Technical Seminar presentation on

Voice morphing

Presented
By
Mradul Singh

Supervised
By
Sir Nishant Gupta

SEMINAR OUTLINES

- What It is?
- Need of Voice Morphing
- Description the Morphing.
- Technical details of Morphing.
- Application areas.

Voice Morphing

- Transition Phenomenon.
- Technology developed at the Los Alamos National Laboratory in New Mexico, USA by George Papcun .

What it actually performs ?

- It is a technique to modify a source speaker's speech to sound as if it was spoken by a target speaker.
- Voice morphing enables speech patterns to be cloned
  - And an accurate copy of a person's voice can be made that can wishes to say, anything in the voice of someone else.

Need of voice morphing

- Text To Speech (TTS)
- In public speech systems
- For special effects (just like video or image morphing is done).
- To diminish Ethnical barriers.

Voice Morphing Process

- Preprocessing or representation conversion.
- Pitch and Envelope analysis.
- Morphing which includes Warping and interpolation.
- Signal re-estimation.

Block Diagram

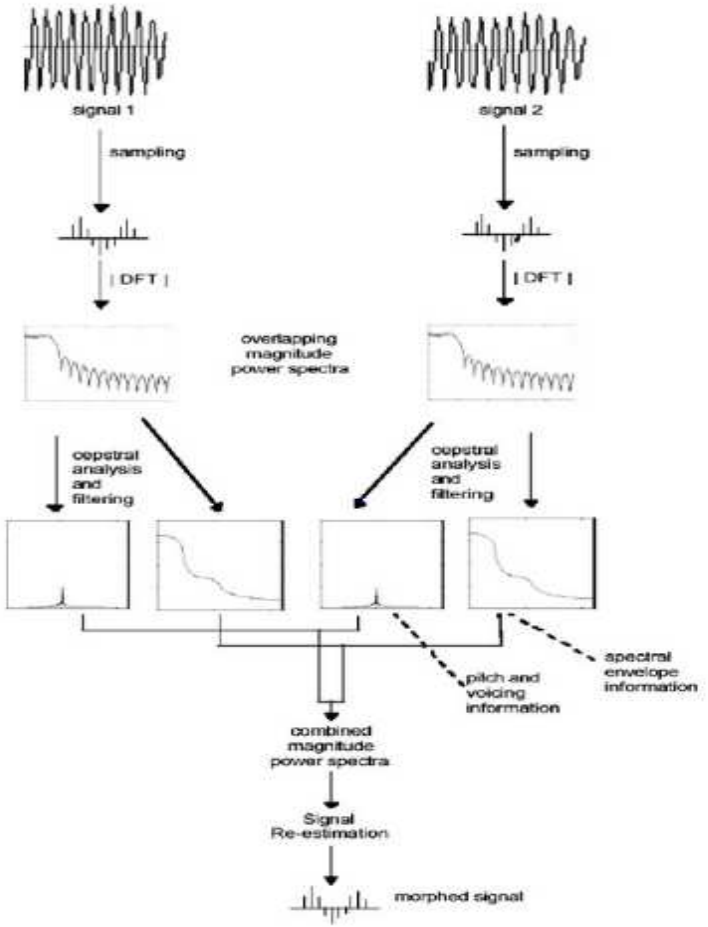


Figure 2.1 Schematic block diagram of the speech morphing process

Pre-Processing

- Involves processes like signal acquisition in discrete form and windowing.

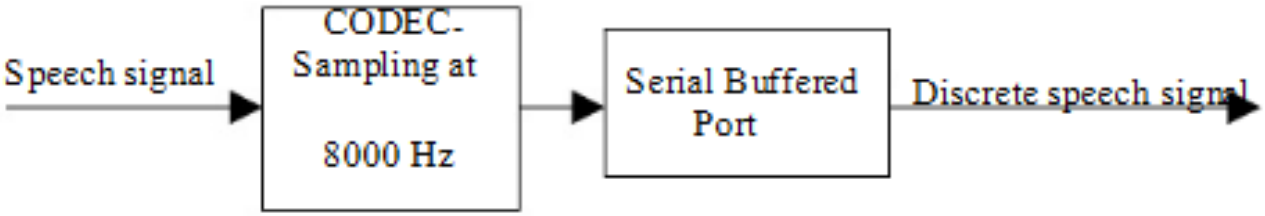


Fig 3.2: Signal acquisition

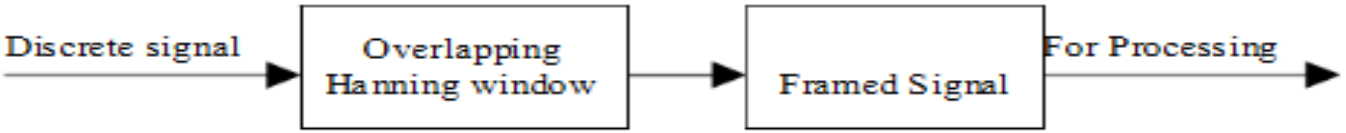


Fig 3.3: Windowing

Pitch And Envelope Analysis

- This process will extract the pitch.
- Formant information in the speech signal.

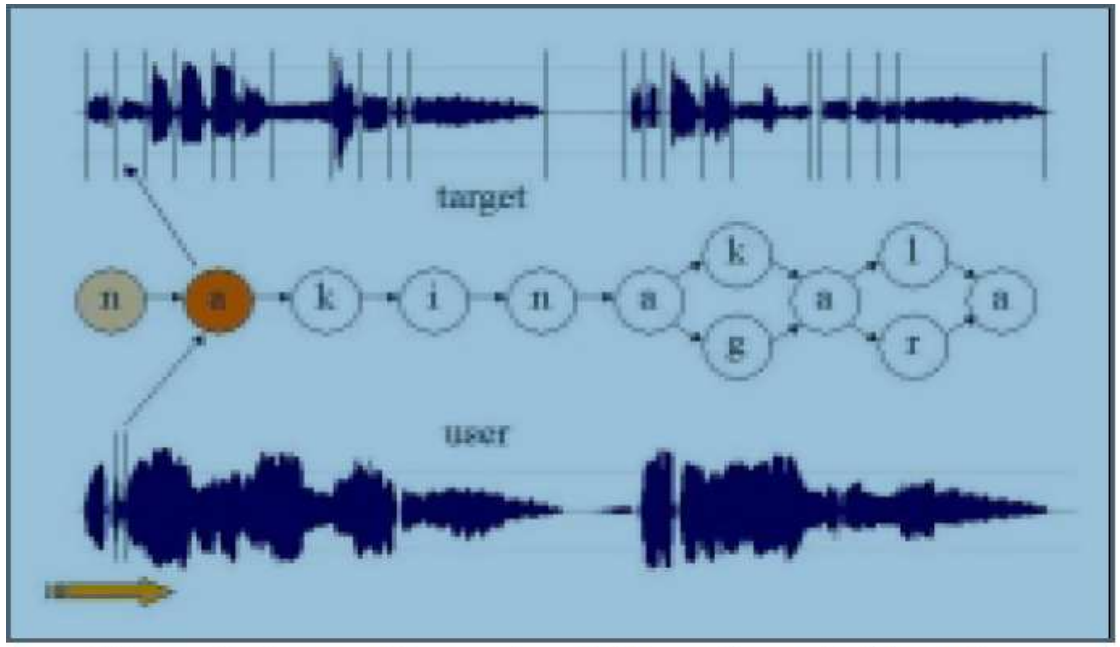


Figure 2. Recognition and matching of morphable units.

Conversion

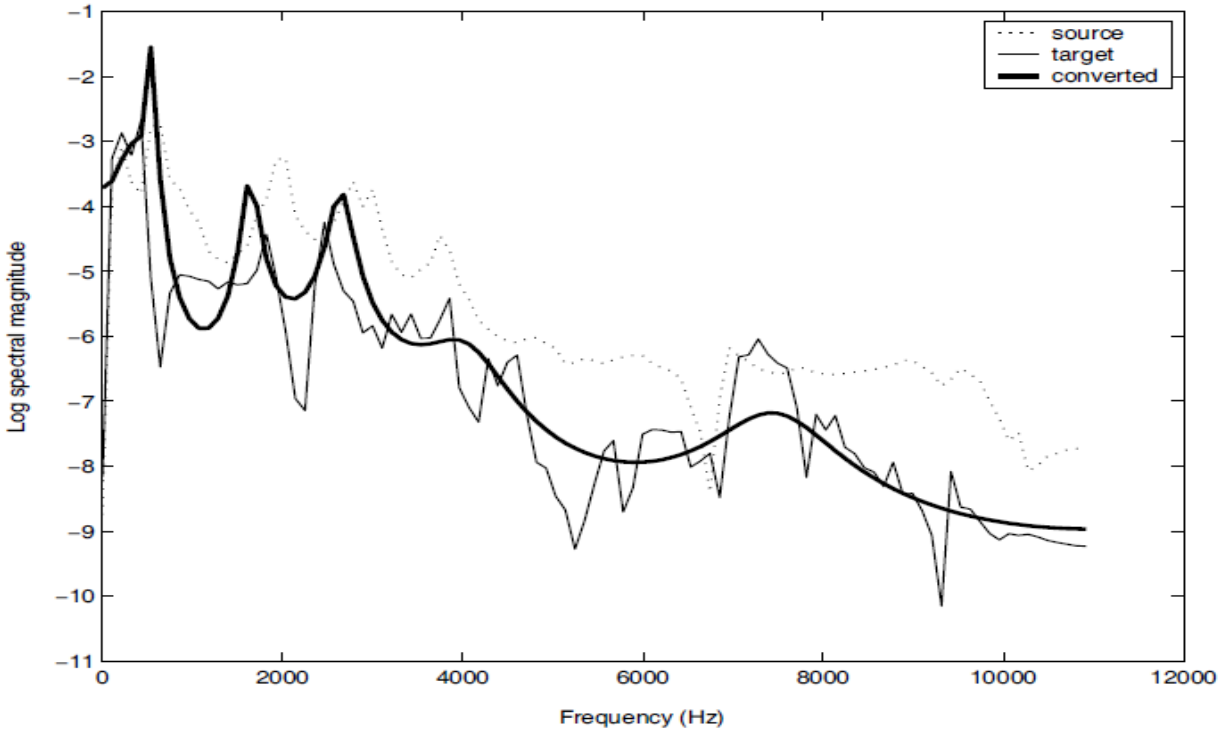


Fig. 1. *Spectral envelope conversion.*

Matching and Warping

- DTW(Dynamic Time Warping)

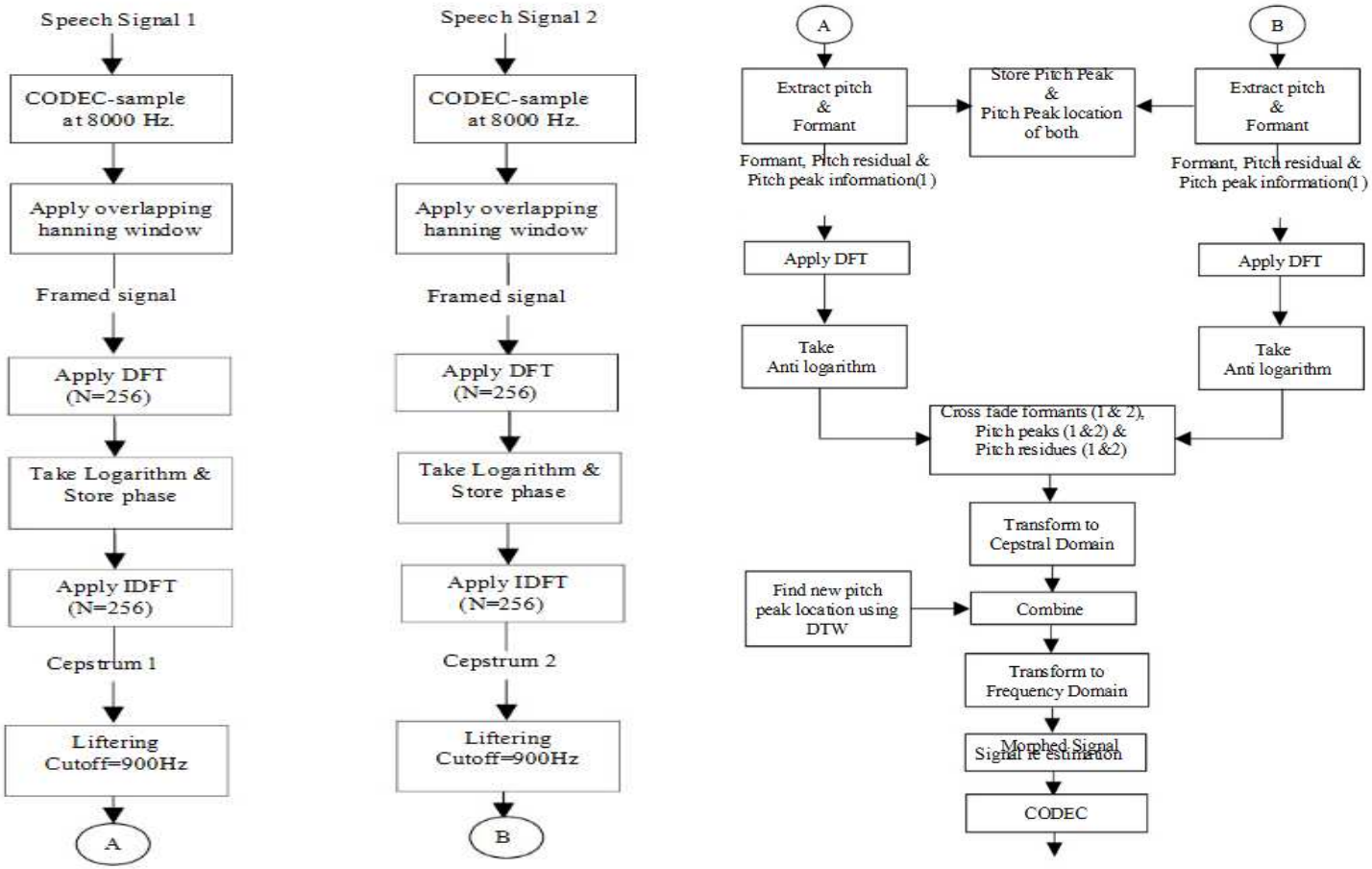
  - Dynamic Time Warping (DTW) is used to find the best match between the pitch of the two sounds.

Signal Re-Estimation

- Loss during Signal re-estimation

  - Due to signals being transformation into the cepstral domain, a magnitude function is used. This results in a loss of phase information in the representation of the data.

Summarized Block Diagram



Limitations

- Lots of normalizing problems.
- Some applications require extensive sound libraries.
- Different languages require different phonetics.
- It is very seldom complete.

Advantages

- Allows speech model to be duplicated and an exact copy of a person's voice.
- Powerful combat zone weapon.

Disadvantages

- Use to pull out the useful information.
- It hides the actual identity of the user.

Conclusion

- The approach we have adopted separates the sounds into two forms:
  - Spectral envelope information
  - Pitch and voicing information.
- Dynamic Time Warping
  - Aligns the sounds with respect to their pitches.
- Signal re-estimation algorithm.
  - Frames are converted back into a time domain waveform.

Application Areas

- Fake telephone conversations as evidence in courts of law.

- Powerful battlefield weapon.

  - Provide fake orders to the enemy's troops, appearing to come from their own commanders.

Future Scope

- Extending the functionality of tool.
  - Create a powerful and flexible morphing tool.

- Increased user interaction.
  - Graphical User Interface could be designed and integrated to make the package more ‘user-friendly’.

Thank you!!!

Questions??