A Microphone Array System for Speech Source Localization, Denoising, and Dereverberation

A thesis presented

by

Scott Matthew Griebel

to

The Division of Engineering and Applied Sciences
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University
Cambridge, Massachusetts

April 2002

Thesis advisor
Michael S. Brandstein

Author
Scott M. Griebel

A Microphone Array System for Speech Source Localization, Denoising, and Dereverberation

Abstract

There is a great deal of potential for advancement in distant-talker speech acquisition research, and a wealth of current and future technology depends upon these advances. The goal of this work is to allow users the opportunity to roam unfettered in diverse environments while still providing a high quality speech signal and a robustness to background noise and reverberation effects. In this thesis, a microphone array speech enhancement system is presented which has three main components: source localization, background noise reduction, and dereverberation. The localization algorithm is effective in the presence of both background noise and reverberations and simultaneously produces relative time delay estimates and a source location estimate. It provides a procedure applicable to all time delay estimators which either maximize or minimize an appropriate objective function, improving the estimators' robustness to environmental degradations. The denoising algorithm is a multi-microphone extension to the Minimum Statistics denoising technique [Martin (2001)]. This algorithm also has an additional and optional SNR-dependent beamforming stage that is shown to be very useful in certain environments. The final component is a multi-channel dereverberation algorithm which models the speech source and room reverberations independently. A weighting function is estimated and applied in the Wavelet Transform domain to de-emphasize portions which are less coherent across microphone signals, an indication of reverberation effects. Results for the various components are provided as proof of the effectiveness of the proposed multi-microphone speech enhancement system.

Acknowledgements

I would like to express my gratitude for the advising and friendship of Prof. Michael Brandstein whose guidance in matters both personal and professional was always helpful. I also thank my thesis committee members Prof. Roger Brockett and Prof. Irvin Schick for their input over the years.

I would like to acknowledge my partners in crime at the Harvard Multi-Media Lab: Daniel Freedman (the smartest man I know), Ce Wang (who would die for his dream), and Daniel Kirsanov (who was always with basketball, but never played basketball). Those outside the lab that were equally appreciated include Soundouss Bouhia, Anthony Volpe, Winston Yu, and Brendan Zinn. They all sparked numerous pointless conversations that were often needed to break up a day's work.

I also thank Homer J. Simpson for his weekly words of wisdom. His knowledge regarding nuclear engineering proved helpful even in my unrelated field of speech signal processing.

I would also like to express my appreciation for my parents for their caring and support throughout my graduate career.

Finally, I thank my wonderful wife Kristen (the smartest woman I know). She waited patiently for me to begin providing some sort of real salary, but always understood that my quest for "higher knowledge" was, to me at least, well worth the trip.

If there's a way to do it better... find it.

Thomas A. Edison

Every time I learn something new, it pushes out something old.

Homer J. Simpson

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.