# **基于MATLAB的实时语音变声系统开发**

## **1.研究背景与意义**

语音变声技术在娱乐、音频处理和语音隐私保护等领域有广泛应用。随着数字信号处理技术的发展，实现高质量的实时语音变声成为可能。本项目旨在利用MATLAB开发一个实时语音变声系统，实现男声女声互换的功能。

## **2.研究内容**

本项目主要包括以下研究内容：

· 语音信号的采集与预处理

· 基频估计算法的实现与优化

· 音高调整算法的设计与实现

· 共振峰频率调整方法的研究

· 变声后音频的输出与保存

## **3.技术路线**

本项目采用MATLAB作为开发平台，利用其强大的信号处理工具箱。主要技术路线如下：

1. 使用audiorecorder函数进行实时语音采集

2. 采用自相关法（NCF）进行基频估计

3. 使用PSOLA算法进行音高调整

4. 通过重采样技术调整共振峰频率

5. 利用sound函数实现音频播放，audiowrite函数保存处理后的音频

## **4. 关键技术点及其实现**

### **4.1 语音信号采集**

使用audiorecorder函数进行语音采集。该函数的使用方法如下：

recObj = audiorecorder(Fs, nBits, nChannels);

参数说明：

· Fs：采样率（Hz）

· nBits：每个样本的位数

· nChannels：通道数（1为单声道，2为立体声）

返回值：

· recObj：音频录制对象

在本项目中，我们使用如下参数：

recObj = audiorecorder(44100, 16, 1);

### **4.2 基频估计**

使用pitch函数进行基频估计。该函数的使用方法如下：

[f0, idx] = pitch(x, fs, 'Method', method, 'WindowLength', winLength, 'OverlapLength', overlapLength);

参数说明：

· x：输入音频信号

· fs：采样率

· 'Method'：基频估计方法（如'NCF'为归一化相关函数法）

· 'WindowLength'：分析窗口长度

· 'OverlapLength'：相邻窗口重叠长度

返回值：

· f0：估计的基频序列

· idx：基频估计的可靠性指标

### **4.3 音高调整**

使用stretchAudio函数进行音高调整。该函数的使用方法如下：

y = stretchAudio(x, r);

参数说明：

· x：输入音频信号

· r：时间拉伸因子（<1为加速，>1为减速）

返回值：

· y：时间拉伸后的音频信号

### **4.4 共振峰频率调整**

使用resample函数进行重采样，从而调整共振峰频率。该函数的使用方法如下：

y = resample(x, p, q);

参数说明：

· x：输入音频信号

· p：上采样因子

· q：下采样因子

返回值：

· y：重采样后的音频信号

## **5.预期成果**

· 实现一个基于MATLAB的实时语音变声系统

· 能够准确识别输入语音的性别，并实现男声女声的互换

· 变声后的音频保持良好的音质和可懂度

· 完成系统的性能评估和优化

## **6.参考资料**

[1]MathWorks.(2023).audiorecorder. <https://www.mathworks.com/help/matlab/ref/audiorecorder.html>

[2]MathWorks.(2023).pitch. <https://www.mathworks.com/help/audio/ref/pitch.htm>l

[3]MathWorks.(2023).stretchAudio. <https://www.mathworks.com/help/audio/ref/stretchaudio.html>

[4] MathWorks. (2023). resample. <https://www.mathworks.com/help/signal/ref/resample.html>

[5] Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. Speech communication, 9(5-6), 453-467.

[6] Stylianou, Y. (2009). Voice transformation: a survey. In 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 3585-3588). IEEE.

[7] Erro, D., Navas, E., & Hernaez, I. (2013). Parametric voice conversion based on bilinear frequency warping plus amplitude scaling. IEEE Transactions on Audio, Speech, and Language Processing, 21(3), 556-566.

**7.附录**

% 录音

recObj = audiorecorder(44100, 16, 1);

disp('开始录音...');

recordblocking(recObj, 5); % 录音5秒

disp('录音结束');

% 获取录音数据

y = getaudiodata(recObj);

fs = 44100;

% 使用自相关法估计基频

frame\_length = round(0.03 \* fs); % 30ms帧长

overlap = round(0.015 \* fs); % 15ms重叠

[f0, ~] = pitch(y, fs, 'Method', 'NCF', 'WindowLength', frame\_length, 'OverlapLength', overlap);

% 计算平均基频

mean\_f0 = mean(f0(f0 > 0));

% 变声处理if mean\_f0 > 150 % 假设150Hz为男女声的分界

% 女声变男声

formant\_shift = 0.8;

pitch\_shift = 0.8;else

% 男声变女声

formant\_shift = 1.2;

pitch\_shift = 1.2;

end

% 使用PSOLA算法进行pitch shifting

y\_pitch\_shifted = stretchAudio(y, 1/pitch\_shift);

% 重采样以改变formants

y\_resampled = resample(y\_pitch\_shifted, round(fs\*formant\_shift), fs);

% 将音频长度调整回原始长度

y\_shifted = resample(y\_resampled, length(y), length(y\_resampled));

% 归一化音量

y\_shifted = y\_shifted / max(abs(y\_shifted));

% 播放原声

sound(y, fs);

pause(6); % 等待原声播放完毕

% 播放变声后的声音

sound(y\_shifted, fs);

% 保存变声后的音频

audiowrite('D:\数信综合实验\shifted\_voice\_improved.wav', y\_shifted, fs);

disp('变声后的音频已保存为 shifted\_voice\_improved.wav');