The Research Reports of the Faculty of Engineering, Kagoshima University, No.53 (2011)

Doctoral Dissertation, Graduate School of Science and Engineering, Kagoshima University, Doctorate No.358

Transformation of Whispering Voice to Pseudo-Real Voice for Unvoiced Telephony and Communication Aid for Voice-Handicapped Persons

Anderson Pierre PASSOS

1. Introduction

For voice handicapped people, an easy to use voicing aid device is wanted. In mobile telephony, so-called non-speaking speech communication is an expected solution for essential privacy as well as for acoustic nuisance prevention. The study introduced here intends to cover both issues, introducing a system where the whispering (non-speaking voice or talk without vocal fold activation) signal is converted to pseudo-real voice signal, which is to be sent to, or heard by, the other party. The study also include validation test with multiple volunteers for its output legibility. Unlike general concept of speech regeneration being inclined to signal recognition or decomposition to text followed by electronic reading (voicing), our system converts it almost directly without recognition or decomposition steps.

2. Background

Prior researchers challenged for text recognition for this sort of signal with some success. However, they didn't try to derive or to regenerate pseudo-real-voice based on this signal. The "recognition-reading" method is generally far from ideal due to its complexity, big computational load and necessary database behind the processing chain.

3. System Design and Method

An important feature in our system is that, instead of injection of the pitch-pulse signal, our system's playback is driven by the pitch-pulse. Basically, on every given pitch pulse the system looks at the current frame, analyzes the level of the input signal and decides among three actions: Playback the autocorrelation of the frame, playback the frame as it is, or just to ignore the frame.

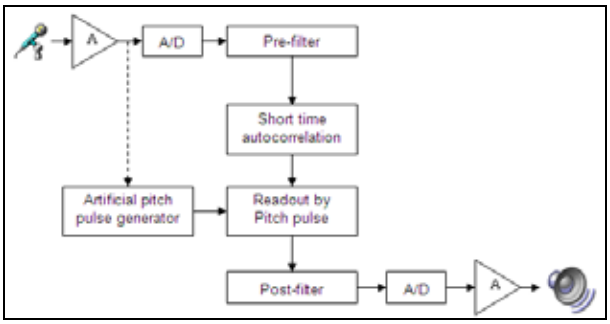

Figure 1 - Proposed schematic of whisper to speech signal conversion system

4. Experiments

To validate the output of the method proposed by this study, two live tests were performed. The first experiment was conducted in our laboratory at Kagoshima University and had as its main objective, to determine the parameters for the filters to be implemented by our proposed system.

The second experiment was done in late 2010, and its main objective was to verify how well the volunteers' perception works when trying to recognize the sound samples. In this second experiment we also tried to evaluate qualitatively the data presented to the volunteers.

5. Conclusions

For the handicapped, we believe our system shows an easy an inexpensive alternative for voice reconstruction. Being able to include the pitch information in the resulting signal was a great achievement and we believe that the method presented here is not only the starting point for other researches targeting whispering voice, but also it and can be improved and used in a near future to convert whispering voice signal into normal speech.

Date of Conferment of the Degree: March 14, 2011
Kagoshima University