
Voice Morphing
Chipmunk
Slow
Helium
C.Chorus
Voice effects

Welcome

Voice Morphing

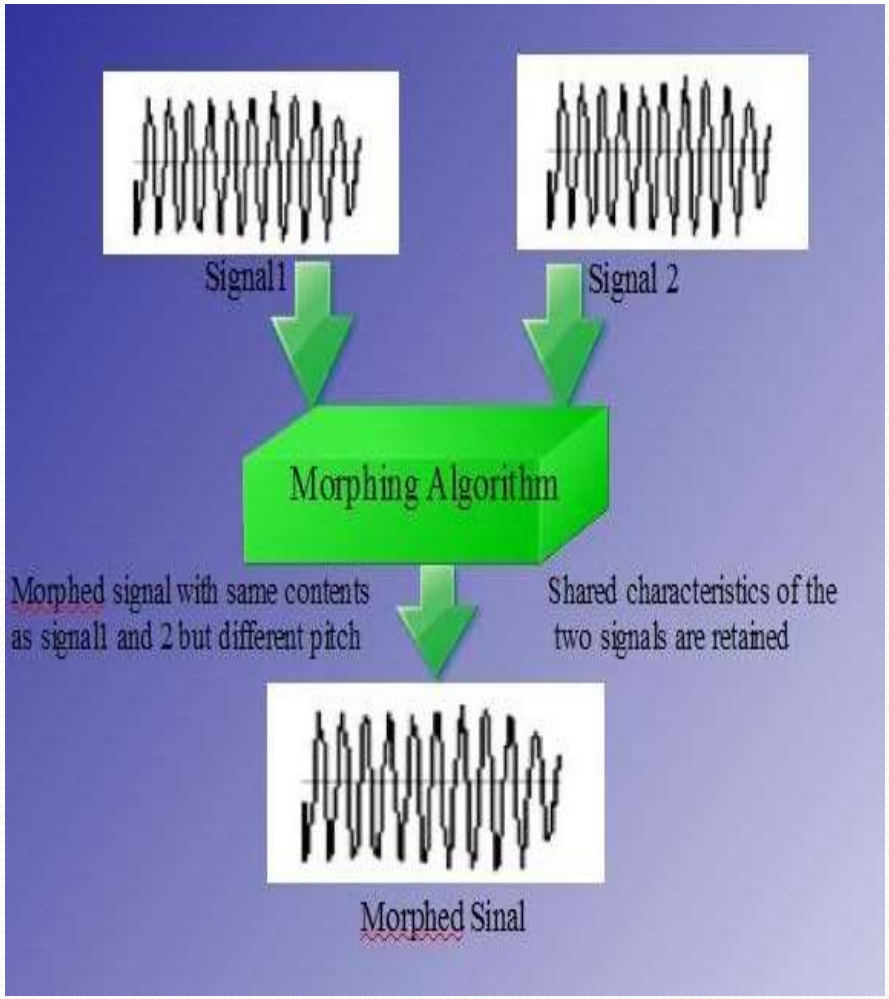
Submitted By
Cicia Paul
S2 MCA
Roll no:22

CONTENT

- INTRODUCTION
- AN INTROSPECTION OF MORPHING PROCESS
- MORPHING PROCESS
- MORPHING STAGE
- SIGNAL RE-ESTIMATION AND LIMITATION
- SUMMARIZED BLOCK DIAGRAM
- APPLICATIONS
- CONCLUSION

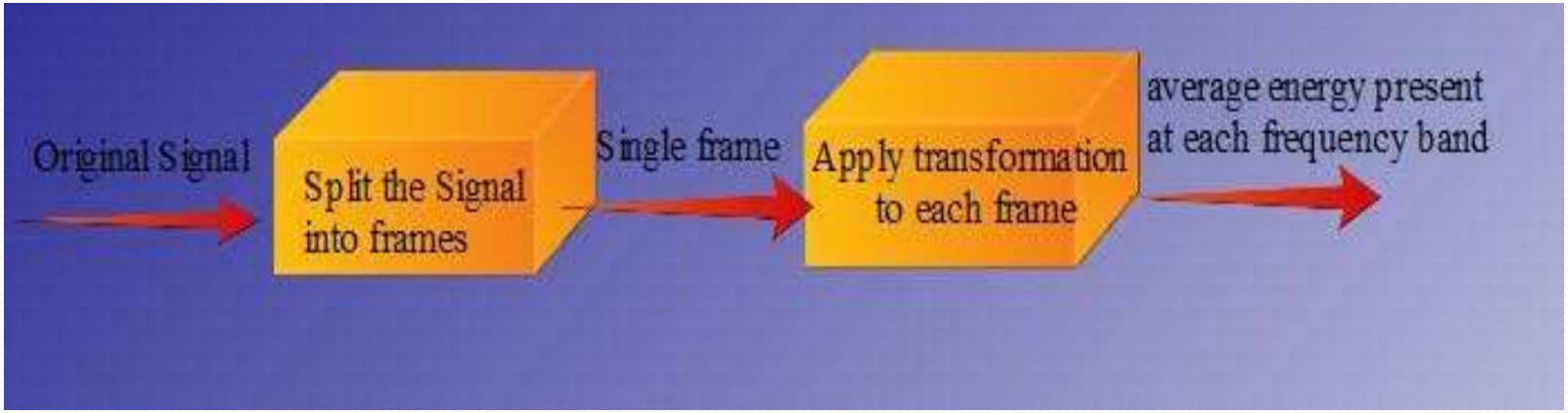
INTRODUCTION

- Voice morphing means the smooth transition of one speech signal to another, keeping the shared characteristics of starting & ending signals
- Pitch and Envelop Information are the two major properties of Speech signal.
- Cepstral analysis is used to extract characteristics.



AN INTROSPECTION OF MORPHING PROCESS

- Speech morphing can be achieved by transforming the signal's representation from the acoustic wave form obtained by sampling of the analog signal, to another representation.



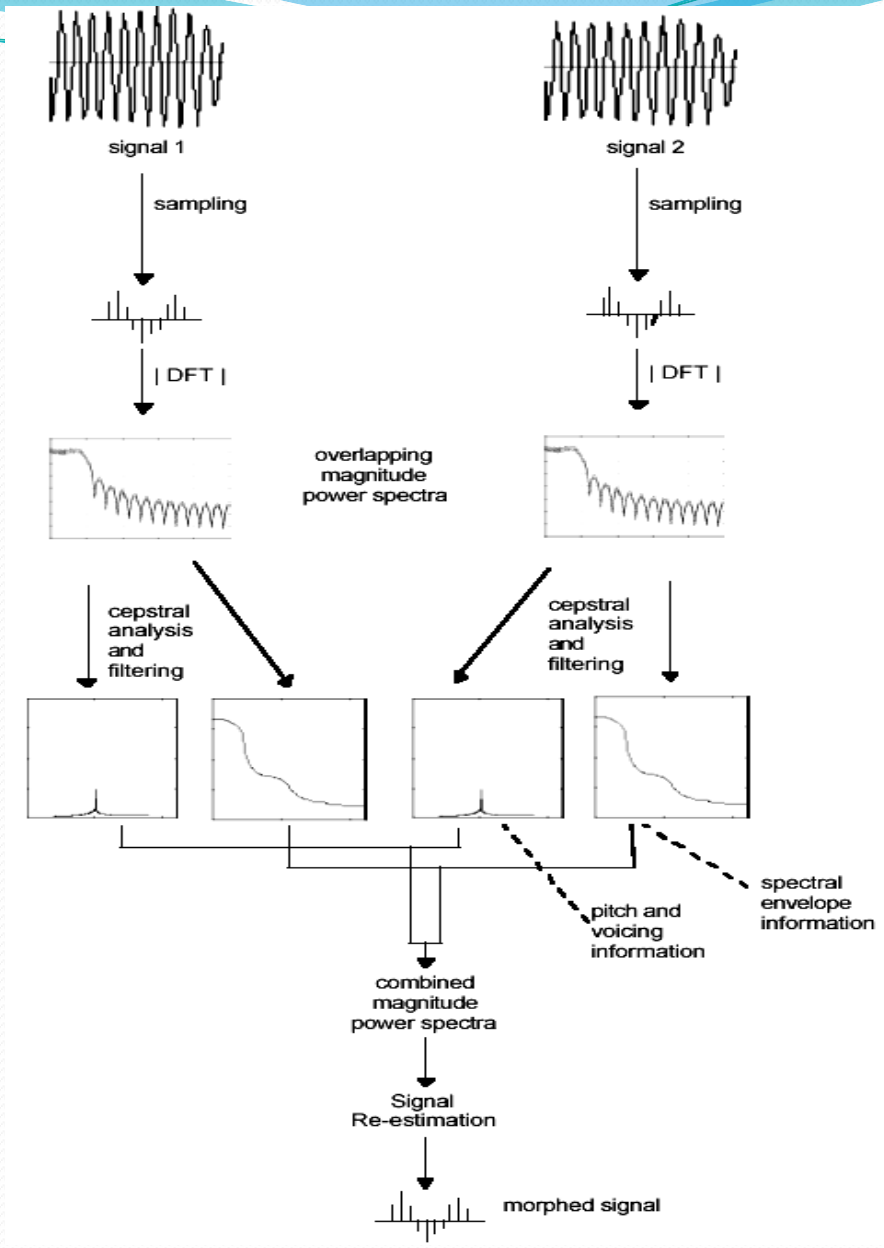
Other necessary processing to obtain the morphed speech signal :

➔ Cross fading of envelope information.

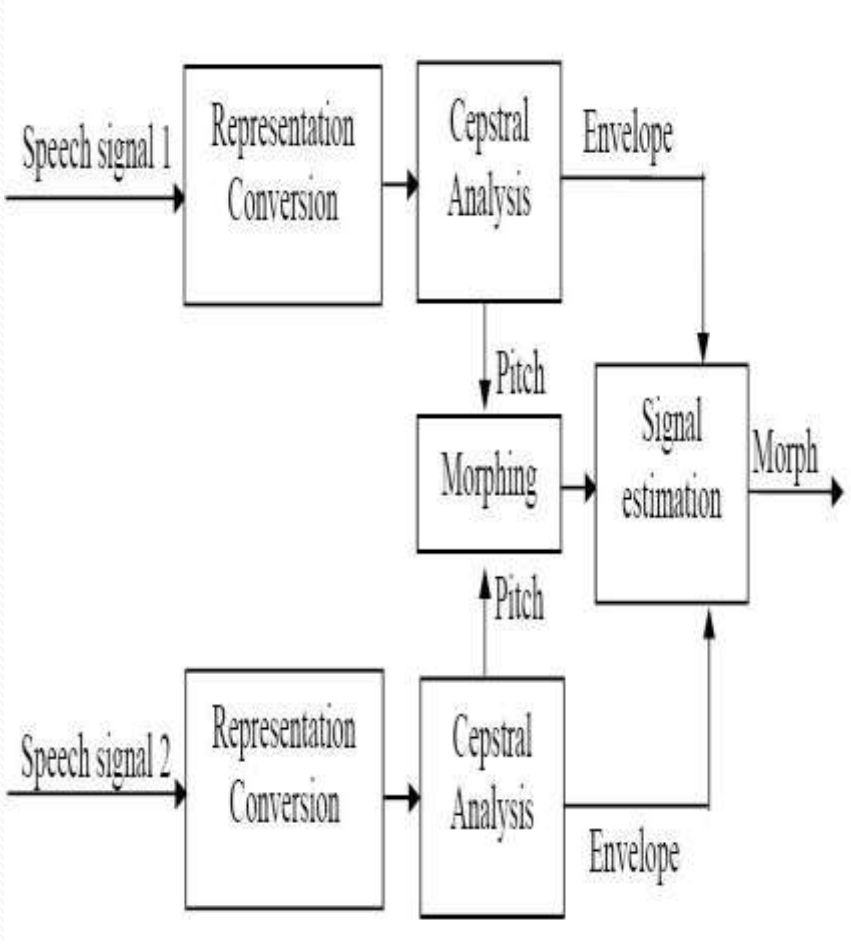
➔ Dynamic Time Warping to match the major signal features (pitch).

➔ Signal Re-estimation to convert the morphed speech signal back into the acoustic waveform.

- 1. Two signals with different pitches were simply **cross-faded which** produces two separate sounds .
- 2. The pitch information of each sound is compared to provide the **best match between the two** signals' pitches.
- To do this match, the signals are **stretched and compressed so that** important sections of each signal match in time.
- 3. The **interpolation of the two** sounds can then be performed which creates the intermediate sounds in the morph.
- 4. The final stage is then to convert the frames back into a normal waveform.
- 5. The information lost has to be **re estimated** for the morphed sound.



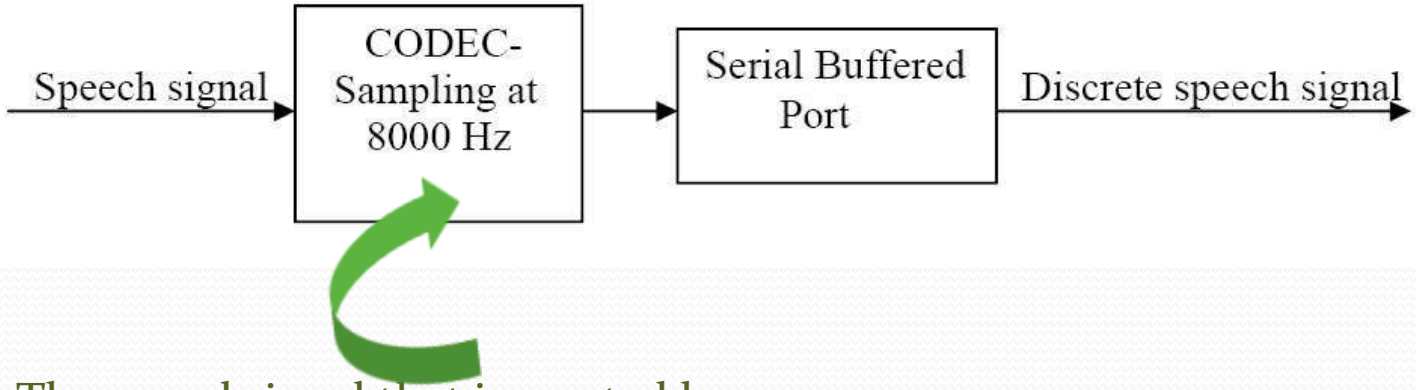
MORPHING PROCESS - Simplified speech morphing algorithm



- The main processes can be categorized as follows:
- **Preprocessing or representation conversion: This involves** processes like signal acquisition in discrete form and windowing.
- **Cepstral analysis or Pitch and Envelope analysis:** This process will extract the pitch and formant information in the speech signal.
- **Morphing which includes** Warping and interpolation.
- **Signal re-estimation.**

Preprocessing:

- Every Speech signal is processed and transformed into a new required representation to affect the morph.
- 1. **Signal Acquisition:**
- The input speech signals are taken using MIC and CODEC. The analog speech signal is converted into the discrete form by the inbuilt CODEC TLC320AD535 present onboard and stored in the processor memory.

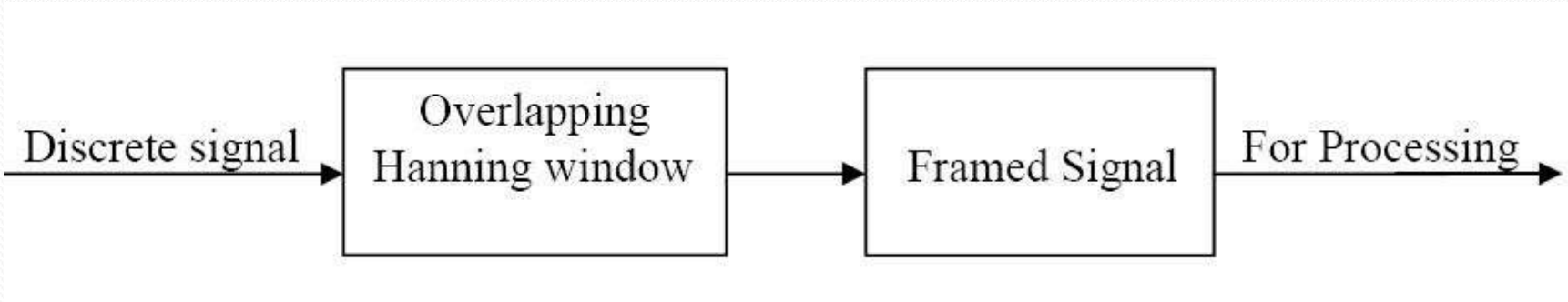


The sound signal that is created by some real-world process has to be ported to the computer by a method which is known as sampling.

Preprocessing:

2. **Windowing:**

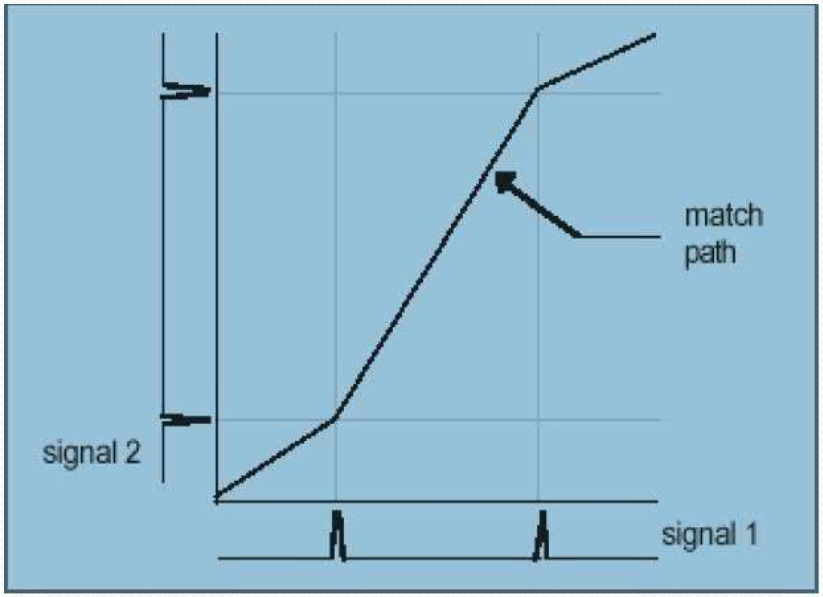
- A DFT can only deal with a finite amount of information. Hence a long signal must be split up into a number of segments called frames.
- Frame should be short enough to make the segment almost stationary and yet long
- enough to resolve consecutive pitch harmonics.(25 to 75 milli seconds).
- The Hamming window is given



- The windowing function splits the signal into time-weighted frames.
- When the frames are put back together, modulation in the signal becomes evident. So a simple method to overcome modulation is to use overlapping windows.
- In overlapping spectra, as one frame fades out, its successor fades in.

Morphing and Warping:

- To smoothly morph the pitch information, the pitch present in each signals needs to be matched and then the amplitude at each frequency cross-faded.
- To perform the pitch matching, a pitch contour for the entire signal is required.
- Pitch peak location in each cepstral pitch slice ---> pitch contour.
- The match path between two signals with differently located features.
- Such a match path is obtained by **Dynamic Time Warping (DTW).**



The amount of movement (or warping) required in order aligning corresponding features in time.

MORPHING STAGE

Why Interpolation is necessary?

- If no interpolation were to occur then this would be equivalent to the warped cross-fade which would still be likely to result in a sound with two pitches.

Properties of the manufactured pitch peak.

- 1. At the beginning of the morph, the pitch peak will take on more characteristics of the signal 1 pitch peak - peak value and peak location - than the signal 2 peak.
- 2.Towards the end of the morph, the peak will bear more resemblance to that of the signal 2 peaks.

Three stages of Morphing:

1. Combination of the envelope information
2. Combination of the pitch information residual – pitch information excluding the pitch peak
3. Combination of the pitch peak information

SIGNAL RE-ESTIMATION

- The conversion of the sound to a representation in which the pitch and spectral envelope can be separated loses some information.
- Therefore, this information has to be re-estimated for the morphed sound.
- This process obtains an acoustic waveform, which can then be stored or listened to.

APPLICATIONS

- It has possibilities in military psychological warfare and subversion, particularly in conjunction with the use of recorded telephone conversations as evidence in courts of law.
- An agency can use voice morphing to provide a fake confession or incriminating evidence appearing to be spoken by a suspect which in reality is fake
- Voice morphing is a powerful battlefield weapon which can be used to provide fake orders to the enemy's troops, appearing to come from their own commanders.
- Cartoons where one person voice of number of characters.

Interesting Fact:

- In 1990, the US department of defense considered using voice morphing to produce a propaganda recording of Iraqi president Saddam Hussein, which could then be distributed throughout the Arab world and Iraq to discredit the Iraqi leader.

CONCLUSION

- The approach separates the sounds into two forms: spectral envelope information and pitch information.
- These can then be independently modified. The morph is generated by splitting each sound into two forms: a pitch representation and an envelope representation.
- Dynamic Time Warping of these contours aligns the sounds with respect to their pitches.
- At each corresponding frame, the pitch and envelope information are separately morphed to produce a final morphed frame.
- These frames are then converted back into a time domain waveform using the signal re-estimation algorithm


Thank
You!