

# 摆脱呜呜祖拉噪声

董峦，胡春华 新疆农业大学 2018

---

## 实验背景

呜呜祖拉（Vuvuzela）是一种长约一米的号角，南非的球迷常在足球比赛上吹奏，它因2010年南非世界杯而为广大中国观众所知。



由于呜呜祖拉的声音令人十分烦躁，影响人们正常欣赏球赛，所以在世界杯期间就有人提出各种的滤除该噪声的方法。在本次实验中，我们通过分析呜呜祖拉声音的频谱，来确定合适的数字滤波器，然后以数字信号处理的方式科学合理的滤除呜呜祖拉噪声。

## 实验内容

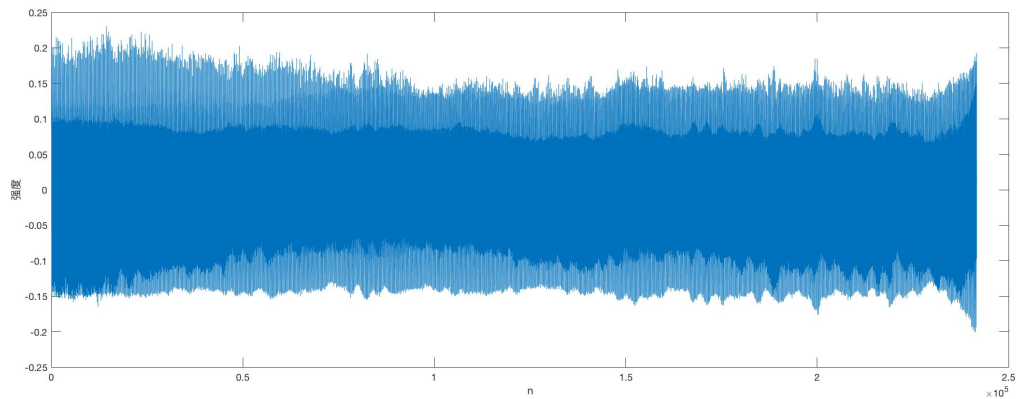
我通过互联网找到了一段呜呜祖拉的声音，同时录制了一段我的讲话声音，两段音频的采样率（sample rate）都是44.1KHz（这是数字音频中常见的一种采样率，关于它的历史可阅读[https://en.wikipedia.org/wiki/44,100\\_Hz](https://en.wikipedia.org/wiki/44,100_Hz)）。先将数据载入Matlab

```
load data.mat
```

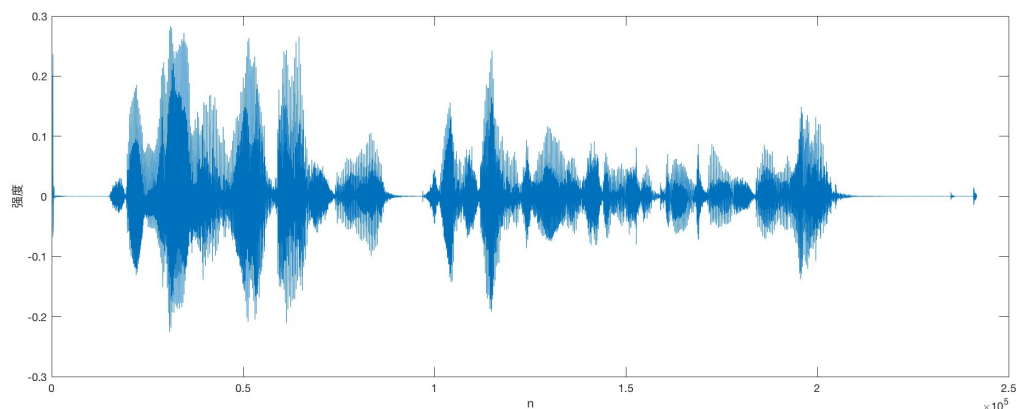
然后播放出来（注意soundsc的第二个参数是采样率，该参数要设置的与音频实际采样率一致）

```
Fs = 44100;           % 采样频率:44.1KHz
soundsc(vuvuzela, Fs);
plot(vuvuzela);
xlabel('n');ylabel('强度');
```

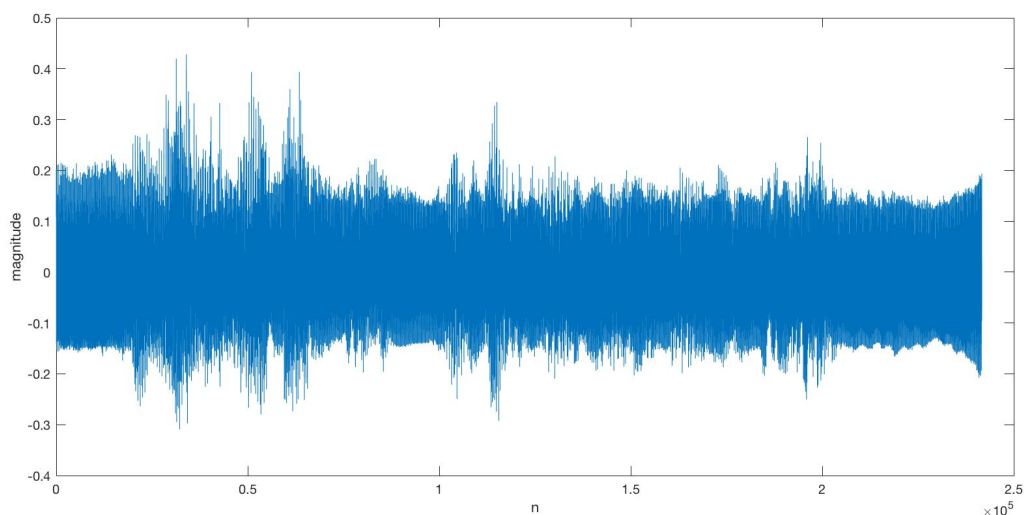
从以下呜呜祖拉声音信号的时域波形中，我们完全看不出头绪（像白噪声一样）



而人声信号则不同，其反映出吐字发音的起承转合



我们将两个信号相加，如下图所示，会发现原来的人声几乎被呜呜祖拉的声音所淹没（2010年在电视机前观看比赛的情景再次浮现）。

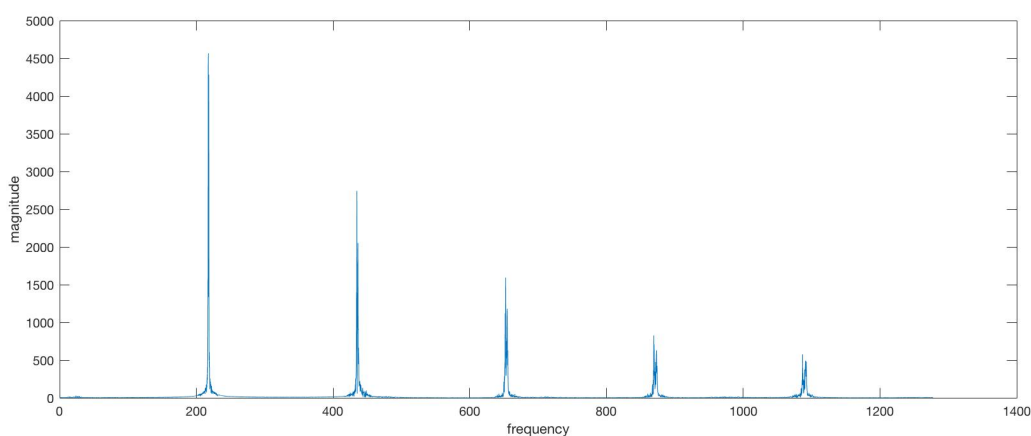


下面我们就要想办法从该混合了呜呜祖拉的声音中恢复原始人声（很多事做加法总是比做减法容易，比如菜做淡了加盐就是了，但做咸了就不好办了）。

作为一个合格的数字信号处理工程师，我们首先要看看信号的频谱是什么样子，我们手头有Matlab提供的快速傅里叶变换工具（fft函数），用它对原始呜呜祖拉声音信号做离散傅里叶变换，就能得到信号的频谱了。

```
N = length(vuvuzela);  
vu_spectrum = abs(fft(vuvuzela)); % 计算频谱的幅值，这里我们不关心相位  
f = Fs * (0:N-1) / N; % 把频点映射到实际频率  
plot(f(1:7000),vu_spectrum(1:7000)); % 观察局部频谱
```

我们从中可以发现呜呜祖拉的频谱是等间隔分布的，即有谐波成分。其基频是217.15Hz，该频率的归一化频率 $\Omega$ 是0.0309 rad ( $\Omega = 2\pi f/f_s$ )。

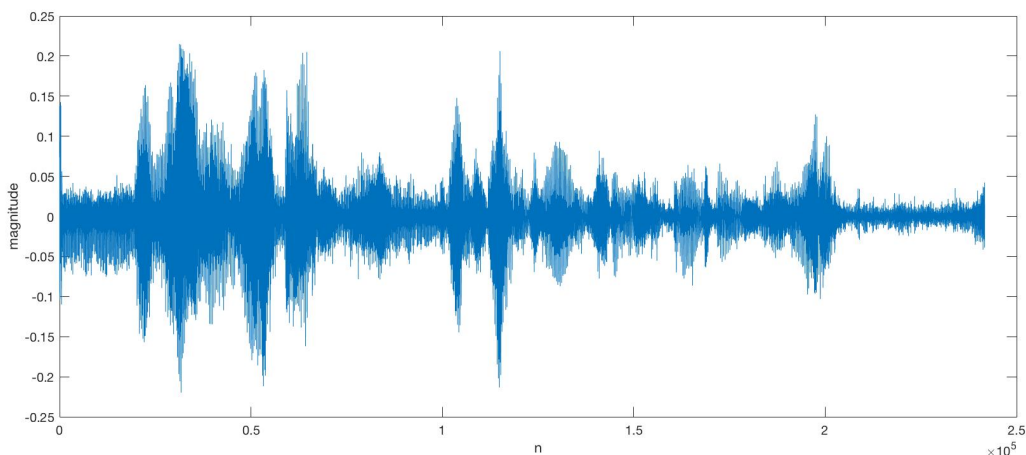


滤除这种有谐波成分的信号用梳状滤波器（comb filter）最合适了（想象一下梳齿的样子）。梳状滤波器的传递函数是：

$$H(z) = \frac{b(1 - z^{-k})}{1 - az^{-k}} \quad (1)$$

其中有三个参数 $a, b, k$ 需要确定取值。我们首先来确定 $k$ ，因为梳状滤波器的零点是在单位圆上均匀分布的，为了让零点与基波和高次谐波对齐，则 $k$ 需要等于 $2\pi/\Omega$ ，取整后 $k = 203$ 。接下来确定 $a$ 的取值，参数 $a$ 影响极点的位置，进而影响阻带带宽，如果我们令阻带带宽 $B = 70\text{Hz}$ ，则在归一化频率下阻带带宽 $B_N = 2\pi B/f_s$ ，根据课堂教学中对点阻滤波器阻带带宽的解释，我们可以计算在上述归一化带宽下极点的模 $M$ ，然后根据该值计算 $a = M^k$ 。最后来确定参数 $b$ 的取值，由于其影响的是滤波器通带的增益，如果我们令通带增益为1，则该参数的值为 $(1 + a)/2$

现在我们可以写出此梳状滤波器的传递函数进而写出差分方程了（有了差分方程就意味着可以在时域处理信号了），你需要实现该差分方程对应的函数 comb\_filter，用该函数处理混有呜呜祖拉的声音信号，然后观察呜呜祖拉噪声是否被滤除。下面是处理后的信号，可以看到呜呜祖拉信号被大幅滤除而人声信号得到了保留。



### 实验要求：

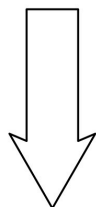
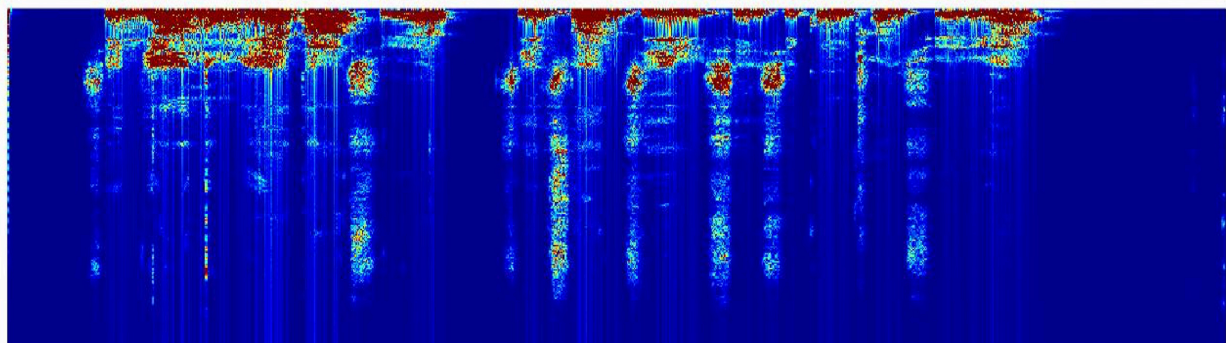
1. 根据 `main.m` 中的注释理解脚本，计算出梳状滤波器的三个参数，写出差分方程。
2. 编辑 `comb_filter.m` 实现 `comb_filter` 函数。
3. 用 `comb_filter` 函数处理混有呜呜祖拉的信号，观察呜呜祖拉是否被有效滤除。

## 小结

信号的谱（spectrum）反映的是信号的成分，梳状滤波器是在频域像剥丝抽茧那样有针对性的抑制某些频率成分（呜呜祖拉的基波和高次谐波），而其它频段的信号则几乎不受影响。本次实验用一种系统化的方法从信号中去除了呜呜祖拉的声音，显示了数字信号处理技术的科学性和有效性。

人声的频谱可以利用傅里叶变换计算并观察。但需要说明的是，人声的音频由于是一连串发音组成的所以不同位置的频谱是不同的，如果用离散傅里叶变换分析整个信号则得到的频谱是信号整体的频谱，并不能反映信号的局部特征，变通的手段是分析一段较短时间窗口内，比如10毫秒内信号的频谱（其假设信号在这样短的时间内性质是稳定的）。

运行 `show_speech_spectrum` 可以得到以下图像，该图像是对10毫秒“窗口”内的信号计算频谱得到的，从中可以看到信号频谱在局部的变化。语音识别技术就是对语音局部信号的处理最终得到文字符号的技术。



新疆农业大学 计算机与信息工程学院