SPEECH ENHANCEMENT IN VEHICULAR ENVIRONMENT

By

Abdul Wahab bin Abdul Rahman

School of Applied Science,

Nanyang Technological University,

Nanyang Avenue,

Singapore 639798

First Year Report for the degree of PhD.

In

Applied Science

Abstract

The increasing demand for digital cellular telephony and other new services including multimedia communications prompted numerous studies on implementing algorithms for low rate speech coding below 4.8 kbits/s using available DSP processors on the market. In addition there are needs to enhance the speech quality subject to both degradations due to road, engine and wind noise and the echoes present in the near-end speaker side--sources effecting the car phone input. All of these tasks must be achieved with a single DSP chip in order for the system to be both cost-effective and power efficient and thus widely accepted. This dissertation research proposes:

1) To study both analytically and experimentally the above degradation.

2) To develop sound signal processing algorithms to combat these imperfections.

3) To address architectures for real-time implementation.

4) To implement them on a DSP platform using state-of-the-art devices and reconfigurable systems.

Firstly, the impediment to the speech quality in a vehicular chamber is the echo generated by the leakage of the far-end speaker and is mixed with the speech from the near-end speaker and transmitted as a composite signal. The first task of the proposed speech enhancer is to adaptively cancel these echoes. This necessitates the inclusion of near-end speaker activity detection.

Secondly, in the vehicular hands-free cellular communication framework, it has been observed that the degradation in the intelligibility and the general quality of the cellular speech due to the engine, road, and wind noise components is equally disturbing as the vehicular echoes. Hence, the second important task of the enhancer is to combat these imperfections in the cellular speech.

This last point, in particular, makes a form of beam forming based on a microphone array followed by an adaptive filtering process as a conceptually sound candidate for our speech enhancer. The most simple form of beamforming is called the delay and sum beamforming, which compensates the delay of the target signal and sums the signals in the beam so that the target signals have the same phase while the interfering signals exhibit different phase. The delay and sum beamforming technique will be used to first follow the genuine speaker. Then it will adaptively cancel the noises coming from the interfering speakers, the engine, the wind, especially critical when the windows of the vehicle are down, and the road noise coming from other vehicles and the road itself. There are a few studies in the literature on this for speech recognition in a hands-free telephone set-up [1-7]. These imperfections would be look into and a unified approach develops to combat all of these different types of degradations.

In addition the research will develop architectures for real-time implementation on a single chip DSP as part of the next generation digital cellular phones operating at 4.8 KB/s or less. At the time of this proposal, the preferred DSP platform is the TMS320C4x family from the Texas Instruments, Inc. However, the algorithms develop will be reconfigurable.

1. Introduction

It has been in the public debate for some time that vehicles in the future would need to detect, process, and communicate significantly more information. They will be between the vehicle and the driver, among the people in the vehicle and between the vehicle and the outside world, including other vehicles, road itself, and the Advanced Traffic Management and Information Systems (ATMIS).

The driver and other passengers may want to communicate with the outside world verbally, or to have a conference call. These activities have been traditionally handled by car phones or short-wave radios, where the underlying signal is the band-limited voice grade waveform. These signals are transmitted over a communication channel, which is extremely corrupted by echoes both in the transmission link and inside the chamber of the vehicle. In addition, there are also natural and man-made noise from numerous sources, and interfering signals from other channels, passengers, and audio information subsystem present. It is commonly accepted that the next generation car telephones will be totally digital cellular and the volume of applications will increase. However, a number of ills will not go away and a speech processing system will be required to tackle them. Some of the tasks for the research will be the noise suppression, echo cancellation, source localisation, speaker identification, speech coding, compression and transmission by digital means.

This report will discuss the spectral dissection of various degradations in vehicular environment followed by proposal for a cost-effective model for the speech enhancer system and the introduction to a re-configurable digital signal-processing concept. The detail discussion on the propose speech enhancer system covers:

1. Identifying the various man-made noises and categorising them into different optimal sub-bands so that noise cancellation and suppression can be accurately achieved.

2. Handling of echoes from the near-end and the far-end speakers to adaptively cancel them.

3. Genuine speaker identification and employing the beam former to follow the genuine speaker and then to adaptively cancel the noise coming from the interfering speaker, the engine, the wind and the road noise.

4. Since these noises can be categorise into optimal sub-bands multi-rate signal processing can be employed to improve the computing performance of the speech enhancer.

Finally, the conclusion, summary and schedule of the proposed speech enhancer system architecture for the PhD research.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.