

自选实验音乐播放器音效指导

1. SoundFont 简介

在我们之前的音频实验中，我们用正弦波来模拟了不同音调。但是，在正常的音乐演奏中往往会需要不同音色的乐器来达到更好的音乐效果。

在计算机中，这类音效大多是采用音色库 (SoundFont) 来实现的。音色库中存储了每种乐器的各个音调的录音，当需要弹奏某个乐器的时候，只要在合适的时间点播放对应的录音即可。这样，计算机不再需要以原始录音的方式来存放乐曲，只需要存放乐谱即可，大大节省了存储空间。这一原理和我们在字符显示中使用的原理是类似的，音色库类似于字体库，而乐谱类似于我们存放 ASCII 码的显存。

2. SoundFont 操作

对实现不同乐曲音效感兴趣的同学可以自行搜索 SoundFont 相关资料。网上有大量的免费 SoundFont 文件 (.sf2) 可供下载。在下载了 sf2 文件后，可以用 Linux 下的命令行工具 sf2extract 将 sf2 文件解包成单个音符的原始音频文件 (.wav)。然后可以自己写一些小程序来将 wav 文件转换成 mif 文件写入 FPGA。

我们的 FPGA 上的内存有限，只有 5.72Mbit。因此，对于 48kHz 采样的原始音频，最多只能存储 $5.72\text{M}/8/48\text{k}=14.9$ 秒。而钢琴的一个音往往可以持续数秒。因此，我们提供的转换好的音色库只包含了 13 个音符的 12k 采样率的波形，每个采样点也只有 10bit（以有符号整数方式实现）。如要实现钢琴音色，首先要将 12k 采样率的音色库转换为 48k，并且将音色库中没有的音用加速播放（减速播放）的方式来补全。钢琴中每个半音对应频率升高 $2^{1/12} \approx 1.06$ ，因此用临近的音略微快放/慢放是可以获取接近的音的。例如，以 1.06 倍速快放 C5 可以得到 C#5。我们提供的音色库包含表 1-1 中的音符。

当然，最终乐曲的动听与否取决于具体实现以及系统能够支持多少个音符同时发音。

表 1-1: 音色库信息 (12000Hz 采样率, 每点 10bit 有符号整数)

音符	起始地址 (十进制)	样本数 (十进制)	时长 (秒)
C#2	0	48000	4
F#2	48000	48000	4
B2	96000	36000	3
E3	132000	36000	3
A3	168000	24000	2
D4	192000	24000	2
G4	216000	24000	2
C5	240000	24000	2
F5	264000	24000	2
A#5	288000	24000	2
D#6	312000	24000	2
G#6	336000	12000	1
C#7	348000	12000	1