

管理音频编解码器不同时钟频率的简单方法

4月 2011



作者

Carlos Azeredo-Leme
新思公司模拟设计部高级工程师

概述

音频处理对移动电话、MP3 播放器和其它许多电子产品应用来说是非常重要的。在市场提出高质量高保真(Hi-Fi)音频功能要求的同时，尺寸和功耗也常常是关键的设计指标。这些趋势推动音频编解码器发展成为了现代系统级芯片(SoC)中的一个嵌入式内核。

音频编解码器在数字主处理器和音频转换器(如麦克风和扬声器)之间创建了一个接口。音频编解码器还负责多种常规的音频功能，因而能减轻主处理器的工作负载。当音频编解码器作为 IP 内核嵌入在 SoC 中时，对内部接口呈现为一个数字模块，用于透明地处理所有片外模拟转换器和输入/输出(I/O)。

在内部数字接口方面，重要的是要理解与音频采样率和时钟有关的知识。音频编解码器中的数据转换器所要求的时钟取决于音频采样率以及主设备和 SoC 上的可用时钟。由于音频采样率和可用的主机时钟多种多样，因此组合情况非常复杂。在音视频(A/V)应用中，更复杂的是音频时钟还需要与视频数据转换器所要求的视频时钟同步。因此，许多设计师在为了尽量降低与时钟产生和兼容多种采样率有关的系统成本而进行折衷决策时面临复杂的选择。

数字滤波器在帮助解决上述问题中扮演着重要的角色，因为它们能够处理数字音频接口和音频数据转换器之间的数字采样值，从而完成采样率转换。

本文将讨论音频编解码器中的数字滤波器功能，并通过几个支持互连的实例介绍如何在有众多采样率和复杂时钟环境中使用这些数字滤波器。

音频处理

音频编解码器的内核由两种数据转换器组成：用于回放的数模转换器(DAC)和用于记录的模数转换器(ADC)。对于立体声或多通道编解码器而言，只是相应地复制这些转换器而已。图 1 显示了音频编解码器的典型功能框图。

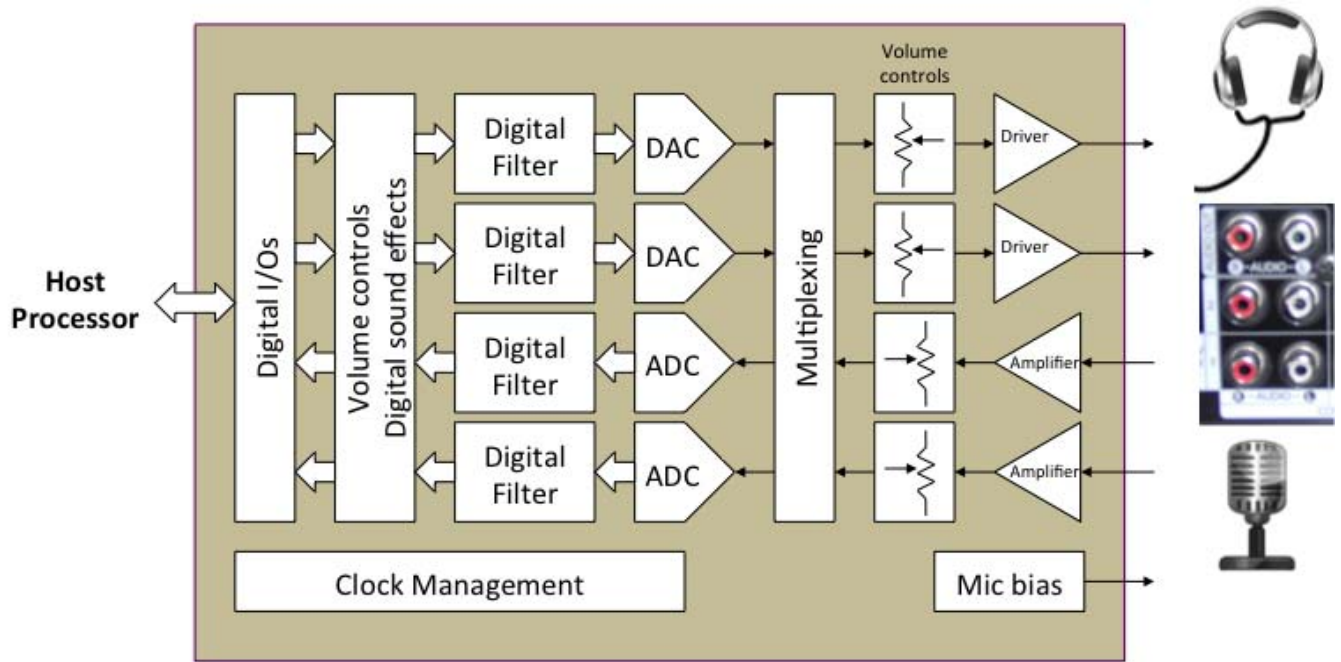


图 1——音频编解码器的功能框图

ADC 输入端使用了带音量控制的放大器，能将微弱的麦克风电平和较大的互连线路电平调整到 ADC 的输入范围。DAC 输出端的输出驱动器可以直接连接耳机或小型扬声器，每个驱动器都有各自的音量控制功能。可以将输入信号直接旁路到输出端而不通过 ADC 与 DAC。图中还有一个低噪声电源用于麦克风偏置。

数字侧有多个模块，其中最重要的是数字音频滤波器，它能将数据速率转换为数据转换器的超采样时钟，并去除音频频带外的高频噪声。时钟管理模块也很重要，它能确保所有多速率模块相互间同步，并支持所有要求的采样率组合。在数字域中还有其它一些模块，如音量控制，可为数字和模拟模块间分配增益/衰减提供额外的灵活性。另外，一些音频编解码器可能还包含用数字方式实现的音效电路，比如低音提升、3D 音效、自动音量控制和参数均衡等。

数字音频滤波器

目前音频编解码器中使用的数据转换器工作在很高的超采样频率，也就是说它们的转换频率要比音频频带高很多，通常在 100 倍以上。举例来说，假设音频数据速率是 44.1kSamples/s(kS/s)，比如来自 Redbook CD 播放器的音频，那么典型的超采样速率是这个值的 128 倍，致使 DAC 的转换速率高达 5.6448MSamples/s(MS/s)。

使用如此高的转换速率的原因与数据转换器设计采用的技术有关。由于所有要求的滤波都在数字域中实现，因此这些技术可以提供相当重要的优势。用数字方式实现滤波器比用模拟方式实现高效得多。在数字实现方法中，转移函数没有失真，左右通道之间能够做到完美匹配。而且在现代半导体技术中，由于工艺节点的不断缩小，数字滤波器的硅片面积和功耗与它们的模拟实现相比已经变得可以忽略不计。

为何音频编解码器要使用滤波器？

音频编解码器使用滤波器的主要目的是为了消除混叠或映像频率。这些频率是在音频采样率倍数附近的信号频率的复制信号，是多种速率工作的结果。例如，44.1kS/s 的音频流超采样到 5.6448MS/s 时会在 88.2kHz、132.3kHz 等处产生重复频谱。这是奈奎斯特采样理论下的必然结果，如图 2 所示。

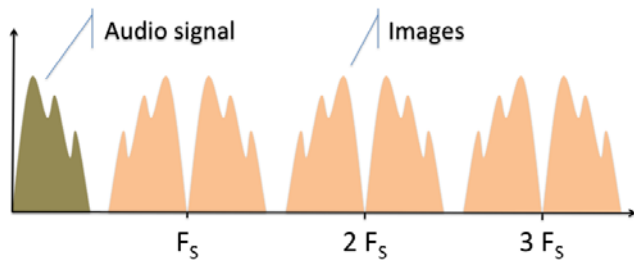


图 2——以频率 F_s 采样的音频信号会在采样频率的倍频处产生重复频谱(橙色)

映像频带会在 DAC 上产生阶梯状波形，见图 3。滤波器可以对这种波形实施平滑处理，减少其高频能量。如果这种高频能量得不到消除，它将浪费能量，并导致输出驱动器中产生互调失真，最终使扬声器中出现可闻噪声。

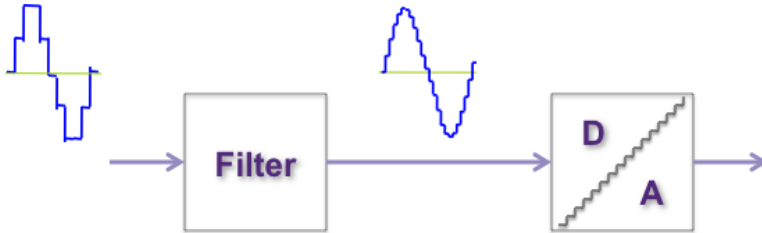


图 3——在将信号送到 DAC 之前，先由数字滤波器对信号波形进行超采样和平滑处理

在 ADC 部分，滤波器可消除输入端拾取的或在 ADC 中产生的任何带外噪声，如图 4 所示。如果这种噪声没有被消除，那么当以标准音频速率重新采样信号时，这种噪声由于混叠效应将被折叠回带内，人耳可以听到。

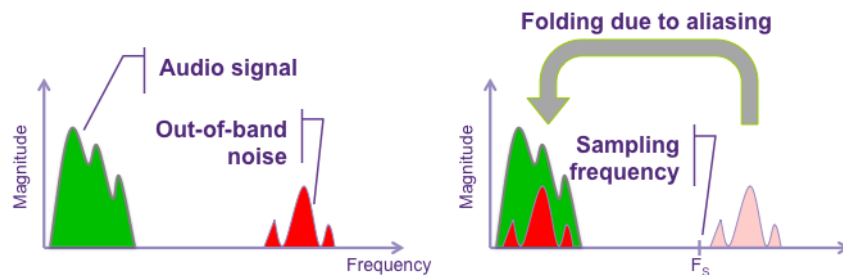


图 4：在 ADC 部分，当采样速率在输出端被重新采样到标准音频速率时，任何带外噪声(左边的红色图形)都将被折叠回信号频带内(右图)

滤波器种类和相位保真度

数字滤波器有两种：

- ▶ 无限冲击响应(IIR)，也称为递归滤波器，因为其输出同时取决于过去的输入值和过去的输出值。
- ▶ 有限冲击响应(FIR)，也称为非递归滤波器，因为其输出只取决于过去的输入值。

IIR 数字滤波器与模拟滤波器非常相似，它们都有极点和零点。极点允许非常高效的转移函数实现。因此实现的阶数或复杂度可以得到有效降低。但极点会招致相位响应失真。

FIR 滤波器没有极点，它们可以做到所有零点都是理想的陷波(在 Z 平面中的单位圆上)。这些滤波器可以保证线性相位响应。然而，要想达到相同的响应指标，FIR 滤波器必须具有比 IIR 等效滤波器更高的阶数。即使实现零点的复杂性要比实现极[this is an translation error, should be pole not zero]点低，但 FIR 滤波器的复杂度一般都要比相应的 IIR 高，因此面积和功耗也较大。

对消费类 Hi-Fi 设备来说线性相位响应是否必要业界尚有争论。有研究表明，人类的耳朵并不能分辨相位失真，除非失真太严重(参考 <http://www.ocf.berkeley.edu/~ashon/audio/phase/phaseaud2.htm>)。一个设计正确并且用相位均衡补偿了的 IIR 滤波器应该还是比较完美的。然而，增加相位均衡器会将复杂性提高到与等效 FIR 滤波器相当的程度。优势是相位补偿 IIR 具有比 FIR 低得多的延时，在游戏等一些应用中这点非常重要。

一般情况下，使用外接电源的 Hi-Fi 设备可以选用 FIR 滤波器以期获得理想的相位响应，而电池供电的消费类设备倾向于使用 IIR 滤波器，因为它具有低得多的功耗和成本，而且人耳基本上听不出相位失真。

时钟和音频采样率

数字音频信号以标准频率采样。鉴于老的 Redbook CD 遗留下来的传统，许多录音设备使用标准的 44.1kS/s 速率。这个尴尬的数字源自早期复用 PAL 录像带设备进行录音的实践。像 DVD 格式的现代音频系统则使用 48kS/s 及其倍数 96kS/s 和 192kS/s。蜂窝电话等语音应用使用 8kS/s 及其倍数 16kS/s 和 32kS/s。一些应用也可能使用 44.1kS/s 的倍数，比如 88.2kS/s 和 176.4kS/s。由于数据转换器必须工作在典型 128 倍或 256 倍的超采样频率，因此驱动数据转换器所需的主时钟频率将落在 5MHz 至 12MHz 范围内。

因此音频编解码器必须支持种类广泛的音频采样速率，并适应各种主时钟频率，从而促进其在应用中的集成。但情况没那么简单，因为组合方式繁多，可能的时钟频率比率也有限制。基于这个原因，数字滤波器必须包含采样速率转换程序。让我们看一个实际的例子，其音频速率为 48kS/s，转换器采样频率为 6.144MS/s。最终的采样速率转换是 128 倍。为了支持实际的晶体频率，使用了二分频电路，因此主时钟频率为 12.288MHz。

为了支持 96kS/s，需要针对 64 倍的采样率转换重新配置滤波器。为了支持 192kS/s，需要针对 32 倍的采样速率转换重新配置滤波器。数据转换器采样频率保持在 6.144MS/s 不变，因为音频频带固定限制为 20kHz。对于 44.1kS/s 的音频速率系列，相应的主时钟频率应为 11.2896MHz [this is a typo in the original english file, should be 11.2896MHz]。

用于音频时钟的锁相环

例如在空间和/或成本受限的情况下，许多应用无法为音频编解码器提供专门的晶振。因此音频编解码器必须能够支持源自用主机主时钟的不同音频速率，而这个主时钟通常是 12MHz 的 USB 时钟或其倍数。然后再用锁相环(PLL)产生所需的音频时钟。但这种 PLL 不是普通 PLL，因为需要非常精细的频率分辨率以支持所有的频率组合，同时为了保证性能而要提供低抖动输出时钟。不需要 PLL 的其它解决方案可能更具优势。

无锁相环(PLL-less)技术

也有采用 PLL-less 技术的替代性解决方案，这种技术可以复用 USB 时钟，从而避免增加音频专用 PLL。USB 是一种非常流行的接口，几乎在任何应用中都普遍存在。USB 使用的 12MHz 或 24MHz 时钟具有相对较低的抖动，而低抖动对音频应用来说是一个重要要求。

24MHz 的 USB 时钟可以支持 48kS/s 的音频速率及其倍频 96kS/s 和 192kS/s，因为两者是整数倍的关系($24000=125 \times 192$)。要想使用这个 USB 时钟，滤波器采样率转换需要从标称的 128 倍重新配置为 125 倍。

然而，44.1kS/s 音频速率只能近似产生。使用 136 倍的采样率转换时，音频时钟可以接近 44.1176kS/s，与标称值稍有不同。但这个差异非常小，几乎可以忽略。实际上，变化程度只有 0.04%，准确地讲不到半音程的百分之一。另外一种正确认识时钟近似效果的方法是回放一首歌曲：对于一首 3 分钟的歌曲来说回放时间快了 0.04%，仅提前了 10ms。

音视频时钟

音视频多媒体设备一般同时产生音频和视频数据流，如 DVD 播放器/刻录机和 MPEG 媒体机。音频和视频的采样率是独立的。视频使用 27MHz 时钟或其倍数。音频时钟必须从 27MHz 得到，并且要支持所有标准，包括基于 44.1kS/s、48kS/s 和 8kS/s 的标准。

这些音频时钟最好是利用将 27MHz 视频时钟的分频作为参考的 PLL 产生。鉴于音频和视频数据之间的同步机制，这里不能用 PLL-less 技术，因为这种技术会改变相对于视频的音频回放节奏，使音频与视频图像对不准，导致口型同步无法实现。

实现与源时钟的同步

在某些应用场合，数据源与设备不在一块儿，比如电视广播(有线电视、地面或卫星广播)、HDMI 电缆和网络数据流。在这些情况下，音频和视频时钟必须与源时钟取得同步。为了阐述这种情况，让我们具体看一个数字电视传输例子。

数字电视信号主要由 MPEG 传输数据流组成。一般来说，传输数据流由数字编码的视频和音频数据组成。在经过解码后，数字信号在音频 DAC 和视频 DAC 中转换为模拟信号。电视台发射的数据在发射站是同步于 27MHz 参考时钟的。由于发射机晶振的

容差和传输期间的多普勒效应，接收站电视天线接收到的信号频率可能有所改变。为了确保数据的同步，在发射机的 MPEG 数据信号中包含有一个编码过的时间戳(CTS)。CTS 可以确保视频和音频数据包的同步，从而在回放过程中实现精确的口型同步。

由于接收到的频率会随时间而改变，数据到达时有快有慢，从而产生相对于接收机自己的 27MHz 时钟的时钟漂移。数据流中的这种变化将导致到达接收机等待处理的数据太多或太少。

对于视频数据来说，这种漂移不是十分严重，可以通过丢弃帧或复制帧来校正。在解码视频时，接收机不断地比较从发射机接收到的时间戳和解码后的视频中的时间戳。当解码器中积累了太多的数据时(也就是说接收机的 27MHz 时钟慢于发射机时钟)，将丢弃一个帧。当等待处理的数据太少时，将复制一个帧。由于丢弃帧或复制帧不是很频繁，因此观看者不会察觉。

音频数据也必须与发射机时钟同步。音频时钟必须从同一个 27MHz 视频时钟得到，并且必须支持所有标准。但丢弃/复制音频样值是不可接受的，因为人类的耳机很容易察觉到这种变化。解决方案是用 PLL 产生音频时钟，使用时间戳信息将 PLL 锁定于接收到的采样数据，如图 5 所示。

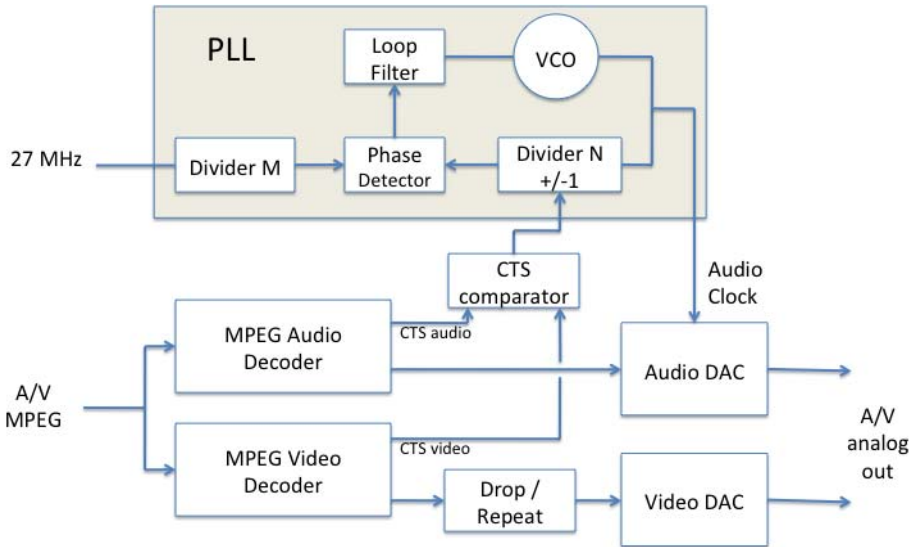


图 5: 使用 PLL 的解决方案框图，该方案使用来自视频解码器的时间戳信息同步音频时钟

就拿对 48kS/s 音频流进行 256 倍过采样的例子来说，PLL 需要产生 12.288MHz。27MHz 和 12.288MHz 的最大公约数是 24kHz，因此 27MHz 时钟必须首先被 $M=1125$ 分频，然后用 PLL 乘以 $N=512$ 。

为了同步音频时钟，CTS 比较器将比较音频和视频的时间戳，相应增加或减少 PLL 的乘数 N。当音频解码太快于 CTS 时，N 将被减少 1 至 511 范围内的一个值。当音频解码太慢于 CTS 时，N 将被增加 1 至 513 范围内的一个值。

结束语

音频编解码器在具有多种标准和音频速率的各种音视频应用中有广泛应用。由于主机设备的时钟产生功能有一定限制，使得支持音频编解码器的情况变得十分复杂。解决之道是充分利用现代音频编解码器中的数字滤波器和时钟管理功能。

通过正确地重新配置数字滤波器、采样率转换和灵活的时钟频率选择，现代音频编解码器可以支持种类繁多的音视频应用。本文提供了一些应用例子和这些应用的解决方案，范围涵盖从支持 USB 时钟的 PLL-less 技术到支持音频和视频时钟间同步的多媒体系统，可以帮助设计师理解折衷决策，最大程度地降低 SoC 成本。

通过成功实现经硅片验证的全功能音频编解码器，新思公司的 DesignWare® IP 有望帮助 SoC 设计师在成本和性能之间取得最佳折衷，同时允许他们专注于实现产品差异化。欲了解有关新思 DesignWare 音频 IP 解决方案的更多信息，请访问：

www.synopsys.com/audio。