AIM (Asynchronous Interpolation Model)

Hamid Mohammadi
CSLU, OHSU
08/19/13
• Asynchronous Interpolation Model (AIM)
• AIM on Formant
• AIM in LSFs
• AIM on Harmonics
• The problems
AIM

- Events: phonemes, allophones, etc
- If no significant event occurs in between the selected events, we can interpolate them
- The core idea of AIM: describe speech region by varying influence of preceding and following events.
AIM

[Kain2010]
AIM on formants
AIM on formants
AIM on formants

• Pros:
  • Features (Formant Frequencies) are highly interoperable
  • Can produce highly intelligible speech

• Cons:
  • Not high-quality (regarding naturalness)
  • Does not reflect all the details of the speakers
AIM on LSFs

- The same procedure can be applied on other features
- We selected to use Line Spectral Features
- Okay interpolation property
AIM on LSFs

- The same procedure can be applied on other features
- We selected to use Line Spectral Features (LSFs)
- Fair interpolation property and Fair quality
AIM on LSFs

• Pros:
  • Robustness (Stable interpolated filters)
  • Fair interpolation property
  • Straight-forward

• Cons:
  • Interpolation does not always work
  • Quality limit: Vocoder Quality
AIM on LSFs
AIM on LSFs
AIM on Harmonics
Conclusion

• Harmonic Vocoder
Harmonic

- Use the AIM on Formants as a “cue” or guide to modify the Harmonic spectrum
Harmonic

• Pros:
  • Potentially high quality
  • keeps most of speaker information

• Cons:
  • Repeating a frame results in low-quality speech
  • Hard to generate natural-like speech with current technology
Future Work

• Problems:
  • LSFs: Resolve the interpolation (One solution is Pole Interpolation)
  • Harmonics: Resolve low quality of frame repetition
  • Optimizing weights
  • Optimizing Event locations
Question?
References