日重采样选取进行加权确定重采样点的幅值的相邻点的个数比线性重采样的点的个数多，而且还可以根据音频信号的瞬变值的大小来灵活的选取窗口的长度。相对于线性重采样，拉格朗日重采样不仅可以更加精确的确定重采样点的幅度值，而且能够根据需要在时间开销和重采样质量之间做出选择，所以拉格朗日重采样相对于线性重采样有所改进。

(3)正弦重采样算法和拉格朗日重采样算法一样，在确定重采样点的幅值之前，首先要确定一个窗长，同样窗口的长度不仅决定了所要重采样点的幅值必须由在此窗长范围内的所有点的幅值进行加权计算，而且在加权之前，必须由这个窗长内的所有点的幅值，通过复杂的数学运算来确定窗长内的每个幅值在进行重采样时所需要的权值，每一个权值的确定都需要这个窗长内的所有点的幅值进行复杂运算。只不过在权值的计算方面，正弦重采样用的是正弦函数的方法来计算所需要的权值。等到确定权值之后，然后将权值和窗长内的幅值进行组合，最终得到所要重采样点的幅度值。这种算法的窗口长度同样是可以进行调节，所以相比较线性重采样来说，也可以灵活的选取窗长以适应不同重采样质量要求。窗口的长度选择的越长，计算的复杂度就越高，所需要重采样点的幅值就会越精确，同拉格朗日重采样一样，窗口长度达到一定长度时，改善的效果将不会太明显；同时在时间开销方面，选择的窗口长度越长，时间开销也就越大。正弦重采样算法相较于线性重采样，其思想复杂度很高，而且窗口的长度可以根据需要进行不同的选取，所以更加具有灵活性，但是相较于拉格朗日重采样算法来讲，正弦重采样的算法复杂度和时间开销都比较小。正弦重采样由于算法思想的复杂性决定了，对于任何音频信号的重采样，其在时间开销方面要比线性重采样大得多。正弦重采样算法也不仅可以使瞬变值很小的音频信号得到很好效果，而且在瞬变值大的音频重采样信号中也可以得到很好的效果，理由和拉格朗日重采样相同。正弦重采样算法同样能够根据需要在时间开销和重采样质量之间做出选择。

4.2.2 信噪比

一个音频信号的有效性，通常采用-信噪比的提高作为评估的标准。信噪比的定义如下：

$$SNR = 10 \log_{10} \left\{ \frac{\sum_{n=0}^M s^2(n)}{\sum_{n=0}^M (s(n) - \hat{s}(n))^2} \right\} \quad (4-1)$$

其中 $s(n)$ 为原始音频信号， $\hat{s}(n)$ 为转换得到的同频率音频信号， M 为音频信号的长度。信噪比可以很好的反映重构音频信号在幅值上与原始音频信号整体上的差异。信噪比越大，说明重构音频信号与原始音频信号整体越接近，重构信号的整体失真越小；信噪比越小，说明重构音频信号与原始音频信号的整体差异越大，重构信号的整体失真越大。

采用不同算法得到的信噪比为表4.3(线性音频重采样算法)，表4.4(正弦音频重采样算法，取插值点前后各10个点重构的音频信号，即 $w=10$)，与取插值点前后各5个点重构的音频信号，即 $w=5$)，表4.5(拉格朗日音频重采样算法，取插值点前后各10个点重构的音频信号，即 $w=10$ ，与取插值点前后各5个点重构的音频信号，即 $w=5$)。

从表4.3可以看出线性重采样在下采样时的信噪比相对于上采样时的信噪比大，随着采样因子的变化，下采样的变化幅度较大，但是信噪比依然比上采样高，这说明线性重采样适合于下采样。

表4.3 线性重采样算法各速率转换时的信噪比(单位：dB)

原采样速率 转换后速率	48kHz	44.1kHz	32kHz	16kHz
48kHz		14.422	10.092	3.341
44.1kHz	21.763		10.941	3.342
32kHz	14.359	12.153		4.485
16kHz	10.475	8.359	10.716	

从表4.4中比较正弦重采样在下采样时的信噪比和上采样时的信噪比，在由同一个采样速率进行上下采样而且上下采样因子相差不大时，其信噪比也相差不大。随着采样因子的增大，上下采样的信噪比都有所下降，但是相比较线性音频重采样，正弦音频重采样在上下采样时，信噪比都比线性重采样的信噪比高，这说明正弦音频重采样比线性音频重采样的效果好。将计算插值点幅值的窗口增大时，信噪比没有明显的变化。这说明正弦重采样在窗口达到一定的长度时，重构信号的质量就不再随着窗长的增加而有明显改善。

表4.4 不同窗口长度正弦重采样算法各速率转换时的信噪比(单位: dB)

原采样速率 转换后速率	48kHz		44.1kHz		32kHz		16kHz	
	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5
48kHz			26.977	24.181	19.686	20.440	7.874	8.203
44.1kHz	16.517	15.609			28.556	26.815	8.873	9.268
32kHz	27.695	27.546	25.151	25.944			10.380	10.728
16kHz	14.163	14.161	12.875	13.564	17.370	17.369		

从表4.5中拉格朗日重采样在下采样时的信噪比和上采样时的信噪比对比来看，上下采样时，拉格朗日重采样和正弦重采样的信噪比相差无几，但是比线性重采样的信噪比高。随着采样因子的增大，上下采样时信噪比也是都有所下降。同样，将计算插值点幅值的窗口增大时，信噪比没有明显的变换。这说明拉格朗日重采样在窗口达到一定的长度时，重构信号的效果就不会再随着窗口长度的增大而有明显改善，所以窗口的长度选取到一定的长度即可。

表4.5 不同窗口长度拉格朗日重采样算法各速率转换时的信噪比(单位: dB)

原采样速率 转换后速率	48kHz		44.1kHz		32kHz		16kHz	
	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5
48kHz			26.324	26.334	20.096	20.115	8.000	8.014
44.1kHz	16.258	16.271			29.114	28.968	9.032	9.049
32kHz	28.200	28.198	25.750	25.755			10.529	10.546
16kHz	14.162	14.162	13.102	13.102	17.369	17.369		

对比表4.3、表4.4、表4.5，上采样时，拉格朗日重采样和正弦重采样的信噪比明显比线性重采样的信噪比大，而且上采样因子变化越大，拉格朗日重采样和正弦重采样的信噪比相对于线性重采样的信噪比优势越明显，这说明拉格朗日重采样和正弦重采样比线性重采样更适合进行上采样；在下采样时，拉格朗日重采样和正弦重采样算法的信噪比相对于线性重采样的优势没有上采样时明显，三种音频重采样算法中线性重采样在下采样因子变化大时，线性重采样的信噪比与其它两种算法相差不多。拉格朗日重采样和正弦重采样在相同窗口长度时，上下重采样时的信噪比区别都不大，但是考虑到时间开销的不同，正弦重采样算法就比拉格朗日算法有优势。

表4.6中数据为三种算法在上采样时的时间开销和信噪比，其中拉格朗日重采样和正弦重采样选取插值点前后各5个点重构的音频信号即 $w = 5$ ；表4.7中数据为三种算法在下采样时的时间开销和信噪比，其中拉格朗日重采样和正弦重采样选取插值点前后各5个点重构的音频信号即 $w = 5$ 。

表4.6三种重采样算法上采样的时间开销与信噪比

原采样速率为 16kHz						
重采样算法 转换后速率	线性重采样		拉格朗日重采 $w=5$		正弦重采样 $w=5$	
	时间开销	信噪比	时间开销	信噪比	时间开销	信噪比
48kHz	296ms	3.341dB	1546ms	8.014dB	796ms	8.203dB
44.1kHz	281ms	3.342dB	1468ms	9.049dB	718ms	9.268dB
32kHz	218ms	4.485dB	1046ms	10.546dB	593ms	10.728dB

表4.7 三种重采样算法下采样的时间开销与信噪比

原采样速率为 48kHz						
重采样算法 转换后速率	线性重采样		拉格朗日重采 $w=5$		正弦重采样 $w=5$	
	时间开销	信噪比	时间开销	信噪比	时间开销	信噪比
44.1kHz	312ms	21.763dB	1578ms	16.271dB	734ms	15.609dB
32kHz	281ms	14.359dB	1093ms	28.198dB	640ms	27.546dB
16kHz	265ms	10.475dB	656ms	14.162dB	468ms	14.161dB

综合时间开销和信噪比来看，线性音频重采样算法的虽然具有最好的实时性，可以很快的处理相关音频信号，但是线性音频重采样的信噪比相对于拉格朗日音频重采样和正弦音频重采样就比较差，所以在实时性要求高，但是重采样信号的质量要求不是过于严格的情况下，可以考虑用线性音频重采样算法来处理相关的音频信号；对比正弦音频重采样算法和拉格朗日音频重采样算法，正弦音频重采样算法在信噪比方面与拉格朗日音频重采样算法相差无几，但是在相同窗口长度时，正弦音频重采样实时性比拉格朗日音频重采样算法要好，所以若要求经处理后的音频信号既要有很好的质量又要很好的实时性时，可以使用正弦音频重采样算法去处理相关的音频信号。

4.2.3 段信噪比

信噪比是在整个音频范围上的测量尺度，是对音频信号的长期测度，不能充分反映音频信号的时变特征。所以引入了段信噪比，段信噪比和信噪比都是建立在度量均方误差的基础上，段信噪比是基于帧的信噪比。通过计算音频信号每一帧的信噪比，最后取平均值得到的。相对于信噪比可以很好的反映重构音频信号整体在幅值上与原始音频信号的差异，段信噪比可以更好的反映每帧重构音频信号与原始每帧音频信号的差异。段信噪比越大，说明重构音频信号与原始音频信号在每帧的幅值差异越小，重构信号每帧的失真越小；段信噪比越小，说明重构音频信号每帧与原始音频信号每帧的幅值差异越大，重构信号的每帧失真越大。

段信噪比计算的公式如下：

$$\text{SegSNR} = \frac{1}{N} \sum_{m=0}^{M-1} \log_{10} \left\{ \frac{\sum_{n=1}^r x_n^2}{\sum_{n=1}^r (x_n - \hat{x}_n)^2} \right\} \quad (4-2)$$

其中 $s(n)$ 为原始音频信号， $\hat{s}(n)$ 为转换得到的同频率音频信号， M 为音频信号的帧数， r 为音频信号每帧的长度。采用不同算法得到的段信噪比为表4.6(线性音频重采样算法)，表4.7(正弦音频重采样算法，取插值点前后各10个点重构的音频信号，即 $w=10$)，与取插值点前后各5个点重构的音频信号，即 $w=5$)，表4.8(拉格朗日音频重采样算法，取插值点前后各10个点重构的音频信号，即 $w=10$ ，与取插值点前后各5个点重构的音频信号，即 $w=5$)。

表4.8 线性重采样算法各速率转换时的段信噪比(单位：dB)

原采样速率 转换后速率	48kHz	44.1kHz	32kHz	16kHz
48kHz		15.623	10.413	2.955
44.1kHz	23.522		11.549	2.955
32kHz	15.824	13.166		4.202
16kHz	10.405	8.195	10.672	

从表4.8可以看出线性重采样在下采样时的段信噪比相对于上采样时的段信噪比高，即每帧的失真较小，随着采样因子的增大，上下采样段信噪比都有所下降。

表4.9 不同窗口长度正弦重采样算法各速率转换时的段信噪比(单位：dB)

原采样速率 转换后速率	48kHz		44.1kHz		32kHz		16kHz	
	$w=10$	$w=5$	$w=10$	$w=5$	$w=10$	$w=5$	$w=10$	$w=5$
48kHz			28.825	24.198	22.027	25.271	7.683	8.176
44.1kHz	18.009	16.402			34.637	28.769	8.744	9.366
32kHz	34.189	32.619	29.631	34.054			10.436	11.034
16kHz	14.239	14.248	12.763	13.764	17.849	17.852		

从表4.9可以看出正弦重采样在上采样时的段信噪比相对于线性重采样在上采样时的段信噪比高，即每帧的失真较小，会有较好的效果。随着采样因子的增大，下采样时段信噪比下降很快；上采样时，段信噪比随着采样因子的变化就不是很明显，说明正弦重采样算法对于上采样时每帧的失真的幅度随着采样因子变化较小，适合于上采样。在窗口长度变大时，段信噪比没有明显的变化，这说明，段信噪比在窗口的长度达到一定范围时，不再随着窗口长度的增加有明显的改善。

表4.10 不同窗口长度拉格朗日重采样算法各速率转换时的段信噪比(单位: dB)

原采样速率 转换后速率	48kHz		44.1kHz		32kHz		16kHz	
	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5	w = 10	w = 5
48kHz			27.109	27.107	23.199	23.202	7.836	7.848
44.1kHz	17.397	17.399			37.111	37.088	8.935	8.948
32kHz	44.488	44.494	33.504	33.509			10.628	10.640
16kHz	14.259	14.259	13.138	13.138	17.868	17.868		

表4.10中拉格朗日重采样在下采样时的段信噪比和上采样时的段信噪比相比来看，拉格朗日重采样和正弦重采样的段信噪比相差无几，但是比线性重采样的段信噪比高。随着采样因子的增大，下采样时段信噪比下降很快；上采样时，段信噪比随着采样因子的变化不是很明显，这说明拉格朗日重采样算法对于上采样每帧失真随着采样因子的变化较小，对于上采样与正弦重采样算法一样有着不俗的效果。同正弦重采样一样，在窗口长度到达一定长度时，段信噪比不会再随着增加窗口长度而有明显改善。

对比表4.8、表4.9、表4.10，上采样时，拉格朗日重采样和正弦重采样的段信噪比明显比线性重采样的段信噪比大，而且上采样因子变大时，拉格朗日重采样和正弦重采样的段信噪比相对于线性重采样的段信噪比优势更大，即每帧的失真随着采样因子的变化小。另一方面，窗口长度相同时，拉格朗日重采样和正弦重采样在上下重采样时的段信噪比相差不大。

表4.11中数据为三种算法在上采样时的时间开销和段信噪比，其中拉格朗日重采样和正弦重采样选取插值点前后各5个点重构的音频信号即 $w = 5$ ；表4.12中数据为三种算法在下采样时的时间开销和段信噪比，其中拉格朗日重采样和正弦重采样选取插值点前后各5个点重构的音频信号即 $w = 5$ 。

表4.11三种重采样算法上采样的时间开销与段信噪比

原采样速率为 16kHz						
重采样算法 转换后速率	线性重采样		拉格朗日重采 $w = 5$		正弦重采样 $w = 5$	
	时间开销	段信噪比	时间开销	段信噪比	时间开销	段信噪比
48kHz	296ms	2.955dB	1546ms	7.848dB	796ms	8.176dB
44.1kHz	281ms	2.955dB	1468ms	8.948dB	718ms	9.366dB
32kHz	218ms	4.202dB	1046ms	10.640dB	593ms	11.034dB

表4.12 三种重采样算法下采样的时间开销与段信噪比

原采样速率为 48kHz						
重采样算法 转换后速率	线性重采样		拉格朗日重采 $w=5$		正弦重采样 $w=5$	
	时间开销	段信噪比	时间开销	段信噪比	时间开销	段信噪比
44.1kHz	312ms	23.522dB	1578ms	17.399dB	734ms	16.402dB
32kHz	281ms	15.824dB	1093ms	44.494dB	640ms	32.619dB
16kHz	265ms	10.405dB	656ms	14.259dB	468ms	14.248dB

综合表4.11、表4.12可以得出，线性音频重采样算法的具有最好的实时性，可以很快的处理相关音频信号；但是在上采样和下采样时，线性音频重采样的段信噪比相对于拉格朗日音频重采样和正弦音频重采样就比较差，即线性重采样算法生成的音频信号每帧的失真较其他两种算法生成的音频信号失真大。所以在实时性要求高，但是重采样信号的质量要求不是过于严格的情况下，可以考虑用线性音频重采样算法来处理相关的音频信号；对比正弦音频重采样算法和拉格朗日音频重采样算法，正弦音频重采样算法在段信噪比方面与拉格朗日音频重采样算法相差无几，但是在相同窗口长度时，实时性比拉格朗日音频重采样算法要好，所以若经处理后的音频信号既要要求有很好的质量又要很好的实时性时，可以使正弦音频重采样算法去处理相关的音频信号。

4.3 小结

本章通过客观测试，对三种算法在时间开销，信噪比、段信噪比等方面进行了比较和分析，考虑到各方面的因素，最终得到的结论是：线性重采样算法比较适合下采样，在下采样时，不仅时间方面有优势，而且可以取得较高的信噪比和段信噪比；正弦重采样算法在上采样时，有较高的段信噪比，比较适合上采样；拉格朗日重采样与正弦重采样类似在上采样时也会有不错的效果。但是在时间开销方面，线性重采样有着明显的优势。

拉格朗日重采样和正弦重采样可以调节窗口长度，在质量和时间方面可以灵活选取，相对于线性重采样是一种改进。

第五章 结论

本文在多抽样率技术的基础上，重要研究了音频重采样以及实现音频重采样的关键技术的具体实现音频重采样的三种算法，对于音频信号实现了 48kHz、44.1 kHz、32 kHz、16 kHz 采样率之间的互相转换，同时分析了三种算法的复杂度以及时间开销，对三种算法在应用于何种音频信号会取得更加好的效果进行了比较和说明。其中具体的工作如下：

(1) 研究了多抽样率技术在多媒体通信和音/视频方面的应用，对于将多抽样率中的技术应用于音频重采样进行了仔细的研究和分析。

(2) 研究了实现音频重采样的关键技术，设计了音频重采样的三种具体的算法：线性重采样、拉格朗日重采样、正弦重采样，分别用三种重采样算法实现了音频信号的重采样，在采样频率 48kHz、44.1kHz、32kHz、16kHz 之间实现了任意两个速率之间的转换。并且每种算法在音频速率转换之间得到的音频信号都取得了不错的效果。

(3) 对研究的三种算法进行了音频测试与性能客观评估工作，并将三种重采样算法之间进行了比较。

从算法设计思想上，首先对三种重采样算法进行了分析，比较了用三种重采样算法实现音频重采样所要进行的计算量，从算法复杂度的方面分析了实现音频重采样在时间开销方面的区别。在理论上，对三种重采样算法要成功实现音频重采样的优劣进行了仔细的分析。

在具体的时间开销上，通过具体的时间开销，列表对三种算法分别实现不同速率转换时所花费的时间进行了比较，分析。总结了每种算法的优缺点和分别适用于具体音频信号的转换情况。在原音频信号瞬时幅值变化较大的时候，适合用拉格朗日重采样和正弦重采样来对音频信号进行重采样；在音频信号的质量要求不是非常严格的时候我们适用于线性重采样算法来对信号进行转换，不仅可以得到不错的效果，而且在时间开销方面会取得很大的优势。

从三种算法生成音频信号的信噪比和段信噪比方面，将具体的数据列表进行对比和分析，对每种算法应用音频重采样时的信噪比和段信噪的具体数据分析比较，总结出了每种重采样算法分别适合于哪种采样率的音频信号转换，结合信噪比、段信噪比和时间开销，总结了三种算法分别在什么情况下适合处理哪类音频信号。

总之，通过本课题的研究开发工作，我有机会对音频以及音频重采样的算法以及如何实现音频信号重采样，有了深入和全面的了解，同时在算法的比较和实现方面有了一定得深入了解和认识，这为我以后走上工作岗位打下了坚实的基础。虽然在课题的研究上有了一定得进展，但是课题还存在很多不足之处，一些地方

仍需要进一步加以完善和改进。

致谢

伴随着学位论文的完成，我两年半的研究生生活也趋近了尾声。在校攻读硕士研究生的这两年半的时光，我受益匪浅，这将是我终生难忘的美好回忆。

我很庆幸我能成为马鸿飞老师的学生，本论文的工作自始至终都是在他的亲切耐心指导下进行的。在此过程中，马老师以他渊博的学识、精湛的学术造诣、敏锐的洞察思考能力给我的课题研究提供大量的建议和帮助，使我在学习和科研工作中不断进步；马老师用他刻苦的学习精神、严谨的工作态度、良好的求学方法和正直的做人原则潜移默化地教会我修身治学的道理，这些都会使我受益终生。马老师平易近人，让教研室如同一个温暖和谐的大家庭。在此，我对马老师表示最为诚挚的谢意。

同时还要感谢关心和帮助我的所有同学和朋友：吴礼仲、徐雅俊、郭小川、梁冬蕾、王辛远、熊静、房川、郭泽华、张恩勇等。感谢他们在科研项目中的协作与帮助，感谢他们在学习和生活上对我的关心和支持，谢谢他们让我体会到同学时光、朋友情谊的可贵。我要深深感谢我的家人，他们无微不至的关怀与支持是我一直努力的动力之源，对于他们不需要说太多感谢，我想我不懈的努力是对他们最好的回报。

最后，感谢所有论文评审的老师和对本文提出宝贵意见的各位专家、教授、老师们！

参考文献

- [1] P.P.Vaidyanathan, SeniorMembeb. “Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial [C]”.Proceedings of the IEEE 1990.pp.56-93.
- [2] Crochiere R E, Rabiner L R. “Interpolation and Decimation of Digital Signals: A tutorial review [C]”, Proceedings of the IEEE, 1981. pp 300-331.
- [3] 宗孔德. “多抽样信号处理(第一版)”，北京：清华出版社，1996.pp 114-150.
- [4] Marin Vetterli, “A Theory of Multirate Filter Banks”, Transactions on Acoustics, Speech and Signal Processing of IEEE 1987. pp 356-372.
- [5] 陶然, 张慧云, 王越.“多抽样信率数字信号处理理论及其应用(第一版)”. 北京：清华出版社，2007. pp 248-270.
- [6] Shailesh B. Nerurkar, Khalid H. Abed. “Low-power decimator design using approximated linear-phase N-band IIR filter”. IEEE Transactions on Signal Processing. 4, April 2006. pp 1550-15531.
- [7] Tim Hentschel, Gerhard Fettweis. “Sample Rate Conversion for Software Radio ”. Software and DSP in Radio of the IEEE,2000. pp 142-150.
- [8] Wen Jin, Michael S. Scordilis, Alexander Iliev. “Comparison and Implementation of a 16-bit Fixed-point Audio Resampler”. Proceedings of the IEEE, 2004. pp1798-1800.
- [9] Martin Raspaud, Sylvain Marchand. “Enhanced Resampling for Sinusoidal Modeling Parameters”.2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. pp 1798-1800.
- [10] Tuncer Can Aysal, Kenneth E. Barner. “Generalized Mean-Median Filtering for Robust Frequency-Selective Applications”. IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 55, NO. 3, MARCH 2007.pp.937-948.
- [11] Jo Yew Tham, Lixin Shen, Seng Luan Lee. “A General Approach for Analysis and Application of Discrete Multiwavelet Transforms”. IEEE Transactions on Signal Processing, VOL. 48, NO. 2, February 2000.pp.457-464.
- [12] Aili Li, Klaus Mueller, Thomas Ernst. “Methods for Efficient, High Quality Volume Resampling in the Frequency Domain”. IEEE Visualization 2004. pp.3-10.
- [13] Mihail Petrov, Manfred Glesner. “A Scalable Resampling Architecture”. Communications Society subject matter experts for publication in the IEEE GLOBECOM 2007 proceedings. pp3102-3106.

- [14] 张辉, 曹丽娜.“现代通信原理与技术(第二版)”.西安: 西安电子科技大学出版社, 2007. pp144-183.
- [15] 樊昌信, 詹道庸, 徐炳祥等.通信原理(第四版).北京: 国防工业出版社, 1995. pp187-243.
- [16] TszLhan Chen and P. P. Vaidyanathan. “Commutativity of D-dimensional decimation and expansion matrices, and application to rational decimation systems”. IEEE,1992. pp637-640.
- [17] Goran RajiC', Hrvoje BabiL'. “Efficient implementation of sharpened CIC decimation filters for software radio receivers”. IEEE,2004. pp1672-1676.
- [18] Max W. Hauser and Robert W. Brodersen: Monolithic decimation filtering for custom delta-sigma A-D converters. CH2561-9/88 of IEEE,1988. pp2005-2008.
- [19] Tasio Saramakl. “A Class of Linear-Phase FIR Filters for Decimation,Interpolation, and Narrow-Band Filtering”. IEEE Transactions on Acoustics, Speech, and Signal Processing, VOL. ASSP-32, NO. 5, OCTOBER 1984. pp1023-1036.
- [20] Ronald E. Crochiere, Lawrence R. Rabiner. “Further Considerations in the Design of Decimators and Interpolators”. IEEE Transactions on Acoustics, Speech, and Signal Processing, VOL. ASSP-24, NO. 4, AUGUST 1976. pp296-311.
- [21] Yao-Wu Liu, Edward Kai-Ning Yung, Kenneth K. Mei. “Interpolation, Extrapolation, and Application of the Measured Equation of Invariance to Scattering by Very Large Cylinders”. IEEE Transaction on Antennas and Propagation, VOL. 45, NO. 9, September 1997. pp1325-1331.
- [22] Julius O. Smith III. “MUS420/EE367A Lecture 4 Interpolated Delay Lines, Ideal Bandlimited Interpolation, and Fractional Delay Filter Design”. Center for Computer Research in Music and Acoustics (CCRMA) Department of Music, Stanford University Department of Music, Stanford University Stanford, California.2008.pp1-44.
- [23] Takanori Nishinot, Shoji Kajita', Kazuya Takedui. “Interpolation Head Related Transfer Functions in The Median Plane”. IEEE Workshop on Applications of Signal Processing on Audio and Acoustics, NPW Pulfz, New York, Oct,1999. pp167-170.
- [24] Markus Biihren ,Marius Pesavento, Johann F. Bohme. “A new approach to array interpolation by generation of artificial shift invariances- interpolated esprit ”. IEEE,2003.pp205-208.
- [25] 湛军. “采样率转换算法的研究与实现”.科技技术信息—机械与电子,2009年第1期. pp528-529.

- [26] 季薇, 杨震. “软件无线电中采样率转换技术的实现及在3G中的应用”.电子工程师,2004年9月, 第30卷, 第9期. pp17-20.
- [27] 王泽华, 王霞. “一种实时语音采样率转换的延迟线算法”.电声技术,2008年32卷, 第2期. pp58-70.
- [28] 朱学芳. “数字语音信号的非均匀采样的实现”.计算机工程与应用,2002.15, . pp31-42.
- [29] <http://ccrma.stanford.edu>.
- [30] Julius O, Smith, Phil Gossett. “A Flexible Sampling.Rate Conversion Method”. IEEE,1984,pp19.4.1-19.4.4
- [31] 常文革, 宋千, 梁甸农. “重采样插值技术实现与应用”.系统工程与电子技术,2000,第22卷, 第5期. pp87-89.
- [32] Shinjiro Ashida, Masaaki Nagahara, Yutaka Yamamoto. “Audio Signal Compression via Sampled-Data Control Theory”.IEEE, SICE Annual Conference in Fukui, August 46,2003,pp1744-1747.
- [33] Ryan S. Prendergast, Bernard C. Levy, Paul J. Hurst. “Multirate Filter Bank Reconstruction of Bandlimited Signals from Bunched Samples”.IEEE, 0-7803-7576-9/02, 2002,pp781-785.
- [34] Walter A. Frank, Imre Varga, “Implicit decimation for FIR systems and its application to acoustic echo cancellation”. IEEE, 1999,pp865-868.
- [35] V. Valimaki' and T. I. Laakso. “Principles of fractional delay filters”. IEEE, 2000,pp3870-3873.
- [36] HAltan Johansson , Lars Wanhammar. “Filter structures composed of allpass and FIR filters for interpolation and decimation with factors of two”. IEEE, 1998,pp45-48.
- [37] Rahul Vanam, Charles D. Creusere.“Evaluating low bitrate scalable audio quality using advanced version of PEAQ and energy equalization approach”.IEEE, ICASSP 2005. pp189-192.
- [38] Antony W. Rix, John G. Beerends etc, “The PESQ Algorithm as the Solution for Speech Quality Evaluation on 2.5G and 3G Networks”. Rev A 2006-01-05 Irina Cotanis,pp1-15.
- [39] Hendrik Knoche, Hermann G. De Meer, David Kirsh. “Utility curves mean opinion scores considered biased”.IEEE,1999. pp12-14.

研究成果

在硕士研究生期间取得的研究成果如下：

一、参加科研情况

研究方向：多媒体通信与信源编码

研究内容：音频信号处理

参加科研：音频重采样算法设计与实现