STRAIGHT as a research tool for L2 study:
How to manipulate segmental and supra-segmental features

Hideki Kawahara (Wakayama Univ., Japan)
Reiko Akahane-Yamada (ATR, Japan)

Results in this presentation are preliminary.
Please contact authors for the latest information.
kawahara@sys.wakayama-u.ac.jp
yamada@atr.jp
For STRAIGHT availability, please mail to the authors.
http://www.sys.wakayama-u.ac.jp/~kawahara/index-e.html
Outline

- Background
- Brief introduction to STRAIGHT
- Morphing as a tool for exploratory study
- Extensions for exploitation
- Examples and compatibility
- Further study
Background

- Human auditory system was designed through long history of evolution
- It is optimized to perceive natural sounds
- It is made from highly nonlinear components
- It is necessary to use relevant stimuli to test such highly nonlinear systems
- Relevant stimuli for human speech perception should be natural speech
Background

- Human auditory system was designed through long history of evolution

But, it is difficult to control acoustic parameters precisely using natural speech stimuli.

- It is necessary to use relevant stimuli to test such highly nonlinear systems

- Relevant stimuli for human speech perception should be natural speech
STRAIGHT

• Architecture: channel Vocoder
• Speech sounds are decomposed into:
  • Fundamental frequency
  • Spectrographic representations of:
    • Power
    • periodicity
AST journal presents a review article on STRAIGHT and morphing applications

- Fundamental frequency
- Spectrographic representations of:
  - Power
  - periodicity
STRAIGHT

• Spectrographic representation of power:
  • F0 adaptive complementary set of windows and spline based optimal smoothing
  • Instantaneous frequency based F0 extraction integrated with correlation based extractor

• Spectrographic representation of periodicity:
  • Harmonic analysis based method is about to be replaced by a prediction based method
STRAIGHT

- Spectrographic representation of power:
  - F0 adaptive complementary set of windows and spline based optimal smoothing

Lay language paper:
Voice quality of artistic expression in Noh

- Spectrographic representation of periodicity:
  - Harmonic analysis based method is about to be replaced by a prediction based method
STRAIGHT

• **Spectrographic representation of power:**
  • F0 adaptive complementary set of windows and spline based optimal smoothing
  • Instantaneous frequency based F0 extraction integrated with correlation based extractor

• **Spectrographic representation of periodicity:**
  • Harmonic analysis based method is about to be replaced by a prediction based method
SFT-based spectrogram
periodic in the **time** domain

periodic in the **frequency** domain
Complementary pitch-synchronous Gaussian window removes the repetitive structure in the time domain
Spline-based optimum smoothing reconstructs the underlying smooth time-frequency representation.
Modification

• Morphing two or more examples:
  • Stimulus continuum can be generated without knowing precise correspondence between physical parameters and perceptual attributes such as emotion
  • Physical parametric differences between stimuli can be precisely manipulated
Emotional morphing

iLoveYouAng2 time span 0 999 (ms) 30–Nov–2006 02:52:39

Anger
Emotional morphing

iLoveYouSadness time span 0 998 (ms) 30–Nov–2006 02:59:55

Sadness
Emotional morphing

Pleasure
Demonstration

Move the cursor and click the mouse to control the "emotions".

カーソルを動かしてマウスをクリック。
感情をコントロールしてみてください。

"STRAIGHT"
Auditory morphing by Hideki Kawahara(Wakayama University/ATR) Interface designed by Takashi Yamaguchi
What morphing does?

example: A

example: B

high-dimensional parameter space
Applications to L2 study

- R-L identification and discrimination task
- Stimulus continuum:
  - Klatt synthesizer
  - Morphing Klatt version of R and L
  - Morphing natural examples
Originals: Klatt synthesis
Originals: natural speech
Originals: natural speech
Extension
interim conclusions

- Independent manipulation of physical parameters
- Time varying morphing rate
- Alternative forms of mapping function
  - piecewise linear function
  - formant shape preservation
Originals: natural speech

AM01IP030Mpart10.png time span 0 519 (ms) 30-Nov-2006 09:08:17

partial morphing: t, F0 fixed
Realtime STRAIGHT
Extension
interim conclusions

• Independent manipulation of physical parameters
• Time varying morphing rate
• Alternative forms of mapping function
  • piecewise linear function
  • formant shape preservation