On Long Term Prediction for Active Cellular Speech Emission Control

Kazuhiro Kondo
Faculty of Engineering, Yamagata University
4-3-15 Jonan, Yonezawa, Yamagata 992-8510, Japan
Email address: kkondo@yz.yamagata-u.ac.jp

Kiyoshi Nakagawa
Faculty of Engineering, Yamagata University
Yonezawa, Yamagata, Japan

ABSTRACT
We investigated on the possibility of an active cancellation system for unnecessary speech radiation control. The intended application of this system is to cancel cellular speech, and also speech input for recognition-based dictation systems. We have shown previously that speech cancellation is possible with a secondary source placed in proximity to the mouth generating linearly predicted phase inverted speech. However, the prediction must also cover the long delay associated with the acoustic to/from electric conversion, as well as A/D, D/A conversions, and all associated processing, which we found to be around 3 [msec].
By using LPC predicted samples recursively to predict further samples, we found that prediction with SNR of about 6 [dB] is possible, even with this long delay. The prediction coefficient update is suppressed during this recursion. Lowering the sampling frequency in order to lower the number of predicted samples at the cost of reduced bandwidth further enhances prediction accuracy. At a sampling frequency of 8 [kHz], speech emission control of about 7 [dB] for female speech and 4 [dB] for male speech was found to be possible.
We plan to investigate further to improve prediction accuracy using prediction coefficient extrapolation. A prototype system implementation using DSPs is also planned.

KEYWORDS: Cellular Speech, Active Control, LPC
INTRODUCTION

Cellular phones have become quite ubiquitous in most developed countries. This situation has created new types of problems: we are often bombarded by speech from people chatting away on their cell phones from all directions. This speech clearly is not intended for us, only to the people on the receiving end of the call, and thus is useless after the microphones in the handsets pick it up. It also creates privacy concerns. Thus it would be beneficial if we could control the radiation of this speech into the surrounding space, or at least if we can mitigate it to some degree.

On the other hand, speech recognition systems have vastly improved these few years. Speech dictation systems with acceptable accuracy have been released, and there is a growing population of regular users. Many of these users will be using these systems in offices, which potentially will have many other users of similar systems, perhaps in neighboring low-partitioned spaces. The speech from the surrounding users would obviously become noise to the dictation system, most likely bringing down the recognition accuracy considerably. It would also undoubtedly create a very user-unfriendly working environment. Thus, it would also be beneficial for this application if we can control the unnecessary radiation of speech into the surrounding environment.

Active Noise Cancellation [1][2] has received great interest this decade with the advancement of Digital Signal Processors. There have been some successful applications of this technology [3], e.g. control of fan noise radiation through ducts, jet engine radiation control, road noise control in automobiles [4] to name just a few. Previously, we investigated on the possibility of applying similar techniques to the control of speech [5].

We have shown that speech cancellation is possible with a secondary source placed in proximity to the mouth generating linearly predicted phase inverted speech. However, the prediction must also cover the long delay associated with the acoustic to/from electric conversion, as well as A/D, D/A conversions, and all associated processing, which we found to be around 3 [msec]. We investigated on the possibility of using linear prediction recursively to predict speech covering this long delay. Use of a lower sampling rate to limit the number of samples which need to be predicted ahead at the cost of limiting the effective bandwidth has also been investigated.

In the next section, we will describe the previously proposed speech cancellation scheme with linear prediction and proximity secondary source, then propose a long term prediction scheme, followed by computer simulation results, and finally conclusions.