Title	
Author(s)	,
Citation	
Issue Date	1998-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1142
Rights	
Description	Supervisor: , ,

Research on speaker individuality included fundamental frequency fluctuation in continuous vowel utterance

Tomoya Minakawa

School of Information Science,
Japan Advanced Institute of Science and Technology

February ,13 ,1998

1 Introduction

Up to now, there were many researches to extract speaker individuality from vocal tract and vocal cord characteristics. For example, individuality in dynamics of fundamental frequencies contours has been analyzed. However, there is few researches on individuality in fundamemtal frequencies fluctuations. Fluctuations of fundamental frequencies are produced at when continuous vowel, words and sentences. This paper investigates individuality in fundamental frequency fluctuations of continuous vowel utterances.

2 About Laryngograph

EGG(Electro-Glotto-Graph) is a signal recorded movements of vocal cords. Laryngograph is a measurement instrument of EGG and can measure impedance variations between electrodes placed at a thyroid cartilage.

Output of Laryngograph, L_x is related to vocal cord vibrations and has the following characteristics.

- open and closure events are periodical.

- To onset of L_x corresponds to closuring of the vocal cord, and offset of L_x crresponds to opening of vocal cord.

- The onset of L_x is more rapid than the offset, because closing of vocal cords is more quickly than opening.

3 Estimation of fundamental frequencies

Fundamental frequencies are frequencies of vocal cord vibrations and reciprocals periods from closing to next closing are regard to instantaneous fundamental frequencies. Hence, the reciprocals are estimated from extraporated instantaneous fundamental frequencies by extracing the instances of vocal cords closing. Thus, fundamental frequencies at any time can be extraporated by applying some process to correct extraordinary values and to interpolate by linear interpolation between each extraporated instantaneous fundamental frequencies. Figure1 shows an example of the extraporated instantaneous fundamental frequency of /a/ uttered by a male speaker. The top of this figure shows the fundamental frequency by using auto-correlation function, and the bottom shows that by Laryngograph. This figure indicates that Larngograph can extract rapid fluctuations of fundamental frequencies.
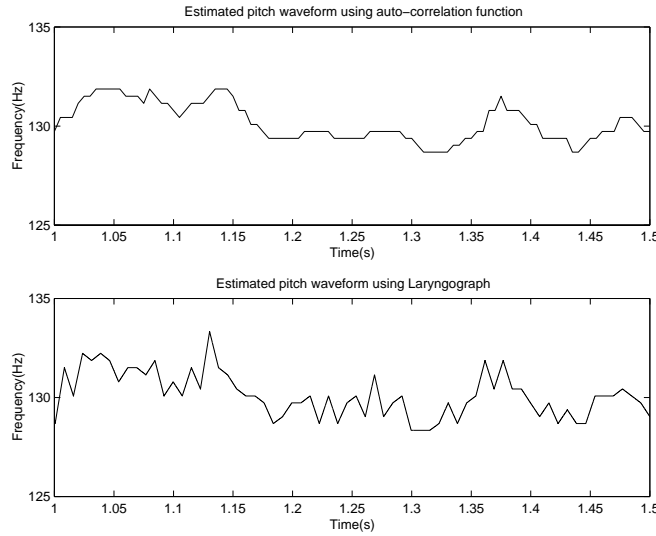

Figure 1: Extraporated value of fundamental frequencies

4 Analysis of rapid fluctuations of fundamental frequencies by histgram

Analyzing data is two second extraporated fundamental frequencies. Total number of these data is forty-five because there are nine speakers and the subject of anlysis is five vowels in japanese.

Then, it is analyzed by histgram that speaker individuality exists in fundamental frequency. From the result of this analysis, it indicates that all of histgram distribution are different each other.Therefore, it is difficult to discriminate a speaker by distribution and magnitude of hostgram.

5 Classification based on fluctuations of fundamental frequencies

There are in fluctuations of estimated fundamental frequencies. It is thought that the whole vertical motion in fluctuations of fundamental frequencies corresponds to low frequency components, closely and repetitive vertical motions corresponds to high frequency components. Thus, it is attempted to classify fundamental frequencies by a certain parameter.

Here after, *rapid variation* indicates the close and repeatedly vertical motions and *slow-variation* indicates the whole vertical motions.

In order to classify fundamental frequencies, it is necessary to extract components of *rapid variation* and *slow variation* from origanal fundamental frequencies. Then, the *rapid variation* and *slow variation* are defined as follows.

- *slow variaton*: Waves consisted of lower components of 10Hz

- *rapid variatoin*: Waves consisted of higher components of 10Hz

The value of 10Hz is decided from power spectrum magnitude in fundamental frequency fluctuations.

The *rapid variation* and *slow variation* are distinguished by the both coefficient of variationl. Fundamental frequencies are classified by the coefficient of variation exceeds a certain threshold or does not exceed the threshold. The threshold of *rapid variation* is 0.0045 and *slow variation* is 0.0075.

Then, fundamental frequencies can be classified into four groups. The four groups are the following.

- First group: The coefficient of *rapid variation* does not exceed the threshold and of *slow variation* also does not exceed the threshold.

- Second group: The coefficient of *rapid variation* exceeds the threshold and of *slow variation* does not exceed the threshold.

- Third group: The coefficient of *rapid variation* does not exceed the threshold and of *slow variation* exceeds the threshold.

- Fourth group: The coefficient of *rapid variation* exceeds the threshold and of *slow variation* also exceeds the threshold.

According to results on classification, the fundamental frequencies in the fourth group do not exist. this is the reason that total number of data on analysis is not enough.

6 Aural evaluation experiments

The first experiment is carried out synthesized sounds are made by using the each group fundamental frequencies and it is estimated that subjects can discriminate the differences between the synthesized sound. The second experiment is performed to investigate when subjects discriminate the differences between fulcutoins in fundamental frequencies of each group, where they percieve the band-width of fluctuations.

The result of the first experiment shows that there are differences which subjects can discriminate between synthesized sounds using fundamental frequencies on each group. Additionally, the results of the second experiment shows the band-width where subjects percieve fluctuaions in fundamental frequencies of each group and that band-width is the following.

- SpeakerA(/a/):First group: subjects are keyed to components in fluctuations of lower 10Hz especially and of 60Hz   100Hz a little

- SpeakerB(/a/):Second group: subjects are keyed to components in fluctuations of 30Hz   60Hz

- SpeakerC(/a/):Third group: subjects are keyed to components in fluctuations of lower 10Hz

7 Conclusion

The results lead to the following three conclusions:

1. According to *rapid variation* and *slow variation*, fundamental frequencies can be classified into four groups.

4

2. Synthesized sound quality between different group is discriminable

3. There are band-width of fundamental frequency fluctuations in each group which strongly influece quality of synthesized sound.

References

[1] A.J.Fourcin Normal and pathological speech:phonetic,acoustic and laryngographic aspects, attached papers with Laryngograph

[2] Koike,Y. Aplplication of some acoustic measures for the evaluation of laryngeal dysfunction: Studia Phonolgica(Kyoto Univ.), 7:pp17–23, 1973.

[3] Hideki Kasuya,Shigeki Ogawa and Yoshinobu Kikuchi, An acoustic analysis of pathlogical voice and its application to the evaluation of laryngeal pathlogy, Speech Communication, 5, pp171–181, 1986.