

Fundamental Frequency of a Human Voice

Introduction

This application predicts the fundamental frequency of a human voice using the [ComplexCepstrum](#) command.

After converting a small window of the audio to the cepstral domain, we find the pitch by noting the maximum "quefrequency" in a carefully selected range.

This loudspeaker component is needed to play the sound:



```
> restart:
with(ColorTools):
with(plots):
with(SignalProcessing):
with(DocumentTools):
with(AudioTools):
```

Import and Visualize Audio

```
> aud := Read(FileTools:-JoinPath([kernelopts(datadir), "audio",
"maplesim.wav"]));
Fs := attributes(aud)[1]
```

```
aud := [
  "Sample Rate"  11025
  "Bit Depth"    16
  "Channels"     1
  "Points/Channel" 8227
  "Duration"     0.75 s
  Fs := 11025
```

(2.1)

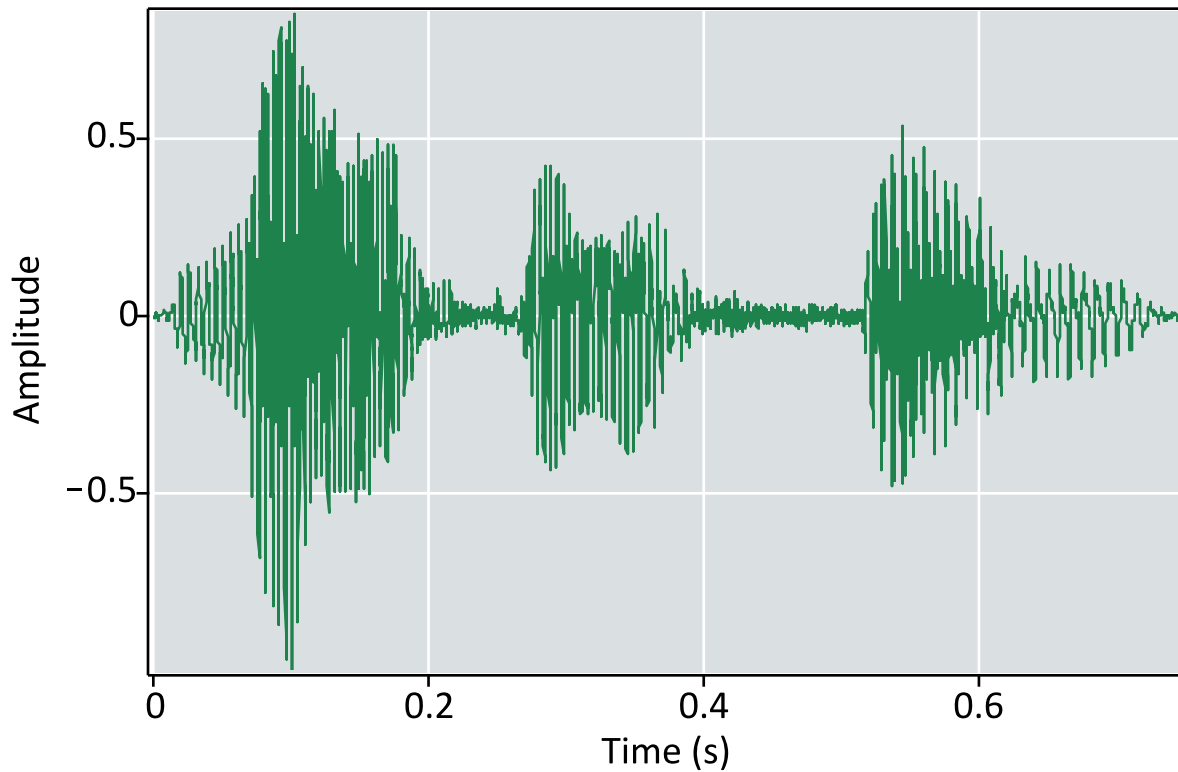
Play the audio

```
> SetProperty("Speaker0", samplerate, Fs);
Play(aud, "Speaker0")
```

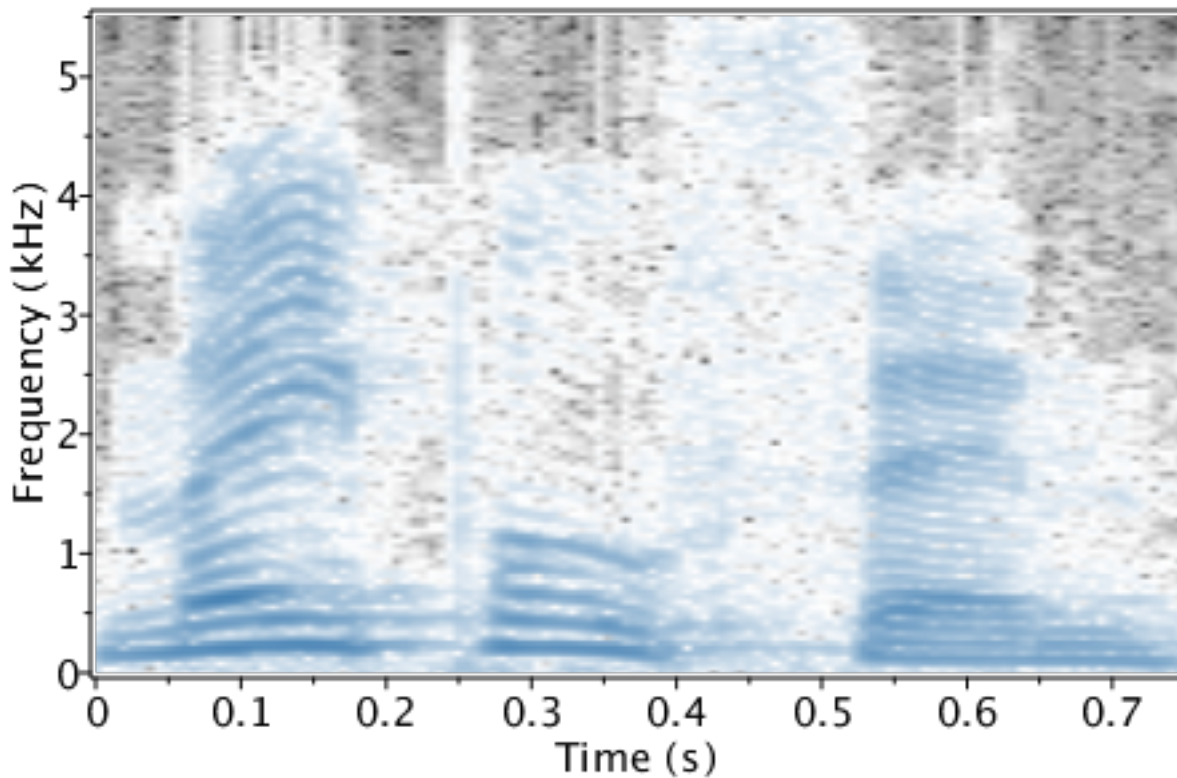
```
> t := Vector(numelems(aud), i -> 1.0* (i - 1) / Fs, datatype =
```

```
float[8]):
```

```
pltOrigAudio := plot(t, aud, thickness = 0, color = Color("RGB",  
[30/255, 130/255, 76/255]), axes = boxed, size = [800,300], font  
= [Calibri], labels = ["Time (s)", "Amplitude"], labeldirections  
= [horizontal, vertical], labelfont = [Calibri], background =  
Color("RGB", [218/255, 223/255, 225/255]), axis = [gridlines =  
[5, color = Color("RGB", [1, 1, 1])]]);
```



```
> Spectrogram(aud, size = [800,300], colorscheme = [black, white,  
"SteelBlue"], overlap = 0.75)
```



Extract Segment of Audio to Analyze

Select the initial syllable of the audio (

```
> dt      := 1. / Fs;
  I0      := round(0.1 / dt) + 1;
  Iend    := round(0.28 / dt) + 1;
  audSeg  := aud[I0 .. Iend];
  Play(audSeg, "Speaker0")
```

```
dt := 0.00009070294785
```

```
I0 := 1104
```

```
Iend := 3088
```

```
audSeg :=
```

"Sample Rate"	11025
"Bit Depth"	16
"Channels"	1
"Points/Channel"	3088
"Duration"	0.18s

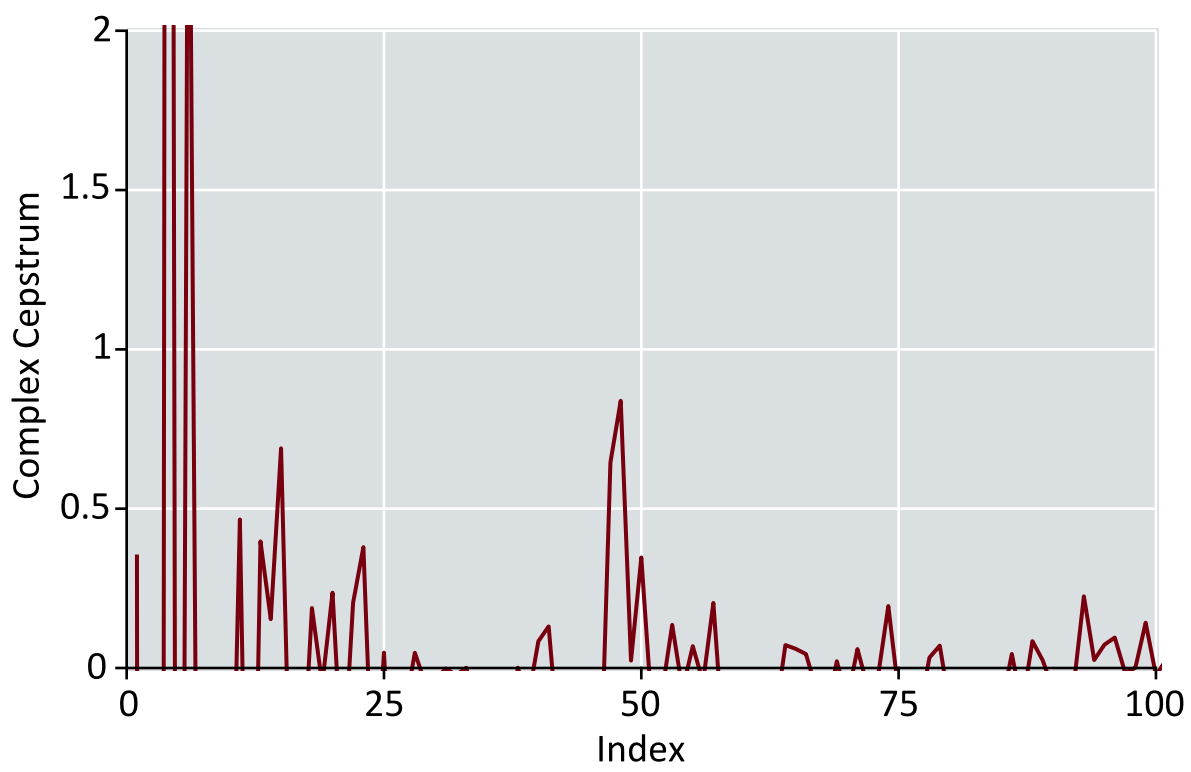
(3.1)

Calculate and Plot the Complex Cepstrum

```
> c := ComplexCepstrum(audSeg)[1]:
  ind := [seq(i, i = 1 .. numelems(c))]:
```

```
plot(ind, c, view = [0 .. 100, 0 .. 2], size = [800,300], font =
  [Calibri], labels = ["Index", "Complex Cepstrum"],
```

```
labeldirections = [horizontal, vertical], labelfont = [Calibri],  
background = Color("RGB", [218/255, 223/255, 225/255]), axis =  
[gridlines = [5, color = Color("RGB", [1, 1, 1])]]
```



A large peak occurs at an index of 48. Hence the fundamental frequency of the speaker is about 230 Hz

> $F_s / 48.$

229.6875000

(4.1)