STUDY OF LINEAR PREDICTION MODEL FOR AUDIO SYNTHESIS

Presenter: Anjini Shukla

MAT-201 Media Signal Processing
Professor: Jerry. D. Gibson
INDEX:

- Introduction
- Implementation
- Results
- Discussion of Results
- Conclusion
Applications of Speech Compression:

Transmission / storage
real-time transmission e.g. mobile phones
archive storage e.g. voicemail

Modification Synthesis
speech synthesis / text-to-speech (change the words)
speech transformation/disguise (change the speaker)

Classification / matching
speech recognition (lexical content)
other signal classification
content-based retrieval
Mathematical Model of LPC:

\[ H(z) = \frac{G}{1 - \sum a_k z^{-k}} \]

LPC Filter == Vocal Tract

Innovation == Air

Voiced (V) == Vocal Chord Vibration
Pitch Period (T) == Vocal Chord Vibration period

Gain (G) == Air Volume
How does LPC Work....?
Voice Excited LPC

- Good reconstruction of the excitation requires the low frequencies
- DCT of the residual signal.
- DCT concentrates most of the energy of the signal in the first few coefficients.

Block diagram of Voice Excited LPC
Implementation

Implementation is done using Matlab: LPC and V-LPC

Performance Evaluation parameters:
- Computational Complexity
- Overall Delay of the System
  
  Time taken after the first speech sample is input in the system to the time first speech sample is synthesised.
- Signal to Noise Ratio
- Bit Rate
## Results

### SNR:

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Guitar</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC</td>
<td>-12.487</td>
<td>-10.648</td>
<td>-10.366</td>
<td>-12.3018</td>
</tr>
<tr>
<td>V-LPC</td>
<td>-3.163</td>
<td>-3.27</td>
<td>-1.6734</td>
<td>-2.668</td>
</tr>
</tbody>
</table>

### Bit Rate:

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Guitar</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>LPC</td>
<td>8.349</td>
<td>8.349</td>
<td>8.393</td>
<td>8.56</td>
</tr>
<tr>
<td>V-LPC</td>
<td>103.44</td>
<td>103.312</td>
<td>103.312</td>
<td>105.21</td>
</tr>
</tbody>
</table>
# Results

Performance of different LPC order(n):

<table>
<thead>
<tr>
<th></th>
<th>n =1</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-LPC</td>
<td>-6.9027</td>
<td>-5.4628</td>
<td>-3.36</td>
<td>-3.2011</td>
<td>-3.163</td>
</tr>
<tr>
<td>Bit Rate LPC</td>
<td>1.5947</td>
<td>2.3893</td>
<td>3.978</td>
<td>5.1707</td>
<td>8.394</td>
</tr>
<tr>
<td>VLPC</td>
<td>96.5573</td>
<td>97.352</td>
<td>98.9413</td>
<td>100.1333</td>
<td>103.43</td>
</tr>
</tbody>
</table>
Conclusion:

- Achieved bit rate in both methods is low.
- Voice-excited LPC coding roughly requires a bandwidth twice as large as the plain LPC coding.
- Computational complexity is same for both methods.
- Voice excited LPC should be improved to work at lower bit rate.
Questions....??