

Production of Various Voice Qualities Using Mechanical Vocal Code Model

Masaaki Honda, Kotaro Fukui, Eiji Shintaku, Yuma Ishikawa,
Nana Sakakibara, Yoshikazu Mukaeda, Atsuo Takanishi
Waseda University

Production of Various Voice Qualities Using Mechanical Vocal Code Model

Masaaki Honda, Kotaro Fukui, Eiji Shintaku,
Yuma Ishikawa, Nana Sakakibara, Yoshikazu
Mukaeda, Atsuo Takanishi
Waseda University

Waseda-Talker Series

WT-4 (1999), WT-5 (2005), WT-6 (2006), WT-7 (2007), WT-7R (2008)

2D Model (1999-2005), 3D Vocal Cord Model (2005-2008), 3D Vocal Tract (Tongue) Model (2006-2008)

Purpose of Talking Robot

- Direct reproduction aero-acoustic phenomenon in speech production process
- Development of humanoid talking robot

Human Speech Production → Modeling → Mechanical Model → Reproduce → Human Speech Production

*Honda, ATR

Speech Production in Anthropomorphic Talking Robot

Palate, Turbulent Flow, Voice, Lip, Teeth, Tongue, Glottal Source, Vocal Cords, Airflow from Lungs

WT-7R

Anthropomorphic Talking Robot WT-7R(Waseda Talker-No.7 Refined) in 2008

Total DOF : 17

Articulators : 13

- Tongue : 7
- Lips : 4
- Jaw : 1
- Velum : 1

Vocal Organs : 4

- Lungs : 1
- Vocal Cords : 3

Topics of MyTalk

- Voice quality generation using mechanical vocal code model Model - Breathy – Creaky
- same control manner as human vocal cord.
- Measurements of the vibration and the aero-acoustic characteristics

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.
OPAL-09, June 26-27, 2009, Vancouver, BC, CA
Copyright remains with the author(s).

Fold Vocal Cord

- Two folds formed by thin cover and hollow inside.
- Made from thermoplastic rubber "Septon"

Sound

Airflow Rate
433~779 [ml/s]

Vocal Cords
Vibration

Decision Experiment of Vocal Cord Shape Methods

Procedure

- Connected to mechanical lung
- Record the sound and take high-speed movie

Variation of Parameters

UC: Thickness of Upper Cover
LC: Thickness of Lower Cover

UC: 0.5, 1.0, 2.0
LC: 0.5, 1.0, 1.5, 2.0

3X4=12 pattern

Decision Experiment of Vocal Cord

UC>LC (UC,LC)=(1.0,0.5), (2.0,0.5), (2.0,1.0)
No Vibration

UC=LC (UC,LC)=(0.5,0.5), (1.0,1.0), (2.0,2.0)
Unstable Vibration (Unbalanced)

UC<LC (UC,LC)=(0.5,1.0), (0.5,1.5), (0.5,2.0)
Stable Vibration

LC>UC is needed (Restitutive force from lower part is important to keep vibration)

Select the parameters as UC=0.5[mm], LC=1.5[mm]

Mechanical Vocal Cord Vibration

1 2 3 4
5 6 7 8

Vibrated in Different Phase

Spectrum of Robot Vocal Cord

Rubber Plate Model

Fold Vocal Cord

Sound of Human Vocal Cords (Ideal)

Insufficient Attenuation

Attenuated in Higher Frequency

Code Tension Control

- Discs attached to the vibrating part directly
- Tensioning vocal cords by rotating the discs

Vocal Cords
Fixed
Tensioning
Pair of Discs

Glottal length is not change

Pitch range is from 130 to 240[Hz]

Glottal Opening Control

- Changing distance of two discs (Same mechanism as human)
- Arm mechanism to control glottal open rate continuously
- Enabled switching voiced/unvoiced sounds by arm

Pair of Discs
Fixed
Glottis Open/Close
Vocal Cords

Control Mechanism of Parameters

Control Units of Vocal Cord Model

— Pitch Control and Glottal Opening —
Pitch Change: 130 → 240 [Hz]

Change glottal width
Change glottal length
Airflow from lungs

Breathy Voice

Physiological Characteristics

No Close

Vocal Cord Vibration of Breathy Voice

Sound Characteristics

Breathy
Modal
Unclear Harmonics on High Frequency Area

Glottal Vibration for Breathy Voice

Movie

Model 2.2 ms / 1 s
Human 2.2 ms / 1 s

※ University of Tokyo

No closure in a vibration cycle

Aero-acoustics of Breathy Voice

Spectrum

Model
Human
Unclear Harmonics

Air Flow

Model
Human
Air Flow > 0

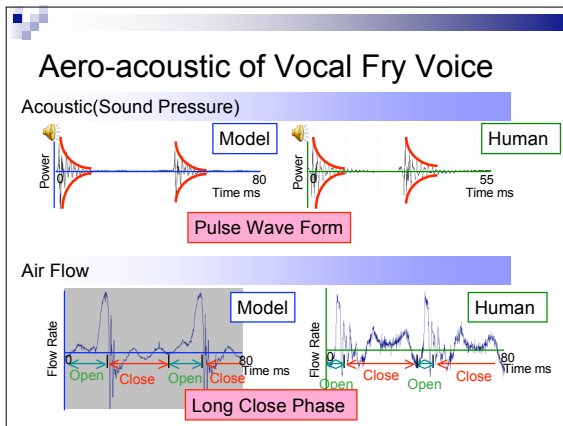
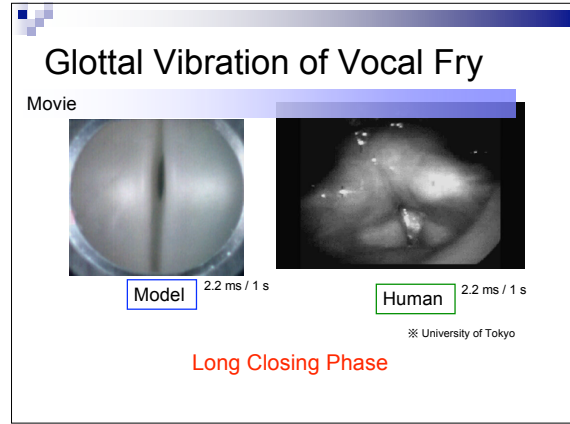
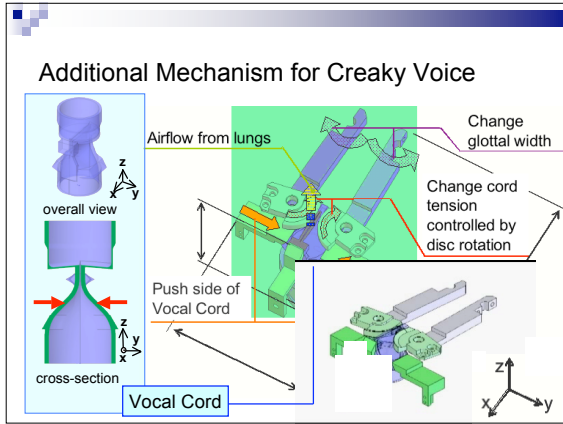
Vocal Fry Voice

Physiological Characteristics

Long Closing Interval in Vibration Cycle

Sound Characteristics

Vocal Fry /a/ 17 ms Long Cycle Pulsive Waveform
Modal /a/ 10 ms



- ### Summary
- Developed mechanical vocal cord model
 - Generated vocal cord vibration of modal, breathy and vocal fry voices
 - Experimental results showed that the model predicts well the motional and the aero-acoustic characteristics of these human voices.